

## Predictability of a Stochastically Forced Hybrid Coupled Model of El Niño

CHRISTIAN ECKERT AND MOJIB LATIF

*Max-Planck-Institut für Meteorologie, Hamburg, Germany*

(Manuscript received 13 May 1996, in final form 5 November 1996)

### ABSTRACT

The El Niño–Southern Oscillation (ENSO) phenomenon is modeled as a stochastically driven dynamical system. This was accomplished by adding to a Hybrid Coupled Model (HCM) of the tropical Pacific ocean–atmosphere system a stochastic wind stress anomaly field that was derived from observations. The model exhibits irregular interannual fluctuations, whose space–time characteristics resemble those of the observed interannual climate variability in this region. To investigate the predictability of the model, the authors performed ensemble integrations with different realizations of the stochastic wind stress forcing. The ensembles were initialized at various phases of the model’s ENSO cycle simulated in a 120-yr integration with a particular noise realization. The numerical experiments indicate that the ENSO predictability is severely limited by the stochastic wind stress forcing. Linear stochastic processes were fitted to the restart ensembles in a reduced state space. A predictability measure based on a comparison of the stationary and the time-dependent probability distributions of the fitted linear models reveals an ENSO predictability limit of considerably less than an average cycle length.

### 1. Introduction

The El Niño–Southern Oscillation (ENSO) phenomenon is the dominant interannual natural climate fluctuation. It results from the interaction of ocean and atmosphere manifesting itself, among other things, in a quasiperiodic large-scale anomalous warming (El Niño) or cooling (La Niña) of surface waters in the eastern and central equatorial Pacific. Neelin et al. (1994) give a review of recent research activities and provide an extensive list of literature as well. Observed indices of ENSO—for example, as derived from the U.K. Meteorological Office (UKMO) Global Ice and Sea Surface Temperature (GISST) dataset, Parker et al. (1994)—show that ENSO extremes occur irregularly in time and with varying amplitudes. The typical recurrence time of El Niño is 2–7 yr, and its amplitudes, measured in terms of sea surface temperature anomalies (SSTA), averaged over the eastern equatorial Pacific NINO3 index region (5°N–5°S and 150°–90°W), typically range from 1° to 2°C. Additionally, ENSO is more or less tightly phase locked to the annual cycle. The degree of phase locking shows marked interdecadal variations as observed by Balmaseda et al. (1995), who performed a statistical analysis of the GISST NINO3 time series.

To explain the irregularity of ENSO, different hypotheses were offered. The first is linked to the characteristics exhibited by nonlinear dynamical systems.

Jin et al. (1994) and Tziperman et al. (1994), for instance, explain the irregularity of ENSO by nonlinear interactions between the annual cycle and the fundamental ENSO mode. Varying parameters of their simulation models yield ENSO frequencies that are rational fractions of the annual frequency. Irregularity, in turn, occurs by the overlapping of frequency locked regimes. Instead of varying internal parameters, Chang et al. (1994) studied the response of their coupled ocean–atmosphere model to different amplitudes of the seasonally varying part of heat flux forcing. Apart from chaotic regimes and transition regimes to chaos, they get three phase-locked regimes with 1-, 2-, and 3-yr period. Consequently, changes in the forcing amplitude with time may render ENSO irregular. A second candidate for rendering ENSO irregular is stochastic forcing. Recently, Kleeman and Power (1994) studied the influence of random wind stress perturbations on a coupled model. Forecast ensembles initialized in the 1970s and 1980s showed considerable error growth after initialization of the individual predictions when random forcing was included. In this paper, we study the influence of stochastic wind stress forcing on ENSO as well. We designed a stochastically forced ocean–atmosphere system that mimics the observed statistics to some extent. This approach is related to Hasselmann (1976), who pointed out the importance of stochastic forcing for inducing climate variability. A similar study to ours was performed by Blanke et al. (1997, this issue). Although they use a different approach to get an estimate of the stochastic wind forcing, their results agree well with ours. In both models, the sensitivity to the random per-

---

*Corresponding author address:* Christian Eckert, Max-Planck-Institut für Meteorologie, Bundesstraße 55, D-20146 Hamburg, Germany.

turbations introduced is quite large. The models' predictability is reduced to less than one average cycle length.

At present, prediction models of different degrees of complexity, ranging from purely statistical to fully coupled ocean–atmosphere general circulation models, are used to forecast ENSO. Barnston et al. (1994) and Latif et al. (1994) give reviews on this topic. The forecast skills of the models indicate that, currently, ENSO is predictable about 1 yr ahead. ENSO predictability is limited by various reasons. Among these are the quality of the forecast models, the ability of getting good initial states, and a good estimate of the influence of random perturbations that disturb the system perpetually. Given credence to the forecast model as such, one may concentrate on the latter two. The amplification of initial errors can be caused by different mechanisms such as unstable modes on which the error structure projects (Goswami and Shukla 1991) or the nonself-adjoint nature of the linearized system dynamics (Blumenthal 1991). The latter yields rapid initial error growth even in asymptotically stable linear dynamical systems. To get more insight into the physical processes that are relevant for initial error growth Moore and Kleeman (1996) and Chen et al. (1997) studied the fastest growing perturbations in coupled models, the singular vectors. Moore and Kleeman ascribe initial error growth to a mechanism that is associated with penetrative convection anomalies in the atmosphere. They also note a seasonal variation of error growth and a dependence on the ENSO cycle itself. Chen et al. observe similar characteristics in their model, although differing considerably in detail from Moore and Kleeman. They propose as a possible physical explanation an ocean–atmosphere instability that is linked to changes in ocean dynamics in response to changes in the atmospheric wind stress field. Certainly, this topic needs further investigation especially with regard to the definition of the error norm and the sensitivity of results to the model physics.

Our approach, like Kleeman and Power (1994), focuses on the random perturbations and complements the studies concerned with uncertain initial conditions. Our aim is to get an estimate of the ENSO predictability limits. We performed ensemble integrations from different phases of the ENSO cycle simulated by our stochastically forced ocean–atmosphere model. The individual ensembles start from the same initial state but were performed with different noise realizations. We approximate the time evolution of the ensemble by a low-dimensional linear stochastic process that is constructed in such a way that it simulates the evolution of the ensemble mean and the ensemble covariance over a finite time period after initialization of the model. A predictability measure can now be defined by comparing the stationary and the time-dependent probability distributions of the fitted linear stochastic process.

The paper is organized as follows. Descriptions of the stochastically forced ocean–atmosphere model and

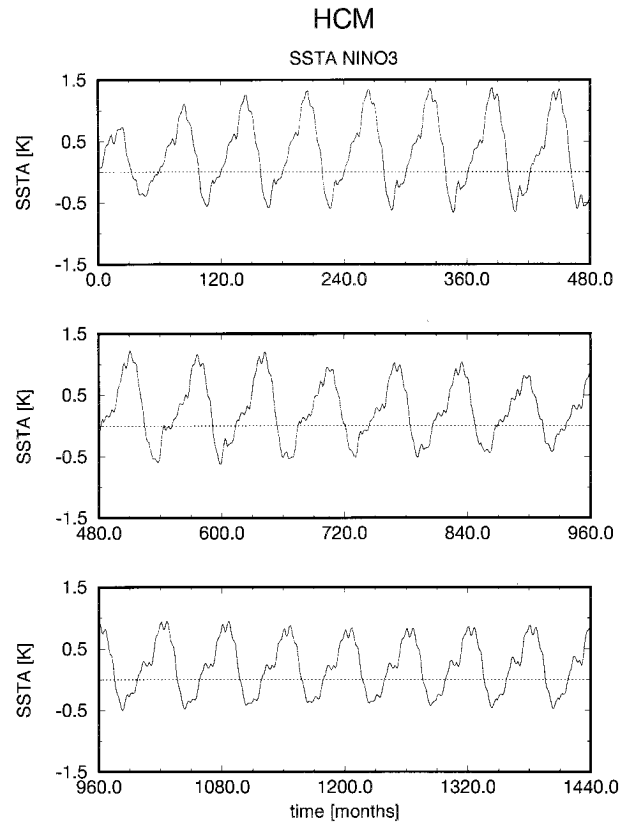


FIG. 1. Time series of anomalous sea surface temperature (SSTA), averaged over the NINO3 region ( $150^{\circ}$ – $90^{\circ}$ W,  $5^{\circ}$ S– $5^{\circ}$ N), obtained by integrating the Hybrid Coupled Model (HCM) for 120 yr.

ensemble integrations performed are given in sections 2 and 3, respectively. The optimal fits of linear stochastic processes in a reduced state space to the ensemble data are presented in section 4. We address the predictability of the fitted linear models in section 5 and conclude the paper with a summary and discussion of our results in section 6.

## 2. The model

The coupled ocean–atmosphere model used to simulate ENSO consists of an ocean general circulation model (OGCM) of the equatorial Pacific and a statistical atmosphere model that was derived from observations. Our modeling approach is similar to the one described by Barnett et al. (1993). However, there are slight differences in the model formulation. Therefore, and for better readability of our paper, we give a short description of the coupled model and its constituents.

