Title

Predicting Critical Transitions in ENSO models. Part II: Spatially Dependent Models

Authors

Mukhin, Dmitry
Kondrashov, Dmitri
Loskutov, Evgeny
et al.

Predicting Critical Transitions in ENSO models. Part II: Spatially Dependent Models

6 authors, including:

Dmitry Mukhin
Russian Academy of Sciences
42 PUBLICATIONS **144** CITATIONS

D. Kondrashov
University of California, Los Angeles
96 PUBLICATIONS **2,369** CITATIONS

Michael Ghil
Ecole Normale Supérieure, Paris
565 PUBLICATIONS **18,081** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Abrupt climate changes View project

Project   ONR-MURI: Extended-Range Prediction with Low-Dimensional, Stochastic-Dynamic Models: A Data-driven Approach View project

Predicting Critical Transitions in ENSO models. Part II: Spatially Dependent Models

DMITRY MUKHIN

Institute of Applied Physics of Russian Academy of Sciences, and Lobachevsky State University of Nizhni Novgorod, Nizhny Novgorod, Russia

DMITRI KONDRASHOV

Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, Los Angeles, California

EVGENY LOSKUTOV, ANDREY GAVRILOV, AND ALEXANDER FEIGIN

Institute of Applied Physics of Russian Academy of Sciences, and Lobachevsky State University of Nizhni Novgorod, Nizhny Novgorod, Russia

MICHAEL GHIL

Geosciences Department, and Laboratoire de Météorologie Dynamique, CNRS and IPSL, École Normale Supérieure, Paris, France, and Department of Atmospheric and Oceanic Sciences and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, Los Angeles, California

ABSTRACT

The present paper is the second part of a two-part study on empirical modeling and prediction of climate variability. This paper deals with spatially distributed data, as opposed to the univariate data of Part I. The choice of a basis for effective data compression becomes of the essence. In many applications, it is the set of spatial empirical orthogonal functions that provides the uncorrelated time series of principal components (PCs) used in the learning set. In this paper, the basis of the learning set is obtained instead by applying multichannel singular-spectrum analysis to climatic time series and using the leading spatiotemporal PCs to construct a reduced stochastic model. The effectiveness of this approach is illustrated by predicting the behavior of the Jin–Neelin–Ghil (JNG) hybrid seasonally forced coupled ocean–atmosphere model of El Niño–Southern Oscillation. The JNG model produces spatially distributed and weakly nonstationary time series to which the model reduction and prediction methodology is applied. Critical transitions in the hybrid periodically forced coupled model are successfully predicted on time scales that are substantially longer than the duration of the learning sample.

1. Introduction and motivation

In two previous papers (Molkov et al. 2012; Mukhin et al. 2015, hereinafter Part I), we formulated an empirical modeling and prediction methodology based on artificial neural networks (ANNs; Hornik et al. 1989). A key difficulty in applying this methodology to construct an empirical, nonlinear stochastic model that helps simulate and predict the real climate system's behavior is the mismatch between the large number of variables by which one wishes to describe the system versus the shortness of the time series of available experimental data.

Observational datasets comprise usually time-varying spatial fields of several climatic variables—such as sea surface temperatures (SSTs), pressure, wind velocity, and humidity—measured at differing levels of detail. For certain purposes, a researcher might study a very large set of time series, the number of which equals the number of nodes in a spatial grid multiplied by the number of climate variables to be considered. For other purposes, however, it is well known [cf. Ghil et al. (2002,

Corresponding author address: Dmitry Mukhin, Institute of Applied Physics of Russian Academy of Sciences, 46 Ulyanova St., Nizhni Novgorod 603950, Russia.
E-mail: mukhin@appl.sci-nnov.ru

and references therein) or Part I] that the short duration, of about 60 years, of measured time series is not sufficient to determine the dynamics of the full system by inverse modeling.

Efficient reduction of the system's dimensionality is thus essential in order to infer an evolution operator for a low-dimensional subsystem that determines the key properties of the observed dynamics. Specifically, in Part I we showed that nonlinear, stochastic modeling in a two-dimensional subspace turns out to be sufficient to predict critical transitions in the setting of conceptual, delay-differential equation models of El Niño–Southern Oscillation (ENSO). In this setting, we applied the methodology to scalar time series obtained from one of three highly simplified but still infinitely dimensional ENSO models.

To extract the robust information needed to reconstruct and predict the dynamics from the spatially distributed time series, the choice of basis functions is important: it determines both the dimension and the geometry of the subspace in which we wish to carry out the reconstruction and prediction (Kravtsov et al. 2009; Strounine et al. 2010). A widely used basis for climate fields is that of empirically orthogonal functions (EOFs; Preisendorfer 1988; Navarra and Simoncini 2010).

Kondrashov et al. (2005) and Kravtsov et al. (2009) have shown the usefulness of this basis in constructing an ENSO model by data-driven empirical model reduction (EMR) from the multivariate time series of the SST field. Moreover, the real-time ENSO prediction by this leading EOF-based EMR model has proved to be highly competitive among other dynamical and statistical ENSO forecasts (Barnston et al. 2012).

Still, preparation of the learning sample by projection onto the leading EOFs has an important drawback: it uses only the instantaneous correlations between points of the spatial grid, and it does not take into account time-lagged relationships. The latter, however, are largely determined by physical processes in the atmosphere–ocean system, which should be included when uncorrelated data patterns are constructed. An expansion of the dataset that does make allowance for the time-lagged correlations between spatially distributed time series can be more informative, since the modes obtained in this way are apt to represent better the physics of the modeled process.

In this paper, we propose using the dynamic variables obtained by projection onto the leading modes of singular-spectrum analysis (SSA; Vautard and Ghil 1989; Vautard et al. 1992; Ghil et al. 2002) for model reduction. Essentially, the SSA method of data-adaptive time series analysis is a generalization of the Mañe–Takens delay method of embedding manifolds into minimal-dimension Euclidean spaces (Mañe 1981; Takens 1981; Sauer et al. 1991).

In practice, SSA is based on extending PCA into the time domain. The temporal EOFs correspond to the eigenvectors of the lagged-covariance matrix. In this case, the resulting variables are uncorrelated linear combinations of the delayed variables. Multichannel, or multivariate SSA (MSSA; Keppenne and Ghil 1993; Plaut and Vautard 1994; Ghil et al. 2002) generalizes univariate, or scalar SSA to the case of multivariate time series: it uses lagged correlations of each scalar time series, as well time-lagged correlations between distinct time series. In the case of interest here, when the time series are spatially gridded values of a climatic field, MSSA allows one to obtain a parsimonious empirical model in terms of the so-called spatiotemporal EOFs (ST-EOFs; Plaut and Vautard 1994; Ghil et al. 2002).

In this paper, we apply the empirical stochastic model of Molkov et al. (2012) and Part I to the prediction of critical transitions in the climate system, when weakly nonstationary time series—which are spatially distributed, rather than merely scalar—are available. As a step toward the use of actually observed climate data, we worked with data from an intermediate-complexity, hybrid coupled ENSO model (Neelin 1991; Jin and Neelin 1993a) over the tropical Pacific Ocean.

The Jin–Neelin–Ghil (JNG) model (Jin et al. 1994, 1996) is a seasonally forced version of the Jin–Neelin model (Jin and Neelin 1993a,b). The latter describes two important mechanisms that contribute to ENSO dynamics: the propagation of planetary ocean waves across the tropical Pacific and the interaction of the ocean dynamics with the atmosphere above through the SST field. This model is a simplification of the more complete atmosphere–ocean model of Zebiak and Cane (1987): it is governed by a set of shallow-water equations for the ocean, coupled with an integral equation for the SST; both the atmospheric and the oceanic fields depend only on time and the longitude along the equator. The coupling parameter between the two media is an important control parameter in the seasonally forced JNG, as well as in the Jin–Neelin model.

Following the approach of Part I, we introduce here an adiabatic trend of this parameter to simulate the system's slow forcing on decadal time scales. As a result, the ENSO attractor undergoes qualitative, as well as quantitative changes in time. These changes include bifurcations, which manifest themselves as critical transitions within the observed dynamics. We consider situations in which the observed time series used for empirical model learning do not contain critical transitions and the transitions between a chaotic regime and periodic oscillations took place beyond the limits of the learning interval.

This paper is structured as follows: in the next section, we describe briefly the methods used to construct the

basis of the phase-space variables from EOF and MSSA expansions, respectively, and then formulate the empirical model used here. This formulation relies on the general stochastic model proposed by Molkov et al. (2012) and applied in Part I to simpler ENSO models, and on this general model's implementation via ANNs.

Section 3 reviews the ENSO model (Jin et al. 1994, 1996) used to generate the data. Section 4 presents our analysis of the time series generated by the JNG model of section 3 by using the empirical model formulated in section 2; this section gives our results on predicting critical transitions and demonstrates the advantages of the MSSA data expansion over the more traditional use of purely spatial EOFs. In the last section, we generalize the results obtained here and discuss the perspectives of empirical modeling in climate dynamics.

