

A Balance Condition for Stochastic Numerical Models with Application to the El Niño—Southern Oscillation

Author(s): Cécile Penland and Ludmila Matrosova

Source: *Journal of Climate*, Vol. 7, No. 9 (September 1994), pp. 1352–1372

Published by: American Meteorological Society

Stable URL: <https://www.jstor.org/stable/26198585>

Accessed: 26-04-2023 09:22 UTC

REFERENCES

Linked references are available on JSTOR for this article:

https://www.jstor.org/stable/26198585?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<https://about.jstor.org/terms>



JSTOR

American Meteorological Society is collaborating with JSTOR to digitize, preserve and extend access to *Journal of Climate*

A Balance Condition for Stochastic Numerical Models with Application to the El Niño–Southern Oscillation

CÉCILE PENLAND AND LUDMILA MATROSOVA

Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado

(Manuscript received 30 April 1993, in final form 16 December 1993)

ABSTRACT

Stochastic forcing due to unresolved processes adds energy to a measurable system. Although this energy is added randomly in time, conservation laws still apply. A balance condition for stochastically driven systems is discussed. This “fluctuation–dissipation relation” may be used either to deduce the geographical properties of the stochastic forcing from data given a model for the evolution of the macroscopic variables or to diagnose energy conservation in a stochastic numerical model.

The balance condition in its first role was applied to sea surface temperatures (SSTs) in the Indo-Pacific basin. A low-dimensional empirical dynamical model of SSTs was generated in such a way that observed statistical properties of the field are preserved. Experiments varying the stochastic forcing in this model indicated how the geographical characteristics of the forcing affect the distribution of variance among the various normal modes, thereby determining the dominant timescales of the SST field. These results suggest that the south Indian Ocean and the equatorial Pacific close to the date line are important to the amplitude and timing of the warm phase of El Niño–Southern Oscillation.

Fourier spectra obtained from output of the stochastically forced linear model were found to agree with those obtained from COADS data when time series of equal length were compared. A discussion of how spectra from a multivariate linear system can be confused with those of a nonlinear system is presented.

1. Introduction

It has become increasingly common in meteorological studies to treat multiscale processes in numerical models by approximating the fast variables as stochastic processes. Several techniques for treating the stochasticity have been introduced. Lau (1985) simply added a stochastic term to the signal in his model of the Southern Oscillation index, although his justification for a stochastic model involved forcings by fast transients. Such a forcing generally precludes separation of the signal into contributions due solely to deterministic dynamics and those due solely to stochastic forcing. Salby and García (1987) and García and Salby (1987) considered excitation of disturbances by tropical stochastic heating. Their linearized version of the primitive equations allowed a purely spectral treatment, thus avoiding many of the pitfalls of trying to generate numerically a physically meaningful stochastic process in the time domain. A similar approach has been used by Kim and North (1993), Kim and North (1992), and Leung and North (1991) in their treatments of stochastically forced energy balance models. Still, the

treatment of nonlinearities is obviously important. Barnett et al. (1993) have driven the Hamburg Ocean General Circulation Model with statistical forcing corresponding to EOFs of observed wind stress. In this latter study, ocean–atmosphere coupling is treated by “enslaving” the wind stress EOFs to SST patterns by means of linear regression. This model is, strictly speaking, deterministic; the stochastic element is averaged out in formulating the dynamical interactions. Any effects due entirely to the stochastic nature of the forcing are thereby ruled out. Other studies (e.g., Fleming 1993; Spall 1993; Syu et al. 1993; Cane et al. 1992; Mikolajewicz and Maier-Reimer 1990; Lin and Koshyn 1987) have attempted to include subscale effects in numerical models by treating the unresolved processes as stochastic forcing in the time domain.

One problem in the implementation of stochastic models is the need to check the dynamical consistency of the model; one must ensure that the model sees dynamically what the modelers think they are imposing. For example, in a coupled ocean–atmosphere model where the atmosphere is parameterized as a random process, one must verify that the energy transferred stochastically from the atmosphere is equivalent to that absorbed by the ocean. As we will see, the approximation of a physical, quickly varying system as a numerical, stochastic process is rather delicate. It is

Corresponding author address: Dr. Cécile Penland, CIRES, University of Colorado, Campus Box 449, Boulder, CO 80309-0449.
E-mail address: mcp@noaacrd.colorado.edu

the purpose of this article to discuss a balance equation, a generalized “fluctuation–dissipation relation” (GFDR), relating the statistical properties of random forcing and the deterministic dynamics. This balance condition serves a dual purpose. If the deterministic dynamics are known but the forcing by unresolved processes is not, the balance can often be used to diagnose the macroscopic, dynamically consistent properties of that forcing. On the other hand, if the stochastic forcing is imposed on a numerical model, the balance can be used to verify conservation of energy in a mean sense. The usefulness of this balance in diagnosing the subscale forcing is illustrated with two runs of an empirical–dynamical model (EDM) of the El Niño–Southern Oscillation (ENSO). The best linear approximation of the deterministic dynamics, as well as the geographical properties of the random forcing, is obtained from sea surface temperature data, and a stochastic control model using these properties is generated and examined. In a second experiment, the effect of varying the geographical properties of the forcing on the empirically obtained normal modes of the system is observed. That details of deterministic forcing are important to details of response is obvious; however, it seems less well known that the random nature of stochastic forcing requires the same care to ensure dynamical balances.

The use of fluctuation–dissipation theory in climate studies is not new. Leith (1971, 1975, 1978) suggested its use in providing estimates of climate response to impulsive external forcing. Bell (1980) investigated the applicability of an FDR to a low-dimensional spectral model of the barotropic vorticity equation. North et al. (1993) showed that the transient behavior of two general circulation models approximated that prescribed by an FDR. The FDR used in these studies was developed specifically in the context of isotropic turbulence and, recognizing the danger of employing a term to describe a general relation when that term has been widely used to describe a special case, we modify it with the word “generalized.”

Section 2 of this article discusses an evolution equation for second moments that must be satisfied by stochastically driven models. Statistical stationarity implies a balance between random forcing and model response, but even when stationarity does not attain, the relation can be used to perform consistency checks. Section 3 presents a simple model of ENSO, with general remarks on the generation of numerical stochastic models. The results of our stochastic model are examined in section 4, and a general review and discussion is presented in section 5. The procedure for obtaining from data the parameters of a linear stochastic process is reviewed in appendix A, and the relation between a continuous linear model and the equivalent discrete model (e.g., Cane et al. 1992) is investigated in appendix B.

2. Fluctuation–dissipation relations

The existence of geophysical processes that coexist and affect each other’s development but have widely different autocorrelation times leads us naturally to the approximation of “fast” (autocorrelation times shorter than a sampling interval) processes as random white noise. The literature discussing how this is to be done in a dynamically consistent manner is fairly recent and can be said to have begun with Wong and Zakai (1965). The following discussion is based on books by Gardiner (1985, especially section 5.3.1), Riskin (1989), and Arnold (1974; especially section 10.3). A rigorous justification of the method is found in Papаниcolaou and Kohler (1974). A clear, heuristic description of an important special case is found in Hasselmann (1976), although the reader should beware the unfortunate confusion between high frequency and low correlation time in that article.

In most cases, the quickly varying fluctuations are the combined effect of weakly coupled factors, allowing one to assume via the central limit theorem that the white noise fluctuations are Gaussian. A general expression for a model of a multivariate, macroscopic (autocorrelation times on the scale of the sampling interval) process \mathbf{X} of dimension d driven by m white noise processes is

$$d\mathbf{X}_i = F_i(\mathbf{X}, t)dt + \sum_{j=1}^m S_{ij}(\mathbf{X}, t)dW_j, \quad (1)$$

where \mathbf{W} is an m -dimensional Wiener process and where $1 \leq i \leq d$. The first term on the right-hand side of Eq. (1) is the deterministic model (e.g., an ocean model), while the second term represents the interaction of the stochastic fluctuations (e.g., the quickly varying component of the wind stress forcing) with the macroscopic variables. The time-dependent probability density $p(\mathbf{X}, t)$ of the system described in Eq. (1) is given by

$$\begin{aligned} \frac{\partial p(\mathbf{X}, t)}{\partial t} = & - \sum_{i=1}^d \frac{\partial}{\partial X_i} \left\{ F_i(\mathbf{X}, t)p(\mathbf{X}, t) \right. \\ & + \frac{1}{2} \sum_{k=1}^d \sum_{j=1}^m \frac{\partial S_{ij}(\mathbf{X}, t)}{\partial X_k} S_{kj}(\mathbf{X}, t)p(\mathbf{X}, t) \Big\} \\ & + \frac{1}{2} \sum_{i=1}^d \sum_{k=1}^d \sum_{j=1}^m \frac{\partial}{\partial X_i \partial X_k} \\ & \times [S_{ij}(\mathbf{X}, t)S_{kj}(\mathbf{X}, t)p(\mathbf{X}, t)]. \end{aligned} \quad (2)$$

This equation describes conservation of probability flux. The first term in curly brackets, which we will call the “deterministic drift,” describes the dynamics of the deterministic system. The second term in curly brackets, called a “noise-induced drift” in the literature, results from approximating a real, continuous system

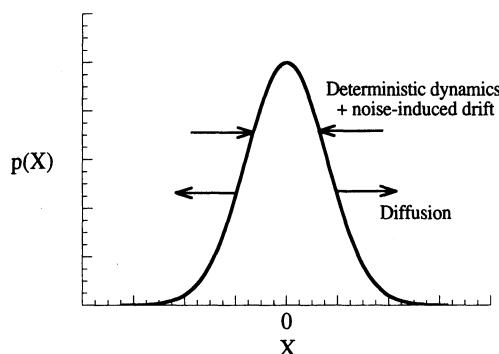


FIG. 1. Schematic diagram showing the roles of the deterministic drift, the noise-induced drift, and the diffusion in maintaining the shape of the probability distribution. This example is appropriate to a stationary system with a single fixed point.

with a small but finite correlation time as white noise (a so-called Stratonovich system) and does not occur in some hydrological applications [“Itô” systems; Rodriguez-Iturbe et al. (1991)], where *discrete*, uncorrelated fluctuations are approximated as *continuous* white noise [see Horsthemke and Lefever (1984, section 5.4.2): “Itô or Stratonovich: A Guide for the Perplexed Modeler”]. In a continuous system where waves are expected, the noise-induced drift must be included. Together, the deterministic drift and noise-induced drift determine the dynamics of the average field, including its stability properties. The noise-induced drift can cause a measurable average effect. For example, waves traveling through a stochastic medium travel with average phase and group velocities different from those they would have in a deterministic medium; adding a zero-mean stochastic component to the propagation medium alters the waves’ *average* dispersion properties (Personick 1971; Marcuse 1972; Penland 1985). The third term on the right-hand side of Eq. (2) is called the “diffusion.” This is the term that gives the probability distribution its width, keeping the distribution of a perfectly specified initial condition from remaining a δ function, which would travel through phase space as dictated by the deterministic dynamics.

It is useful to compare Eq. (2) with the Liouville equation [e.g., Epstein (1969), his Eq. (4)] used in stochastic dynamic prediction (SDP). Equation (2) collapses to a Liouville equation when $\mathbf{S}(\mathbf{X}, t)$ is set to zero. In SDP, the initial conditions are stochastic but the dynamics are deterministic; the deterministic drift $\mathbf{F}(\mathbf{X}, t)$ causes the distribution of a cloud of initial states to evolve according to the Liouville equation, thus exhibiting the sensitivity of the deterministic dynamics to initial conditions. If, in addition to (or instead of) uncertainty about initial conditions, the distribution is affected by unresolved forcings, the evolution equation for the distribution is modified by the noise-induced drift and the diffusion as given in Eq. (2). A schematic of the interaction between the drift

and diffusion terms is presented in Fig. 1 for a stationary system with one stable fixed point. The diffusion constantly and randomly perturbs the system and the drift tries to drive these fluctuations toward the stationary state. In contrast, a Wiener process has a variance that increases linearly with time; it has no drift to limit the growth of its variance, and the width of its distribution grows without bound.

By multiplying Eq. (2) by $X_p X_q$ and integrating over the domain of the system, a relation between the covariance structure of the system and its dynamics is obtained. That is,

$$\begin{aligned} \frac{\partial \langle X_p X_q \rangle}{\partial t} &= \langle F_p(\mathbf{X}, t) X_q \rangle + \langle X_p F_q(\mathbf{X}, t) \rangle \\ &+ \frac{1}{2} \sum_{k=1}^d \sum_{j=1}^m \left\{ \left\langle \frac{\partial S_{pj}(\mathbf{X}, t)}{\partial X_k} S_{kj}(\mathbf{X}, t) X_q \right\rangle \right. \\ &\quad \left. + \left\langle X_p \frac{\partial S_{qj}(\mathbf{X}, t)}{\partial X_k} S_{kj}(\mathbf{X}, t) \right\rangle \right\} \\ &+ \sum_{j=1}^m \langle S_{pj}(\mathbf{X}, t) S_{qj}(\mathbf{X}, t) \rangle, \quad (3) \end{aligned}$$

where angle brackets indicate ensemble averages. Equation (3) gives the full nonlinear, nonstationary relation between the statistics and the dynamics of a physical system driven by white stochastic forcing. If the statistics are stationary, then the left-hand side of Eq. (3) is zero. Stationary statistics is a consequence of equilibrium in thermodynamic systems. If the driving noise does not depend on the state of the system, then there is no noise-induced drift; that is, the double summation on the right-hand side of Eq. (3) does not contribute. The last term on the right-hand side is the covariance structure of the stochastic forcing. As discussed above, if the macroscopic system is fluctuating around a stable attractor, then any random kick away from this attractor will be followed by a decay toward it as dictated by the deterministic dynamics combined with noise-induced drift, which masks itself as deterministic dynamics. That is, the dissipation of the random forcing is contained in the drift terms. Rapid decays require large kicks to maintain a stable covariance structure. Because the relation between this dissipation and the driving fluctuations is explicit in Eq. (3), it is known as the generalized fluctuation–dissipation relation of the system when the statistics are stationary.