### a. The ocean model

The OGCM is a further development of the model described in Latif (1987). The domain extends from  $30^{\circ}$ N to  $30^{\circ}$ S and from  $130^{\circ}$ E to  $70^{\circ}$ W. The vertical

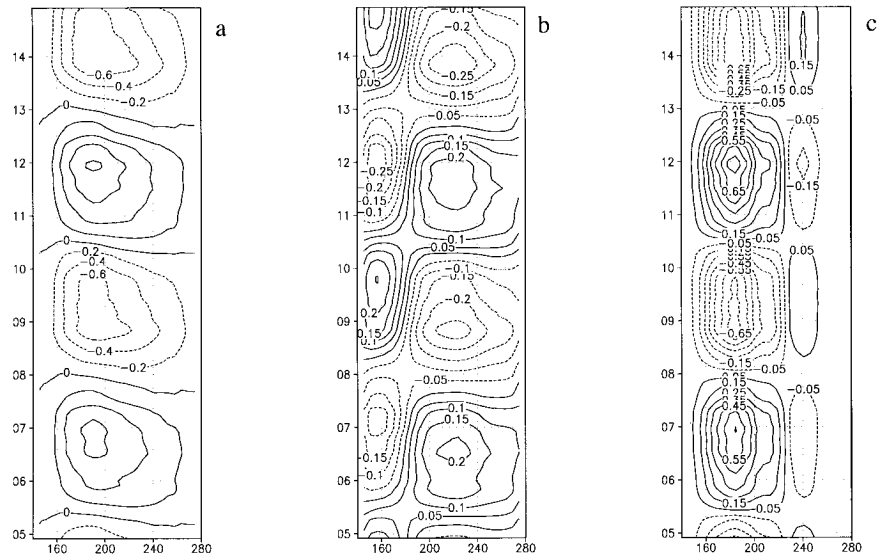


FIG. 2. Longitude time section along the equator from year 5 to 15 during the 120-yr integration of the HCM. The combined scaled fields of SSTA, anomalous sea level, and anomalous zonal wind stress were subjected to a POP analysis. Shown are the projections of the individual fields onto the dominant POP mode. The leading POP mode has a period of 62 months, an  $e$ -folding time of 450 months, and accounts for 83% of the variance of the dataset. To get physical units, the depicted fields have to be multiplied by factors of (a) 3.46 K, (b) 36.6 cm, and (c)  $7.3 \times 10^{-2}$  Pa, respectively.

mixing is Richardson number dependent. Prognostic variables are the fields of horizontal currents, the three-dimensional temperature, and sea surface elevation. The model is based on the primitive equations, formulated

on an equatorial  $\beta$  plane, and simplified by making use of the hydrostatic and Boussinesq approximations. The bottom is flat at a depth of 4000 m. Eastern and western coastlines are realistically chosen, whereas the southern

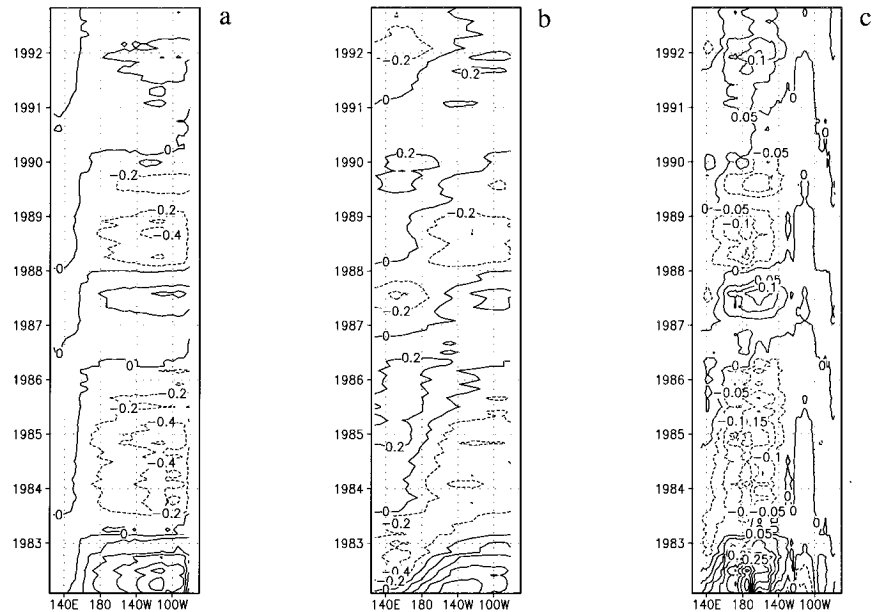


FIG. 3. Longitude time section analogous to Fig. 2. In this case, the NMC reanalysis data for SSTA and anomalous sea level were combined with the zonal wind stress anomalies of the FSU dataset. The individual fields were scaled with the factors given below before they were pooled. The dominant POP mode has a period of 52 months, an  $e$ -folding time of 16 months, and accounts for 15% of the variance of the dataset. The individual fields have to be multiplied by factors of (a) 7.68 K, (b) 51.5 cm, and (c)  $2.5 \times 10^{-1}$  Pa to get physical units.

**THE STOCHASTICALLY FORCED HYBRID COUPLED MODEL**

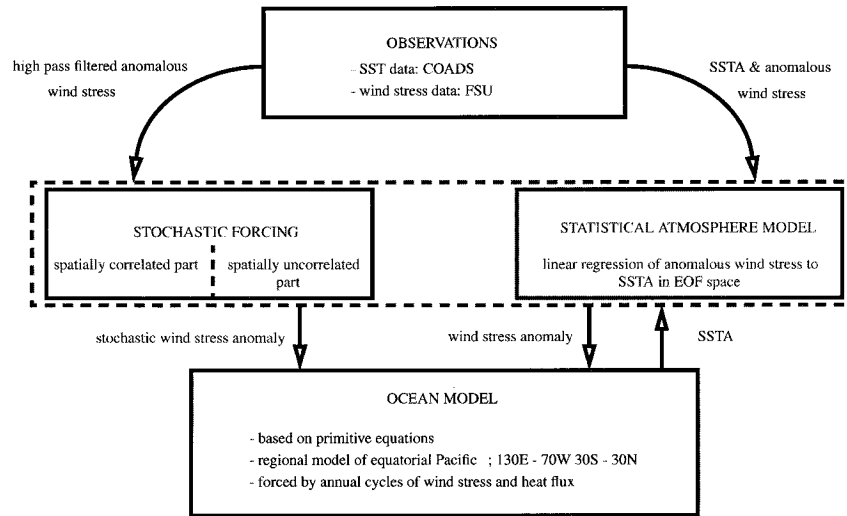


FIG. 4. Schematic picture of the stochastically forced HCM we use in our predictability study.

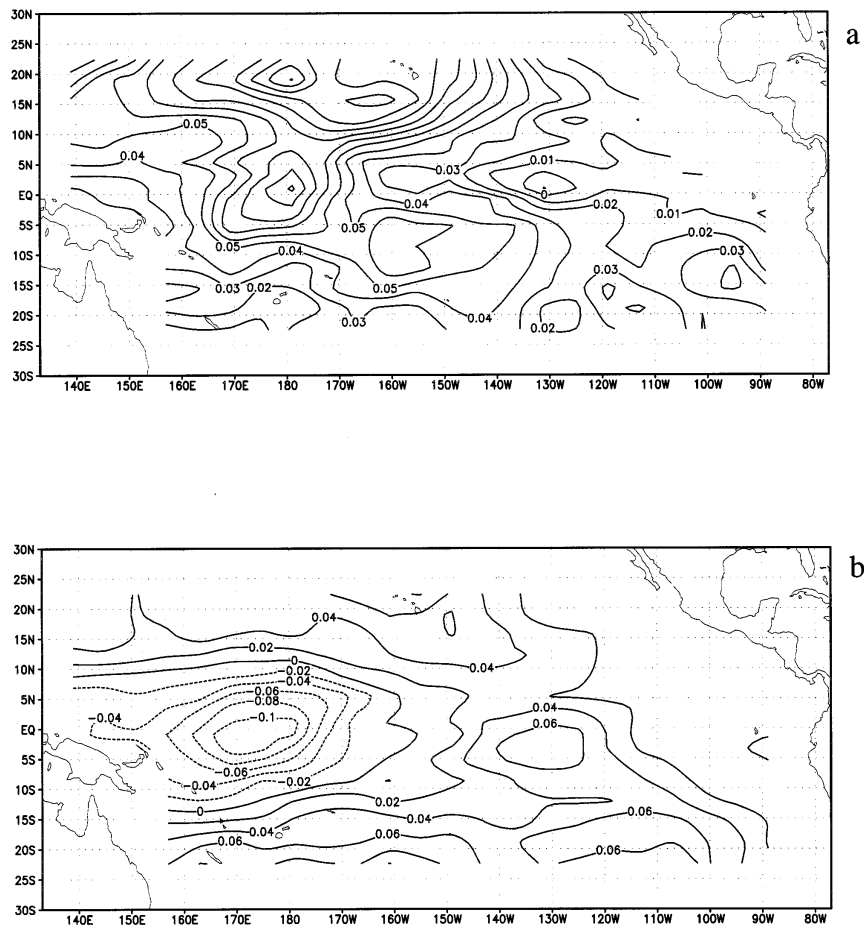


FIG. 5. The first (a) and the second (b) EOF in units of Pa of the high-pass filtered wind stress anomalies of the FSU dataset. They explain 10.3% and 7.0% of the variance, respectively.

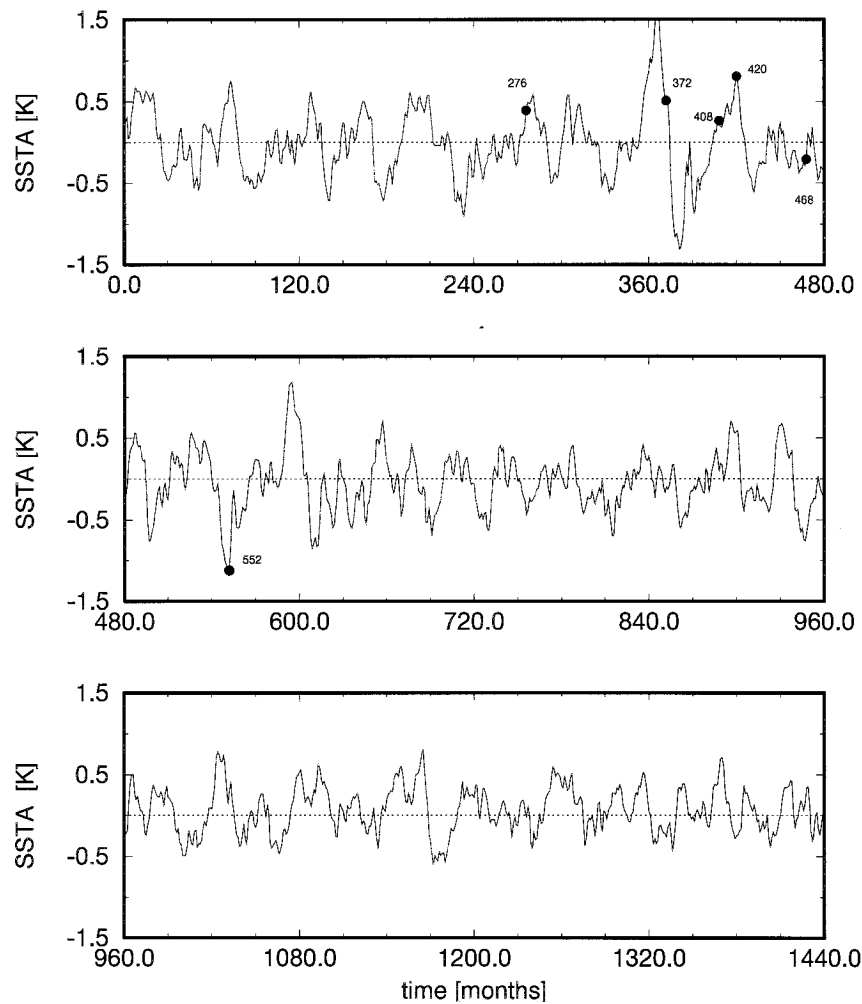


FIG. 6. Time series of SSTA averaged over the NINO3 index region obtained by integrating the stochastically forced HCM for 120 yr. The labeled dots mark the restart months we used to generate ensembles of trajectories with different realizations of the stochastic part.

and the northern boundaries are implemented as solid walls. The zonal resolution is 670 km, and the meridional resolution varies from 25 km at the equator to 400 km at the southern and northern boundaries. In the vertical, the model has 13 levels, with 10 in the upper 300 m. The ocean model is driven by annual cycles of wind stress and heat flux. The latter enters as a Newtonian relaxation term to climatology with a time constant of 30 days.

