An Experimental Stochastic Model of the El Niño – Southern Oscillation System at Climatic Time Scales

Victor Privalsky^{1,*}, Sergey Muzylev²

¹N.N. Zubov State Oceanographic Institute, Moscow, 119034, Russian Federation

²P.P. Shirsov Institute of Oceanology, Russian Academy of Sciences, Moscow, 117218, Russian Federation

*Corresponding Author: vprivalsky@oceanography.ru

Copyright © 2013 Horizon Research Publishing All rights reserved.

Abstract An explicit bi-variate stochastic model of the El Niño – Southern Oscillation system (ENSO) is built for the first time directly from observations of the Southern Oscillation Index (SOI) and sea surface temperature (SST) in the Niño region 3.4 (1876-2011). According to the resulting stochastic difference equation, the system remembers its past for two years, has the eigenfrequency 0.24 year^{-1} and the damping factors 0.5 and 0.05. Statistics of SST variations, including properties of El Niño and La Niña, are also estimated through a Monte Carlo experiment. The time-domain model is used to describe ENSO in the frequency domain, including maximum entropy estimates of spectra, coherent spectra, coherence function, and frequency response function. The spectra are peaked at 0.22 year^{-1} ; the coherence increases from 0.6 at the lowest frequencies to over 0.9 between 0.18 year^{-1} and 0.33 year^{-1} exceeding 0.8 at all frequencies above 0.12 year^{-1} . This means that SST and SOI determine each other by up to 85% everywhere but at the lowest frequencies. The gain factor is peaked at 0.24 year^{-1} while the phase factor's magnitude is close to π everywhere. Therefore, the statistical predictability of the bi-variate ENSO time series does not improve much over the scalar case predictability.

Keywords ENSO, Multi-variate Time Series, Time and Frequency Domain Analysis

1. Introduction

The El Niño – Southern Oscillation (ENSO) system is regarded as “the strongest internal climate mode on interannual time scales” [1] and “as one of the best studied climate phenomena, both observational and theoretical” [2]. Yet, in the multitude of publications on ENSO we could not find an explicit experimental (based upon analysis of observations) description of ENSO as a bi-variate stochastic system in the time and frequency domains. Our paper fills this gap.

Using the actual time series of mean annual Southern

Oscillation Index (SOI) and sea surface temperature (SST) in the ENSO region 3.4, we will construct a stochastic model of the bi-variate ENSO time series and study properties of sea surface temperature in the time and frequency domains using observations as well as Monte Carlo simulations of the system's behavior over long time intervals. Our goal is to study statistical properties of the ENSO system – the sea surface temperature and, to a lesser extent, the Southern Oscillation Index – as components of a bi-variate random process; studying properties of El Niño and La Niña events constitutes only a part of this work.

Statistical properties of the ENSO system are examined here at climatic time scales, which begin, roughly, with year-to-year variability; this is why the unit time interval Δt is selected to be 1 year. Also, at $\Delta t = 1$ month the energy of both SST and SOI oscillations at frequencies exceeding 0.5 year^{-1} are orders of magnitude smaller than at the lower frequencies.

We believe that our detailed description of the ENSO stochastic system based on experimental data rather than on conceptual theoretical constructions can be used for validation of both theoretical models and the results of numerical simulations of climate within the IPCC program.

2. Data and Methods

The initial data for the study include the time series of SOI mean annual SST within the ENSO region 3.4 (averaged from 5°S to 5°N and from 120°W to 170°W), both from 1876 through 2011. The SOI data are published by the Australian Bureau of Meteorology [3] while the SST data were extracted from the file HadCRUT3 [4].

The observed time series

$$\mathbf{x}_n = [x_{1,n}, x_{2,n}]', n = 1, \dots, N,$$

Where $N = 136$ (and the strike means transposition), is regarded here as a sample of a bi-variate linearly regular random process, and our task is to evaluate its statistical properties in the time and frequency domains. Being a sample of a linearly regular process, the time series $\mathbf{x}_n$ can be

presented as the output of a linear filter with a bi-variate white noise at the input. Specifically, it is assumed here that $\mathbf{x}_n$ is a sample of a zero mean autoregressive process of order p , or AR(p):

$$\mathbf{x}_n = \sum_{j=1}^p \Phi_j \mathbf{x}_{n-j} + \mathbf{A}_n, \quad (1)$$

where

$$\Phi_j = \begin{bmatrix} \phi_{11}^{(j)} & \phi_{12}^{(j)} \\ \phi_{21}^{(j)} & \phi_{22}^{(j)} \end{bmatrix} \quad (2)$$

are matrix AR coefficients and $\mathbf{A}_n = [a_{1,n}, a_{2,n}]'$ is a bi-variate innovation (white noise) sequence with the covariance matrix

$$\mathbf{R}_A = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}. \quad (3)$$