Although the GFDR has been introduced in the context of exterior stochastic forcing, relations such as Eq. (3) must also be obeyed by internal random fluxes (Landau and Lifschitz 1954; the entire turbulence literature; García and Penland 1991). In that case, however, the noise matrix $\langle \mathbf{S} \mathbf{S}^T \rangle$ and noise-induced drift contain spatial derivatives corresponding to the nature of the internal fluctuations.

For modelers, Eq. (3) represents a relation needed to verify conservation laws in a stochastically driven model, or in a model where unresolved transients are treated as random fluxes. The balance in Eq. (3) may be (and usually is) complicated if the stochastic terms depend on the macroscopic variables, such as when wind stresses depend on sea surface temperature. Also, nonstationary dynamics may necessitate averaging multiple runs of a GCM to evaluate individual terms in Eq. (3). Nevertheless, even if Eq. (3) cannot be solved directly, the model output $\mathbf{X}(t)$ and the imposed deterministic dynamics $\mathbf{F}(\mathbf{X}, t)$ can in principle be used to calculate the individual terms to perform a consistency check on the statistics of the random forcing $\mathbf{S}(\mathbf{X}, t)d\mathbf{W}$.

As an example of a system where the GFDR can be solved, consider a linear model driven by additive, stationary white noise ξ :

$$\frac{\partial X_i}{\partial t} = \sum_{j=1}^d B_{ij}X_j + \xi_i, \quad (4)$$

where \mathbf{B} is a constant matrix. The white noise in Eq. (4) is related to the random forcing in Eq. (1) as $\mathbf{S}d\mathbf{W} = \xi dt$, where, in this case, \mathbf{S} depends on neither \mathbf{X} nor t . Equation (4) is more general than the stochastic model considered by Mori (1965) and Leith (1971) in the sense that the change in X_i depends upon the other components of the system; it is more restricted in the sense that \mathbf{B} is constant in time. Here, the dimensionality of the macroscopic system is equal to the dimensionality of the Wiener process, and the covariance structure of the forcing is $\langle \xi_i(t + \tau)\xi_k(t) \rangle = Q_{ik}\delta(\tau)$. For this case, the statistics are stationary and the GFDR (3) reduces to

$$0 = \sum_{j=1}^d B_{ij}\langle X_j X_k \rangle + \sum_{j=1}^d \langle X_i X_j \rangle B_{kj} + Q_{ik}. \quad (5)$$

In this linear case, the eigenstructure of the deterministic matrix \mathbf{B} , and hence \mathbf{B} itself, can be obtained directly from data (Penland 1989; García and Penland 1991), after which the spatial characteristics of the random fluctuations can be calculated from Eq. (5) (linear inverse modeling: appendix A). The eigenvectors of \mathbf{B} are called principal oscillation patterns (POPs; Hasselmann 1988) or empirical normal modes (ENMs). The term POP was originally intended to be used when a single modal pair with its characteristic decay time and frequency dominates the dynamics (von Storch 1990, personal communication). However, recent papers (e.g., Xu 1993) have essentially used the two terms interchangeably. The term "empirical normal mode" is general enough that it might be appropriate for several types of analysis, and we have learned that Brunet (1994) has independently employed the term for modes obtained in a completely

different manner. However, there should be no confusion in terminology when the research is taken in context.

Before proceeding, it is interesting to relate the continuous linear model [Eq. (4)] to the discrete model used in some studies (e.g., Von Storch et al. 1988). Briefly, the procedure used in those studies for obtaining the characteristic timescales of linear systems assumes that their model is a discretization of Eq. (4). Therefore, the random noise term in the discrete model involves a convolution of the white stochastic forcing with the deterministic dynamics; the explicit relation is derived in appendix B.

3. An empirical-dynamical model of ENSO

a. Determination of the parameters

We consider sea surface temperatures (SSTs) from the Comprehensive Ocean-Atmosphere Data Set (COADS) (Slutz et al. 1985) in the tropical Indo-Pacific basin 30°N - 30°S , 30°E - 70°W , which were consolidated onto a $4^\circ \times 10^\circ$ grid. Monthly data between January 1950 and December 1990 were available. Anomalies were calculated by centering the multivariate time series to their lengths (making them zero mean) with the annual cycle removed, and then were subjected to a three-month running mean. This light filtering eliminated some noise in the data, due to data gaps, for example, while minimizing the distortion of statistically determined parameters inherent in low-pass filtering. In addition, filtering by a coarse-graining technique like this one does not directly contradict the assumption of a linear stochastic model to which the data are to be fit. The anomalies were projected onto a subset of empirical orthogonal functions (EOFs) as in Penland and Magorian (1993; PM93 hereafter) and the parameters of the linear model were obtained through evaluation of the empirical normal modes as explained in appendix A. Although PM93 found that 10 EOFs containing 61% of the variance were sufficient to describe the evolution of SSTs in the Niño 3 region (90° - 150°W , 6°N - 6°S), we found that more degrees of freedom were necessary to accurately predict the different kinds of ENSO patterns as exemplified by the 1972-73 event, when warm SST anomalies propagated east to west, and the 1982-83 event, when eastward propagating warm anomalies joined equatorward-traveling anomalies that originated off Baja California. The data were therefore projected onto $d = 20$ EOFs containing 73% of the variance. All calculations were performed in the space spanned by the 20 EOFs and then transformed into geographical space for interpretation.

There would be little meaning to conclusions based on this linear model if it could not make accurate predictions. Predictions are made by applying a "Green function" matrix to a given initial condition (appendix

A and PM93). PM93 found that a similar prediction model of Niño 3 SST anomalies δT_3 involving 10 EOFs had an average prediction error of about 0.5°C (about 70% of the rms value) at a lead time of nine months. However, examination of individual cases indicates that this average error does not increase uniformly with lead time; rather, the predictions are either very good, such as those initialized during the onset phase of a warm event, or very bad, such as those initialized during the persistent phase of a warm event. These conclusions were found to hold when the number of retained EOFs was increased from 10 to 20. The parameters of the propagator, or Green function $\mathbf{G}(\tau) = \exp(\mathbf{B}\tau)$, were based on statistics calculated with a lag $\tau_0 = 7$ months on the training period February 1950–January 1985, and verifications were performed using subsequent months in the dataset as initial conditions. The average *e*-folding time (see below) of all the modes is about 10 months, so that predictions using these modes would not be expected to be very good after nine or ten months even if the linear Markov model were perfect.

Five predictions of δT_3 and subsequent verifications are shown in Fig. 2. The prediction error at $\tau = 0$ is nonzero because the predictor/predictand field has been projected onto 20 EOFs, but we have included all EOFs in the verification field (i.e., Fig. 2 includes the truncation error). The predictions are excellent during periods of rapidly rising SST anomalies, even up to a lead time of about 18 months (Figs. 2a,d). The linear model has difficulty describing the duration of a warm event (Fig. 2b). Also, although the decline of a warm event is fairly well described up to a lead time of a few months (Fig. 2c), the inferior prediction obtained when the initial SST anomaly in Niño 3 is above one standard deviation is again clear. The model was successful in predicting not only the occurrence but also the duration of the 1988–89 cold event when the initial condition was chosen six months later, in April 1988, when the SSTs had fallen to nearly climatological conditions (Fig. 2d). Finally, the model performed well when required to predict the period of near-climatology SSTs following that cold event (Fig. 2e). Note that the quality of the predictions is more dependent on the phase of the ENSO cycle than it is on the annual cycle; in particular, accurate predictions through the spring season are common. The average prediction error in Niño 3 SST anomalies during the verification period as a function of lead time is given in Fig. 2f, both for predictions based on the model and predictions based on a persistence forecast. The prediction error for the model approaches the rms SST anomaly of about 0.7°C , as it must, while the persistence error is much larger. These statistics are based on 46 to 70 verifications, depending on the lead time, which varied between 1 and 24 months. The long training period needed to evaluate the long periods of some empirical

normal modes left little of the available data to use for verifications, and general statements made on the basis of a few case studies should be taken with some caution. Still, it is impractical to wait for several more ENSO cycles before attempting progress, and besides, predictions based on initial conditions during the training period show the same general behavior. Based on these results and those of PM93, we conclude that our stochastically forced linear model exhibits most characteristics of tropical Indo-Pacific SST anomalies while underestimating the duration of the warmest phase of the ENSO cycle.

The deterministic matrix of grid-resolvable dynamics, \mathbf{B} , is related to the empirical normal modes \mathbf{u}_α and their adjoints \mathbf{v}_α as follows:

$$\mathbf{B} = \sum_{\alpha=1}^d \mathbf{u}_\alpha \beta_\alpha \mathbf{v}_\alpha^T, \quad (6)$$

where β_α is the α th eigenvalue of \mathbf{B} . Each eigenvalue is related to a period $2\pi/(\text{Im}\beta_\alpha)$ and an “*e*-folding,” or “decay,” time $-1/(\text{Re}\beta_\alpha)$ of the empirical normal mode. This last quantity is related to how important random forcing is to the maintenance of that mode, rather than to that mode’s importance to the entire field. The timescales determined from our analysis of the training period are given in Table 1. Ten complex modes (mode pairs 4/5, 9/10, 11/12, 14/15, and 19/20) have ENSO-like patterns and are correlated with each other at more than 50%. Other complex modes (mode pairs 2/3, 6/7, 17/18) have some characteristics of a classic ENSO pattern but correlate with the others at less than 50%. Modes 4 and 14 along with their adjoints are shown in Figs. 3 and 4.

Given the feedback matrix \mathbf{B} and the covariance structure of the data, the noise matrix \mathbf{Q} was evaluated using Eq. (5). The eigenvectors of \mathbf{Q} are EOFs of the noise covariance matrix, and their corresponding eigenvalues should have the general properties of EOF eigenvalues; that is, they should be positive and indicative of the amount of variance in the corresponding eigenvector. An eigenvector analysis of \mathbf{Q} yielded 18 positive eigenvalues and two spuriously negative eigenvalues, probably due in part to the short length of our training period, but also suggesting the presence of nonlinear dynamics important to the duration of warm events. If we consider the total noise variance to be contained in the first 18 eigenvectors, or “forcing functions” $\{\psi_\alpha\}$, of \mathbf{Q} , then more than two-thirds of that variance is contained in the first six eigenvectors (Fig. 5).

Examining Fig. 5, we see that although the two leading forcing functions are dominated by loadings in the eastern Pacific, most of the important forcing functions have large loadings in the southern Indian Ocean. Large loadings in the southern Indian Ocean are also characteristic of the adjoints of all ten ENSO modes.

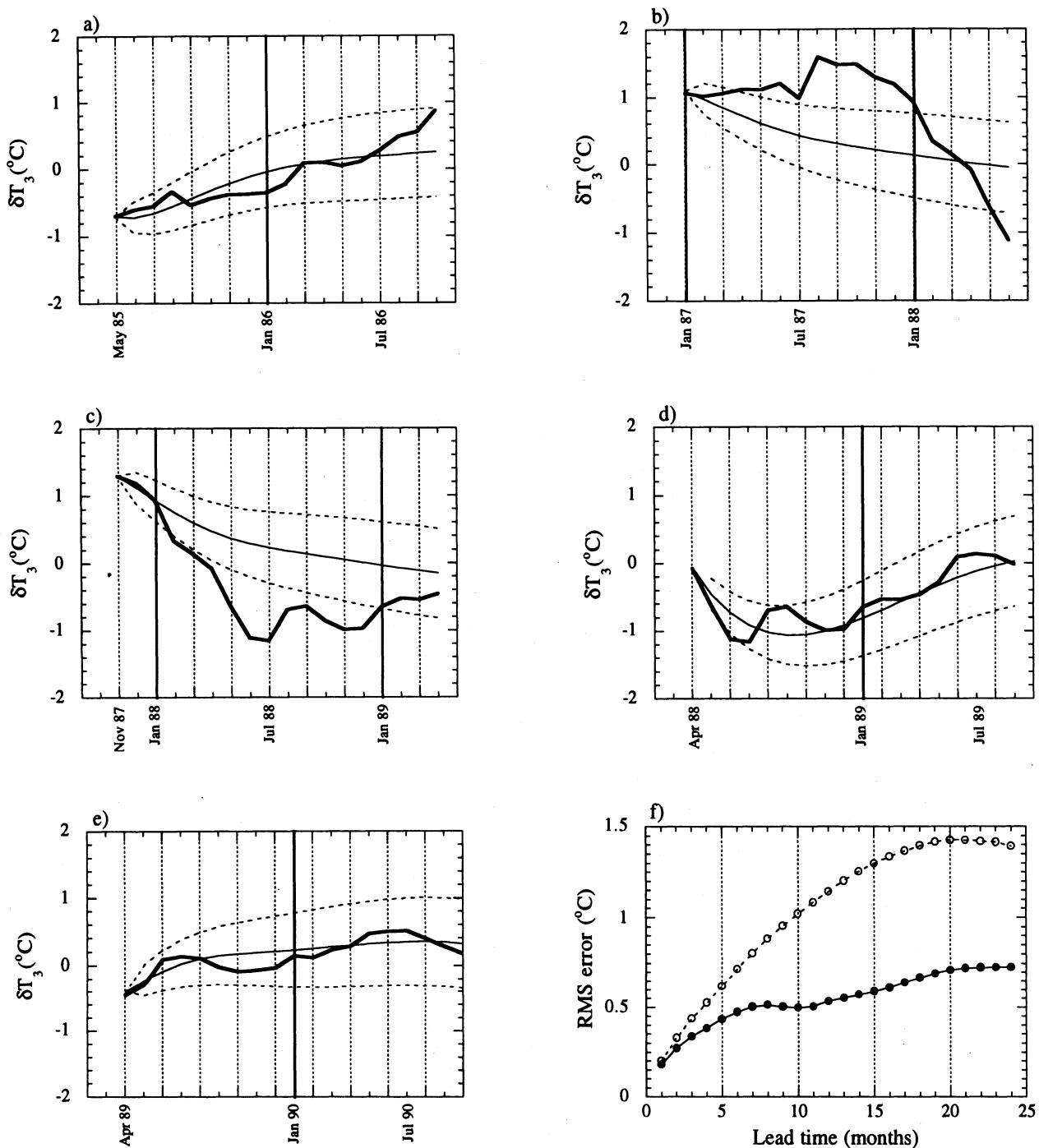


FIG. 2. Prediction (light solid line) and verification (bold solid line) of Niño 3 SST, based on principal component (PC) field of (a) May 1985, (b) January 1987, (c) November 1987, (d) April 1988, and (e) April 1989. Dashed lines show 95% confidence interval. (f) The rms error in Niño 3 SST anomalies evaluated throughout the verification period based on predictions using the linear model (filled symbols) and persistence (open symbols).