2. Phase space reduction and empirical model

a. Expansion in spatial EOFs

Let us represent the spatial field of the climatic variable X, measured at discrete locations, as the time-dependent vector $\mathbf{X}^k := \mathbf{X}(t_k)$, $\mathbf{X}^k \in \mathbb{R}^D$, $k = 1, \ldots, N$; here, the jth component of the vector X_j^k corresponds to the jth node of the spatial measurement grid, which consists of D nodes, and observations are available at N successive, equidistant times. In what follows, we will consider normalized data: that is, those reduced to the form

$$\langle X_j^k \rangle_k = 0 \quad \text{and} \quad \langle X_j^k X_j^k \rangle_k = 1, \qquad (1)$$

where the operation $\langle \cdot \rangle_k$ denotes averaging with respect to the time index k, and replaces in practice the expectation operator.

Standard expansion of the vectors $\mathbf{X}^k$ into spatial EOFs $\{\mathbf{v}_i: i = 1, \ldots, D\}$ is given by

$$\mathbf{X}^k = \sum_{i=1}^{D} Y_i^k \mathbf{v}_i, \quad \mathbf{v}_i \in \mathbb{R}^D, \qquad (2)$$

where the principal components (PCs) $\mathbf{Y}^k$ do not correlate with each other; that is, the covariance matrix $\mathbf{C}^Y = \langle \mathbf{Y}^k \mathbf{Y}^{k\mathrm{T}} \rangle_k$ is diagonal, and its diagonal elements are equal to the variances in the $\mathbf{v}_i$ directions of the series $\{\mathbf{Y}^k\}$.

The vectors $\mathbf{Y}^k$ are derived from the initial vectors $\mathbf{X}^k$ by using the orthonormal transformation $\mathbf{Y}^k = \mathbf{V}^{\mathrm{T}} \mathbf{X}^k$, where $\mathbf{V}$ is the unitary matrix whose columns are the EOFs $\mathbf{v}_i$, and

$$\mathbf{C}^Y = \mathbf{V}^{\mathrm{T}} \langle \mathbf{X}^k \mathbf{X}^{k\mathrm{T}} \rangle_k \mathbf{V}. \qquad (3)$$

Thus, the EOF matrix $\mathbf{V}$ rotates the initial data $\mathbf{X}^k$ to the eigenvectors of their covariance matrix, $\mathbf{C}^X = \langle \mathbf{X}^k \mathbf{X}^{k\mathrm{T}} \rangle_k$, and $\mathbf{V}^{\mathrm{T}} = \mathbf{V}^{-1}$.

Note that, since the data are reduced to the normalized form of Eq. (1), $\mathbf{C}^X$ is a correlation, rather than a covariance matrix, with respect to the original, unnormalized data. Therefore, the above-described EOFs are called, strictly speaking, correlation EOFs (Jolliffe 1986). Actually, the basis $\{\mathbf{v}_i\}$ depends on the scaling of the initial spatial field, and one can specify the problem of finding an optimal scaling function [i.e., one that yields relationships differing from Eq. (1)] as studied by Kravtsov et al. (2009) and Strounine et al. (2010), among others. Here we consider, for simplicity, only correlation EOFs and normalized-data compression based on Eq. (1), since our emphasis is on including temporal correlations along with the spatial ones used so far.

b. Spatiotemporal expansion

As already mentioned in the introduction, an empirical model that uses spatiotemporal information already in the data compression stage might be better at simulating and predicting the behavior of a system in which variables depend on time, as well as space. This leads us to the idea of using an orthonormal basis extended to the time domain for the learning subspace. Without leaving the framework of linear transformations, we can construct such basis functions by relying on time-lagged, as well as instantaneous spatial correlations between the nodes of our spatial grid. In other words, in order to derive the variables of our low-order empirical model, we propose to apply MSSA analysis (Keppenne and Ghil 1993; Plaut and Vautard 1994; Ghil et al. 2002) to the initial multivariate time series $\{\mathbf{X}^k\}$.

We thus consider lagged correlations of variables within the limits of the specified time lag T, which corresponds to the time scales under consideration. To do so, we evaluate the mutual covariations of the elements of the $D \times T$ trajectory matrix $\mathbf{Q}^k$ (Broomhead and King 1986a,b). The columns of the matrix are the vectors $\mathbf{X}^k$, shifted relative to each other at every time k:

$$\mathbf{Q}^k = (\mathbf{X}^k, \mathbf{X}^{k+1}, \ldots, \mathbf{X}^{k+T}). \qquad (4)$$

The covariance matrix we seek is the time-averaged tensor product $\mathbf{C}^J = \langle \mathbf{J}^k \mathbf{J}^{k\mathrm{T}} \rangle_k$ of the vectors $\mathbf{J}^k$ formed by arranging the matrix columns $\mathbf{Q}^k$ in one column of length DT:

$$\mathbf{J}^k = \mathrm{vec}(\mathbf{Q}^k), \quad \mathbf{J}^k \in \mathbb{R}^{DT}. \qquad (5)$$

Orthogonal rotation of the vectors $\mathbf{J}^k$ above, similar to Eq. (3), ensures diagonalization of the covariance matrix $\mathbf{C}^J$ and makes it possible to perform the transition to mutually uncorrelated coordinates:

$$\mathbf{Y}^k = \mathbf{V}^{\mathrm{T}} \mathbf{J}^k. \qquad (6)$$

In this multivariate case, too, the columns of the $DT \times DT$ matrix $\mathbf{V}$ form the orthonormal basis $\{\mathbf{v}_i\}$, where each

component can be represented now as a matrix $\mathbf{E}_i$ such that $\mathbf{v}_i = \text{vec}(\mathbf{E}_i)$: that is, a matrix with the spatiotemporal structure of the indices that is similar to Eq. (4). One thus obtains a set of basis functions of space and time and, therefore, time series for Eq. (6) of the spatiotemporal PCs (ST-PCs) Y_i^k, which are found by projecting the data field on these functions. Note that these ST-PCs are only uncorrelated at zero lag (Plaut and Vautard 1994; Ghil et al. 2002).

c. Construction of the empirical model

Following the PCA logic of sections 2a or 2b, we construct a model of the evolution operator in the space of the leading PCs or ST-PCs $\mathbf{Y}^k$, respectively (i.e., those that are associated with the largest variances). This means that we will use the modes that provide the principal contribution to the dynamics of the measured values. Actually, we will model the evolution operator of the shortened vector $\mathbf{U}^k = (Y_1^k, \ldots, Y_d^k)$ for $d < D$, which consists of the d leading components of $\mathbf{Y}^k$ arranged in the order of decreasing variations. Selection criteria for an optimal d in various settings are discussed by Ghil and Robertson (2002), Ghil et al. (2002), and Molkov et al. (2009), respectively, along with many references.

We modify here somewhat the discrete stochastic model described in Part I, in order to adapt it to modeling interannual climate variability in the setting of spatially distributed data. The modified model is given by

$$\mathbf{U}^{k+1} = \mathbf{f}(\mathbf{U}^k, \mathbf{U}^{k-1}, \ldots, \mathbf{U}^{k-m}; \boldsymbol{\mu})$$
$$+ \mathbf{g}(\mathbf{U}^k, \mathbf{U}^{k-1}, \ldots, \mathbf{U}^{k-m}; \boldsymbol{\eta})\boldsymbol{\xi}^k, \qquad (7)$$

and details appear in appendix A. Here $\mathbf{f}$ and $\mathbf{g}$ are vector and matrix functions, respectively, represented as ANNs and depending on the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\eta}$, while $\boldsymbol{\xi}$ is Gaussian uncorrelated noise with unit variance. As in Part I, we represent the $\mathbf{f}$ and $\mathbf{g}$ functions by ANNs formulated as 3-layer perceptrons with a hyperbolic tangent used as the activation function (Hornik et al. 1989). The dimensionality of this model equals $d \times m$, and it is determined by the number of PCs or ST-PCs taken into account and the number of the consecutive states of the system used to predict the following state. This model is trained by means of the Bayes procedure described in Molkov et al. (2011, 2012) and in Part I of this work.

Grieger and Latif (1994) already constructed an ANN-based model of ENSO in the subspace of four leading EOFs that were derived from a combined dataset of SSTs, surface wind stress, and upper-ocean heat content and compared it with a linear model based on principal oscillation patterns (Hasselmann 1976;

Penland 1996). These authors showed that the linear model only exhibits oscillations that decay slowly to a single fixed point, while the ANN model recovers a stable limit cycle in its four-dimensional phase space. Here we go beyond the work of Grieger and Latif (1994) in three important ways: (i) we show the superiority of ST-EOFs over the usual, purely spatial EOFs; (ii) we recover a chaotic attractor; and (iii) we carry out a predictability study of critical trasitions.

3. The JNG model

We use the intermediate coupled JNG model as our generator of spatiotemporal climate variability. A brief description of this model is provided in appendix B; see Jin and Neelin (1993a,b), Sun et al. (2002), and Kondrashov et al. (2008) for further details on the original Jin–Neelin model with time-independent forcing and Jin et al. (1994, 1996) for the model formulation with seasonal forcing. Chekroun et al. (2014) have called a further simplification of the JNG model the forced JN (fJN) model.