Here R_{11} and R_{22} are the variances and $R_{12} = R_{21}$ covariances, of the white noise components $a_{1,n}$ and $a_{2,n}$. This is a proper and efficient approach in time series analysis, especially in situations when the number of observations is limited, as in our case. Note also that the AR models are capable of presenting random processes with practically any given spectrum. Estimates of the AR order p , matrix AR coefficients Φ_j , $j = 1, \dots, p$, and the white noise covariance matrix $\mathbf{R}_A$ are required in order to describe properties of the time series $\mathbf{x}_n$ in the time and frequency domains. For the sake of definiteness, we will be regarding the time series of SST $x_{1,n}$ and SOI $x_{2,n}$ as the output and input, respectively, of a linear system, which is described in the time domain with the above given stochastic difference equation (1). The matrices (2) and (3) are estimated here by using the multi-variate version of the Levinson's algorithm while the optimal order p of autoregression is selected on the basis of four criteria: Akaike's AIC, Parzen's BIC, Hennen-Quinn's CAT, and Schwarz-Rissanen's Ψ [5,6].

The spectral matrix $\mathbf{s}(f)$ that corresponds to Eq. (1) is

$$\mathbf{s}(f) = \frac{2|\mathbf{R}_A|\Delta t}{\left| \mathbf{I} - \sum_{j=1}^p \Phi_j \exp(-i2\pi jf\Delta t) \right|^2}, \quad 0 \leq f \leq 1/2\Delta t,$$

where $\mathbf{I}$ is the identity 2×2 matrix, $|\mathbf{R}_A|$ the determinant of the matrix $\mathbf{R}_A$, f cyclic frequency (year^{-1}), and $i = \sqrt{-1}$. The elements of the spectral matrix

$$\mathbf{s}(f) = \begin{bmatrix} s_{11}(f) & s_{12}(f) \\ s_{21}(f) & s_{22}(f) \end{bmatrix} \quad (4)$$

are the spectral densities $s_{11}(f)$ and $s_{22}(f)$ of the output $x_{1,n}$ (SST 3.4) and the input $x_{2,n}$ (SOI), respectively, and $s_{12}(f) = \overline{s_{21}(f)}$ are the complex-valued cross-spectra (the bar means complex conjugation). These quantities are used to calculate other functions that describe the response of the output component $x_{1,n}$ of the ENSO system to the input $x_{2,n}$ in the frequency domain:
Coherence function

$$Co_{12}(f) = |s_{12}(f)| [s_{11}(f)s_{22}(f)]^{-1/2},$$

$$\text{coherent spectrum } Cs_{12}(f) = Co_{12}^2(f)s_{11}(f),$$

$$\text{gain factor } G_{12}(f) = |s_{12}(f)| / s_{11}(f), \text{ and}$$

$$\text{phase factor } \phi_{12}(f) = \tan^{-1}[\text{Im } s_{12}(f) / \text{Re } s_{12}(f)].$$

The confidence intervals for these spectral characteristics can be calculated in accordance with the number of degrees of freedom ν for the spectral estimates [7]; for the case of autoregressive analysis of bi-variate time series, ν can be assumed equal, as a first approximation, to $N/2p$ [8]. Spectral characteristics obtained in this manner for Gaussian time series with a properly chosen AR order present maximum entropy estimates.

The autoregressive approach to the analysis of multi-variate time series has at least two advantages over the non-parametric approach. Firstly, it results in an explicit description of the time series in the time domain with bi-variate (in our case) stochastic difference equations. Secondly, through assuming a specific parametric form for the time series, one can obtain relatively reliable estimates of statistics for short time series whose length is comparable to the time scales of interest.

Figure 1. Observed annual values of SOI and SST in the Niño region 3.4, 1876-2011.