Before generating the numerical model, we considered what effect adding surface wind anomalies to the analysis did to the predictions of SSTs. These wind anomalies (from COADS) were sampled in the same region as the SSTs and subjected to a three-month running mean after removal of the annual cycle. In fact,

it was found that adding winds to this analysis made very little difference in the predictions. Figure 6 shows bubble plots of the Green function matrices in EOF space (obtained with $\tau_0 = 3$ mo), which operate on an initial condition vector of 8 SST principal component time series (PCs), 8 zonal wind (U) PCs, and 8

TABLE 1. Timescales of empirically derived normal modes.

Mode number	Decay time (mo)	Period (mo)
1	30	N/A
2/3	19	242
4/5	13	40
6/7	11	100
8	11	N/A
9/10	8	20
11/12	7	22
13	7	N/A
14/15	6	79
16	5	N/A
17/18	5	17
19/20	3	58

meridional wind (\mathbf{V}) PCs at some time t to obtain the most probable prediction of the 24-dimensional vector at lead times of $\tau = 3, 6$, and 9 mo. The SST principal components are given far more weight than any of the wind PCs in determining the future of the SST–wind system. This property that the winds do not improve SSTA prediction is stable when either τ_0 or the number of retained EOFs is varied, although the actual weights themselves do change due to the importance of nonlinearities to the propagation of wind anomalies.

This result does not indicate that the winds are unimportant. First, the signal of wind forcing is found in the SSTs themselves and, as long as a dynamical forcing can be parameterized by SST, the empirical analysis quite willingly mixes forcing and response. Much of the wind forcing is already taken into account, although indirectly, by an empirical analysis of the SSTs alone; that is, it is often only partially correct to call the empirically derived deterministic matrix a “feedback” matrix. Repeating the coupled wind–SST analysis with

$\tau_0 = 7$ mo gave the SST–wind system enough time to feed back on itself so that those Green functions (not shown) at $\tau = 3, 6$, and 9 mo gave more weight to winds, especially \mathbf{V} , affecting both wind components than did the calculation using $\tau_0 = 3$ mo. However, even in the $\tau_0 = 7$ mo case, SST PCs were still more important than winds in the prediction of any of the three fields. Based on these considerations, the number of degrees of freedom necessary to adequately describe the SST–wind system, and the shortness of our training period, it was decided to base the empirical–dynamical model on SSTs alone.

b. Generation of the stochastic model

The numerical simulation of stochastic differential equations is well documented in Rümelin (1982). However, Rümelin’s results are valid in the limit of vanishing time step. For a simulation of a fluid described by Eq. (4) using a finite time step, preserving the nonzero equal-time correlation between the signal and noise inherent in a linear Stratonovich system [$\langle x_i(t)\xi_j(t) \rangle + \langle x_j(t)\xi_i(t) \rangle = Q_{ij}$; García et al. (1987)], a modification of Rümelin’s method must be employed (Mansour, Baras, and García 1993, personal communication). The appropriate method of integrating Eq. (4) is a two-step process:

$$Y_i(t + \Delta)$$

$$= Y_i(t) + \sum_{j=1}^d B_{ij} Y_j(t) \Delta + \sum_{\alpha=1}^m \psi_{i\alpha} \sqrt{q_\alpha \Delta} \mathcal{R}_\alpha, \quad (7a)$$

$$X_i(t + \Delta/2) = \{Y_i(t) + Y_i(t + \Delta)\}/2, \quad (7b)$$

where Y_i is an intermediate variable, X_i is the principal component corresponding to the i th SST EOF, q_α is the α th eigenvalue of \mathbf{Q} , and $m = 18$ since two eigen-

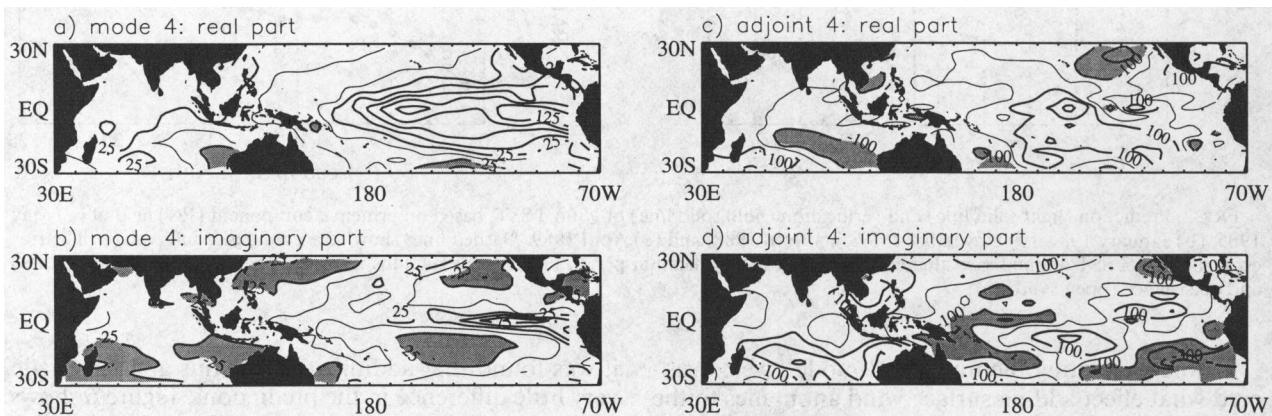


FIG. 3. Eigenvector 4 of $\mathbf{G}(\tau, \tau_0 = 7$ mo) having decay time of 13 mo and an oscillation period of 40 mo. Also shown is the corresponding modal adjoint. Contour labels are multiplied by 1000. In (a) and (b) the contour interval is 0.025 and shading indicates loadings less than -0.025 . (a) Mode 4: real part. (b) Mode 4: imaginary part. In (c) and (d), the contour interval is 0.100, and shading indicates loadings less than -0.100 . (c) Modal adjoint 4: real part. (d) Modal adjoint 4: imaginary part.

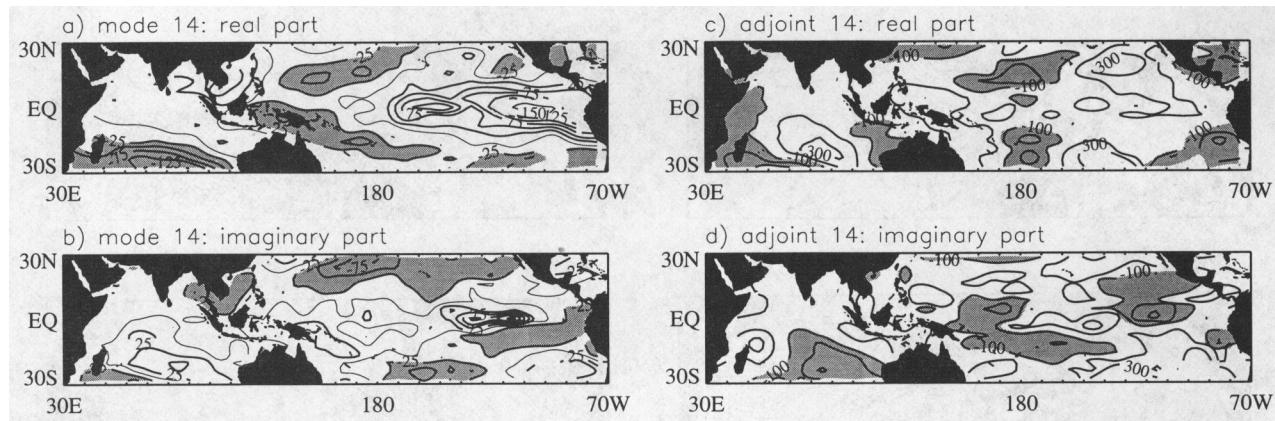


FIG. 4. As in Fig. 3 but for the empirically determined normal mode 14. In (c) and (d) contour interval is 0.200, and shading indicates loadings less than -0.200.

vectors of \mathbf{Q} are spurious. Because two eigenvalues are negative, the sum of positive q_α is somewhat greater (about 12%) than the sum of the diagonal elements of \mathbf{Q} . In this model, each positive q_α was proportionately reduced so that the noise variance in the model was equal to that of \mathbf{Q} . The m -dimensional vector \mathcal{R} of independent Gaussian deviates, each having unit variance, must be generated and then transformed using the noise eigenvectors $\{\psi_\alpha\}$ at each time step. Generating these random numbers within the sum over α results in unphysical random phases that interfere to produce a diagonal \mathbf{Q} , regardless of the space in which the calculation is being performed. Note that the deterministic feedback is updated with the time step while the stochastic forcing is proportional to $\sqrt{\Delta}$ (Garcia et al. 1987). This is related to the fact that the variance of the Wiener process increases linearly with time.

The deterministic term may be treated in a modified forward-stepping scheme because \mathbf{Q} is independent of \mathbf{X} in this case. If the stochastic forcing were dependent on the state of the system, a predictor-corrector method would be necessary to accurately simulate the effects of the noise-induced drift, but schemes of higher order than two are usually no more accurate than the second-order Runge-Kutta method for generating continuous Markov models (Rümelin 1982).

To obtain a "control run" time series, we chose a time step Δ of six hours and sampled data every month. That is, $\Delta = 1/120$ mo. The model run was initiated at the peak of the multivariate marginal distribution ($\mathbf{X} = 0$) and exercised for 1000 months. Memory of the initial condition was lost after about two years, well before any model output was recorded. The 2000-year run was performed on a Sun 4/690. Upon completion of the run, the output multivariate time series was subjected to a consistency check. Figures 7a,b are bubble plots of the \mathbf{B} and \mathbf{Q} in EOF space obtained from the data and used to generate the model. Figures 7c,d are bubble plots of \mathbf{B} and \mathbf{Q} as determined from the model

output. The correspondence is good; the two spurious eigenvectors were safely neglected in our generation of the driving noise.

4. Analysis of model output

a. Time domain results

By projecting the model output into geographical space, the time series of the average model SST anomaly δT_3 in the Niño 3 region (6°N – 6°S , 150° – 90°W) was calculated and used to diagnose the sensitivity of the output to various aspects of the generating scheme. The standard deviation (SD) of δT_3 obtained from the COADS data is 0.71°C . The difference in the standard deviation of the model's δT_3 (0.72°C) when the average of Y_i was taken between successive time steps [Eq. (7b)] from that obtained by simply setting $X_i(t + \Delta) = Y_i(t + \Delta)$ was completely negligible ($\sim 0.01\%$) for the 6-hour time step. The size of the time step was found to be more important. When a time step of one month was used in the two-step procedure, the standard deviation of δT_3 was 0.88°C , an error of 24%; neglecting Eq. (7b) increased the error by an additional 2%. The matrix \mathbf{Q} determined from this large-time step run violated the necessary conditions of being a covariance matrix (i.e., $Q_{ii} > 0$; $|Q_{ij}| < \sqrt{Q_{ii}Q_{jj}}$ for all i, j). Some of this error may be due to the deterministic timescales. However, the reason the chosen time step is on the order of hours, rather than weeks, is because simulating a continuous stochastic forcing with a discrete random number generator results in a time series noisier than is dynamically consistent (see section 2) if the time step is too large.