The Jin–Neelin model (Jin and Neelin 1993a,b) is essentially a further idealization of the coupled ocean–atmosphere model of Zebiak and Cane (1987). The vertical mean motions above the thermocline are governed by linearized shallow-water equations—forced by the wind stress—on an equatorial β-plane in the long-wave approximation (Neelin 1991). The resulting currents drive an advection equation describing the SST at Earth's equator. The major simplification is to treat explicitly only the zonal dependence of the SST field over an equatorial strip, while the meridional structure of the associated atmospheric forcing is given as a Gill-type model (Gill 1980) for the wind stress anomalies. It is the latter model that establishes a diagnostic relation between the SST anomalies and the wind stress anomalies.

In the JNG model, the ENSO oscillation arises via destabilization of a seasonally modulated basic state (i.e., a periodic solution of period 1 yr). This loss of stability leads to a rich variety of possible ENSO dynamics, from frequency-locked regimes to chaotic ones via a quasi-periodic route to chaos (Jin et al. 1994, 1996). This rich dynamics results from the nonlinear interactions between the externally forced seasonal cycle and the internally generated ENSO cycle.

The nature of the intrinsic ENSO cycle depends on the values of two control parameters: namely, the coupling coefficient μ between the atmosphere and ocean and the surface-layer coefficient δ_s; see appendix B. The parameter δ_s varies from zero to one and controls the intensity of the anomalous surface-layer currents as a function of the wind stress anomalies. When δ_s is close to unity (i.e., in the case of strong surface-layer feedback) stronger

vertical and advection anomalies add to the rate of SST change (Jin et al. 1996) and place the JNG model in a chaotic regime that is also strongly mixing.

Strong mixing, also referred to as rapid mixing in dynamical systems theory (Chekroun et al. 2014), corresponds to chaotic behavior that is, in addition, characterized by fast decay of lagged autocorrelations. In the case of the JNG model, the strongly mixing behavior also manifests itself by moderately energetic peaks in the power spectrum, as opposed to the nearly periodic behavior that is associated, in this model, with weak mixing [i.e., slow decay of correlations; see Fig. 2 in Chekroun et al. (2014)]. The choice of $\delta_s = 0.95$ in this paper thus corresponds to putting the model in a less periodic and thus less predictable regime.

We shall study here the JNG model for such a strong mixing regime, at first with a fixed coupling between the oceanic and atmospheric components, with $\delta_s = 0.95$ and $\mu = 1.3$. Next, we will analyze the model's behavior by changing another key model parameter, namely δ. This parameter affects the travel time of the equatorially trapped waves (Jin and Neelin 1993a,b; Jin et al. 1996), which play an essential role in ENSO dynamics. Hence, changes in δ will lead to the critical dynamical transitions that we are interested in simulating and predicting.

4. Prediction of critical transitions

We consider the time series of SST anomalies $\{\mathbf{X}^k\}$ generated by the JNG model with a prescribed adiabatic linear trend in the control parameter δ. The SST field consists of 24 time series, each of which corresponds to a grid point in a uniform grid along the equator within the Pacific basin, from Indonesia to South America. Note that this is only a small fraction of the total number of degrees of freedom of the numerically discretized JNG model; the latter number equals over 400, as indicated in the appendix B. The length of the series is 15 000 yr, with a sampling step of half-a-month; the model parameter δ changes linearly from 1.26 at the beginning toward 1.36 at the end of the time interval.

A short segment of the SST time series produced by the JNG model is plotted in Fig. 1. This solution corresponds to a standing wave, with a period of roughly 5 yr; that is, it matches fairly well the 4–5-yr period of the observed quasi-quadrennial (QQ) ENSO mode (Jiang et al. 1995; Ghil et al. 2002).

As in the examples from Part I, a priori information about the system can be taken into account during the initial processing of the time series. Here, this information consists in the dependence of the JNG model's time-continuous evolution operator on periodic external forcing, and it helps one reconstruct the
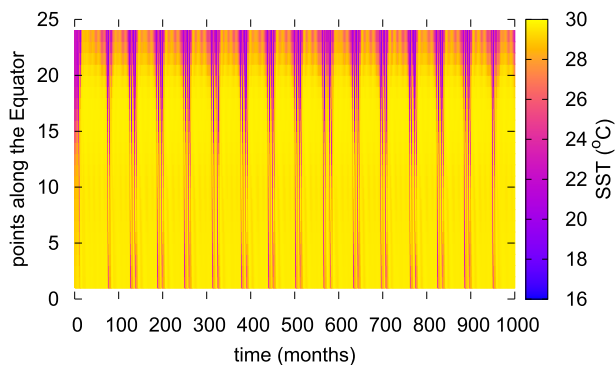

FIG. 1. Segment of the SST time series produced by the JNG model and used here for training the empirical ANN model. The points along the equator, shown on the ordinate, are numbered from west to east. This JNG model solution corresponds to a standing wave, with a period of roughly 5 yr; see text for details.

discrete evolution operator as a Poincaré map, which relates model states separated by a time interval of 1 yr.

Figures 2a,b show an example of a JNG model solution, sampled once a year at two grid points from the eastern and central parts of the Pacific basin. One can see that the model's attractor undergoes qualitative changes in time, which we associate with critical transitions. These transitions correspond to the disappearance of the chaotic regime and the consequent birth of periodic regimes, whose periods are multiples of the seasonal cycle, as well as with the change in the multiplicity of the periodic regimes. Note that all qualitative transitions represented in the two time series in Figs. 2a,b take place in the series of all the other components of the vector $\{\mathbf{X}^k: k = 1, \ldots, N\}$, since they result from the bifurcations occurring in the JNG model as the parameter δ crosses certain threshold values in time. Such bifurcations are part of the well-known quasi-periodic route to chaos, as discussed by Jin et al. (1994, 1996).

As a learning sample, we take a 24-dimensional time series over the first 4000 years, sampled with half-a-month time steps. Our task is to train an empirical model of the form given by Eq. (7) on this series and then predict the sequence of critical transitions that occur beyond the training interval. We will train the empirical model in two ways, based on the two approaches to variables, constructing the basis of the learning subspace (i) using purely spatial EOFs and (ii) using ST-EOFs, as discussed in sections 2a and 2b, respectively.

a. Results using a spatial EOF basis

Figure 3 shows the four leading EOFs, which capture 49%, 21.7%, 16.5%, and 5.7% of variance, respectively, for a total of over 90%. The spatial EOFs here were computed from the original SST time series, not the subsampled one. We then used subsampled PCs for model learning.
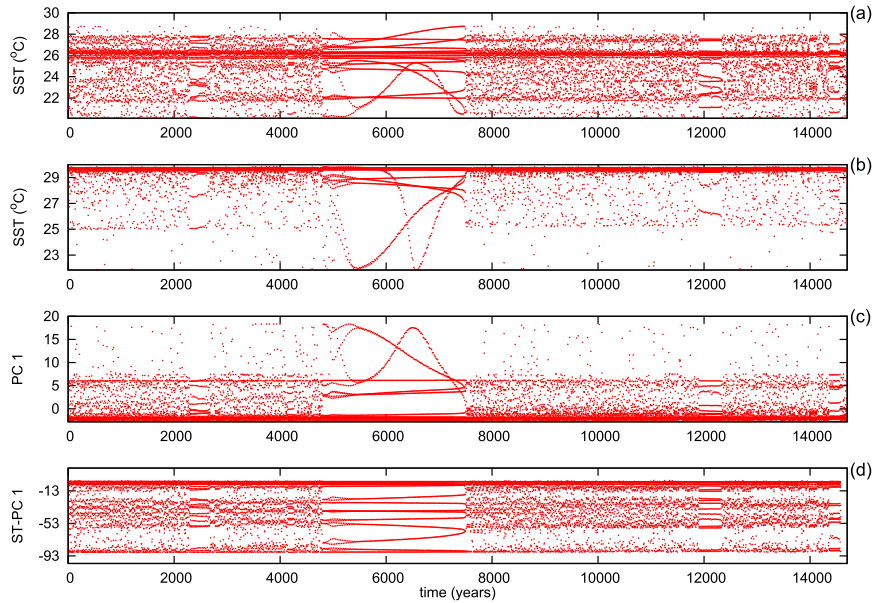
FIG. 2. Poincaré map of the JNG model solution in several variables, with a 1-yr time step. The SST time series at two points along the equator: (a) eastern and (b) central Pacific basin. The leading (c) spatial PC and (d) ST-PC time series.

The top panel in Fig. 4, as well as Fig. 2c, shows the leading PC $\{Y_1^k\}$ at 1-yr intervals. In Fig. 4, the black color marks the time interval used to learn the ANN model, and the red color marks the unknown behavior of the JNG model that should be predicted. The empirical model of Eq. (7) was learned in accordance with the procedure described in section 2 of Part I for different values of the dimension d and the maximum lag m, with $1 \le d \le 15$. The numbers m_f and m_g of neurons in the ANN that approximated the functions $\mathbf{f}$ and $\mathbf{g}$ also varied from 1 to 15.