3. Analysis and Discussion

3.1. Time domain

As can be seen from Figure 1, the behavior of ENSO from 1876 through 2011 does not seem to contradict the hypothesis of stationarity for the random process that generated the time series $\mathbf{x}_n$. If the time series is split in halves, each containing 68 values, the statistical moments such as mean values, medians, variances, correlation

functions, and spectral densities estimated for these shorter time series do not differ significantly from each other or from respective estimates for the entire time series. The estimates of skewness and excess for the halves of the time series do not differ much from each other and satisfy the Gaussianity conditions. Thus, the time series of SOI and SST can be regarded as samples of Gaussian stationary processes.

When approximating the bi-variate ENSO time series $\mathbf{x}_n$ with autoregressive models of orders from one through five the AR order $p = 2$ has been selected by all four criteria, with the respective linear stochastic difference equation

$$\begin{aligned} x_{1,n} &= 0.55x_{1,n-1} + 0.03x_{2,n-1} - 0.37x_{1,n-2} - 0.01x_{2,n-2} + a_{1,n} \\ &\quad (0.15) \quad (0.01) \quad (0.15) \quad (0.01) \\ x_{1,n} &= -2.57x_{1,n-1} + 0.06x_{2,n-1} + 5.09x_{1,n-2} + 0.16x_{2,n-2} + a_{2,n} \\ &\quad (1.73) \quad (0.16) \quad (1.73) \quad (0.16) \end{aligned} \quad (5)$$

and the white noise covariance matrix

$$\mathbf{R}_A = \begin{bmatrix} 0.33 & -3.07 \\ (0.04) & (0.43) \\ -3.07 & 41.61 \\ (0.43) & (4.79) \end{bmatrix}. \quad (6)$$

The numbers in parentheses in Eqs. (5) and (6) show the rms errors of respective estimates. (Only two decimal digits are kept here.)

The system (5) can be re-written as

$$\begin{aligned} x_{1,n} &= 0.55x_{1,n-1} + 0.03x_{2,n-1} - 0.37x_{3,n-1} - 0.01x_{4,n-1} + a_{1,n} \\ x_{2,n} &= -2.57x_{1,n-1} + 0.06x_{2,n-1} + 5.09x_{3,n-1} + 0.16x_{4,n-1} + a_{2,n} \\ x_{3,n} &= x_{1,n-1} \\ x_{4,n} &= x_{2,n-1} \end{aligned} \quad (7)$$

where $x_{3,n}$ and $x_{4,n}$ are new variables. In a matrix form:

$$\mathbf{X}_n = \Psi \mathbf{X}_{n-1} + \tilde{\mathbf{A}}_n,$$

where

$$\mathbf{X}_n = [x_{1,n}, x_{2,n}, x_{3,n}, x_{4,n}]',$$

$$\tilde{\mathbf{A}}_n = [a_{1,n}, a_{2,n}, 0, 0]',$$

and, according to Eqs. (5) and (7)

$$\Psi = \begin{bmatrix} 0.55 & 0.03 & -0.37 & -0.01 \\ -2.57 & 0.06 & 5.09 & 0.16 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

The respective system of homogeneous equations has four linearly independent solutions λ_1^n , λ_2^n , λ_3^n , and λ_4^n , where λ_i , $i = 1, \dots, 4$, is the i -th eigenvalue of the matrix Ψ (e.g., [9]). In our case,

$$\lambda_{1,2} = 0.046 \pm i0.500, \lambda_3 = 0.048, \lambda_4 = 0.474.$$

The complex-conjugated roots can be written as

$$\lambda_{1,2} = \rho_{1,2}(\cos 2\pi f_e \Delta t \pm i \sin 2\pi f_e \Delta t),$$

where

$$f_e = \tan^{-1}(\text{Im } \lambda_{1,2} / \text{Re } \lambda_{1,2}) / 2\pi \approx 0.24.$$

The damping factors $\rho_i = |\lambda_i|$, so that $\rho_{1,2} \approx 0.5$, $\rho_3 \approx 0.05$, and $\rho_4 \approx 0.5$.

This means that the bi-variate ENSO time series presents an oscillatory system with the eigenfrequency $f_e \approx 0.24$ year⁻¹ and the damping factors 0.05 and 0.5. The real roots λ_3 and λ_4 correspond to transient exponentially-decaying processes inside the ENSO system.

Note that according to this statistically-optimal bi-variate model of ENSO the current values of SOI and SST depend upon a linear combination of SOI and SST values for the two previous years. This means that the model (7) that was obtained directly from analysis of observations is more complicated than conceptual physical or stochastic models developed under assumptions about the oscillatory form of the basic equations that describe the ENSO system (e.g., [10-12]).