Warm events in the model output were identified by upward excursions of the δT_3 time series across the 1 std dev threshold, with the proviso that δT_3 must dip below -0.4°C between events. This criterion when applied to data would allow, for example, the 1982–83 event to be considered as separate from the 1986–87

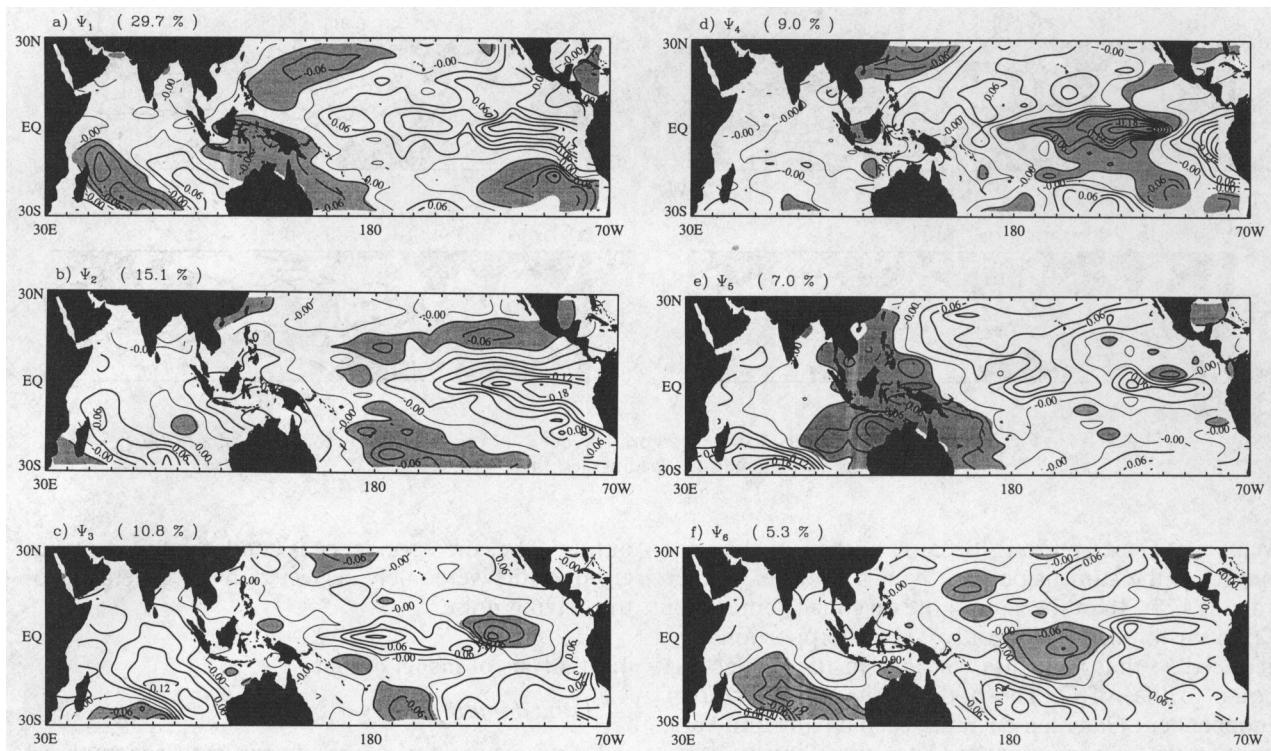


FIG. 5. Eigenfunctions of \mathbf{Q} ; Contour interval: 0.01, and areas with loadings less than -0.01 are shaded. In (a) ψ_1 explains 29.7% of the noise variance, in (b) ψ_2 explains 15.1% of the variance, in (c) ψ_3 explains 10.8% of the variance, in (d) ψ_4 explains 9.0% of the variance, in (e) ψ_5 explains 7.0% of the variance, and in (f) ψ_6 explains 5.3% of the variance.

event. The distribution of times between warm events in the control run is shown in Fig. 8, indicating that, as observed, most events occur two to five years after the last one, with the mean time between events being 47.9 months. This is significantly smaller than the average time between events (52.4 months) obtained when the same analysis was applied to observed SST anomalies between 1950 and 1989, consistent with the observation that predictions based on the linear stochastic model underestimate the duration of warm events. The model's composite event is shown in Fig. 9. The narrow latitudinal extent of the cold anomalies seen in the composite a year after maximum warmth is an artifact of the different lengths of time the warm events persist; a composite of cold events (not shown) is very similar to the warm pattern at $t = 0$ with the signs reversed.

The mode pair 4/5 was the most regularly evolving structure in the composite, but no single ENSO-like mode dominated. Figure 10 shows the evolution of the composite warm event with the contribution from the mode pair 4/5 removed. Although the patterns are significantly altered (cf. Fig. 9), a recognizable warm event still obtains. This result is consistent with PM93, who found that combinations of modes are necessary in order to accurately describe an ENSO event.

To clarify the role of the random forcing, the forcing functions $\{\psi_a\}$, weighted by $\sqrt{q_a}$, were projected onto the adjoints of the empirically derived modes. The most predictable of the ENSO-type normal modes, mode 4/5, is not as sensitive to the most important forcing function as it is to ψ_3 , ψ_5 , and ψ_6 (Fig. 11), which have large loadings in the south Indian Ocean and in the equatorial Pacific in the region of the date line. Other modes, however, do show sensitivities to ψ_1 . The EDM was run with the contribution to the forcing from ψ_3 , ψ_5 , and ψ_6 removed; however, the variance of the other forcing functions was increased by a proportional amount so that the total variance of the forcing remained the same. Therefore, differences between this EDM and the previous one are attributable completely to the geographic distribution of the random forcing. The standard deviation of the Niño 3 SST is lower (0.63°C vs 0.72°C) in the second experiment, a difference of about 11%. There are more warm events (547 vs 500) than when all forcing functions are used, partly because the standard deviation of the field is smaller, and so they are closer together (Fig. 12 vs Fig. 8). The mean time between events is 43.8 mo, a difference of four months from the first experiment and almost ten months smaller than observed. A composite warm event (not shown) was calculated using the pro-

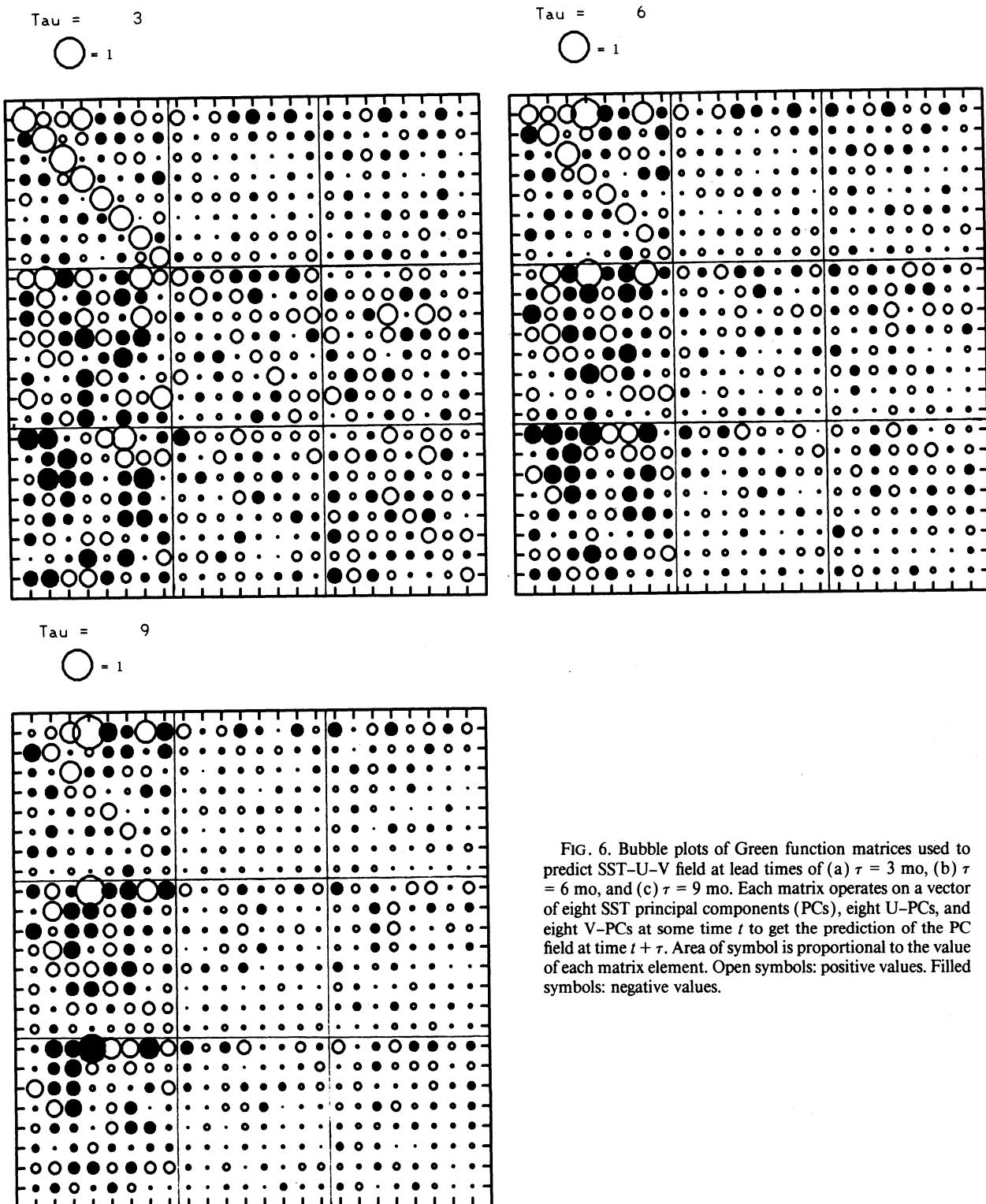


FIG. 6. Bubble plots of Green function matrices used to predict SST-U-V field at lead times of (a) $\tau = 3$ mo, (b) $\tau = 6$ mo, and (c) $\tau = 9$ mo. Each matrix operates on a vector of eight SST principal components (PCs), eight U-PCs, and eight V-PCs at some time t to get the prediction of the PC field at time $t + \tau$. Area of symbol is proportional to the value of each matrix element. Open symbols: positive values. Filled symbols: negative values.

cedure described above. The amplitude of u_4 in this composite is reduced by about 40% from what it is when all forcing functions are used (Fig. 13). The amplitude of other modes, which are more sensitive to the augmented forcing functions, are enhanced al-

though most modal amplitudes are reduced, as reflected in the smaller rms value of δT_3 . The largest difference in SST at $t = 0$ is about 0.35°C (Fig. 14). This difference is about half the rms value of δT_3 in the control model. Still, the composites from the two experiments

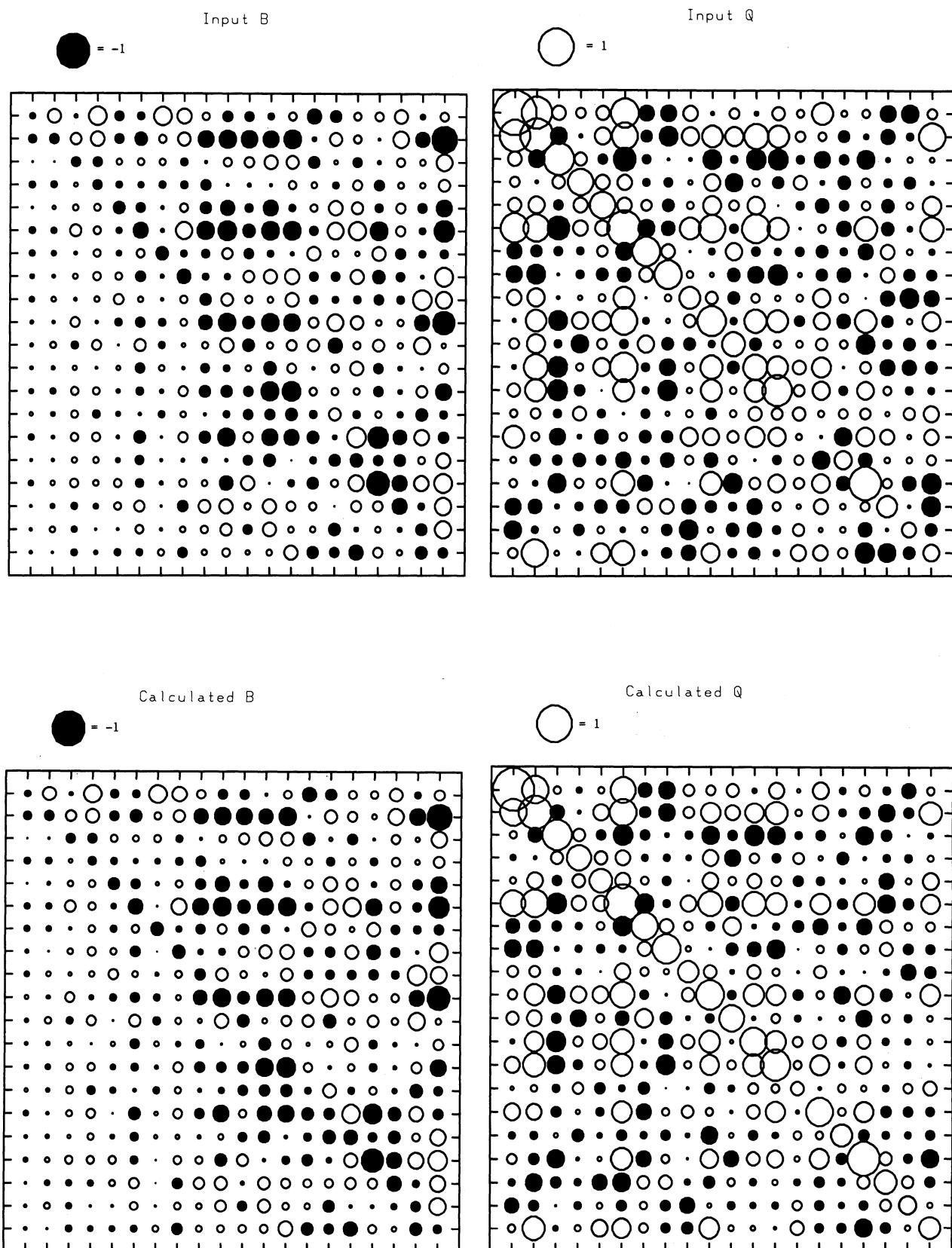


FIG. 7. Bubble plots of deterministic matrices \mathbf{B} and forcing covariance matrices \mathbf{Q} defining the linear stochastic model. (a) Input \mathbf{B} ; (b) input \mathbf{Q} ; (c) \mathbf{B} as derived from EDM output; (d) \mathbf{Q} as derived from EDM output. Open symbols: positive values. Filled symbols: negative values.