#### b. The statistical atmosphere model

The statistical atmosphere model relates sea surface temperature anomalies (SSTA) linearly to wind stress anomalies via regression. The regression matrix is derived from observations during the period 1970–85. The SST data were taken from the dataset of Reynolds (1988) and the wind stress from The Florida State University (FSU) dataset (Goldenberg and O'Brien

1981). The linear regression is done in EOF space, retaining the first 5 EOFs of the anomalies, which account for 76%, 41%, and 35% of the field variance in the SST, zonal, and meridional wind stress anomalies, respectively. The dominant modes of interannual variability are found to be well captured by the EOF truncation. To compensate for systematic deficiencies of the OGCM, its simulated SSTA is corrected using a regression matrix based on the leading 5 EOFs of observed and simulated SSTA, the latter being obtained by forcing the ocean model by the observed wind stresses.

Let the observed and simulated SSTA be denoted by  $T_{\text{obs}}(x, t)$  and  $T_{\text{sim}}(x, t)$ , respectively, and a component of the observed wind stress anomalies, either zonal or meridional, by  $\tau_{\text{obs}}(x, t)$ . The estimated component of this wind stress anomaly is denoted by  $\tau_{\text{est}}(x, t)$ . The EOF expansions of the observed and simulated fields are

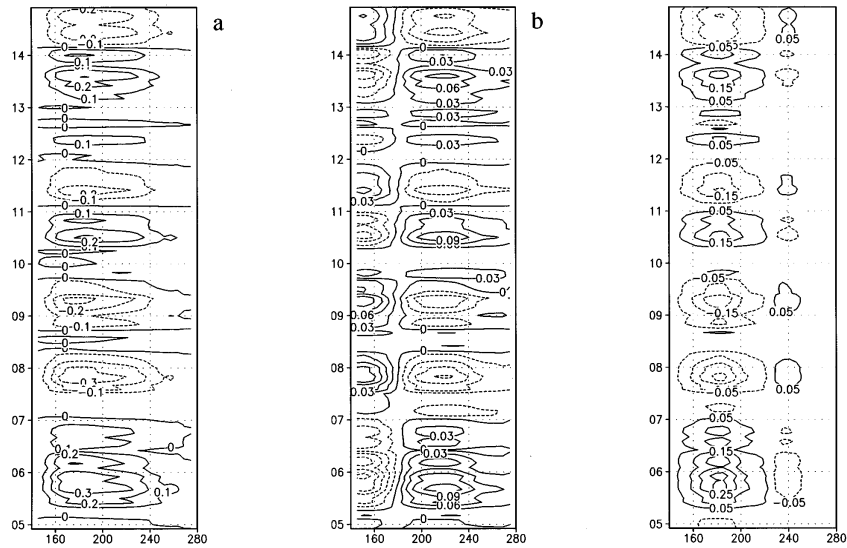


FIG. 7. Longitude time section analogous to Fig. 2 using the data of the 120-yr integration of the stochastically forced Hybrid Coupled Model. The individual fields were scaled with the factors given below before they were pooled. The dominant POP mode has a period of 61.9 months, an  $e$ -folding time of 22.6 months, and accounts for 32.4% of the variance of the dataset. The individual fields have to be multiplied by factors of (a) 4.74 K, (b) 49.3 cm, and (c)  $1.3 \times 10^{-1}$  Pa to get physical units.

$$\begin{aligned}
 T_{\text{obs}}(x, t) &= \sum_n \alpha_n(t) e_n(x), \\
 \tau_{\text{obs}}(x, t) &= \sum_m \beta_m(t) f_m(x), \quad \text{and} \\
 T_{\text{sim}}(x, t) &= \sum_l \gamma_l(t) g_l(x).
 \end{aligned}
 \tag{1}$$

$$\mathbf{C}_{mn}^{(1)} = \frac{\langle \beta_m \alpha_n \rangle}{\langle \alpha_n^2 \rangle} \quad \text{and} \quad \mathbf{C}_{nl}^{(2)} = \frac{\langle \alpha_n \gamma_l \rangle}{\langle \gamma_l^2 \rangle}.
 \tag{2}$$

Angle brackets signify time averages. The estimated wind stress anomaly may now be obtained by the simulated SSTA via

$$\tau_{\text{est}}(x, t) = \sum_m \beta_{\text{est},m}(t) f_m(x)
 \tag{3}$$

and

$$\beta_{\text{est},m}(t) = \sum_{n,l} \mathbf{C}_{mn}^{(1)} \mathbf{C}_{nl}^{(2)} \gamma_l(t),
 \tag{4}$$

Finally, we multiply the estimated wind stress anomaly by a factor, usually termed coupling strength, which was assigned a value of 1.4. One reason for doing so is the relatively weak dynamical response of our ocean model to an imposed wind stress forcing, especially in the eastern equatorial Pacific. Most important, however, we chose the specific value of 1.4 to run the coupled model in a regime

The linear regressions yield matrices  $\mathbf{C}^{(1)}$  and  $\mathbf{C}^{(2)}$  whose entries are given by

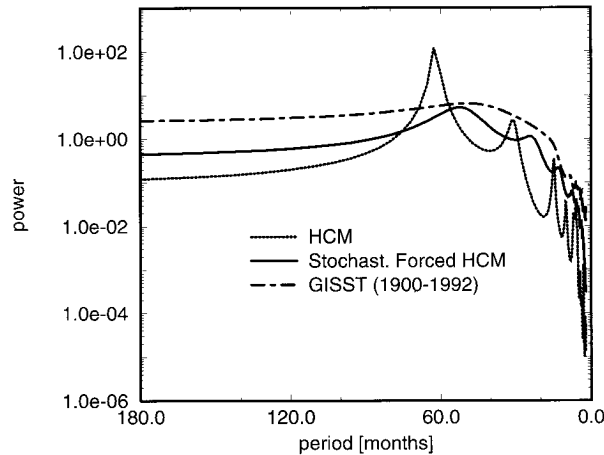


FIG. 8. Maximum entropy spectra of the NINO3 time series of the UKMO GISST dataset lasting from 1900 to 1992, the Hybrid Coupled Model (using the data corresponding to Fig. 1), and the stochastically forced Hybrid Coupled Model (using the data corresponding to Fig. 6).

TABLE 1. List of the restart months corresponding to Fig. 6. The stochastically forced Hybrid Coupled Model was reinitialized at these states and integrated forward for 72 months with different realizations of the stochastic part.

Restart month	Phase of model ENSO cycle
276	Onset of El Niño
372	El Niño to La Niña
408	Onset of El Niño
420	El Niño extreme
468	Quiet phase
552	La Niña extreme

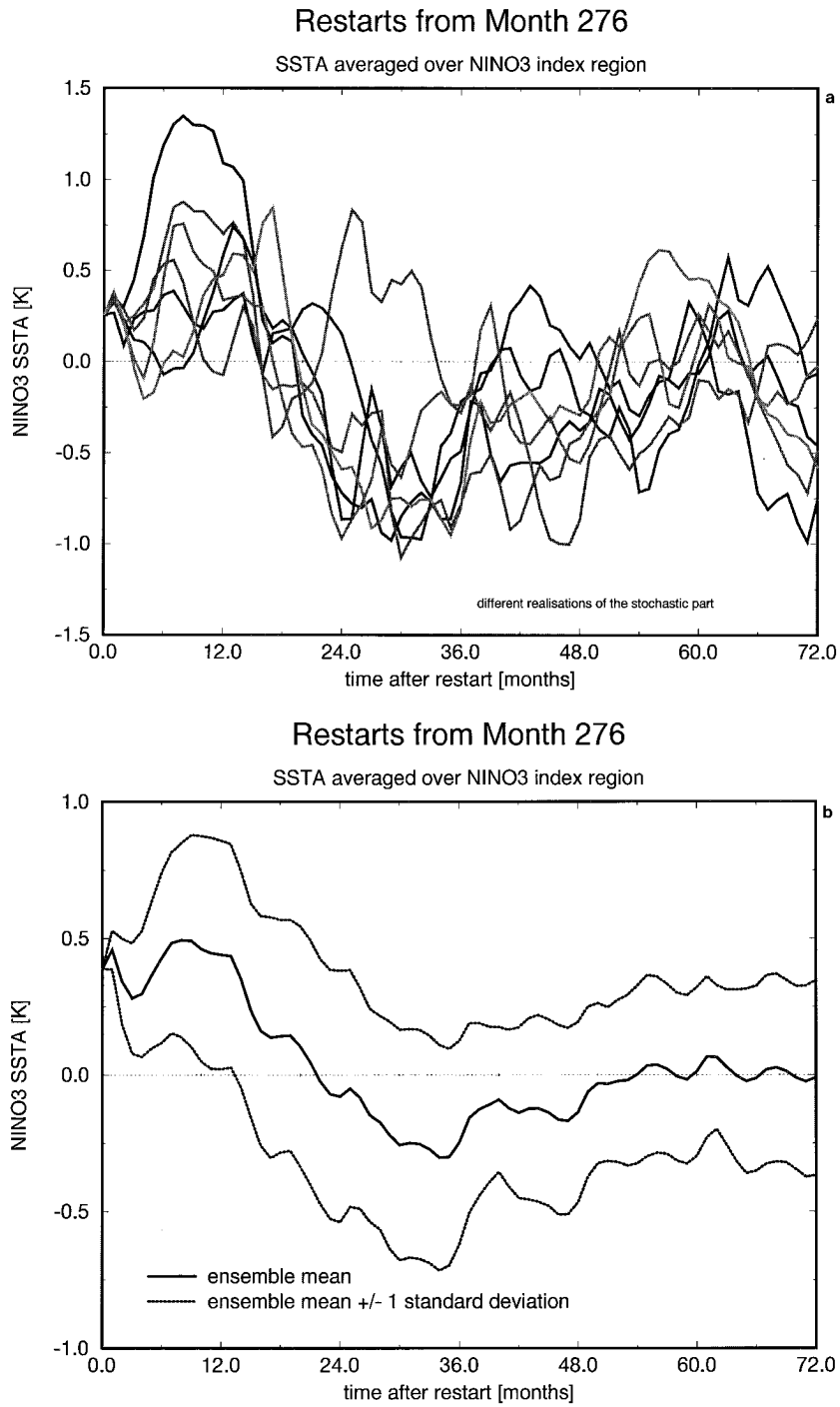


FIG. 9. NINO3 time series of SSTA of the restart experiment initialized at month 276 of the control integration of the stochastically forced HCM. In (a) we show some individual realizations, (b) depicts the ensemble mean and the ensemble mean  $\pm$  1 standard deviation, and in (c) the standard deviation is shown separately.