The presence of a feedback loop in the stochastic equations (5) that describe annual variations of ENSO can lead to mutual modulation of SST by SOI and SOI by SST. Examples are given in Figure 2.

This oscillatory behavior of the SST–SOI system obtained from our experimental stochastic model is similar to the behavior of the bi-variate system SST–thermocline depth in the equatorial Pacific as described with the conceptual model developed in Jin[10]. According to Jin's conceptual model, the ENSO system has a robust oscillatory mode in the range from 3 to 5 years; according to our observation-based model, the period of oscillations is $T_e = 1/f_e \approx 4.2$ years. In the recent publication [13] dedicated to the effects of random forcing upon the behavior of SST at smaller time scales, the eigenfrequency of the underlying oscillation was assumed to be 0.25 year⁻¹; this assumption is quite close to our experimental result.

The statistical predictability of interannual variations of SST as a scalar process is known to be low [14]. Our results show that it does not become much better in the bi-variate case. As seen from the white noise covariance matrix $\mathbf{R}_A$ above, the one-step relative SST prediction error $d(1)$, which is equal to the ratio R_{11}/σ_T^2 (where σ_T^2 is the SST variance), amounts to 0.84, and, of course, the error $d(\tau)$ cannot improve as the lead time τ grows. Note, however, that the quality of prediction of SST as a scalar process (independent of SOI) is even worse: $d(1) = 0.90$. In other words, the transition from a scalar to a bi-variate model does not significantly improve the statistical predictability of SST within the Kolmogorov-Wiener theory. As the ENSO time series can be regarded as Gaussian, this poor statistical predictability statement holds for nonlinear approaches to statistical extrapolation of ENSO (or any other Gaussian process) as well [15].

Figure 2. Examples of damped periodic oscillations in observed (1) and simulated (2) bi-variate ENSO time series.

All in all, the description of the ENSO system in the time domain with the bi-variate stochastic difference equation (5) shows that the system

- is inertial (the dependence on the past extends for two years at $\Delta t = 1$ year);
- has an oscillatory mode with the characteristic time scale of about 4.2 years and the damping factors 0.05 and 0.5;
- contains a feedback loop which seems to be able to excite damped periodic oscillations within the SOI/SST system though the main contribution to SST variations comes from the external stochastic forcing;
- has poor statistical predictability properties.

Equations (5) and (6) can be used to simulate the bi-variate time series $\mathbf{x}_n$ for arbitrarily long time intervals and thus to study its statistical properties including those that cannot be deduced with sufficient reliability from the observed data (such as higher statistical moments, the range, etc.). An example of simulated SST data is given in Figure 3. The analysis of a long (50,000 years) time series of SST simulated in accordance with Eqs. (5) and (6) showed that the simulated series has the same parameters – autoregression order, coefficients and the noise covariance matrices – as the observed data.

We will concentrate on the analysis of the SST component of the time series. As seen from Table 1, the first four statistical moments of the simulated SST do not differ significantly from those obtained for the observed time series. The Gaussian approximation for the probability density is

acceptable for both the observed and simulated data. This agreement between observations and simulations allows one to study additional statistical characteristics of SST using simulated data.

Figure 3. Simulated (1) and observed (2) annual SST.

The two columns of the table describe statistically equivalent (at the 95% confidence level) sets of random variables whose probability densities are either Gaussian or close to it.

Table 1. Major statistical properties of observed and simulated annual SST

Quantity	Observations	Simulations
N , years	136	50,000
Mean value, °C	25.62	25.62
Standard deviation, °C	0.63	0.62
Standardized skewness	1.78	0.10
Standardized excess	-0.45	-0.69
Minimum, °C	24.44	22.92
Maximum, °C	27.42	28.17
Range, °C	2.98	5.25

The results of analysis of El Niño and La Niña maximum and minimum temperatures in the actual and simulated data are shown in Table 2. It is assumed here that we have an El Niño (La Niña) event if the annual temperature exceeds the temperature averaged over the entire SST time series by at least 0.5°C (drops below the average by at least 0.5°C).

Table 2. Maximum (minimum) temperature anomalies during El Niño (La Niña) years.