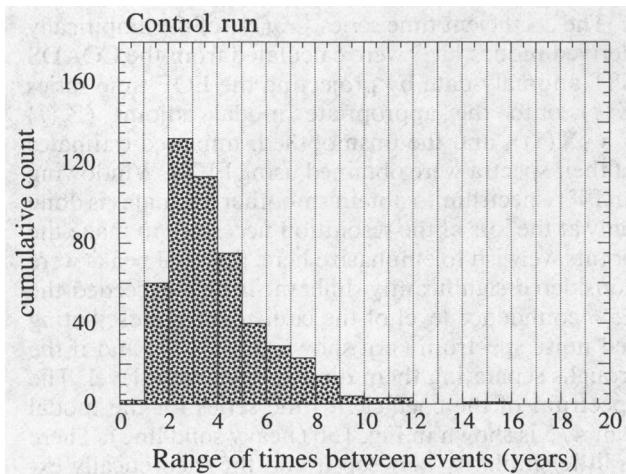


FIG. 8. Histogram of times between warm events (control run of empirical dynamical model).

are similar even though their modal makeup and, consequently, their strengths and timescales are different.

b. Fourier spectral analysis

As a propagating mode evolves, stochastic forcing adds random phases to the wave velocity, which have the effect of introducing larger contributions to frequencies adjacent to the modal angular frequency $\text{Im}\beta_\alpha$. This spread in the spectrum of the empirical normal mode is highly affected by the size of $\text{Re}\beta_\alpha$; the spectral density $dP_\alpha(\omega)$ of the time series coefficient of \mathbf{u}_α is equivalent to a first-order maximum entropy method spectrum (Penland et al. 1991):

$$dP_\alpha(\omega) = \frac{\mathbf{v}_\alpha^T \mathbf{Q} \mathbf{v}_\alpha^*}{(\omega^2 - 2 \text{Im}(\beta_\alpha)\omega + \beta_\alpha \beta_\alpha^*)}. \quad (8)$$

Note that since the empirical normal modes occur in complex conjugate pairs, the spectrum of one member of the pair is peaked while the spectrum of the other is not ($\omega > 0$). Although the variance of the coefficient time series of one member of the pair is equal to that of the other, this variance is not simply related to the spectrum of that mode alone. However, we find that the usual relation, involving a number of samples, between the variance of the time series and the spectrum as calculated by fast Fourier transforms (FFTs) is reproduced if the spectrum is calculated as the average contribution from both members of the modal pair. If the ratio $\text{Im}(\beta_\alpha)/\text{Re}(\beta_\alpha)$ is small enough, the spectrum of a complex mode pair has no peak.

The spectrum of the modal pair 4/5 [cf. Eq. (8)] based on the empirically obtained parameters is shown in Fig. 15a (heavy solid line). The spectrum of the modal pair 14/15 has no peak, and its variance is somewhat larger than that of the modal pair 4/5. Al-

though the spectral density of mode 14 alone (heavy dashed line) is peaked at the appropriate period (i.e., given by the respective complex eigenvalue; cf. Table 1), the spread in the spectrum of that mode pair, due to the smaller decay time, indicates that it relies more heavily on the unpredictable random forcing to maintain its variance than does the mode pair 4/5. The mode pair 14/15 has an amplitude comparable to that

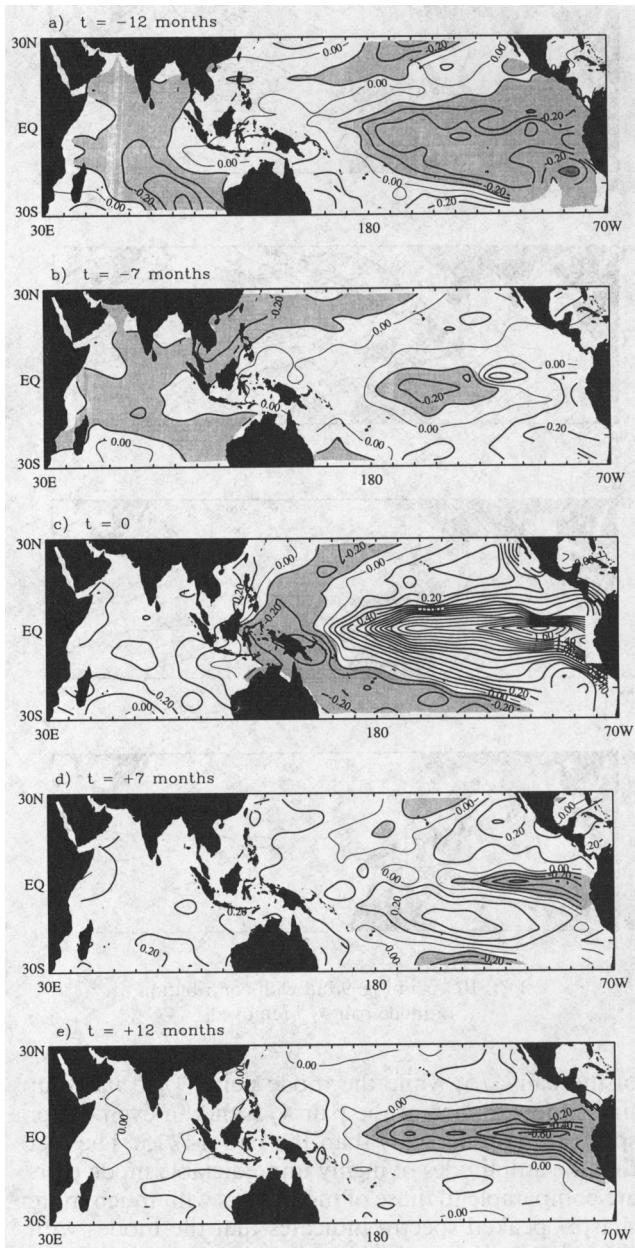


FIG. 9. Composite of warm events with $\delta T_3 > \sigma$. As in Figs. 10 and 14, time is measured with $t = 0$ at the SST maximum in Niño 3; (a) $t = -12$ mo, (b) $t = -7$ mo, (c) $t = 0$, (d) $t = +7$ mo, and (e) $t = +12$ mo. Contour interval is 0.1°C , and temperature anomalies less than -0.1°C are shaded.

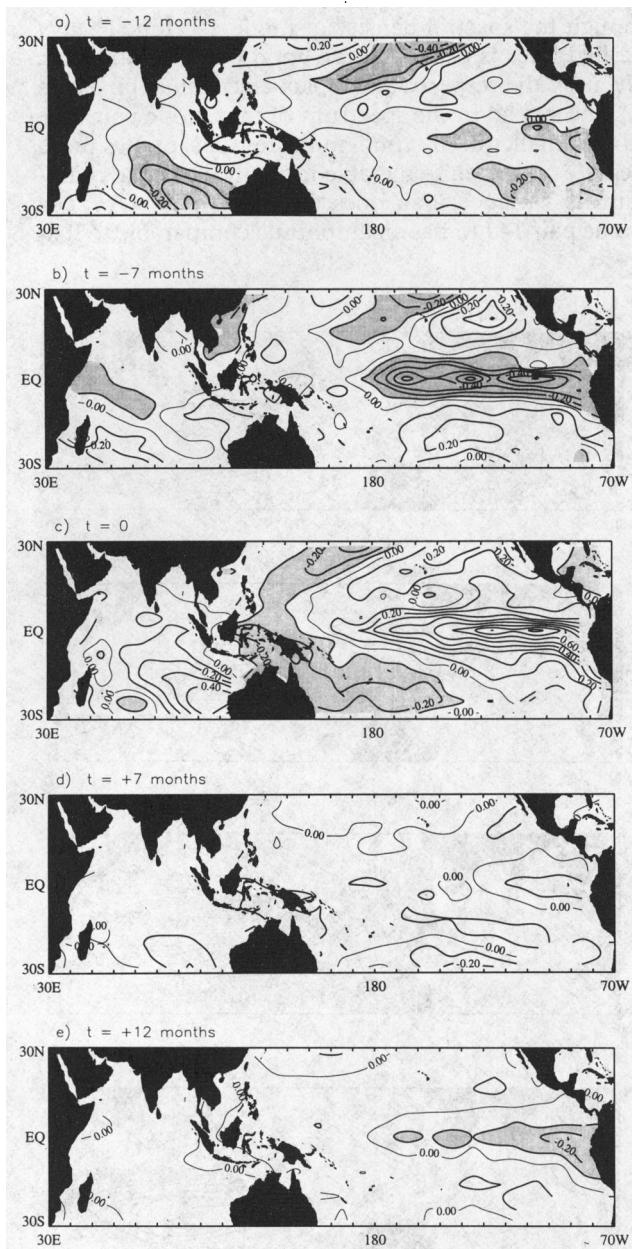


FIG. 10. As in Fig. 9 but with contribution of mode pair 4/5 removed.

of the pair 4/5, while the mode pair 19/20 has four times the variance of the pair 4/5 and an even flatter spectrum (not shown) than the pair 14/15. The fact that the amplitudes of highly unpredictable mode pairs are comparable to those of mode pairs with much more sharply peaked spectra indicates that the modes with short decay times must be included in our model if the empirically obtained statistics of the SSTs are to be reproduced. Also, the large overlap in modal spectra precludes separation of these modes by bandpass filtering.

The coefficient time series $\{Z_\alpha(t)\}$ of the empirically derived modes $\{\mathbf{u}_\alpha\}$ were calculated from the COADS SST anomaly data by projecting the EOF time series $\mathbf{X}(t)$ onto the appropriate modal adjoint ($Z_\alpha(t) = \mathbf{v}_\alpha^T \mathbf{X}(t)$), and the unsmoothed, unbiased estimates of their spectra were obtained using FFTs. Windowing an FFT spectrum to obtain smoother estimates is done only at the loss of the resolution necessary to make the points we wish to emphasize here. Spectral peaks were considered significantly different if they exceeded the 95% confidence level of the equivalent nonoscillating red noise spectrum (not shown in Fig. 15) and if the troughs separating them dipped below that level. The spectrum of the coefficient time series for the modal pair 4/5 is shown in Fig. 15b (heavy solid line). There is little similarity between it and the theoretically expected spectrum shown in Fig. 15a, and the multiple significant peaks at 58–90 mo, 37–54 mo, 25–36 mo, 22–25 mo, 20–22 mo, and 17–20 mo in a spectrum, supposed to have a single broad peak at 40 mo might suggest that nonlinear interactions, assumed to be negligible in this study, actually dominate the dynamics.

In fact, nonlinear dynamics are not required for this result. The coefficient time series obtained from the data contains only 490 samples. When the spectrum of the first 500 samples of the corresponding coefficient time series in the 2000-year (24 000 samples) control run was calculated, significant peaks at similar frequencies (45–55 mo, 27–45 mo, 21–27 mo, and 16–19 mo) were found (Fig. 15b, light solid line with symbols), even though the generating model was linear by construction. The 80-month peak is missing in this realization of the model time series, but it shows up in other 500-sample chunks. An ensemble spectrum of 48 realizations (Fig. 15a, light solid line with symbols), each calculated from 500 samples of the control run, is much closer to the theoretical spectrum, although peaks corresponding to those found in the data are still visible. Ensemble spectra with up to 100 realizations (not shown) indicate that the spectrum of the mode pair 4/5 as calculated in the linear model does eventually converge to the theoretical spectrum, but this convergence is slow.

The high-resolution technique was chosen over a more reliable windowing technique for calculating the spectra to show that the significant peaks found in the data are close to oscillation periods of other ENSO-like modes. These modes are highly correlated; the absolute value of the correlations between mode 4 and modes 11 and 14 are more than 60%, and those between mode 4 and modes 9 and 20 are more than 50%. The adjoints of their respective modal pairs are also correlated at 30% to 50%. Although the modes and their adjoints form a biorthogonal set, the short 35-year training period allows large uncertainties in the set of 20 empirically derived complex modes, their corresponding adjoints, and the appropriate timescales.

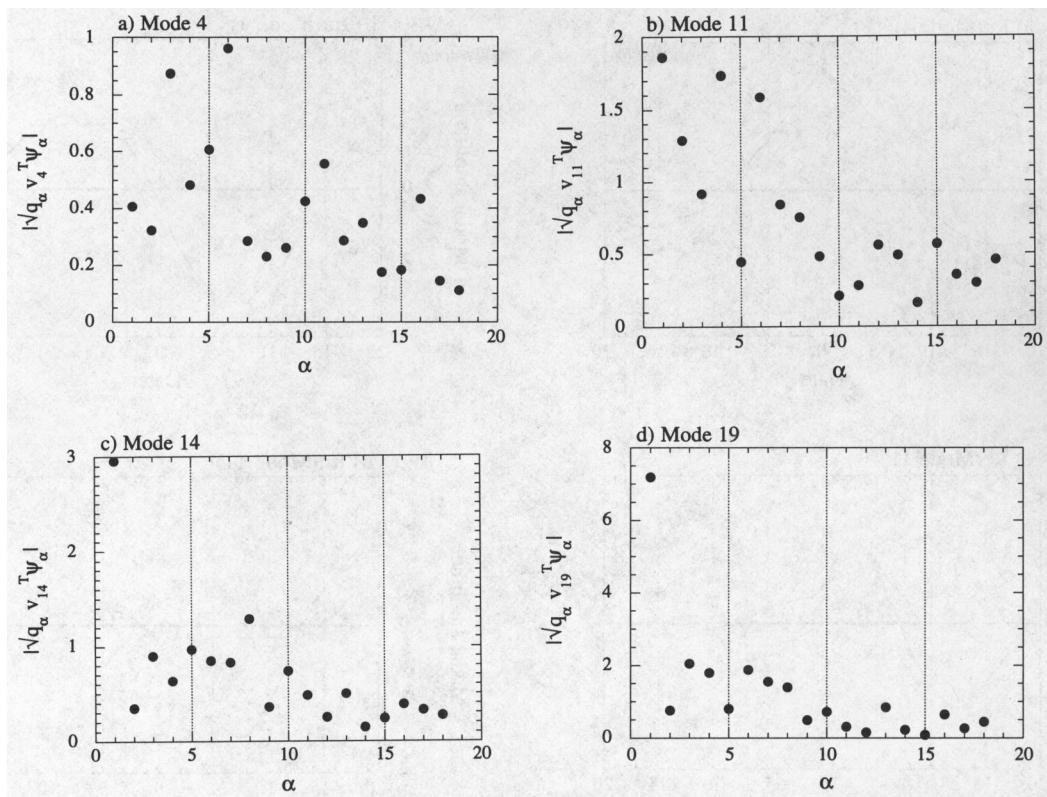


FIG. 11. Magnitude of the projection of forcing functions, weighted by $\sqrt{q_\alpha}$, onto modal adjoints for (a) mode 4, (b) mode 11, (c) mode 14, and (d) mode 19.