We will not dwell here on the behavior of empirical models with different dimensionalities and values of d and m or different numbers of neurons. It suffices to show only the characteristic pattern of the predictions made on the basis of spatial EOFs alone. Accordingly, in the two lower panels of Fig. 4 are plotted predictions of the leading PC shown in the upper panel, using empirical models with ($d = 1$, $m = 5$) and ($d = 3$, $m = 5$), respectively.

It is evident from these two plots that the spatial EOF-based models failed to reconstruct the JNG model's behavior; not only are the future critical transitions predicted incorrectly, but the empirical model fails to simulate correctly the JNG model behavior during the learning interval.

b. Results using spatiotemporal EOFs

Let us now consider the results of empirical modeling when using a basis of spatiotemporal EOFs, as in Eq. (6).

As in the previous subsection, spatiotemporal EOFs were computed solely from a learning interval of 4000 yr. Next, the whole dataset was projected onto these EOFs to obtain the associated PCs that cover both the learning and the verification intervals.

The time lag T used to construct the matrix $\mathbf{Q}^k$ in Eq. (4) was equal to 20 months. Thus, the number of columns of the covariance matrix $\mathbf{C}^J$ used to obtain the ST-EOF basis is $40 \times 24 = 960$. Figure 5 shows the first six of these—namely, the first six matrices $\mathbf{E}_i$, $i = 1, \ldots, 6$, from
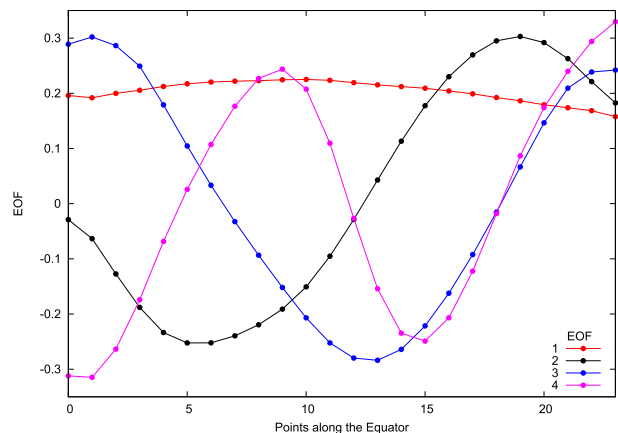

FIG. 3. The four leading spatial EOFs computed from the SST time series of the JNG model and used to develop the empirical model. The EOFs are ranked by variance captured, in decreasing order: 49%, 21.7%, 16.5%, and 5.7%, respectively.
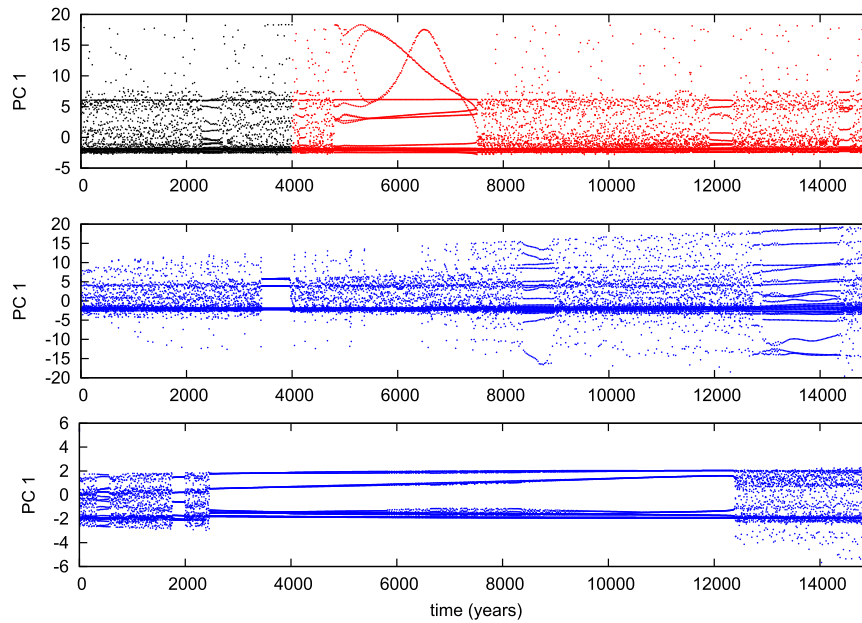
FIG. 4. Predictions of JNG model behavior by using empirical models governed by Eq. (7) and trained on the purely spatial PCs. (top) The time series of the JNG model's leading EOF. The black color marks the learning interval used to construct the empirical models, and the red color corresponds to the JNG model behavior that needs to be predicted empirically. The prediction results obtained by using ANN models with (middle) $d = 1$ and $m = 5$ and (bottom) $d = 3$ and $m = 5$.

section 2b—arranged in the order of decreasing variance captured by the ST-PCs Y_i^k.

These six leading ST-EOFs capture respectively 21.8%, 12.5%, 7.4%, 6.8%, 5.9%, and 5.7% of variance, for a total of just over 60%. They are pairwise orthogonal and do fall into oscillatory pairs, although the variance of EOF 1 is almost twice that of EOF 2: EOFs 1 and 2 capture the dominant QQ mode, already visible directly in Fig. 1, while EOFs 3 and 4 appear to be associated with the first harmonic of the seasonal cycle. In fact, it is known that this 6-month periodicity is important near the equator, since the solar heating is a maximum twice per year at the equator: once when the Sun crosses it southward and then again when it returns northward. Finally, EOFs 5 and 6 capture the annual cycle itself. The nonlinear interaction of these three modes collectively describes the complex chaotic behavior of the JNG model.

The leading ST-PC of the JNG model is shown in both the top panel of Fig. 6 and in Fig. 2d. Note that the evolution of the model's behavior, including its critical transitions, will not differ qualitatively when plotting either the components of the original variables $\mathbf{X}^k$ or the components of the transformed variables $\mathbf{Y}^k$, because of the linear character of the transformation

between the two vectors. In other words, critical transitions are represented equally well in the time series that are directly simulated here by the JNG model—or observed, in other potential applications of the proposed methodology—as well as in the purely spatial PCs and the ST-PCs used to train the model. This theoretical statement can be immediately confirmed by comparing the time series of the JNG model's variables, shown in Figs. 2a,b, and the time series of the PCs, shown in Figs. 2c,d. Therefore, correct prediction of the qualitative behavior of the variables $\mathbf{Y}^k$ given by Eq. (6) also implies correct prediction of the behavior in the original variables $\mathbf{X}^k$.

As with the spatial EOFs in the previous subsection, we performed experiments in which empirical models of the form Eq. (7), with different parameters d and m, were learned. The spatiotemporal transformation [Eq. (6)] proved to be quite efficient, as shown by the fact that the model with $d = 1$ and $m = 2$ turned out to successfully predict the JNG model's qualitative behavior during a time interval exceeding the duration of the learning sample. In the lower panel of Fig. 6, we plot the time series generated by this empirical model, in which $m_f = 9$ and $m_g = 5$ neurons were used in the ANN that approximated the functions $\mathbf{f}$ and $\mathbf{g}$, respectively.
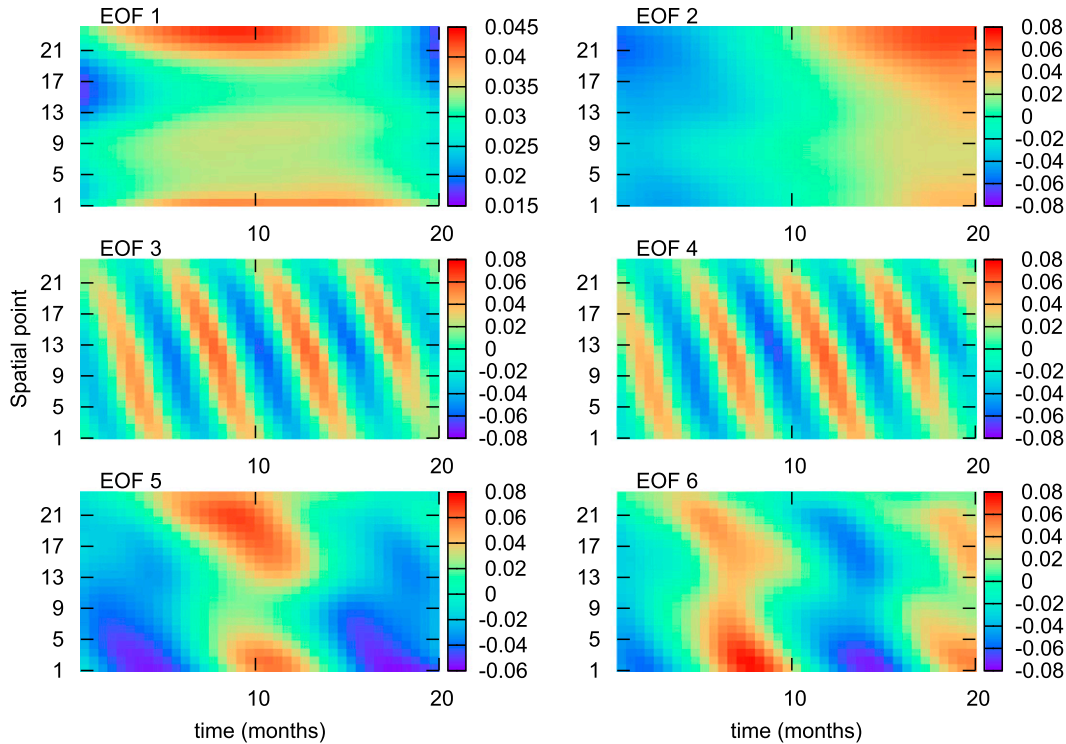
FIG. 5. The six leading ST-EOFs computed from the SST series produced by the JNG model and used to develop the empirical model. The EOFs are ranked by variance captured in decreasing order: 21.8%, 12.5%, 7.4%, 6.8%, 5.9%, and 5.7%, respectively.