Quantity	Maximum (El Niño)		Minimum (La Niña)	
	Observed	Simulated	Observed	Simulated
Number of occasions	20	7261	19	7173
Mean, °C	0.99	0.90	-0.86	-0.91
Standard deviation, °C	0.34	0.31	0.19	0.31
Standardized skewness	1.56	35.5	0.34	-38.2
Standardized excess	0.09	15.6	-0.91	23.9
Minimum, °C	0.52	0.50	-1.18	-2.70
Maximum, °C	1.80	2.55	-0.56	-0.50
Range, °C	1.28	2.05	0.62	2.20

Table 3. Statistics of intervals between successive El Niño (La Niña) events

Quantity	El Niño's		La Niña's	
	Observed	Simulated	Observed	Simulated
Number of occasions	19	7260	18	7172
Mean, yrs	6.95	6.89	6.67	6.97
Standard deviation, yrs	4.20	4.38	3.91	4.42
Standardized skewness	2.33	66.9	1.81	68.9
Standardized excess	1.01	97.2	0.21	107.9
Minimum, yrs	2	2	2	2
Maximum, yrs	17	49	16	47
Range, yrs	15	47	14	45

As seen from the table, the numbers of occurrences of El Niño and La Niña events and their major properties are rather symmetric with respect to the mean SST. This should have been expected because SST variations are Gaussian.

The higher statistical moments of respective probability densities of the simulated data – skewness and excess – have much larger magnitudes than for the observed data and indicate that the probability distributions are by no means Gaussian as they are bounded on one side by 0.5 (or by -0.5). Further analysis shows that maximum (minimum) temperatures during El Niño (La Niña) occurrences can be approximated rather satisfactorily with Gumbel maximum (minimum) probability distribution.

The results of analysis of intervals between successive El Niño (La Niña) occurrences are shown in Table 3. According to both the actual and simulated data, the average interval between successive El Niño's (or La Niña's) amounts to 7 years; it can be as long as 17 (16) years (observations) and almost 50 years (simulations). The probability densities of the lengths of intervals between successive El Niño's and between successive La Niña's are best approximated with a Poisson distribution with parameter $\lambda = 7$. The estimates of probability density types for extreme temperatures and

intervals between events for the observation data are unreliable because of the small number of data (20 points or less).

The longest run with El Niño events in the 50,000 years simulated data set was 8 consecutive years (7 years for La Niña). However, the probability of such extreme events is negligibly small ($\approx 10^{-4}$) while most El Niño's and La Niña's runs contain one or two years (66% – 67% and 27% – 28%, respectively); events that occur three years in a row have probabilities about 0.05. Having in mind the short length of the observed time series, these results do not seem to disagree with observations.

3.2. Frequency Domain

The maximum entropy estimates of SOI and SST spectra are shown in Figure 4. Note first that the results of spectral analysis of the simulated ENSO data (shown with crosses) that were discussed earlier practically coincide with respective spectra estimated from the observed time series. This statement is true of all other frequency domain characteristic and it proves the high quality of stochastic simulations of the ENSO system.

Figure 4. Spectral densities of SST and SOI (1) with approximate 90% confidence limits (2). Results for simulated data ($N = 50,000$) are shown with crosses here and in all other figures.

The spectra shown in Figure 4 reveal two striking features. Firstly, their general shape differs from the red-noise model characteristic of many if not most processes at climatic time scales. (Examples include the World Ocean SST, global, hemispheric and land surface temperature, Pacific Decadal Oscillation, Atlantic Multi-Decadal Oscillation, North Atlantic Oscillation (white noise), river flow, global atmospheric pressure, etc.) Secondly, the spectra contain a wide peak at intermediate frequencies, which should have been expected because of the presence of the damped periodic oscillation in the bi-variate ENSO system.

The presence of maxima somewhere between 0.15 year^{-1} and 0.3 year^{-1} in observation-based estimates of SST and SOI spectra in this area of the Pacific is well-known (e.g., [16,17]). We can see here that there is just one maximum at about 0.22 year^{-1} . This and the behavior of the spectral density at lower frequencies makes the ENSO spectra rather unique in climate.