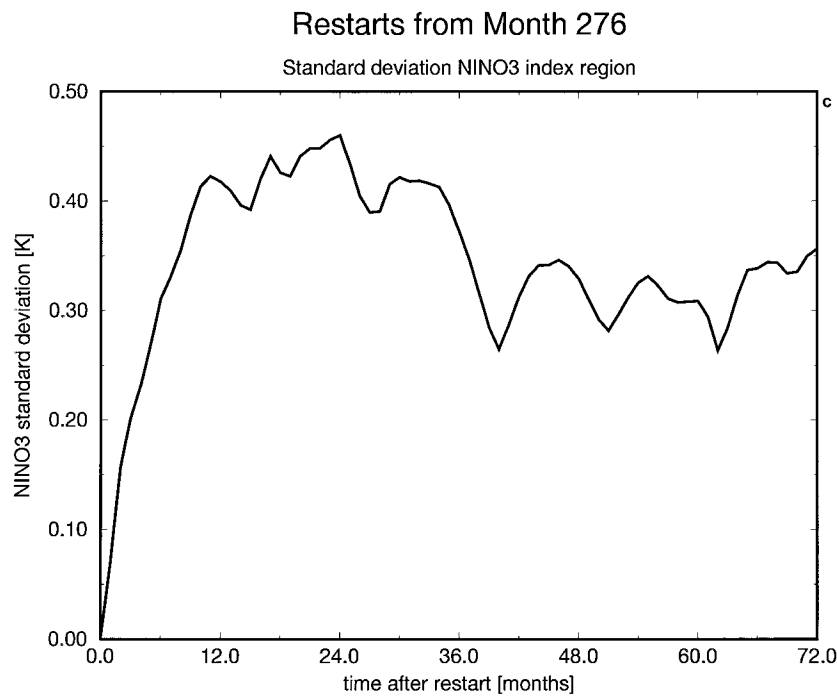


FIG. 9. (Continued)

where it exhibits a regular oscillation. Coupling is done on a monthly basis with linear interpolation in between. Forced by observed SSTA, the above statistical atmosphere model predicts wind stress anomalies that are well correlated to wind stress anomalies observed during the same time interval. In the central equatorial Pacific, the correlations are highest, with values up to 0.8 in the date-line region (Flügel 1994). Before these correlations were computed, the observed wind stress anomalies were filtered with a 5-month running mean.

A coupled control integration yields a self-sustained oscillation with a period of approximately 60 months. In Fig. 1 we display the time series of SSTA averaged over the NINO3 index region. The model underestimates the amplitudes of the extremes slightly. The spatial characteristics of the oscillation are revealed by a POP (Principal Oscillation Patterns) analysis of the combined fields of SSTA, anomalous zonal wind stress, and anomalous sea surface elevation, which is a good proxy for upper-ocean heat content. The POP analysis technique was originally proposed by Hasselmann (1988). For a review of the technique, its applications, and possible extensions, see von Storch et al. (1995). In essence, a multivariate first-order Markov process is fitted to the dataset, minimizing the one time step prediction error. The normal modes of the process serve as a new basis to describe the data. They are characterized by an  $e$ -folding time and an oscillation period in the case of a complex eigenvalue of the fitted linear system, or an  $e$ -folding time only if the corresponding eigenvalue is real.

The leading POP mode derived from the control integration has an oscillation period of 62 months, an  $e$ -folding time of 450 months, and accounts for 83% of the variance in the combined dataset. The individual fields were scaled by the maximum value attained during the control integration before they were pooled. The result of this POP analysis indicates that practically all information contained in the dataset is well modeled by one oscillatory mode. Figure 2 shows longitude time sections along the equator of the anomaly fields projected onto the leading POP mode. SSTA and anomalous zonal wind stress are characterized by standing patterns, whereas the anomalous sea level propagates slowly from west to east. This behavior is in accordance with the delayed action oscillator scenario proposed by Schopf and Suarez (1988), in which equatorial wave propagation plays an important role. The ocean is not in equilibrium with the wind stress forcing (Cane and Sarachik 1981), which adds a delayed response in the time evolution of SST. Many wave modes contribute to the ocean adjustment, giving rise to the relatively slow phase propagation.

For comparison, we performed the same analysis with observational data. We combined the SST and sea level anomalies of the National Meteorological Center (NMC, now known as the National Centers for Environmental Prediction) reanalysis dataset (see Ji et al. 1995) with the FSU zonal wind stress anomalies. The NMC data were available for the time period February 1982–December 1992 only. The leading POP mode derived from the observations has an  $e$ -folding time of



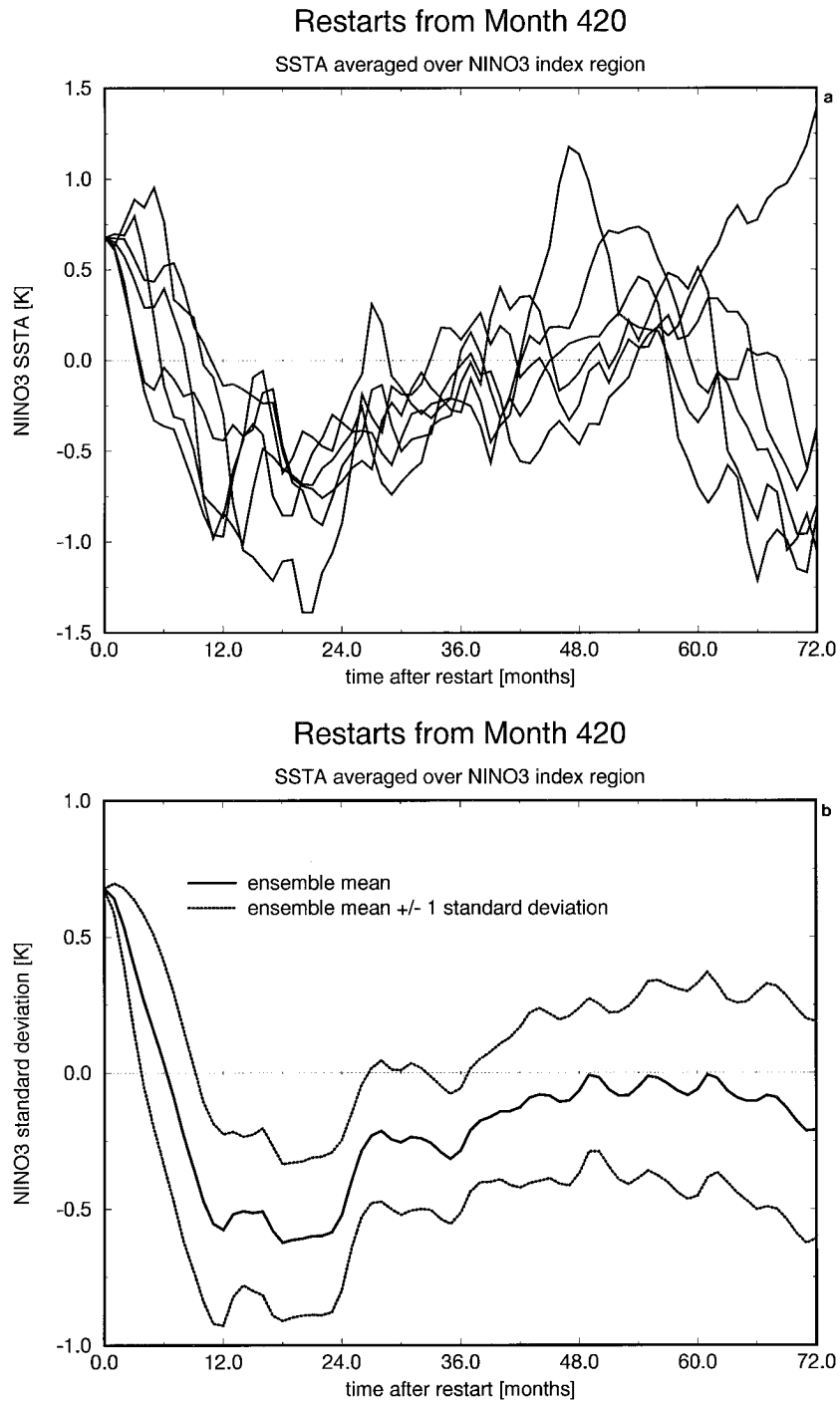


FIG. 10. NINO3 time series of SSTA analogous to Fig. 9 for restart month 420: (a) individual realizations, (b) ensemble mean  $\pm$  1 standard deviation, (c) the standard deviation of the ensemble. This restart is characterized by a rapid initial increase in variance (standard deviation).