One might expect, therefore, the projection of the data time series onto the normal modes to be imperfect and that highly correlated modes would contaminate each other's spectra.

This does not explain why the same phenomenon is seen in the model output. In the model, each mode

is excited by random forcing, which is temporally white but spatially coherent (cf. Fig. 5). A complex spatial filter applied to this real forcing changes the amplitude and phase of the white noise at each geographical location so that interference with frequency components in the deterministic feedback can occur. This is how the adjoint of a particular normal mode, acting as a spatial filter, projects out that mode's particular spatial pattern with its particular spectral properties. The geographically varying phase applied by the adjoint is not affected by the spatially constant, arbitrary phase allowed by the normalization, which cancels with its complex conjugate multiplying the modal pattern. When several modes are highly correlated, many realizations are necessary to smear out competing frequencies into the broad peak characteristic of a particular modal spectrum. Of course, this process obtains in the real ocean as well as in the model, but one suspects that uncertainty in the modal structure also affects the spectra of oceanic data.

5. Summary and conclusions

Checks for conservation of quantities such as energy and momentum are standard procedures in the diagnosis of deterministic models (e.g., Klinker and Sar-

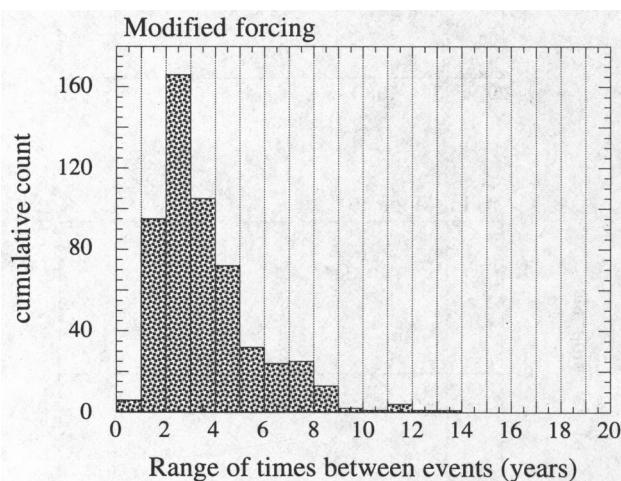
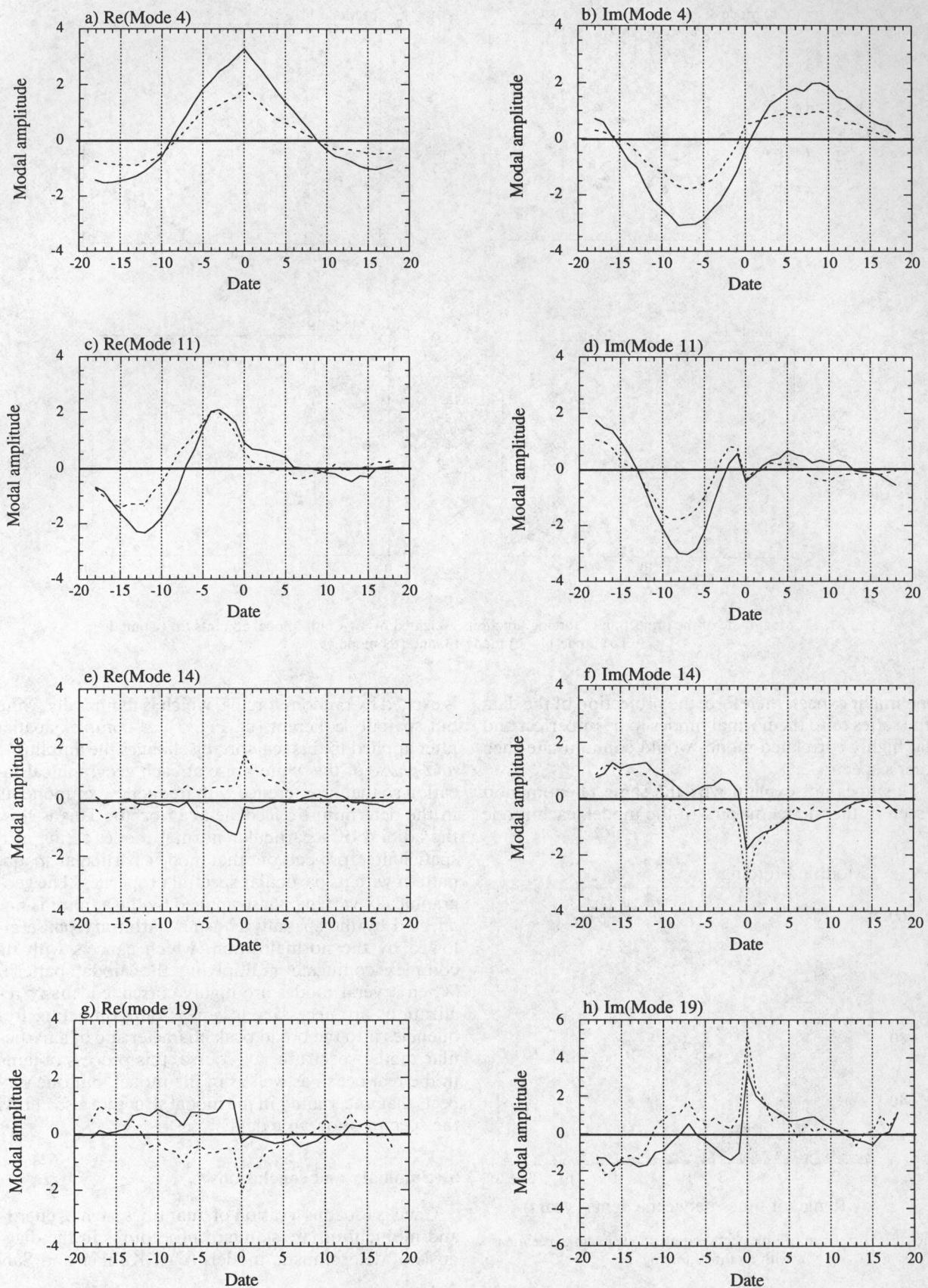


FIG. 12. Histogram of times between warm events in experiment with modified forcing.



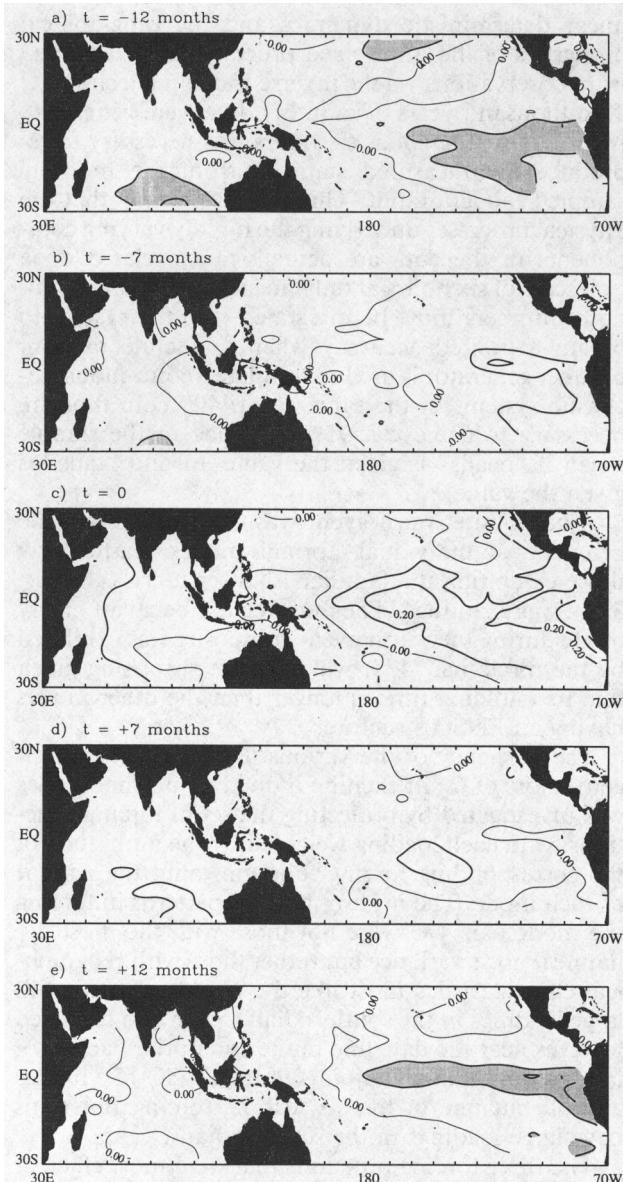


FIG. 14. Difference maps between the warm composite shown in Fig. 9 and that obtained using modified forcing at (a) $t = -12$ mo, (b) $t = -7$ mo, (c) $t = 0$, (d) $t = +7$ mo, and (e) $t = +12$ mo. Contour interval is 0.1°C , and temperature anomalies less than -0.1°C are shaded.

deshmukh 1992). It has not been obvious, however, what dynamical constraints ought to be imposed on stochastic models. A balance condition based on the well-known fluctuation-dissipation relation between

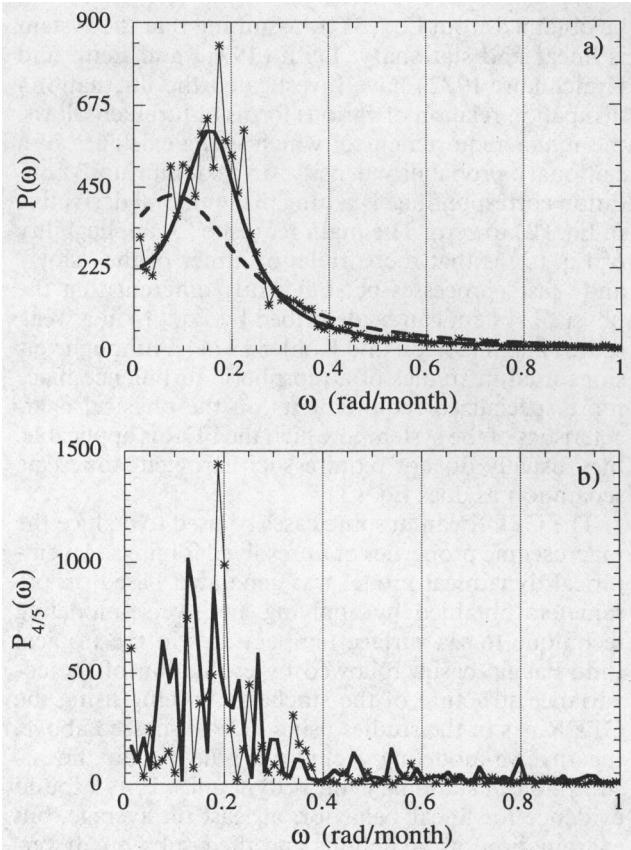


FIG. 15. Fourier spectra of the coefficient time series. (a) Theoretical spectra of modal pair 4/5 (heavy solid line) and of mode 14 alone (heavy dashed line). The ensemble spectrum of the coefficient time series of the modal pair 4/5 (light solid line with symbols) was obtained using 48 realizations, each consisting of a 500-sample chunk from the control run. (b) Spectrum of the coefficient time series of the modal pair 4/5 as obtained from COADS data (heavy solid line), and spectrum of the coefficient time series of the modal pair 4/5 as obtained from a single 500-sample realization of the control run (light solid line with symbols).

the covariance structure of a measured quantity and that of its stochastic forcing has been introduced. It is usually possible to simplify the full nonlinear, nonstationary version of this relation, but even if that is not the case, and even if it is not actually solvable for the second moments, it is still possible to use it to check a stochastic model for dynamic consistency.

It is interesting to compare Eq. (3) with the FDR used by Leith (1975, 1978), Bell (1980), and North et al. (1993). The FDR used in those studies was developed in the context of isotropic turbulence and can

FIG. 13. Composite warm event behavior of empirical normal modes in control run (solid lines) and with modified forcing (dashed lines): (a) Re (mode 4), (b) Im (mode 4), (c) Re (mode 11), (d) Im (mode 11), (e) Re (mode 14), (f) Im (mode 14), (g) Re (mode 19), and (h) Im (mode 19).

be obtained from Eq. (3) by assuming that the system is linear and stationary. Leith (1971) and Leith and Kraichnan (1972) have investigated the fluctuation-dissipation relation of various forms of turbulent flows, the major requirement of which is the existence of a stationary probability density, an "equilibrium" condition corresponding to setting the temporal derivative in Eq. (2) to zero. The main requisite for applicability of Eq. (3) is that the correlation times of the "slow" and "fast" processes be sufficiently different that the physical system can be described by Eq. (1) in a weak sense (Papanicolaou and Kohler 1974). Although versions used in studies of atmospheric turbulence place more specialized requirements on the physical characteristics of the system to which the FDR is applicable, they usually do not require such stringent timescale separation as does Eq. (3).