The correlation coefficient between SOI and SST is known to be high (about -0.83 [e.g., 18,19]), which is unusual for climatic scale processes. However, the correlation coefficient should not generally be used as an indicator of the degree of linear dependence between random processes. High values of linear correlation do mean that the time series are closely related (as in our case) but the dependence of the relation on frequency remains unknown. Low values of correlation do not necessarily mean the lack of a strong linear dependence. A simple example would be the correlation coefficient or even the entire cross-correlation function between a stationary random process and its first increment: it can be close to zero everywhere in spite of the strictly linear relation between the two processes. Relations

between time series should be described in the frequency domain using the coherence function, coherent spectra, and the frequency response function with its gain and phase factors. An example of such calculations for the Southern Oscillation Index contribution to the global temperature time series was given in Privalsky and Jensen [20].

The coherence function for the ENSO system (Figure 5) stays statistically significant everywhere, exceeds 0.9 within the frequency band between 0.18 year^{-1} and 0.33 year^{-1} , and never goes below 0.81 at frequencies above approximately 0.12 year^{-1} . This means, in particular, that the linear relation between SOI and SST is responsible for at least 65% and up to 85% of the power of their variations at $f > 0.12 \text{ year}^{-1}$ and in the band between 0.18 year^{-1} and 0.33 year^{-1} , respectively. The smallest but still significant coherence (0.56 and higher) occurs at the lowest frequencies.

Figure 5. Coherence functions $Co_{12}(f)$: SST-SOI (1) with approximate 90% confidence limits (2) and the approximate upper 90% limit for the true zero coherence (3); SST-global surface temperature (4); SOI-Pacific Decadal Oscillation (5)

Such behavior is unusual for coherence function estimates at climatic scales. To the best of the authors' knowledge, this issue has not been studied yet but it looks like the "normal" behavior of the coherence function between climatic processes would be a more or less monotonic decrease with frequency (see examples in Figure 5). The ENSO system seems to be an exception.

The high coherence between SOI and SST explains why the coherent spectrum of SST (the spectrum of the part of SST strictly coherent to SOI) is so close to the full spectrum, especially at frequencies above 0.15 year^{-1} (Figure 6). Note also that the largest share of the SST spectrum which is not related at all to SOI belongs to the lowest frequencies.

Figure 6. Coherent SST spectrum $C_{S12}(f)$ (1) with approximate 90% confidence limits (2); SST spectrum (3); non-coherent SST spectrum (4).

The gain factors $G_{12}(f)$ and $G_{21}(f)$ shown in Figure 7 describe how SOI variations are transformed into SST variations and vice versa at different frequencies; in agreement with the behavior of the spectra and the coherence function, they increase from $f=0$, reach the maximum at $f \approx 0.24 \text{ year}^{-1}$ and then stay high at frequencies above 0.15 year^{-1} .

Figure 7. Gain factors SOI-SST (1) and SST-SOI (3) with approximate 90% confidence limits (2, 4).

As seen from the figure, the gain factor $G_{21}(f)$ that describes how SST is transformed into SOI is quite similar to the SOI-SST gain factor $G_{12}(f)$ which means that the components of the bi-variate ENSO process excite each other in the same manner.

The phase factor $\varphi_{12}(f)$ (Figure 8) is close to $-\pi$ practically everywhere, which means that the oscillations of SOI and SST occur mostly in opposite directions at all frequencies: SST increases when SOI decreases and vice versa. In the frequency band from 0.2 year^{-1} to 0.4 year^{-1} , the phase factor may differ statistically significantly from $-\pi$ but this does not help with forecasting SOI or SST because the phase difference in that band, when transformed into time units, amounts to less than 1 year, the unit time step. This is another feature that confirms the low statistical predictability within the system. Obviously, the phase factor SOI-SST will be a mirror image of the factor SST-SOI with respect to π .

Figure 8. Phase factor SST-SOI (1) with approximate 90% confidence limits (2).

The above empirical results seem to agree with the recent theoretical stochastic model of ENSO suggested by Kleeman [12] and based upon the assumption that the ENSO system behaves as a bi-variate Ornstein-Uhlenbeck process. Firstly, the shape of the spectra in Figure 4 is similar to the theoretical spectrum suggested in that work for the dissipation parameter values between 0.2 and 0.8. Secondly, Kleeman's model explains the very high coherence observed between the components of the ENSO bi-variate time series (see Figure 5). On the other hand, the Ornstein-Uhlenbeck process is a Markov process, which corresponds to a bi-variate autoregressive process of order $p = 1$. In our discrete-time bi-variate case the AR order $p = 2$ which means that the observation-based model (5) is not Markovian. Its Ornstein-Uhlenbeck (Markovian) analog would be a four-variate AR process of order $p = 1$ [see Eq. (7)].