16 months, an oscillation period of 52 months, and accounts for 15% of the variance. Longitude time sections analogous to those shown in Fig. 2 for the HCM data are displayed in Fig. 3. The HCM simulation is

in qualitative agreement concerning the spatial characteristics. The SST anomalies, however, are simulated too far west. Note that the time interval for the analysis of the observations leads presumably to a biased result

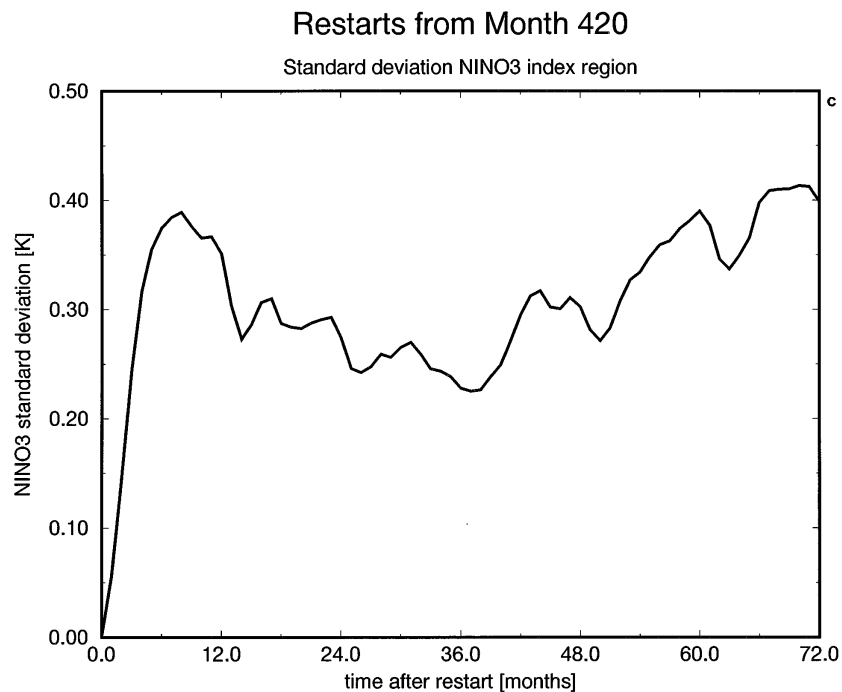


FIG. 10. (Continued)

due to the regular occurrence of ENSO extremes in the 1980s.

### c. The stochastically forced HCM

The FSU wind stress dataset shows a high degree of variability at timescales that are short compared to a typical ENSO period (e.g., Kleeman and Power 1994). To investigate the effects of stochastic forcing on the coupled model variability, we added wind stress noise to the HCM in a second integration. A schematic diagram of the extension of the HCM by the stochastic component is given in Fig. 4. The noise was constructed as follows. The observed wind stress anomalies were high-pass filtered, retaining variability with timescales less than 12 months fully and with decreasing weights up to 16 months to reduce end effects. As noted by Kleeman and Power (1994), the high-pass filtered part of the anomalous wind stress field is consistent with white noise that was truncated in this manner. Next, we combined the 10 leading EOFs of the high-pass filtered wind stress anomaly linearly with random amplitudes according to the variance of their corresponding principal components. The residual part of the high-pass filtered wind stress anomaly—that is, the sum of the remaining EOFs multiplied with their expansion coefficients—is used to construct an additional part of the noise estimate. We assume this field to be uncorrelated in space and time. Consequently, we add a further local random wind stress anomaly. Its variance equals the variance of the residual part. To emphasize the way the

random wind stress forcing was constructed, the stochastic wind stress forcing box in Fig. 4 is split into two parts. We chose this approach because the leading EOFs still exhibit considerable spatial correlations (see Fig. 5, which shows the first 2 EOFs). Both the first and second EOF exhibit large-scale structures in areas that are known to be important for the generation of ENSO.

The model oscillates irregularly when the noise forcing is included, as witnessed by the NINO3 time series of SSTA (Fig. 6), and the amplitudes of the extremes vary considerably. This 120-yr integration with a particular noise realization will henceforth be termed noise run. We infer the spatial characteristics again by performing a POP analysis of the combined fields of SSTA, anomalous sea level, and anomalous zonal wind stress, and subsequent projection onto the leading POP mode. Figure 7 shows the corresponding longitude time section along the equator. The leading POP mode has a period of 61.9 months, an  $e$ -folding time of 22.6 months, and accounts for 32.4% of the variance of the dataset. In the noise run, the propagation characteristics of the sea level anomalies are, however, less pronounced and more consistent with the NMC reanalysis (Fig. 3).

Comparing the spectra of the NINO3 time series obtained from the coupled experiments with a long observational record reveals the improvement that is obtained by adding random anomalies to the wind stress field (Fig. 8). The SST observations were taken from the GISST dataset, covering the period 1900–92. We

estimated the spectra by applying the maximum entropy method from the SSA toolkit, Dettinger et al. (1995), which yielded stable spectral estimates for a range of orders, typically 30–40, of the fitted autoregressive models. The spectral peaks of the HCM are broadened substantially when noise is added. The spectrum of the noise run NINO3 time series is in qualitative agreement with the GISST record. However, it does not simulate enough variability at decadal timescales, which may be explained by the lack of interaction with the global-scale ocean circulation. This problem, however, should not affect the result of our predictability study.

### 3. Restart experiments

To study the model variability further, we reinitialized the coupled model at selected months of the noise run, corresponding to different phases of the model ENSO cycle. The restart months are listed in Table 1 and marked in Fig. 6. We generated ensembles of 216 trajectories for each restart month by integrating the model forward in time for 72 months with different realizations of the stochastic part. We display in Fig. 9 the time evolution of one particular ensemble initialized at month 276, which corresponds to an onset of a model El Niño event. Six particular NINO3 SSTA trajectories are shown in Fig. 9a, whereas Fig. 9b shows the mean and standard deviation of the complete ensemble. The standard deviation quickly saturates after about 12 months (Fig. 9c). The time evolution of the ensemble mean and the ensemble standard deviation depend on the restart month. To illustrate this, Fig. 10 shows an ensemble initialized at the peak phase of a model El Niño. All the ensemble members evolve relatively quickly into La Niña events. However, they are shifted in phase with respect to each other. This leads initially to a strong increase in the standard deviation, while it drops again after all the realizations have reached the cold state. After this, the spread among the ensemble members evolves in the same manner as in the previous case.

In the following sections, we will use these ensemble data to infer the predictability limits of ENSO as derived from our stochastically forced coupled ocean–atmosphere model.

### 4. Linear stochastic process

To quantify the influence of the random perturbations on the model's interannual variability by means of a simple approximating system, we use the ensemble data to derive a linear stochastic model for each restart ensemble separately. Such an approach is justified by the success of the previous POP analyses in extracting the dominant variability modes in the observations and the model data.

The time evolution of a linear stochastic system is given by the stochastic differential equation

$$\dot{x}(t) = \mathbf{A}x(t) + \mathbf{B}\eta(t). \quad (5)$$

The vector  $x(t)$  represents the state of the  $n$  dimensional system at time  $t$ ,  $\mathbf{A}$  and  $\mathbf{B}$  denote  $n \times n$  matrices, and  $\eta(t)$  is meant to be an  $n$  dimensional Gaussian random vector with mean 0 and the identity matrix as covariance matrix. In our application,  $\mathbf{A}$  and  $\mathbf{B}$  are assumed to be constant in time. This implies that the diffusion matrix of the random process is given by  $\mathbf{B}\mathbf{B}^T$ , the superscript denoting the transpose of  $\mathbf{B}$ . The time evolution of the ensemble mean  $\langle x(t) \rangle$  and the ensemble covariance matrix  $\mathbf{C}_{ij}(t) = \langle [x_i(t) - \langle x_i(t) \rangle] [x_j(t) - \langle x_j(t) \rangle] \rangle$  then obey the differential equations

$$\langle \dot{x} \rangle = \mathbf{A}\langle x \rangle, \quad \langle x(0) \rangle = x_0 \quad (6)$$

and

$$\dot{\mathbf{C}} = \mathbf{A}\mathbf{C} + \mathbf{C}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T, \quad \mathbf{C}(0) = \mathbf{C}_0, \quad (7)$$

not explicitly denoting the time dependence of  $\langle x \rangle$  and  $\mathbf{C}$  for notational convenience. For a derivation of Eqs. (6) and (7) see, for instance, van Kampen (1992). Here,  $\mathbf{C}(0)$  vanishes since the ensemble members start from the same initial state. To obtain an optimal linear model in a least squares sense, we define a cost function  $F(\mathbf{A}, \mathbf{B})$ , which measures the misfit between the linear model characterized by the entries of  $\mathbf{A}$  and  $\mathbf{B}$  and the ensemble data over a time interval  $[1 \dots T^*]$ :

$$\begin{aligned} F(\mathbf{A}, \mathbf{B}) = & \sum_{t=1}^{T^*} [(\langle x \rangle - \langle x \rangle_{\text{ens}})^T \mathbf{M}_{1,t} (\langle x \rangle - \langle x \rangle_{\text{ens}}) \\ & + (\mathbf{C} - \mathbf{C}_{\text{ens}})^T \mathbf{M}_{2,t} (\mathbf{C} - \mathbf{C}_{\text{ens}})]. \quad (8) \end{aligned}$$

Here,  $\mathbf{C}$  denotes a column vector with its entries given by the main diagonal and upper triangle of the covariance matrix. The subscript ens signifies the estimates for the mean and covariance obtained from the restart ensembles. The weighing matrixes  $\mathbf{M}_{1,t}$  and  $\mathbf{M}_{2,t}$  were derived from the ensemble assuming normal variates

$$\begin{aligned} \mathbf{M}_{1,t} &= N\mathbf{C}_{\text{ens}}^{-1} \\ \mathbf{M}_{2,t} &= \text{cov}(\mathbf{C}_{\text{ens},ij}, \mathbf{C}_{\text{ens},kl})^{-1} \\ &= \left[ \frac{1}{N} (\mathbf{C}_{\text{ens},il} \mathbf{C}_{\text{ens},jk} + \mathbf{C}_{\text{ens},ik} \mathbf{C}_{\text{ens},jl}) \right]^{-1}, \quad (9) \end{aligned}$$

where  $N$  is equal to the ensemble size. For a derivation of the second matrix see Kendall et al. (1983). To minimize  $F$  in parameter space, one can choose different methods. We chose a software tool, developed by Giering (1995) and Giering and Kaminski (1996). This tool generates the adjoint of the above model, which provides the gradient of  $F$  with respect to the entries of  $\mathbf{A}$  and  $\mathbf{B}$ . A subsequent minimization algorithm based on a standard software package (Numerical Algorithms Group 1991) determines the minimum of the cost function.