The GFDR can in some cases be used to deduce the macroscopic properties of unresolved forcings. An empirical dynamical model was generated based on parameters obtained by applying an inverse-modeling technique to sea surface temperatures in the tropical Indo-Pacific basin, followed by evaluation of the covariance structure of the stochastic forcing using the GFDR. As in the studies using FDR discussed above, the inverse-modeling technique relies upon the assumption of stationary linear dynamics. PM93 found evidence for linear behavior, at least on average, but examination of predictions and their subsequent verifications shows that the assumption breaks down during the warmest phase of a warm event. However, the linear model was successful in predicting the behavior of SSTs during the other phases of the events. It was concluded, therefore, that a numerically generated linear stochastic model would reproduce the salient features of SST evolution.

The successful predictions of this simple model are probably due in part to the projection of the field onto a number of EOFs sufficient to describe the evolving SST patterns but small enough to eliminate much unpredictable noise. That such a projection increases predictability is well known (Mo 1988; Tracton et al. 1989; Branstator et al. 1993) and this procedure has been used successfully in past inverse-modeling projects (Blumenthal et al. 1991; PM93; Penland and Ghil 1993; Y. Xue 1993, personal communication). Adding to the analysis the leading EOFs calculated from wind fields in the Indo-Pacific region, however, did not significantly improve the forecast skill. At the timescales of interest, much of the effect of wind forcing is indirectly taken into account by considering SST alone since the signal of wind forcing is found in the SSTs themselves. Other wind effects are accounted for by parameterizing the fast processes as white noise, which is also integrated into the SST.

Two numerical experiments were performed. The first, the control run, was designed to reproduce the

linear deterministic dynamics and the geographical properties of the unresolved processes as determined from observations via the inverse-modeling technique. Two thousand years of monthly data were generated. A time step as small as six hours was necessary to reproduce the measured statistics from the monthly sampled model output. This does not imply that the physical processes underlying the rapidly varying components in the data are actually uncorrelated on a timescale of six hours; it only means that so many random numbers must be integrated into the system to obtain a desired accuracy when a discrete random number generator is used to simulate a continuous stochastic system. In the same spirit, 400 coin flips are necessary to achieve a 95% accuracy in the sample mean if "heads" is given the value +1 and "tails" is given the value -1.

A composite warm event was constructed and the behavior of individual normal modes, particularly those exhibiting the familiar ENSO pattern (cf. Figs. 3, 4), was examined. The most regular behavior of any mode during the warm event composite was exhibited by the mode pair 4/5, which is not surprising given that its *e*-folding time is longer than the other modes having an ENSO structure.

The efficiency of the various forcing functions (eigenvectors of \mathbf{Q}) in exciting individual normal modes was investigated by projecting the set of forcing functions, with each loading weighted by the amplitude of the corresponding forcing function, onto the adjoint of each mode. The most important patterns in forcing the mode pair 4/5 were not those with the most dynamical noise variance but rather those with geographical characteristics most like the modal adjoint, with large loadings in the south Indian Ocean and localized features near the date line in the equatorial Pacific. In fact, we find that all adjoints of the ENSO-like modes and all but one of the six leading forcing functions have large loadings in the south Indian Ocean.

The three forcing functions that were most efficient in exciting the mode pair 4/5 were removed in the second numerical experiment. The other forcing functions were amplified proportionally so that the total amount of variance remained the same as in the first experiment. Differences between the two experiments are therefore due solely to the geographical distribution of the stochastic forcing rather than to its strength. The warm event composites from the two experiments were similar, with maximum discrepancies of about 0.35°C at the peak of the event. However, the energy distribution among the normal modes was quite different. The amplitudes of most modes in the composite were reduced in the second experiment, while the amplitudes of some higher-indexed modes, which rely on random forcing more than their own free evolution to maintain their variance, were increased. ENSO events were closer together than in the control run both because the

threshold value of the SST anomaly defining an event was somewhat smaller and easier to attain and because modes whose amplitudes were enhanced were more susceptible to the white noise forcing and, therefore, varied more rapidly than the important modes in the control run. These results suggest that the belt reaching diagonally from the equatorial Pacific in the region of the date line to the south Indian Ocean is important to the amplitude and timing of ENSO events. Weickmann (1993, personal communication) has found evidence for similar conclusions. Studies currently under way address what role, if any, anomalous convection there plays in triggering ENSO events.

The theoretically expected spectra of modal coefficient time series bear little similarity to those calculated from the data, and one might have interpreted the multiple significant peaks found in those spectra as evidence for nonlinearity [compare spectral peaks with those of Ghil and Vautard (1991)]. However, similar results were found by analyzing the output of the model, which is linear by construction, and where the timescales associated with the modes are perfectly prescribed. The multiple peaks correspond to frequencies characterizing normal modes that are highly correlated. The adjoints of highly correlated modal pairs are also substantially correlated. The complex adjoints adjust the amplitude and phase of the (real) random forcing at each geographical point so that the white noise forcing is effectively filtered; constructive and destructive interference between the white noise as modified by the adjoints, and the frequency components present in the deterministic feedback, tends to favor the characteristic frequency of an oscillating normal mode. However, if modes and their adjoints are highly correlated, other frequencies are also favored, although to a lesser extent. A long enough time series combined with a spectral technique that allows many random amplitudes of the adjoint-induced phases to enter the calculation is necessary to reproduce the spectrum of a single normal mode as expected from theory. Another interpretation, applicable to the oceanic data but not to the model, is at least partially correct; uncertainties in the empirically obtained modes and adjoints do not allow a perfect projection of the data onto the normal modes, so that timescales of several correlated modes appear in any one coefficient time series.

The frequencies corresponding to significant peaks really do exist in a modal coefficient time series of finite length. These frequency components are excited randomly so that, *on average*, the spectrum consists of a single broad peak. This theoretically expected spectrum is an average of *infinitely many* ensembles, and the dynamical nature of the system is what determines what finite number is sufficient to approximate infinity. Unfortunately, it is often the nature of the dynamical system that is unknown, which one tries to diagnose via Fourier analysis. It is clear from this study, given

the high correlation of the normal modes and its effect on the spectra, that Fourier spectral analysis alone cannot distinguish between a multivariate linear system and a nonlinear one.

Acknowledgments. The authors are pleased to acknowledge conversations with Klaus Weickmann, David Neelin, Prashant Sardeshmukh, and Yan Xue. Comments on the manuscript by Marty Hoerling and both referees resulted in a (hopefully) clearer exposition. Mark Cane is acknowledged for useful discussions pertaining to appendix B. Particular thanks are due to Chuck Leith, for pointing out several important references, and Alejandro García, for allowing us to use and exhibit the modification to Rümelin's method prior to its publication in *Journal of Statistical Physics*. This work was supported by NOAA Grant NA16RC0115-01.

APPENDIX A

Review of Inverse Modeling Technique

Consider an d -dimensional vector $\mathbf{x}(t)$, the i th component $x_i(t)$ of which represents a zero-mean observation at location i and time t and obeying the linear stochastic differential equation

$$\frac{d\mathbf{x}}{dt} = \mathbf{B}\mathbf{x} + \boldsymbol{\xi}, \quad (\text{A1})$$

where \mathbf{B} is constant and deterministic, while $\boldsymbol{\xi}$ is a white noise vector with covariance \mathbf{Q} . The i th component of $\boldsymbol{\xi}$ represents the random excitation of the field at location i . The matrices \mathbf{B} and \mathbf{Q} are to be determined from the data.

Given Eq. (A1), the Green function \mathbf{G} ,

$$\mathbf{G}(\tau) = \langle \mathbf{x}(t + \tau) \mathbf{x}^T(t) \rangle \langle \mathbf{x}(t) \mathbf{x}^T(t) \rangle^{-1}, \quad (\text{A2a})$$

is equivalent to

$$\mathbf{G}(\tau) = \exp(\mathbf{B}\tau). \quad (\text{A2b})$$

The probability density $p(\mathbf{x}, t + \tau | \mathbf{x}', t)$ of finding the pattern $\mathbf{x}(t + \tau)$ at some time $t + \tau$ when the pattern $\mathbf{x}'(t)$ is known at some earlier time t is a multivariate Gaussian (Risken 1989) centered on $\mathbf{G}(\tau)\mathbf{x}'(t)$ with covariance matrix

$$\sigma(\tau) = \langle \mathbf{x}(t) \mathbf{x}^T(t) \rangle - \mathbf{G}(\tau) \langle \mathbf{x}(t) \mathbf{x}^T(t) \rangle \mathbf{G}^T(\tau). \quad (\text{A3})$$

Thus, the most probable pattern at time $t + \tau$ given $\mathbf{x}'(t)$ is simply

$$\hat{\mathbf{x}}(t + \tau) = \mathbf{G}(\tau)\mathbf{x}'(t), \quad (\text{A4})$$

justifying the name given to $\mathbf{G}(\tau)$.

Empirical normal modes (ENMs) $\{\mathbf{u}_\alpha\}$ are the time-independent eigenfunctions of $\mathbf{G}(\tau)$ and, hence, of \mathbf{B} :

$$\mathbf{G}(\tau)\mathbf{u}_\alpha = \mathbf{u}_\alpha g_\alpha(\tau), \quad (\text{A5})$$

$$\mathbf{B}\mathbf{u}_\alpha = \mathbf{u}_\alpha \beta_\alpha, \quad (\text{A6})$$

where $g_\alpha(\tau) = \exp(\beta_\alpha \tau)$ for the linear system. Here $\mathbf{G}(\tau_0)$ is calculated from the data at some lag τ_0 , and the eigenvalues β_α of \mathbf{B} are obtained from the eigenvalues $g_\alpha(\tau_0)$ of $\mathbf{G}(\tau_0)$. The eigenvalues β_α may be complex, indicating standing or traveling waves, or real, indicating exponentially decaying structures. The assumption of stationarity implies that the real part of β_α is negative, as must be the case for a stable system. The analysis cannot identify an unstable system; the requirement that the diagonal elements of $\sigma(\tau_0)$ be positive implies that the real part of $\exp(\beta_\alpha \tau_0)$ must be less than 1.

The transposed matrices $\mathbf{G}^T(\tau)$ and \mathbf{B}^T have the same eigenvalues as $\mathbf{G}(\tau)$ and \mathbf{B} , but different eigenfunctions $\{\mathbf{v}_\alpha\}$. If \mathbf{u}, \mathbf{v} are matrices, the α th column of which is $\mathbf{u}_\alpha, \mathbf{v}_\alpha$, respectively, then \mathbf{u} and \mathbf{v} form a biorthogonal set (Risken 1989):

$$\mathbf{u}\mathbf{v}^T = \mathbf{u}^T\mathbf{v} = \mathbf{1}, \quad (\text{A7})$$

where $\mathbf{1}$ is the unit matrix. The spectral decompositions of $\mathbf{G}(\tau)$ and \mathbf{B} are

$$\mathbf{G}(\tau) = \sum_{\alpha=1}^d \mathbf{u}_\alpha \exp(\beta_\alpha \tau) \mathbf{v}_\alpha^T \quad (\text{A8a})$$

and

$$\mathbf{B} = \sum_{\alpha=1}^d \mathbf{u}_\alpha \beta_\alpha \mathbf{v}_\alpha^T. \quad (\text{A8b})$$

Note that the matrix $\mathbf{G}(\tau)$ need not be calculated from the covariance structure of the data for every lag τ ; once $\mathbf{G}(\tau_0)$ has been determined, the spectral decomposition yields the Green function for any other value of τ :

$$\mathbf{G}(\tau) = \sum_{\alpha=1}^d \mathbf{u}_\alpha [g_\alpha(\tau_0)]^{\tau/\tau_0} \mathbf{v}_\alpha^T. \quad (\text{A9})$$

If the system represented by the data is described by linear dynamics (as is assumed by this analysis), the expression (A9) is independent of τ_0 . This property provides a test for linear dynamics: nonlinear dynamics will, in general, cause $\mathbf{G}(\tau)$ obtained with one value of τ_0 to vary from $\mathbf{G}(\tau)$ calculated with another value of τ_0 (see also Doobs 1942).

After \mathbf{B} is generated from the data, the noise covariance matrix \mathbf{Q} is obtained from the GFDR:

$$\mathbf{B}\langle \mathbf{x}(t)\mathbf{x}^T(t) \rangle + \langle \mathbf{x}(t)\mathbf{x}^T(t) \rangle \mathbf{B}^T + \mathbf{Q} = \mathbf{0}. \quad (\text{A10})$$

APPENDIX B

Continuous and Discrete Linear Markov Models

The basic assumption for principal oscillation pattern analysis and related analyses is that the dynamics

is governed by a linear Markov process. Both the continuous version (Hasselmann 1988; Penland 1989),

$$\frac{\partial X_i}{\partial t} = \sum_{j=1}^d B_{ij} X_j + \xi_i, \quad (\text{B1})$$

and the discrete version (e.g., von Storch et al. 1988),

$$X_i(t + \tau_0) = \sum_{j=1}^d A_{ij} X_j + \xi_i, \quad (\text{B2})$$

have been advanced as starting points in the development of the analysis. In fact, by identifying periods and e -folding times with the logarithms of the eigenvalues of \mathbf{A} , one has already assumed that Eq. (B2) is a discretization of (B1). In that case the residual term ξ in (B2) cannot be white. However, it is shown here that although ξ is strictly interpreted as the error between $\mathbf{X}(t + \tau_0)$ and its most probable prediction, it is closely related to the stochastic forcing and, for values of τ_0 small compared with the smallest decay time determined from \mathbf{A} (or, equivalently, \mathbf{B}), the noise matrix $\mathbf{Q} = \langle \xi \xi^T \rangle dt$ can be obtained from the average residual in Eq. (B2).