The figure shows that we managed to predict the critical transitions connected with the decay of the chaotic regime and the birth of the periodic regime at $t = 4800$ while the rebirth of the chaotic regime from the periodic one at $t = 7800$ is predicted by the empirical model at $t = 6600$. A subsequent, narrower "laminar window" (Kadanoff 1983) in the ANN model's prediction, at $8100 \leq t \leq 4800$ is actually absent in the full JNG model's simulation, while a broad window in the lower panel, near $t = 12\,000$, corresponds to a narrower one in the upper panel. Thus, prediction of critical transitions for the spatially dependent, intermediate JNG model appears to be more difficult than for the scalar models in Part I, but still possible and thus promising.

Note that all the predicted critical transitions take place in every leading ST-PC (not shown), not just the first one, since they correspond to bifurcations of the JNG model's attractor. Hence, the problem of predicting a dynamical model's critical transitions can be reduced to predicting them in any single ST-PC or subset thereof.

Figure 7 displays power spectra computed by the Blackman–Tukey correlogram method for two intervals of the JNG model's simulation: the training interval $0 \leq t \leq 4000$ (top panel) and $8000 \leq t \leq 11\,000$ (bottom panel). In both panels the JNG model simulation is compared with the empirical model simulation (top panel) and prediction (bottom panel).

All four spectra capture, to various degrees, the dominant QQ mode and the lesser QB mode. The figure shows that the dynamic properties of the JNG model—and of ENSO observations—have been captured reasonably well by the empirical model (blue curve in both panels) for both the training interval and the predicted one; in particular the period and amplitude of the dominant QQ mode is well captured by the leading ST-PC.

The relative success of the prediction of critical transitions in the JNG model by using a low-dimensional empirical model with a small number of parameters, as seen in Fig. 6, can be explained by the apposite choice of the basis functions for the empirical model by means of the MSSA transformation [Eq. (6)]. The parameter values used in the empirical model, $d = 1$ and $m = 2$, mean that the function $\mathbf{f}$ in the model of Eq. (7) should approximate the dependence of the ST-PC Y_1^k at time k on the two preceding values of Y_1, namely Y_1^{k-1} and Y_1^{k-2}; the latter two values are plotted as the (x, y) coordinates in Fig. 8, while Y_1^k is plotted as the z coordinate.
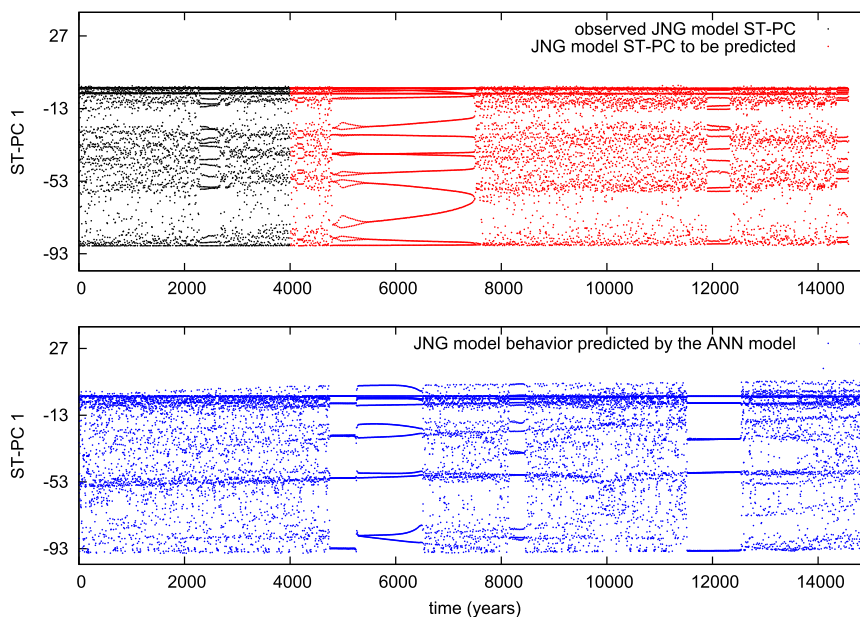
FIG. 6. As in Fig. 4, but for an empirical model governed by Eq. (7) and trained on the ST-PCs obtained by MSSA. (top) The leading ST-PC of the JNG model. Color conventions are as in Fig. 4. (bottom) The prediction results of the ANN model with $d = 1$ and $m = 2$.

The dependence of Y_1^k on (Y_1^{k-1}, Y_1^{k-2}), which is constructed during the learning interval, is shown by the black curve in the middle panel of Fig. 8. One notices that this dependence can be described fairly accurately by a single-valued function in a wide domain of the plane spanned by Y_1^{k-1} and Y_1^{k-2}. This fact plays a key role in obtaining the correct reconstruction. The red curve in the bottom panel marks a similar dependence constructed for the leading ST-PC of the JNG model during a later time interval, $8000 \leq t \leq 11\,000$.

In both these panels, the blue curve of the empirical model's solution stays quite close to that of the JNG model. Increasing either d or m does not improve the fit and tends to make the empirical model less robust.

The top panel of Fig. 8 displays, by contrast, the situation for the ANN model that used spatial EOFs as basis functions. Clearly, the black curve of JNG model behavior is not single valued, and the blue curve of the ANN model does not approximate it well at all.

5. Conclusions

In this paper, we introduced spatiotemporal basis functions as an important tool in implementing the empirical approach to climate system modeling and prediction. The results were compared with those obtained by the usual data compression method using purely spatial EOFs. Our spatiotemporal EOFs

(ST-EOFs) were obtained by an application of MSSA to data fields that vary in both space and time (see section 2).

The empirical ANN modeling introduced in Part I of this work for scalar time series was thus extended to simulate and predict the solutions of an intermediate, seasonally dependent coupled model of the tropical Pacific, the JNG model [see Jin and Neelin (1993a,b); Jin et al. (1994, 1996); and section 3 herein]. Specifically, we demonstrated that the empirical, ST-EOF-based model is capable of predicting critical transitions in the full JNG model's dynamics (see section 4).

The improvement in the simulation and prediction of the JNG model's behavior when comparing the ST-EOF-based empirical model with the one based on the usual spatial EOFs is striking (see especially Fig. 6 versus Fig. 4, as well as Fig. 8). Still, both the simulation during the training interval, $0 \leq t \leq 4000$ yr, and the prediction for $t > 4000$ yr is only approximate, and several ways of improving the ANN model's skill need to be considered.

These ways should include an optimization of the dimension d of the subspace in which the empirical model is developed (see, e.g., the two bottom panels of Fig. 4) along with the number m of delays taken into account in Eq. (7) and the maximum lag T considered in computing the ST-EOFs. Clearly, these parameters of the empirical model depend on the characteristic
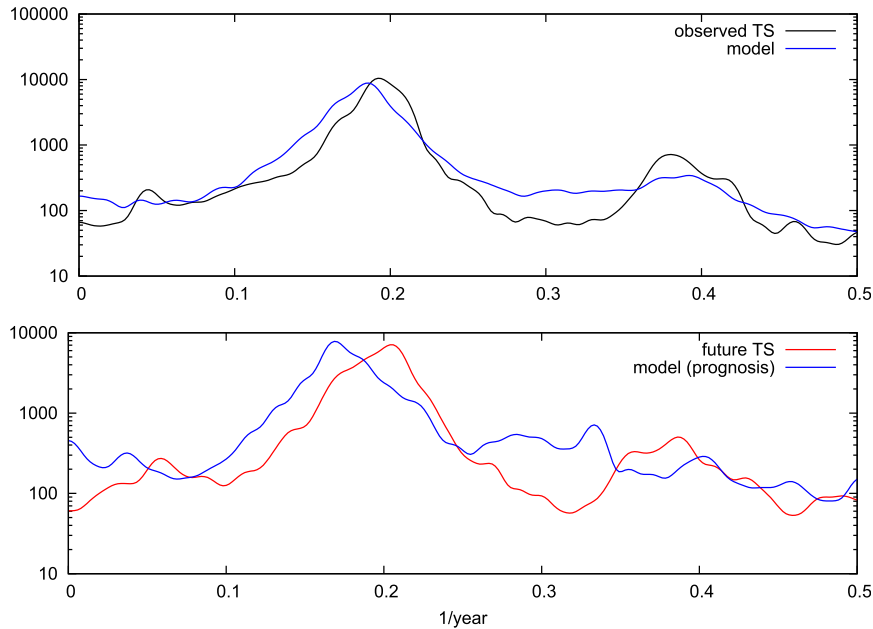
FIG. 7. Comparison of the power spectra of the JNG model's leading ST-PC and of the empirical model derived from it via Eq. (7). The spectra are computed by the windowed Blackman–Tukey correlogram (Kay and Marple 1981; SSA-MTM Group 2013) with a Bartlett window of size 70. (top) The black curve shows the spectrum of the leading ST-PC of the JNG model, marked in red in Fig. 6; (bottom) the red curve marks the spectrum of the same ST-PC for the time interval $8000 \leq t \leq 11\,000$ in Fig. 6. The blue curves in (top) and (bottom) mark the empirical model's spectra within the corresponding time intervals.

times and the complexity of the dynamics that we wish to simulate and predict; hence, no general recipes can be prescribed.