It remains to specify the dimension of the state space for the linear model. To get a reduced state space representation of the stochastically forced HCM, we per-

TABLE 2. Result of the EOF analysis of the combined fields of SSTA, anomalous sea level, and anomalous zonal wind stress for the 120-yr integration of the stochastically forced HCM. The individual fields were scaled with the maximum absolute value they attained during the noise run before they were pooled.

EOF No.	Eigenvalue	Explained variance	Cumulative expl. var.
1	8.26	35.91	35.91
2	1.21	5.25	41.16
3	1.08	4.68	45.84
4	0.78	3.38	49.22
5	0.62	2.68	51.90
6	0.52	2.27	54.17

formed a combined EOF analysis of SSTA, sea level anomalies, and zonal wind stress anomalies of the noise run. Before pooling, the individual anomaly fields were scaled by their maximum absolute value attained during the noise run. We retain four, five, or six EOFs to derive the linear model. Table 2 lists the eigenvalues and explained variances of the leading 6 EOFs. The first two EOFs contain practically all the information about the interannual variability of the model. Therefore, in Fig. 11 we show their expansion coefficients resulting from a fit to a linear stochastic model of the above kind. In the appendix we list the optimization results for all the six restart experiments. We fitted various time spans  $T^*$  and dimensionalities (4, 5, 6) for each restart ensemble separately, to obtain fits as close as possible to the individual ensemble data. We list the dominant eigenmode of the fitted system matrices  $\mathbf{A}$  and compare the variances simulated to those of the stochastically forced HCM.

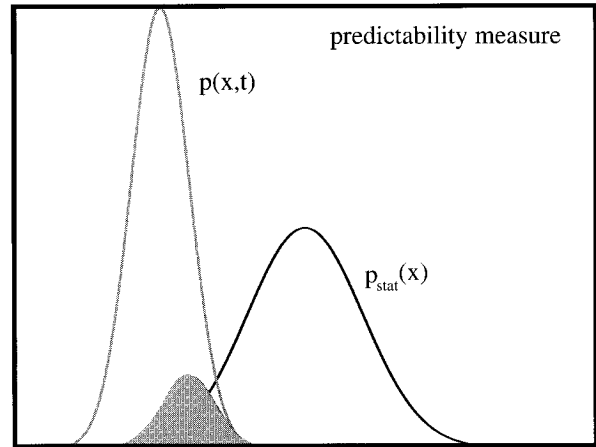


FIG. 12. Illustration of the predictability measure. The shaded area equals the quantity  $1 - s(t)$  according to Eq. (15).

To summarize, the fitted linear stochastic models are able to extract the information about ensemble averages and ensemble covariances contained in the individual restart ensembles. In the next section we will use the linear stochastic models to get more insight into the predictability of the stochastically forced HCM.

### 5. Predictability

The aim of this section is to quantify the predictability of the linear stochastic models by defining a predictability measure that will be based on the knowl-

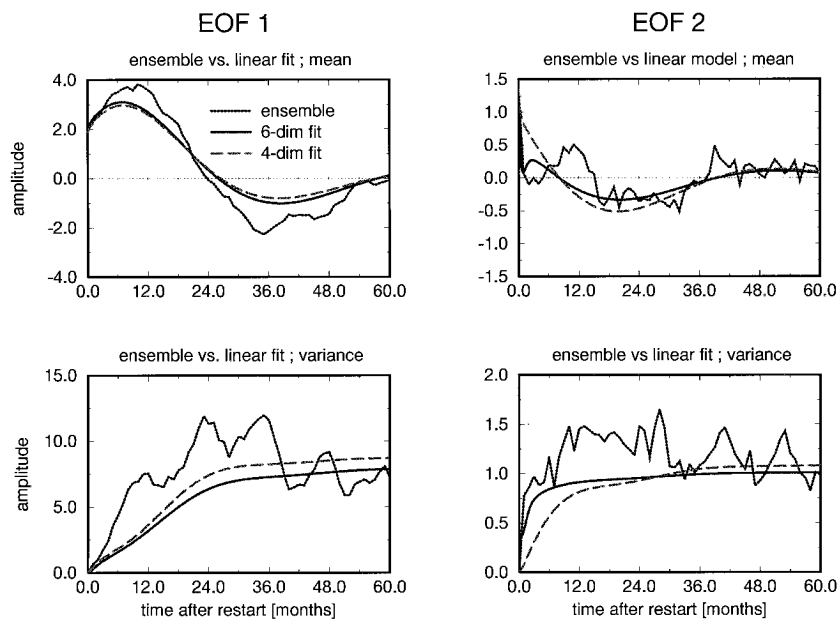


FIG. 11. Ensemble data vs the fitted linear stochastic processes for restart month 276. We show the first and second coefficients of the ensemble mean and variance vs fitted linear stochastic models of dimensionalities 4 and 6. The linear processes were fitted for a time interval of 60 months.

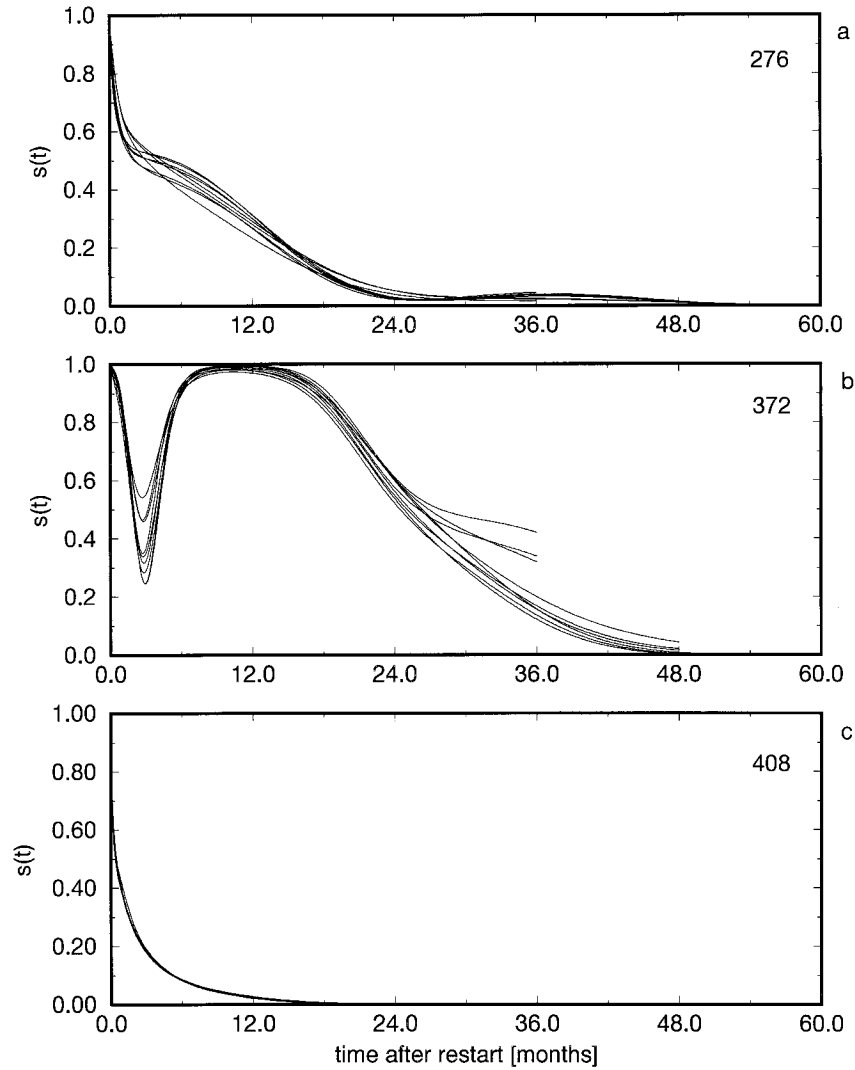


FIG. 13. Predictability measures we obtained for the restart experiments. The restart months are depicted in the upper right of each picture.

edge of the stationary and time-dependent probability distributions. Initially, due to the perfect knowledge of the initial states, the time-dependent probability densities are  $\delta$ -function like. They spread during the course of time and asymptote into the stationary probability density functions at times long compared to the typical timescales of diffusion induced by the stochastic part and damping time of the dynamical part. Consequently, one would require an appropriate predictability measure to decrease from an initial value, which we will require to be equal to unity, to zero for long times.

As a motivation for our definition of the predictability measure, we start with a one-dimensional linear stochastic process, the Ornstein–Uhlenbeck process

$$\dot{x} = -\alpha x + \sigma \eta \text{ and } x(0) = x_0. \quad (10)$$

Its stationary and time-dependent probability density functions are given by

$$p_{\text{stat}}(x) = \sqrt{\frac{\alpha}{\pi\sigma^2}} \exp\left(-\frac{\alpha x^2}{\sigma^2}\right) \text{ and} \quad (11)$$

$$p(x; t) = \sqrt{\frac{\alpha}{\pi\sigma^2(1 - e^{-2\alpha t})}} \exp\left[-\frac{\alpha(x - x_0 e^{-\alpha t})^2}{\sigma^2(1 - e^{-2\alpha t})}\right], \quad (12)$$

respectively. This follows from Eqs. (6) and (7) of the previous section. For some intermediate time, Fig. 12 depicts both probability densities. As a measure of predictability  $s_{OU}$  at time  $t$  we use the overlap of  $p_{\text{stat}}(x)$  and  $p(x; t)$ , measured in terms of their product integrated