The discretization of Eq. (B1), in matrix form, is

$$\begin{aligned} \mathbf{X}(t + \tau_0) &= \exp(\mathbf{B}\tau_0)\mathbf{X}(t) + \exp[\mathbf{B}(\tau_0 + t)] \\ &\quad \times \int_t^{\tau_0+t} \exp(-\mathbf{B}t') \xi(t') dt'. \end{aligned} \quad (\text{B3})$$

This is equivalent to Eq. (B2) with the identification $\mathbf{A} = \exp(\mathbf{B}\tau_0)$ and with the residual ξ given by the integral term in Eq. (B3). The matrix \mathbf{A} is equivalent to the Green function that generates the most probable prediction of the system $\mathbf{X}(t + \tau_0)$ given $\mathbf{X}(t)$ (Penland 1989). The covariance matrix $\sigma = \langle \xi \xi^T \rangle$ of the residual is

$$\begin{aligned} \sigma &= \exp[\mathbf{B}(\tau_0 + t)] \left\{ \int_t^{\tau_0+t} dt' \int_t^{\tau_0+t} dt'' \right. \\ &\quad \times \exp(-\mathbf{B}t') \langle \xi(t') \xi^T(t'') \rangle \exp(-\mathbf{B}^T t'') \Big\} \\ &\quad \times \exp[\mathbf{B}^T(\tau_0 + t)]. \end{aligned} \quad (\text{B4})$$

Using the delta-correlated property of $\langle \xi(t') \xi^T(t'') \rangle = \mathbf{Q}\delta(t' - t'')$ and then substituting the GFDR (5),

$$\begin{aligned} \sigma &= \exp[\mathbf{B}(\tau_0 + t)] \left\{ \int_t^{\tau_0+t} dt' \exp(-\mathbf{B}t') [-\mathbf{B} \langle \mathbf{X} \mathbf{X}^T \rangle \right. \\ &\quad \left. - \langle \mathbf{X} \mathbf{X}^T \rangle \mathbf{B}^T] \exp(-\mathbf{B}^T t') \right\} \exp[\mathbf{B}^T(\tau_0 + t)] \\ &= \exp[\mathbf{B}(\tau_0 + t)] \left\{ \int_t^{\tau_0+t} dt' \frac{\partial}{\partial t'} [\exp(-\mathbf{B}t') \langle \mathbf{X} \mathbf{X}^T \rangle \right. \\ &\quad \left. \times \exp(-\mathbf{B}^T t')] \right\} \exp[\mathbf{B}^T(\tau_0 + t)] \\ &= \langle \mathbf{X} \mathbf{X}^T \rangle - \exp(\mathbf{B}\tau_0) \langle \mathbf{X} \mathbf{X}^T \rangle \exp(\mathbf{B}^T \tau_0). \end{aligned} \quad (\text{B5})$$

This is the covariance matrix of the error between the most probable prediction of $\mathbf{X}(\tau_0 + t)$ and the subsequent verification, given $\mathbf{X}(t)$ and the linear model, and correctly goes to zero with τ_0 (Penland 1989). However, for values of τ_0 small compared with the smallest decay time of \mathbf{B} (or, equivalently, \mathbf{A}),

$$\sigma \approx \langle \mathbf{XX}^T \rangle - \langle \mathbf{XX}^T \rangle - [\mathbf{B} \langle \mathbf{XX}^T \rangle + \langle \mathbf{XX}^T \rangle \mathbf{B}^T] \tau_0 + O(\mathbf{B}^2 \tau_0^2). \quad (\text{B6})$$

Again using the GFDR (5),

$$\sigma \approx \mathbf{Q} \tau_0. \quad (\text{B7})$$

Unfortunately, this method of extracting the covariance properties of the random forcing requires all decay times of the deterministic normal modes to be much larger than the lag τ_0 at which the analysis is performed. For most real systems, unless the decay rates are greatly distorted by filtering, this is not the case even when optimal accuracy of parameters is sacrificed to minimize τ_0 . [If the uncertainty in $\mathbf{G}(\tau_0)$ depends only weakly on τ_0 , an eigenvalue of \mathbf{B} is most accurately determined when its *e*-folding time is approximately equal to τ_0 ; see Penland and Ghil (1993).] In addition, calculating the covariance matrix of the residual involves much more computational effort than the two matrix multiplications required by Eq. (5). However, this appendix shows that the covariance properties of the driving noise may under certain conditions be found by examining the residual, if one chooses to follow that course.

REFERENCES

- Arnold, A., 1974: *Stochastic Differential Equations: Theory and Applications*. Wiley and Sons, 228 pp.
- Barnett, T. P., M. Latif, N. Graham, M. Flügel, S. Pazan, and W. White, 1993: ENSO and ENSO-related predictability. Part I: Prediction of equatorial Pacific sea surface temperature with a hybrid coupled ocean-atmosphere model. *J. Climate*, **6**, 1545–1566.
- Bell, T. L., 1980: Climate sensitivity and fluctuation dissipation: Some simple model tests. *J. Atmos. Sci.*, **37**, 1700–1707.
- Blumenthal, M. B., Y. Xue, and M. A. Cane, 1991: Predictability of an ocean/atmosphere model using adjoint model analysis. *New Developments in Predictability, ECMWF Workshop Proc.*, Reading, United Kingdom.
- Branstator, G., A. Mai, and D. Baumhefner, 1993: Identification of highly predictable flow elements for spatial filtering of medium- and extended-range numerical forecasts. *Mon. Wea. Rev.*, **121**, 1786–1802.
- Brunet, G., 1994: Empirical normal mode analysis of atmospheric data. *J. Atmos. Sci.*, **51**, 932–952.
- Cane, M. A., S. E. Zebiak, and Y. Xue, 1992: Model studies of the long term behavior of ENSO. *Proc. Workshop on Decade to Century Time Scales of Natural Climate Variability*, Washington DC, Natl. Acad. Sci.
- Doob, J. L., 1942: The Brownian movement and stochastic equations. *Ann. Math.*, **43**, 351.
- Epstein, E. S., 1969: Stochastic dynamic prediction. *Tellus*, **21**, 739–759.
- Fleming, R. J., 1993: The dynamics of uncertainty: Application to parameterization constants in climate models. *Climate Dyn.*, **8**, 135–150.
- García, A. L., and C. Penland, 1991: Fluctuating hydrodynamics and Principal Oscillation Pattern analysis. *J. Stat. Phys.*, **64**, 1121–1132.
- , M. M. Mansour, G. C. Lie, and E. Clementi, 1987: Numerical integration of the fluctuating hydrodynamic equations. *J. Stat. Phys.*, **47**, 209–228.
- García, R. R., and M. L. Salby, 1987: Transient response to localized episodic heating in the tropics. Part II: Far-field behavior. *J. Atmos. Sci.*, **44**, 499–530.
- Gardiner, C. W., 1985: *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. Springer Verlag, 442 pp.
- Ghil, M., and R. Vautard, 1991: Interdecadal oscillations and the warming trend in global temperature time series. *Nature*, **350**, 324–327.
- Hasselmann, K., 1976: Stochastic climate models. Part I. Theory. *Tellus*, **28**, 474–485.
- , 1988: PIPs and POPs—A general formalism for the reduction of dynamical systems in terms of Principal Interaction Patterns and Principal Oscillation Patterns. *J. Geophys. Res.*, **93**, 11 015–11 020.
- Horsthemke, W., and R. Lefever, 1984: *Noise-Induced Transitions: Theory and Applications in Physics, Chemistry and Biology*. Springer-Verlag, 318 pp.
- Kim, K. Y., and G. R. North, 1992: Seasonal cycle and second-moment statistics of a simple couple climate system. *J. Geophys. Res.*, **97**, 10 069–10 081.
- , and —, 1993: EOF analysis of surface temperature field in a stochastic climate model. *J. Climate*, **6**, 1681–1690.
- Klinker, E., and P. D. Sardeshmukh, 1992: The diagnosis of mechanical dissipation in the atmosphere from large-scale balance requirements. *J. Atmos. Sci.*, **49**, 608–627.
- Landau, L. D., and E. M. Lifshitz, 1959: *Fluid Mechanics*. Pergamon, 536 pp.
- Lau, K.-M., 1985: Elements of a stochastic-dynamical theory of the long-term variability of the El Niño-Southern Oscillation. *J. Atmos. Sci.*, **42**, 1552–1558.
- Leith, C. E., 1971: Atmospheric predictability and two-dimensional turbulence. *J. Atmos. Sci.*, **28**, 145–161.
- , 1975: Climate response and fluctuation dissipation. *J. Atmos. Sci.*, **32**, 2022–2026.
- , 1978: Predictability of climate. *Nature*, **276**, 352–355.
- , and R. H. Kraichnan, 1972: Predictability of turbulent flows. *J. Atmos. Sci.*, **29**, 1041–1058.
- Leung, L., and G. R. North, 1991: Atmospheric variability on a zonally symmetric land planet. *J. Climate*, **4**, 753–765.
- Lin, J. C., and J. N. Koshyn, 1987: A nonlinear stochastic low-order energy balance climate model. *Climate Dyn.*, **2**, 101–116.
- Marcuse, D., 1972: Derivation of coupled power equations. *Bell Sys. Tech. J.*, **51**, 229–237.
- Mikolajewicz, U., and E. Maier-Reimer, 1990: Internal secular variability in an ocean general circulation model. *Climate Dyn.*, **4**, 145–156.
- Mo, K., 1988: The predictability of low-frequency patterns in the NMC MRF. *Preprints, Eighth Conf. on Numerical Weather Prediction*, Baltimore, Amer. Meteor. Soc., 628–633.
- Mori, H., 1965: Transport, collective motion, and Brownian motion. *Progr. Theor. Phys.*, **33**, 423–455.
- North, G. R., R. E. Bell, and J. W. Hardin, 1993: Fluctuation dissipation in a general circulation model. *Climate Dyn.*, **8**, 259–294.
- Papanicolaou, G., and W. Kohler, 1974: Asymptotic theory of mixing stochastic ordinary differential equations. *Commun. Pure Appl. Math.*, **27**, 641–668.
- Penland, C., 1985: Acoustic normal mode propagation through a three-dimensional internal wave field. *J. Acoust. Soc. Amer.*, **78**, 1356–1365.

- , 1989: Random forcing and forecasting using principal oscillation pattern analysis. *Mon. Wea. Rev.*, **117**, 2165–2185.
- , and M. Ghil, 1993: Forecasting Northern Hemisphere 700-mb geopotential heights using principal oscillation patterns. *Mon. Wea. Rev.*, **121**, 2355–2372.
- , and T. Magorian, 1993: Prediction of Niño 3 sea surface temperatures using linear inverse modeling. *J. Climate*, **6**, 1067–1076.
- , M. Ghil, and K. Weickmann, 1991: Adaptive filtering and maximum entropy spectra, with application to changes in atmospheric angular momentum. *J. Geophys. Res.*, **96**, 22 659–22 671.
- Personick, D., 1971: Time dispersion in dielectric waveguides. *Bell Syst. Tech. J.*, **50**, 843–859.
- Risken, H., 1989: *The Fokker-Planck Equation, Methods of Solution and Applications*. Springer-Verlag, 472 pp.
- Rodriguez-Iturbe, I., D. Entekhabi, and R. Bras, 1991: Nonlinear dynamics of soil moisture at climate scales. 1. Stochastic analysis. *Water Resour. Res.*, **27**, 1899–1906.
- Rümelin, W., 1982: Numerical treatment of stochastic differential equations. *SIAM J. Numer. Anal.*, **19**, 604–613.
- Salby, M. L., and R. R. García, 1987: Transient response to localized episodic heating in the tropics. Part I: Excitation and short-time near-field behavior. *J. Atmos. Sci.*, **44**, 458–498.
- Slutz, R. J., S. J. Lubker, J. D. Hiscox, S. D. Woodruff, R. L. Jenne, D. H. Joseph, P. M. Steurer, and J. D. Elms, 1985: *Comprehensive Ocean-Atmosphere Data Set*. Release 1. NOAA Environmental Research Laboratory, Boulder, CO, 268 pp.
- Spall, M. A., 1993: Variability of sea surface salinity in stochastically forced systems. *Climate Dyn.*, **8**, 151–160.
- Storch, H. von, T. Bruns, I. Fischer-Bruns, and K. Hasselmann, 1988: Principal Oscillation Pattern analysis of the 30–60 day oscillation in a GCM equatorial troposphere. *J. Geophys. Res.*, **93**, 11 021–11 036.
- Syu, H.-H., D. Neelin, W. Weibel, and D. Gutzler, 1993: Tropical ocean-atmosphere interaction in a hybrid coupled GCM: Seasonal cycle and interannual variations. *Proc. Fourth Symp. on Global Change Studies*, Anaheim, CA, Amer. Meteor. Soc., 402–407.
- Tracton, M. S., K. Mo, W. Chen, E. Kalnay, R. Kistler, and G. White, 1989: Dynamical extended range forecasting (DERF) at the National Meteorological Center. *Mon. Wea. Rev.*, **117**, 1604–1635.
- Wong, E., and M. Zakai, 1965: On the convergence of ordinary integrals to stochastic integrals. *Ann. Math. Stat.*, **36**, 1560.
- Xu, J.-S., 1993: The joint modes of the coupled atmosphere-ocean system observed from 1967 to 1986. *J. Climate*, **6**, 816–838.