Another important ingredient for the practical application of the method proposed herein is the evaluation of its skill in predicting critical transitions. In previous papers (Molkov et al. 2011; Feigin et al. 2002), we proposed to calculate the probability of occurrence of a certain dynamical regime at a certain time in the future. The probability of the regime that actually occurs at some future time can be used as a measure of skill in predicting this regime. This probability can be estimated by a statistical analysis of the posterior distribution of the empirical model's parameters, as done in Eq. (7) of Part I; usually, we applied a Markov chain Monte Carlo (MCMC) method for sampling such posterior probability density functions (PDFs).

To do so, however, it is necessary first to identify the dynamical regime of interest in a precise, quantitative way: for instance, one can define the regime by the number of statistically significant maxima in the invariant measure (Ghil and Robertson 2002), the number of significant peaks in its spectral density (Ghil et al. 2002), and so on. Generally speaking, such a Monte Carlo analysis requires a large number of numerical computations. Since the aim of this two-part paper is just to demonstrate the possibility of empirical forecasting of critical transitions in climate dynamics, we simply used visual comparisons between observed versus reproduced behavior and plotted the output of the most probable model (i.e., the model whose parameter values maximize the posterior PDF). A more comprehensive analysis is left for future work devoted to concrete climate applications.
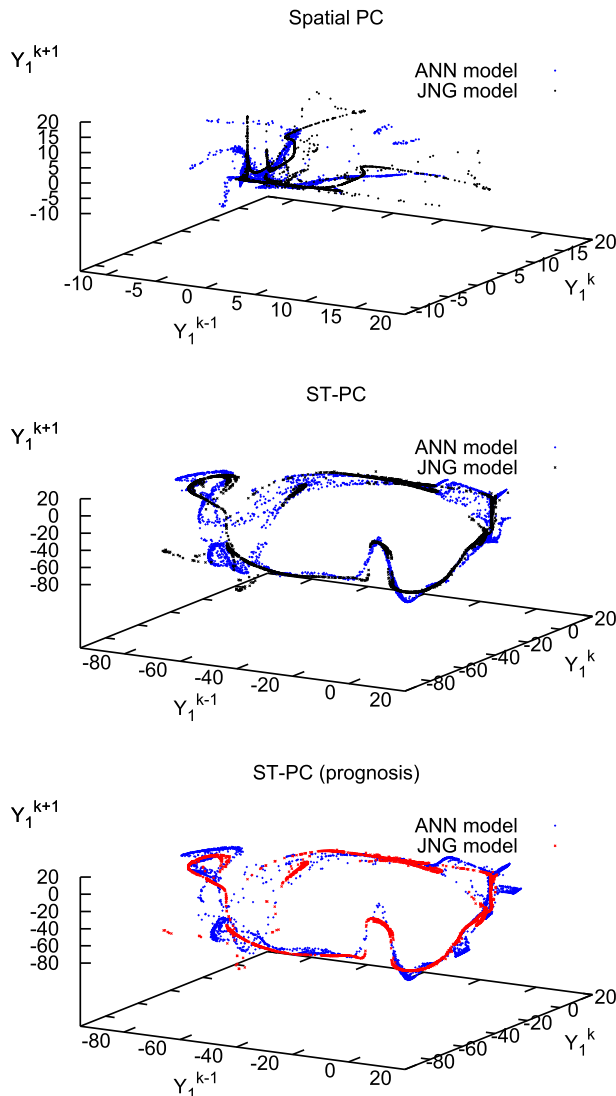
FIG. 8. Comparison of the phase portraits of the JNG model and the empirical model. The portraits are displayed in a three-dimensional subspace spanned by the leading ST-PC at three successive instants, $k - 1$, k, and $k + 1$. (top) ANN model based on spatial EOFs; (middle) ANN model based on ST-EOFs for the training interval $0 \leq t \leq 4000$; and (bottom) As in (middle), but for the time interval $8000 \leq t \leq 11\,000$. Color convention as in the previous figures: black for JNG model during training; red for the same model beyond the training interval, $t > 4000$; and blue for the ANN model.

APPENDIX A

Stochastic Model for Multivariate Time Series

This appendix explains the model form Eq. (7) that is well adapted to the use of multivariate climate time series, as considered in the paper.

We study a multivariate time series $\mathbf{U}(t) \in \mathbb{R}^d$. The components of the vector $\mathbf{U}$ can be the climate variables measured on a spatial grid or a set of leading PCs that capture much of the observed variability on that grid; in this paper, they are the set of PCs or ST-PCs obtained from the space-distributed SST time series. To model the evolution operator governing the dynamics of $\mathbf{U}(t)$, we have to define an embedding (i.e., the set of phase variables $\mathbf{W}$ that determines the model states). The simplest way is to complement the vector $\mathbf{U}$ by a number of time-lagged vectors that, together with $\mathbf{U}$ itself, yield an augmented vector $\mathbf{W}(t) = \{\mathbf{U}(t), \mathbf{U}(t + \tau), \ldots, \mathbf{U}[t + (m - 1)\tau]\}$; the new components W_i span a phase space of dimension $d \times m$. Such an embedding algorithm (Mañe 1981; Takens 1981) provides us with many possibilities of projecting the full, unknown phase space of the actual system under study onto a reduced subspace. An important problem, then, is to find optimal parameters d and m, along with the delay τ, that provide the simplest structure of the evolution operator in such a projection.

An obvious step in analyzing interannual climate variability is to use the Poincaré map that is naturally associated with the seasonal periodicity of the variables. In practice, we select an arbitrary phase of the seasonal cycle (i.e., a particular day, week, or month) and consider $\mathbf{W}(t)$ only at time instants that correspond to that phase and, consequently, are separated by one year. This step achieves two purposes: it reduces the dimension of the attractor and, simultaneously, excludes the zero Lyapunov exponent that is hard to capture in a discrete model.

Next, according to Part I, we construct the stochastic model of the map so obtained, in the form

$$\mathbf{W}(t_{k+1}) = \mathbf{f}[\mathbf{W}(t_k)] + \mathbf{g}[\mathbf{W}(t_k)]\boldsymbol{\zeta}(t_k), \quad \mathbf{W} \in \mathbb{R}^{dm}; \quad \text{(A1)}$$

here $\mathbf{f}: \mathbb{R}^{dm} \rightarrow \mathbb{R}^{dm}$ and $\mathbf{g}: \mathbb{R}^{dm} \rightarrow \mathcal{L}_{dm}(\mathbb{R})$ are matrix functions of the vector variable $\mathbf{W}$, time instants t_k are separated by one year, and $\boldsymbol{\zeta}(t_k)$ is white normalized Gaussian noise.

The choice of the delay $\tau = 1$ yr in Eq. (A1) may seem arbitrary, since there are many methods to estimate it (Casdagli et al. 1991; Buzug and Pfister 1992; Kantz and Schreiber 1997). In this paper, given that interannual ENSO dynamics has characteristic time scales of 2–7 yr, a value of τ of the an order of 1 yr is intuitively the proper choice for it. Especially so when using ST-PCs for our basis, as in Eq. (4) when applying MSSA, since the time lag T was taken larger than 1 yr (see section 4b); consequently, the leading ST-PCs have characteristic time scales that exceed τ.

It is easy to show that the choice of τ exactly equal to 1 yr allows one to simplify substantially the architecture of the model in Eq. (A1). Indeed, if

$$\begin{aligned}
\mathbf{U}(t_{k+1}) &= \mathbf{U}(t_k + \tau), \\
\mathbf{U}(t_{k+1} + \tau) &= \mathbf{U}(t_k + 2\tau), \\
&\vdots \\
\mathbf{U}[t_{k+1} + (m-1)\tau] &= \mathbf{f}\{\mathbf{U}(t_k), \mathbf{U}(t_k + \tau), \dots, \mathbf{U}[t_k + (m-1)\tau]\} \\
&\quad + \mathbf{g}\{\mathbf{U}(t_k), \mathbf{U}(t_k + \tau), \dots, \mathbf{U}[t_k + (m-1)\tau]\}\boldsymbol{\xi}(t_k).
\end{aligned} \tag{A2}$$

we rewrite this equation in terms of $\mathbf{U}$ for $\tau = 1$ yr, we obtain $m-1$ trivial couplings between the variables:

Thus, the last equation in the system of Eq. (A2) is equivalent to the model of Eq. (7). Note that now $\mathbf{f}: \mathbb{R}^{dm} \to \mathbb{R}^d$, and $\mathbf{g}: \mathbb{R}^{dm} \to \mathcal{L}_d(\mathbb{R})$ so that the target spaces of these vector functions are much smaller than in Eq. (A1).