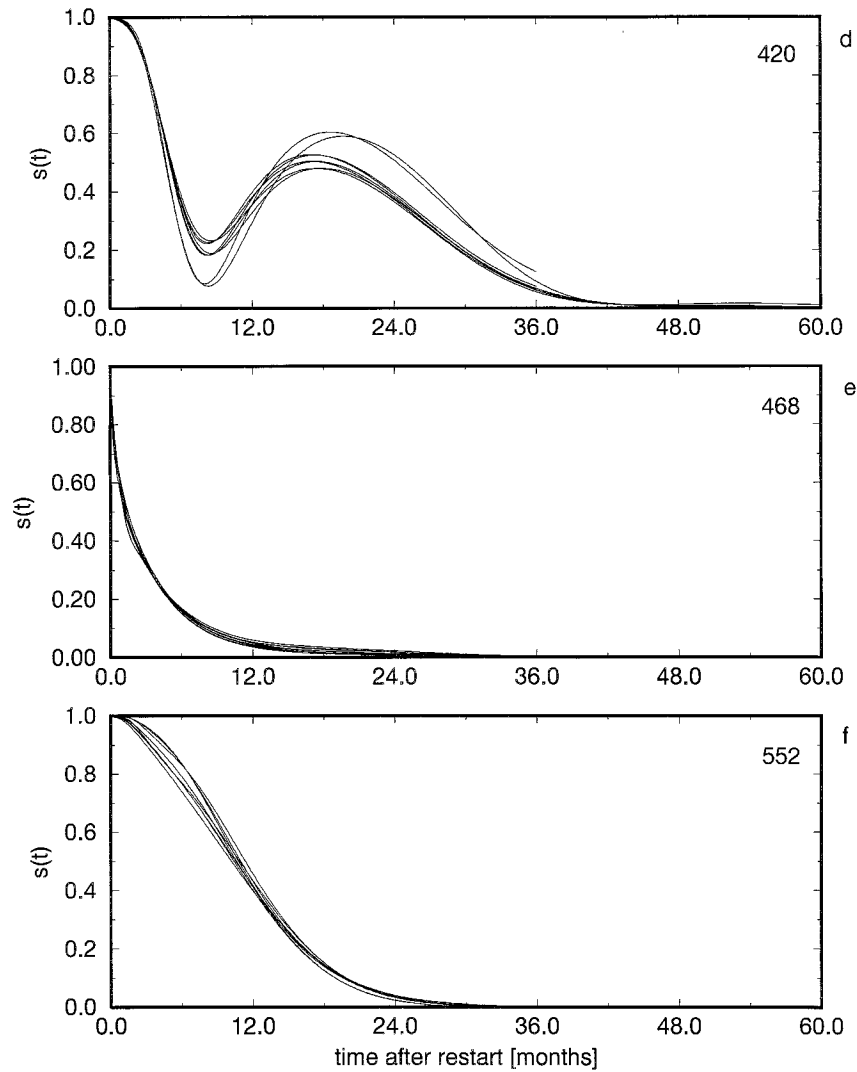


FIG. 13. (Continued)

over  $x$ , and an appropriate normalization to yield  $s_{OU}(\infty) = 0$  and  $s_{OU}(0) = 1$ , as discussed above.

Therefore, we define

$$s_{OU}(t) = 1 - 2\sqrt{\pi\sigma\sigma(t)} \int dx p(x; t) p_{\text{stat}}(x), \quad (13)$$

with  $\sigma(t) = \sigma\sqrt{(1 - e^{-2\alpha t})}$ . Inserting (11) and (12) into (13) gives

$$s_{OU}(t) = 1 - \sqrt{\frac{2(1 - e^{-2\alpha t})^{1/2}}{2 - e^{-2\alpha t}}} \exp\left[-\frac{\alpha x_0 e^{-2\alpha t}}{\sigma^2(1 - e^{-2\alpha t})}\right]. \quad (14)$$

The second term on the left-hand side of Eq. (14) corresponds to the shaded area in Fig. 12. The generaliza-

tion of this predictability measure to an  $n$ -dimensional linear stochastic process (5) is given by

$$s(t) = 1 - (2\sqrt{\pi})^n [\det(\mathbf{C}(t)\mathbf{C}_{\text{stat}})]^{1/4} \times \int dx p(x; t) p_{\text{stat}}(x), \quad (15)$$

with  $\mathbf{C}(t)$  and  $\mathbf{C}_{\text{stat}}$  denoting the covariance matrix at time  $t$  and the stationary covariance matrix, respectively.

In Fig. 13 we summarize the time evolution of  $s(t)$  for all the restart experiments performed. Before the predictability measure was calculated, we projected the fitted stochastic processes onto the leading two EOFs. This projection reduces the initial values of  $s(t)$  slightly, but it has little influence after a few months, since the third and higher principal components saturate quickly

to their stationary values. Thus, they contain little information relevant to characterize the asymptotic behavior of  $s(t)$ . In all cases, the predictability is lost well before a typical cycle period has elapsed. The ensembles initialized at states evolving into La Niñas exhibit a quick drop of the predictability measure right after the start. Thereafter, it increases again and eventually, after having reached a relative maximum, it asymptotes to zero.

## 6. Summary and discussion

We have used a hybrid coupled ocean–atmosphere model to investigate the importance of noise in limiting ENSO predictability. Adding a stochastic wind stress anomaly field yielded a more realistic behavior relative to an unperturbed control run, inducing the irregular occurrence and varying amplitudes of the coupled model’s ENSO extremes. We performed restart experiments from several phases of the model’s ENSO cycle with different realizations of the stochastic part. The spread among the individual trajectories comprising a restart ensemble depends on the initialization date. Those trajectories initialized at an El Niño extreme (420), or shortly after it (372), evolve quickly into La Niña states, but due to stochastic forcing they are shifted in phase with respect to each other. This leads to a rapid increase of the corresponding ensemble variance immediately after the start. When the realizations have reached the cold state, the ensemble variance drops again, and the subsequent spread is similar to that obtained from the other restart ensembles. To quantify the predictability of the coupled model further, we approximated the stochastically forced HCM by linear stochastic processes and defined a predictability measure for the latter. The predictability measure characterizes the overlap of the time-dependent and the stationary probability distributions of the fitted stochastic process. In each case studied, it reaches “critical” values before a typical cycle period has elapsed. The above-mentioned characteristics of the ensembles initialized at an El Niño extreme (420), or shortly after it (372), manifest themselves as an initial drop of the predictability measure and a subsequent increase, before decreasing again for larger times after the restart.

In our study, we concentrated on the perpetual action of random perturbations during the course of time. Additional initial errors would limit the system’s predictability still further. Initial errors would add a term  $\mathbf{C}_{in}(t) = \exp(\mathbf{A}t)\mathbf{C}(0)\exp(\mathbf{A}t)^T$  to the time evolution of the ensemble covariance matrix, which can yield considerable initial error growth due to the nonselfadjointness of  $\exp(\mathbf{A})$ , the one time step propagator. This was discussed by Blumenthal (1991), using an estimate of the propagator obtained from the cyclostationary POP method, which takes into account the seasonality in the underlying model. Here, we estimated the propagator and the driving noise simultaneously by fitting a linear process

locally—that is, over a short time interval—to a restart ensemble. As such, it takes into account the model’s actual state. The fits are optimized to monitor the time evolution of the underlying system for a finite time span. Our results suggest that in making actual ENSO predictions, ensemble forecasts that include noise forcing will provide valuable estimates of the forecasts’ uncertainty. The time evolution of the predictability measure  $s(t)$  indicates a clear dependence on the system’s state and may therefore be used as an objective measure of the expected forecast quality. There are, however, some caveats. The estimate of the stochastic forcing may suffer from errors in the underlying dataset and the way of constructing the random perturbation cannot be justified rigorously. This implies that we may have over- or underestimated the noise variance, which in turn affects the coupled model’s predictability. However, the fact that our results are consistent with those of Blanke et al. (1997) strengthens the reliability of the conclusions drawn in both papers. The dominant period of the model ENSO mode is longer than the observed one. Further, the approximating linear systems do not fit the stochastically forced HCM perfectly, since they are formulated in reduced state spaces, and misfits remain after optimization.

We tested the sensitivity of the stochastically forced HCM to different noise amplitudes (not shown) and coupling strengths. The reduction of the noise amplitude induced a more regular behavior, eventually reproducing the control run. Reducing the coupling strength at a constant noise level yielded an increasingly irregular behavior. At zero coupling strength and nonvanishing noise forcing, the NINO3 time series did not exhibit marked interannual fluctuations any more. Our choice of the coupling strength and the noise level, although admittedly subjective, was guided by the spatio-temporal characteristics of the model simulations.

Nevertheless, we believe that our main results remain valid. The stochastic forcing is important to model ENSO realistically, and the ENSO predictability is limited by it leading to a loss of predictability before a typical cycle period has elapsed.

*Acknowledgments.* We would like to thank Dr. Matthias Münnich for his valuable comments on an earlier version of this manuscript, Dr. Moritz Flügel for providing the Hybrid Coupled Model, Dr. Ralf Giering for providing the optimization software, and Prof. Dr. Klaus Hasselmann for many fruitful discussions. This work was supported by the European Union under Grant ENV4-CT95-0109. The computations were performed at the Deutsches Klimarechenzentrum (DKRZ).

## APPENDIX

### Fitted Linear Stochastic Processes

The following two tables list the results of the fit procedure outlined in section 4. Table A1 contains the

TABLE A1. Periods and  $e$ -folding times of the interannual eigenmodes of the fitted system matrixes.

Restart month	Dimension	Time interval	Period	$e$ -folding time
276	4	36	60.9	25.5
		48	60.7	26.6
		60	62.6	24.8
276	5	36	59.8	33.4
		48	62.1	30.2
		60	63.0	27.2
276	6	36	60.3	37.1
		48	62.6	31.8
		60	63.5	27.4
372	4	36	27.3	8.0
		48	23.8	6.4
		60	23.0	7.1
372	5	36	27.4	7.4
		48	23.6	7.3
		60	24.1	7.6
372	6	36	27.3	6.6
		48	23.8	6.8
		60	24.4	6.8
408	4	36	47.1	11.4
		48	50.8	12.1
		60	52.1	11.5
408	5	36	45.8	10.0
		48	50.3	11.8
		60	50.2	10.5
408	6	36	46.6	10.8
		48	56.6	12.7
		60	51.4	11.2
420	4	36	51.4	11.9
		48	57.0	12.4
		60	66.1	17.4
420	5	36	59.7	13.7
		48	61.7	13.1
		60	63.9	16.1
420	6	36	59.2	13.1
		48	61.5	12.9
		60	64.7	16.0
468	4	36	112.0	11.3
		48	66.8	14.1
		60	58.6	16.6
468	5	36	124.3	11.9
		48	69.1	15.4
		60	60.5	19.0
468	6	36	121.3	12.9
		48	67.2	18.0
		60	61.2	21.0
552	4	36	302.1	7.1
		48	99.7	13.8
		60	102.5	15.1
552	5	36	233.6	7.2
		48	107.6	15.5
		60	112.3	16.6
552	6	36	102.5	9.3
		48	99.1	15.7
		60	103.9	16.9

TABLE A2. Variances simulated by the fitted stochastic processes (sim.) and variances of the input data (dat.).

Rest. month	Dim.	Time interv.	EOF1		EOF2		Dim. EOFs	
			(sim.)	(dat.)	(sim.)	(dat.)	(sim.)	(dat.)
276	4	36	11.80	10.66	1.17	1.21	14.81	13.85
		48	9.24	7.93	1.11	1.18	12.16	11.05
		60	8.41	7.03	1.07	1.10	11.29	10.04
276	5	36	10.79	10.66	1.11	1.21	14.37	14.49
		48	8.47	7.93	1.04	1.18	11.94	11.65
		60	7.84	7.03	1.02	1.10	11.24	10.59
276	6	36	10.80	10.66	1.01	1.21	14.86	15.03
		48	8.44	7.93	1.03	1.18	12.39	12.18
		60	7.79	7.03	1.01	1.10	11.66	11.10
372	4	36	2.49	2.62	0.98	1.09	4.21	5.65
		48	3.27	4.75	0.77	1.00	5.26	7.45
		60	3.98	6.81	0.78	1.08	5.91	9.82
372	5	36	2.91	2.62	0.95	1.09	5.71	6.21
		48	4.00	4.75	0.92	1.00	6.82	8.22
		60	5.02	6.81	0.92	1.08	7.88	10.38
372	6	36	2.75	2.62	0.95	1.09	6.00	6.70
		48	3.76	4.75	0.91	1.00	7.04	8.76
		60	4.73	6.81	0.91	1.08	8.04	10.89
408	4	36	7.09	8.21	1.03	1.10	9.97	11.22
		48	7.06	7.32	1.05	1.16	9.95	10.42
		60	7.00	7.41	1.05	1.14	9.90	10.50
408	5	36	6.90	8.21	1.00	1.10	10.25	11.78
		48	6.91	7.32	1.03	1.16	10.30	10.99
		60	6.84	7.41	1.01	1.14	10.21	11.05
408	6	36	6.91	8.21	1.00	1.10	10.74	12.30
		48	6.91	7.32	1.03	1.16	10.77	11.48
		60	6.82	7.41	1.02	1.14	10.64	11.54
420	4	36	3.86	3.50	1.04	1.06	5.64	6.46
		48	4.17	5.45	1.01	1.00	5.93	8.32
		60	4.37	7.25	1.00	1.08	6.12	10.29
420	5	36	4.93	3.50	0.97	1.06	8.02	7.02
		48	5.16	5.45	0.97	1.00	8.30	8.83
		60	5.21	7.25	0.96	1.08	8.37	10.87
420	6	36	4.81	3.50	0.98	1.06	8.37	7.52
		48	5.06	5.45	0.97	1.00	8.68	9.35
		60	5.11	7.25	0.95	1.08	8.71	11.36
468	4	36	8.46	9.55	0.99	1.19	11.25	12.66
		48	8.40	9.61	1.03	1.19	11.26	12.75
		60	8.11	7.96	1.06	1.20	11.01	11.13
468	5	36	8.57	9.55	0.97	1.19	11.84	13.24
		48	8.44	9.61	1.01	1.19	11.79	13.39
		60	8.12	7.96	1.04	1.20	11.50	11.69
468	6	36	8.57	9.55	0.96	1.19	12.31	14.05
		48	8.49	9.61	1.01	1.19	12.32	13.90
		60	8.13	7.96	1.04	1.20	11.98	12.22
552	4	36	3.13	5.88	0.92	0.99	5.86	8.83
		48	4.21	7.63	0.94	1.14	6.99	10.72
		60	4.82	9.18	0.96	1.13	7.62	12.24
552	5	36	3.10	5.88	0.90	0.99	6.31	9.41
		48	4.49	7.63	0.93	1.14	7.78	11.33
		60	5.23	9.18	0.96	1.13	8.56	12.87
552	6	36	3.16	5.88	0.89	0.99	6.81	9.89
		48	4.28	7.63	0.92	1.14	8.00	11.87
		60	4.93	9.18	0.94	1.13	8.69	13.38

periods of the interannual eigenmode of the system matrix and its  $e$ -folding time for the different time spans and dimensionalities of the fitted linear stochastic models. Table A2 contains for the same experiments the simulated variances (sim.) compared with the input data (dat.). The variances were obtained by averaging over the last 12 months of each dataset.

REFERENCES

Balmaseda, M. A., M. K. Davey, and D. L. T. Anderson, 1995: Decadal and seasonal dependence of ENSO prediction skill. *J. Climate*, **8**, 2705–2715.  
 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.
- Barnston, A. G., and Coauthors, 1994: Long-lead seasonal forecasts—Where do we stand? *Bull. Amer. Meteor. Soc.*, **75**, 2097–2114.
- Blanke, B., J. D. Neelin, and D. Gutzler, 1997: Estimating the effect of stochastic wind stress forcing on ENSO irregularity. *J. Climate*, **10**, 1473–1486.
- Blumenthal, B., 1991: Predictability of a coupled ocean–atmosphere model. *J. Climate*, **4**, 766–784.
- Cane, M. A., and E. S. Sarachik, 1981: The response of a linear baroclinic equatorial ocean to periodic forcing. *J. Mar. Res.*, **39**, 651–693.
- Chang, P., B. Wang, T. Li, and L. Ji, 1994: Interactions between the seasonal cycle and the Southern Oscillation—Frequency entrainment and chaos in an intermediate coupled ocean–atmosphere model. *Geophys. Res. Lett.*, **21**, 2817–2820.
- Chen, Y.-Q., D. S. Battisti, T. N. Palmer, J. Barsugli, and E. S. Sarachik, 1997: A study of the predictability of tropical Pacific SST in a coupled atmosphere/ocean model using singular vector analysis: The role of the annual cycle and the ENSO cycle. *Mon. Wea. Rev.*, **125**, 831–845.
- Dettinger, M. D., M. Ghil, C. M. Strong, W. Weibel, and P. Liou, 1995: Software expedites singular spectrum analysis of noisy time series. *Eos, Trans. Amer. Geophys. Union*, **76** (2), 12–21.
- Flügel, M., 1994: Untersuchungen zu Mechanismus und Vorhersagbarkeit von ENSO mit einem vereinfachten gekoppelten Ozean–Atmosphäre-Modell. Univ. Hamburg, 67 pp. [Available from Universität Hamburg, Fakultät für Geowissenschaften, Bundesstraße 55, D-20146 Hamburg, Germany.]
- Giering, R., 1995: The AMC manual. Max-Planck-Institut für Meteorologie Internal Rep., 21 pp. [Available from Max-Planck-Institut für Meteorologie, Bundesstraße 55, D-20146 Hamburg, Germany.]
- , and T. Kaminski, 1996: Recipes for adjoint code construction. Max-Planck-Institut für Meteorologie Rep. 212, 35 pp. [Available from Max-Planck-Institut für Meteorologie, Bundesstraße 55, D-20146 Hamburg, Germany.]
- Goldenberg, S., and J. O'Brien, 1981: Time and space variability of the tropical Pacific wind stress. *Mon. Wea. Rev.*, **109**, 1190–1207.
- Goswami, B. N., and J. Shukla, 1991: Predictability of a coupled ocean–atmosphere model. *J. Climate*, **4**, 3–22.
- Hasselmann, K., 1976: Stochastic climate models. Part I: Theory. *Tellus*, **28**, 289–305.
- , 1988: PIPs and POPs: The reduction of complex dynamical systems using principal interaction and oscillation patterns. *J. Geophys. Res.*, **93**, 11 015–11 021.
- Ji, M., A. Leetma, and J. Derber, 1995: An ocean analysis system for seasonal to interannual climate studies. *Mon. Wea. Rev.*, **123**, 460–481.
- Jin, F.-F., D. Neelin, and M. Ghil, 1994: El Niño on the Devil's staircase: Annual subharmonic steps to chaos. *Science*, **266**, 70–72.
- Kendall, M., A. Stuart, and J. K. Ord, 1983: *The Advanced Theory of Statistics*. 4th ed. Vol. 3, Charles Griffin, 283 pp.
- Kleeman, R., and S. B. Power, 1994: Limits to predictability in a coupled ocean–atmosphere model due to atmospheric noise. *Tellus*, **46A**, 529–540.
- Latif, M., 1987: Tropical ocean circulation experiments. *J. Phys. Oceanogr.*, **17**, 246–263.
- , T. P. Barnett, M. A. Cane, M. Flügel, N. E. Graham, H. von Storch, J.-S. Xu, and S. E. Zebiak, 1994: A review of ENSO prediction studies. *Climate Dyn.*, **9**, 167–179.
- Moore, A., and R. Kleeman, 1996: The dynamics of error growth in a coupled model of ENSO. *Quart. J. Roy. Meteor. Soc.*, **122**, 1405–1446.
- Neelin, J. D., M. Latif, and F.-F. Jin, 1994: Dynamics of coupled ocean–atmosphere models: The tropical problem. *Annu. Rev. Fluid Mech.*, **26**, 617–659.
- Numerical Algorithms Group, 1991: NAG Fortran Library Manual Mark 15. [Available from NAG, Wilkinson House, Jordan Hill Road, Oxford OX2 8DR, United Kingdom.]
- Parker, D. E., C. K. Folland, A. Bevan, M. N. Ward, M. Jackson, and K. Maskell, 1994: Marine surface data for analysis of climatic fluctuations on interannual to century time scales. *Natural Climate Variability on Decade to Century Time Scales*, D. G. Martinson, K. Bryan, M. Ghil, M. M. Hall, T. R. Karl, E. S. Sarachik, S. Soroostian, and L. F. Talley, Eds., National Academy Press, 241–250.
- Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. *J. Climate*, **1**, 75–86.
- Schopf, P. S., and M. J. Suarez, 1988: Vacillations in a coupled ocean–atmosphere model. *J. Atmos. Sci.*, **45**, 549–566.
- Tziperman, E., L. Stone, M. A. Cane, and H. Jarosh, 1994: El Niño chaos: Overlapping of resonances between the seasonal cycle and the Pacific Ocean–atmosphere oscillator. *Science*, **264**, 72–74.
- van Kampen, N. G., 1992: *Stochastic Processes in Physics and Chemistry*. North-Holland, 210 pp.
- von Storch, H., G. Bürger, R. Schnur, and J.-S. von Storch, 1995: Principal oscillation patterns: A review. *J. Climate*, **3**, 377–400.