The price to pay is that the model of Eq. (7), working as a Poincaré map, has an obvious drawback: it can reproduce data only with the resolution in time imposed by the 1-yr time step of the Poincaré map used for the reduction. This restriction makes it impossible, in particular, to use reconstructed components (Ghil and Vautard 1991; Vautard et al. 1992; Feliks et al. 2013), whose phases are fixed to within the sampling step of the time series, as opposed to PCs, whose phases are indeterminate to within the window width of the analysis (Vautard et al. 1992; Ghil et al. 2002). As shown here, though, the ST-PCs capture well the main critical transitions in the observed behavior. Hence, modeling and forecasting their behavior suffices for the purposes stated in this paper.

APPENDIX B

Brief Description of the JNG Model

a. The ocean

The coupled model's ocean dynamics is governed by linear shallow-water equations for the currents and a nonlinear equation for the SST field. The dynamical variables are the three velocity components (u, v, w) and the thermocline depth anomaly h. We describe here only the model's main features and key parameters. A more complete summary of the model derivation is given in appendix A of Sun et al. (2002), and all the parameter values are listed in Table 1 of that paper.

1) THE SEA SURFACE TEMPERATURE EQUATION

The SST in an equatorial band satisfies the semilinear, hyperbolic partial differential equation in the zonal direction

$$\frac{\partial T}{\partial t} + u_1 \frac{\partial T}{\partial x} + \mathcal{H}(w)\frac{w}{H_{1.5}}[T - T_{\text{sub}}(h)]$$

$$- \mathcal{H}(-v_N)\frac{2v_N}{L_y}(T - T_N) + \epsilon_T(T - T_0) = 0; \tag{B1}$$

Here, T is the temperature of the surface mixed layer, u_1 is the zonal and w the vertical velocity in this surface layer, while v_N is the meridional surface current at the northern boundary of the equatorial box. Symmetry of SST and antisymmetry of v_N about the equator are assumed.

In Eq. (B1), the Newtonian damping time is denoted by ε_T, and its value is set at $(90 \text{ days})^{-1}$; L_y denotes the width of the box; and T_N is the off-equatorial SST at a distance L_y from the equator. The depths H_1 and H_2 of the two layers are taken here to be 50 and 100 m, while $H_{1.5} = 75$ m is the depth scale that characterizes upwelling of the subsurface temperature T_{sub}. An analytical, smooth version $\mathcal{H}(x)$ of the Heaviside function is used in the terms of Eq. (B1) that represent upstream differencing of meridional and vertical advection into the equatorial surface strip.

The meridional velocity v_N is obtained by finite differencing the continuity equation, while w and u_1 each are a sum of three parts: a climatological, seasonally varying basic state; anomalous vertical mean currents above the thermocline obtained from the shallow-water equations [Eqs. (B2)–(B3) below]; and anomalous oceanic shear currents. The subsurface temperature T_{sub} is parameterized as a nonlinear function of thermocline depth anomaly h by assuming that a deeper thermocline is associated with warmer upwelled waters.

2) THE OCEANIC CURRENTS

The vertical mean motions above the thermocline are governed by the linearized, reduced-gravity shallow-water equations on a β-plane in the long-wave approximation. A standard semispectral discretization is used, with parabolic cylinder functions as basis functions in latitude; it leads to a truncated model, including the first Kelvin mode and the first 15 symmetric Rossby modes. Note that this represents a higher resolution than typically used for this model: Jin and Neelin (1993a) retained a total of eight ocean modes, followed by Jin et al. (1994, 1996) and Kondrashov et al. (2008).

The resulting equations for the 16 oceanic wave coefficients $\{q_n : n = 0, \dots, 15\}$ are

$$\left(\delta \frac{\partial}{\partial t} + \epsilon\right) q_0 + \frac{\partial q_0}{\partial x} = \tau_0 \qquad \text{(B2)}$$

for the amplitude q_0 of the Kelvin wave and

$$(n-1)\left(\delta \frac{\partial}{\partial t} + \epsilon\right) q_n - \frac{\partial q_n}{\partial x}$$
$$= n\tau_n - [(n-1)]^{1/2}\tau_{n-2} \qquad \text{(B3)}$$

for the 15 Rossby waves $\{n = 2p : p = 1, \dots, 15\}$ while τ_n is the zonal wind stress projected onto oceanic mode n. The original variables h, u, and v (i.e., the thermocline depth anomalies and anomalous vertical mean currents above the thermocline) are obtained by back transformation as appropriate linear combination of the q_ns. The relative adjustment coefficient δ measures the ratio of the time scale of adjustment by oceanic dynamics to the net time scale of SST change in Eq. (B1). This parameter affects the travel time of the equatorially trapped waves produced by the JNG model, and it helps determine the dynamical regime resulting from the interactions between the model's internal variability and the seasonal forcing (Jin et al. 1996).

The SST Eq. (B1) and Eqs. (B2) and (B3) for the mean zonal currents give a total of 17 equations for the model's 17 variables: the SST, one Kelvin mode, and 15 Rossby modes. These 17 equations are numerically discretized using 24 grid points along the equator, which in turn gives a total of $17 \times 24 = 408$ degrees of freedom.

The horizontal components u_s and v_s of the vertical shear currents are governed by steady-state equations dominated by damping due to interfacial stress between the layers (Zebiak and Cane 1987). The vertical velocity w in the surface layer enters into the SST Eq. (B1), and it equals w_s. This vertical component of the shear currents can be calculated from the continuity equation using u_s and v_s.

b. The atmosphere–ocean coupling

The climatological basic state with a seasonal cycle is constructed as a forced solution of the uncoupled ocean model with a wind stress $\bar{\tau}$ that resembles the annually observed wind stress in the Pacific Ocean along the equator. The oceanic component is then coupled with the atmospheric model to obtain the deviations from this basic state; it is these deviations that are referred to as anomalies. The coupled system is set up using one-way flux correction (Jin and Neelin 1993a), with total wind stress τ given by $\tau = \bar{\tau} + \tau'$, and the wind stress anomaly τ' is derived from the atmospheric response $T' = T - \bar{T}$ to SST anomalies. The response of the zonal wind stress anomalies τ' to the SST anomalies T' at the equator is assumed to be steady, with the parameter μ controlling ocean–atmosphere coupling.

A relative surface-layer parameter δ_s, varying from zero to unity, is introduced; it controls the intensity of the anomalous surface-layer currents as a function of the wind stress anomalies without affecting the climatology. Sensitivity studies with respect to this parameter were carried out in the Jin–Neelin model with time-constant forcing (Neelin 1991; Jin et al. 1996), as well as for the JNG model used here, with a seasonal cycle in its forcing (Jin et al. 1996). As indicated already in section 3, chaotic model regimes are obtained for certain ranges of δ_s values; furthermore, the mixing properties for these regimes, in the model's phase space are strongly affected by δ_s (cf. Chekroun et al. 2014). In the present paper, $\delta_s = 0.95$ is close to unity, so the correlation decay of the simulated Niño-3 index is faster than for small values of δ_s. The resulting behavior is less periodic and, thus, less predictable.

REFERENCES

Barnston, A. G., M. K. Tippett, M. L. L'Heureux, S. Li, and D. G. DeWitt, 2012: Skill of real-time seasonal ENSO model predictions during 2002–2011: Is our capability increasing? *Bull. Amer. Meteor. Soc.,* **93,** 631–651, doi:10.1175/BAMS-D-11-00111.1.

Broomhead, D. S., and G. P. King, 1986a: Extracting qualitative dynamics from experimental data. *Physica D,* **20,** 217–236, doi:10.1016/0167-2789(86)90031-X.

——, and ——, 1986b: On the qualitative analysis of experimental dynamical systems. *Nonlinear Phenomena and Chaos,* S. Sarkar, Ed., Adam Hilger, 113–144.

Buzug, T., and G. Pfister, 1992: Optimal delay time and embedding dimension for delay-time coordinates by analysis of the global static and local dynamical behavior of strange attractors. *Phys. Rev.,* **45A,** 7073–7084, doi:10.1103/PhysRevA.45.7073.

Casdagli, M., S. Eubank, J. D. Farmer, and J. Gibson, 1991: State space reconstruction in the presence of noise. *Physica D,* **51,** 52–98, doi:10.1016/0167-2789(91)90222-U.

Chekroun, M. D., J. D. Neelin, D. Kondrashov, J. C. McWilliams, and M. Ghil, 2014: Rough parameter dependence in climate models and the role of Ruelle–Pollicott resonances. *Proc. Natl. Acad. Sci. USA,* **111,** 1684–1690, doi:10.1073/pnas.1321816111.

Feigin, A. M., Y. I. Molkov, D. N. Mukhin, and E. M. Loskutov, 2002: Investigation of nonlinear dynamical properties by the observed complex behaviour as a basis for construction of dynamical models of atmospheric photochemical systems. *Faraday Discuss.,* **120,** 105–123, doi:10.1039/B102985C.

Feliks, Y., A. Groth, A. W. Robertson, and M. Ghil, 2013: Oscillatory climate modes in the Indian monsoon, North Atlantic, and tropical Pacific. *J. Climate,* **26,** 9528–9544, doi:10.1175/jcli-d-13-00105.1.

Ghil, M., and R. Vautard, 1991: Interdecadal oscillations and the warming trend in global temperature time series. *Nature,* **350,** 324–327, doi:10.1038/350324a0.

——, and A. W. Robertson, 2002: ''Waves'' vs. ''particles'' in the atmosphere's phase space: A pathway to long-range forecasting? *Proc. Natl. Acad. Sci. USA,* **99,** 2493–2500, doi:10.1073/pnas.012580899.

——, and Coauthors, 2002: Advanced spectral methods for climatic time series. *Rev. Geophys.,* **40,** 1003, doi:10.1029/2000rg000092.

Gill, A., 1980: Some simple solutions for heat-induced tropical circulation. *Quart. J. Roy. Meteor. Soc.,* **106,** 447–462, doi:10.1002/qj.49710644905.

Grieger, B., and M. Latif, 1994: Reconstruction of the El Niño attractor with neural networks. *Climate Dyn.,* **10,** 267–276, doi:10.1007/BF00228027.

Hasselmann, K., 1976: Stochastic climate models. Part I: Theory. *Tellus,* **28,** 473–485, doi:10.1111/j.2153-3490.1976.tb00696.x.

Hornik, K., M. Stinchcombe, and H. White, 1989: Multilayer feedforward networks are universal approximators. *Neural Networks,* **2,** 359–366, doi:10.1016/0893-6080(89)90020-8.

Jiang, N., J. D. Neelin, and M. Ghil, 1995: Quasi-quadrennial and quasi-biennial variability in the equatorial Pacific. *Climate Dyn.,* **12,** 101–112, doi:10.1007/BF00223723.

Jin, F.-F., and J. D. Neelin, 1993a: Modes of interannual tropical ocean–atmosphere interaction—A unified view. Part I: Numerical results. *J. Atmos. Sci.,* **50,** 3477–3503, doi:10.1175/1520-0469(1993)050<3477:MOITOI>2.0.CO;2.

——, and ——, 1993b: Modes of interannual tropical ocean–atmosphere interaction—A unified view. Part III: Analytical results in fully coupled cases. *J. Atmos. Sci.,* **50,** 3523–3540, doi:10.1175/1520-0469(1993)050<3523:MOITOI>2.0.CO;2.

——, ——, and M. Ghil, 1994: El Niño on the Devil's staircase: Annual subharmonic steps to chaos. *Science,* **264,** 70–72, doi:10.1126/science.264.5155.70.

——, ——, and ——, 1996: El Niño/Southern Oscillation and the annual cycle: Subharmonic frequency-locking and aperiodicity. *Physica D,* **98,** 442–465, doi:10.1016/0167-2789(96)00111-X.

Jolliffe, I., 1986: *Principal Component Analysis.* 2nd ed. Springer-Verlag, 488 pp.

Kadanoff, L. P., 1983: Roads to chaos. *Phys. Today,* **36,** 46–53, doi:10.1063/1.2915388.

Kantz, H., and T. Schreiber, 1997: *Nonlinear Time Series Analysis.* Cambridge University Press, 388 pp.

Kay, S. M., and S. L. Marple, 1981: Spectrum analysis—A modern perspective. *Proc. IEEE,* **69,** 1380–1419, doi:10.1109/PROC.1981.12184.

Keppenne, C. L., and M. Ghil, 1993: Adaptive filtering and prediction of noisy multivariate signals: An application to

subannual variability in atmospheric angular momentum. *Int. J. Bifurcation Chaos,* **3,** 625–634, doi:10.1142/S0218127493000520.

Kondrashov, D., S. Kravtsov, A. W. Robertson, and M. Ghil, 2005: A hierarchy of data-based ENSO models. *J. Climate,* **18,** 4425–4444, doi:10.1175/JCLI3567.1.

——, C. Sun, and M. Ghil, 2008: Data assimilation for a coupled ocean–atmosphere model. Part II: Parameter estimation. *Mon. Wea. Rev.,* **136,** 5062–5076, doi:10.1175/2008MWR2544.1.

Kravtsov, S., D. Kondrashov, and M. Ghil, 2009: Empirical model reduction and the modelling hierarchy in climate dynamics and the geosciences. *Stochastic Physics and Climate Modelling,* T. N. Palmer and P. Williams, Eds., Cambridge University Press, 35–72.

Mañe, R., 1981: On the dimension of the compact invariant sets of certain non-linear maps. *Dynamical Systems and Turbulence,* D. A. Rand and L.-S. Young, Eds., Lecture Notes in Mathematics, Vol. 898, Springer-Verlag, 230–242.

Molkov, Y. I., D. N. Mukhin, E. M. Loskutov, A. M. Feigin, and G. A. Fidelin, 2009: Using the minimum description length principle for global reconstruction of dynamic systems from noisy time series. *Phys. Rev.,* **80E,** 046207, doi:10.1103/PhysRevE.80.046207.

——, ——, ——, R. I. Timushev, and A. M. Feigin, 2011: Prognosis of qualitative system behavior by noisy, nonstationary, chaotic time series. *Phys. Rev.,* **84E,** 036215, doi:10.1103/PhysRevE.84.036215.

——, E. M. Loskutov, D. N. Mukhin, and A. M. Feigin, 2012: Random dynamical models from time series. *Phys. Rev.,* **85E,** 036216, doi:10.1103/PhysRevE.85.036216.

Mukhin, D., E. Loskutov, A. Mukhina, A. Feigin, I. Zaliapin, and M. Ghil, 2015: Predicting critical transitions in ENSO models. Part I: Methodology and simple models with memory. *J. Climate,* **28,** 1940–1961, doi:10.1175/JCLI-D-14-00239.1.

Navarra, A., and V. Simoncini, 2010: *A Guide to Empirical Orthogonal Functions for Climate Data Analysis.* Springer-Verlag, 151 pp.

Neelin, J. D., 1991: The slow sea surface temperature mode and the fast-wave limit: Analytic theory for tropical interannual oscillations and experiments in a hybrid coupled model. *J. Atmos. Sci.,* **48,** 584–606, doi:10.1175/1520-0469(1991)048<0584:TSSSTM>2.0.CO;2.

Penland, C., 1996: A stochastic model of IndoPacific sea surface temperature anomalies. *Physica D,* **98,** 534–558, doi:10.1016/0167-2789(96)00124-8.

Plaut, G., and R. Vautard, 1994: Spells of low-frequency oscillations and weather regimes in the Northern Hemisphere. *J. Atmos. Sci.,* **51,** 210–236, doi:10.1175/1520-0469(1994)051<0210:SOLFOA>2.0.CO;2.

Preisendorfer, R. W., 1988: *Principal Component Analysis in Meteorology and Oceanography.* Developments in Atmospheric Sciences, Vol. 17, Elsevier, 444 pp.

Sauer, T., J. A. Yorke, and M. Casdagli, 1991: Embedology. *J. Stat. Phys.,* **65,** 579–616, doi:10.1007/BF01053745.

SSA-MTM Group, cited 2013: SSA-MTM Toolkit for Spectral Analysis, Version 4.4. [Available online at http://www.atmos.ucla.edu/tcd/ssa/.]

Strounine, K., S. Kravtsov, D. Kondrashov, and M. Ghil, 2010: Reduced models of atmospheric low-frequency variability: Parameter estimation and comparative performance. *Physica D,* **239,** 145–166, doi:10.1016/j.physd.2009.10.013.

Sun, C., Z. Hao, M. Ghil, and J. Neelin, 2002: Data assimilation for a coupled ocean–atmosphere model. Part I: Sequential state

estimation. *Mon. Wea. Rev.,* **130,** 1073–1099, doi:10.1175/1520-0493(2002)130<1073:DAFACO>2.0.CO;2.

Takens, F., 1981: Detecting strange attractors in turbulence. *Dynamical Systems and Turbulence,* D. A. Rand and L.-S. Young, Eds., Lecture Notes in Mathematics, Vol. 898, Springer-Verlag, 366–381.

Vautard, R., and M. Ghil, 1989: Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D,* **35,** 395–424, doi:10.1016/0167-2789(89)90077-8.

——, P. Yiou, and M. Ghil, 1992: Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Physica D,* **58,** 95–126, doi:10.1016/0167-2789(92)90103-T.

Zebiak, S. E., and M. A. Cane, 1987: A model El Niño–Southern Oscillation. *Mon. Wea. Rev.,* **115,** 2262–2278, doi:10.1175/1520-0493(1987)115<2262:AMENO>2.0.CO;2.