4. Conclusions

In this work, we applied methods of autoregressive analysis of multi-variate time series and maximum entropy

spectral estimation to observations of the annual Southern Oscillation Index and the annual sea surface temperature in the Niño region 3.4 for 1876-2011.

Briefly, the major results obtained here through analysis of observation data can be formulated as follows.

- It was found that the observation data are best approximated with a second-order bi-variate stochastic difference equation; its coefficients and the white noise covariance matrix were estimated along with respective rms errors; the equation describes a damped oscillatory system with the eigenfrequency $\sim 0.24 \text{ year}^{-1}$ and the damping factors 0.05 and 0.5.

- Statistics of SST as a component of the above bi-variate model, including properties of El Niño and La Niña phenomena were estimated through the Monte Carlo approach.

- The time-domain model was used to give a comprehensive description of the ENSO bi-variate process in the frequency domain; the description includes not just the SOI and SST spectra (their behavior had been generally known before) but also the coherence function, coherent spectra, and the frequency response function.

The time-domain model describes the “deterministic” relations between the two components of the ENSO system expressed in the form of a linear dependence between them, plus a substantial contribution from the external stochastic forcing. The second order of the equation means that statistically significant links between SOI and SST exist at time lags of up to two years. The connections do not play a dominant role in the ENSO system predictability properties which depend mostly upon the unpredictable stochastic forcing.

The existence of the eigenfrequency at 0.24 year^{-1} in the ENSO system can be regarded as an experimental confirmation of the theoretical model proposed by Jin in [10,11]. The linear feedback between the ENSO components along with the input white noise can excite the damped periodic oscillations in the system. As in the Jin’s model, oscillations can be generated by the deterministic part of the linear system and by the stochastic forcing. According to our results, the latter play a more important role.

Variations of SOI and SST were simulated over a long time interval (50,000 years) as components of a bi-variate random process and the simulated data were used to determine statistical properties of ENSO variability including those that cannot be estimated reliably on the basis of the observed time series. It was shown that

- extreme temperatures during El Niño (La Niña) events can be as high as 28°C (as low as 23°C),

- the average interval between successive El Niño’s or La Niña’s amounts to 7 years, while the maximum interval between successive El Niño (La Niña) events can be as long as almost 50 years;

- though individual El Niño’s or La Niña’s can last up to 8 or 7 years, 95% of events last 1 or 2 years (67% and 28% cases, respectively).

The behavior of the ENSO time series in the frequency

domain reveals a number of properties uncommon for atmospheric, oceanic, or land processes at climatic time scales:

- the spectra of both SOI and SST do not decrease with growing frequency;

- the spectra have a single broad maximum centered at $f \approx 0.22 \text{ year}^{-1}$; this maximum reflects the existence of the damped periodic oscillation obtained from the solution of the system of stochastic linear difference equations that describe the observed bi-variate ENSO time series;

- the linear interdependence between the components of the ENSO time series is shown to be the strongest at intermediate and high frequencies;

- the coherence between SOI and SST is very high, especially in the band from 0.18 year^{-1} to 0.33 year^{-1} where it reaches 0.93; this means that the interaction between SOI and SST is responsible for up to 85% of the variance of respective processes in that frequency band; the coherence does not go below 0.8 within the entire frequency band from 0.12 year^{-1} to 0.50 year^{-1} ;

- the part of the spectrum of SST variations strictly coherent to variations of SOI is close to the full spectrum of SST, especially at frequencies where the coherence is particularly high; a similar statement holds for the SOI spectrum;

- the respective gain factors show that the mutual transfer of energy between SOI and SST is stronger at frequencies above approximately 0.15 year^{-1} rather than at lower frequencies;

- the phase factor magnitude is close to π , which means that the presence of the damped periodic oscillation in ENSO cannot substantially improve its statistical predictability.

All these conclusions hold for Niño regions SST12, SST3, and SST4, with some variations in the numerical values of the eigenfrequency, maximum coherence, and some other statistical characteristics.

Years ago, it had been shown both theoretically [21] and experimentally [22-24] that the spectra of climatic processes generally decrease as frequency increases and do not contain significant peaks. The ENSO system behaves in a different manner. The theoretical dynamic and stochastic models developed, respectively, by Jin [10,11] and Kleeman [12], lead to the ENSO system’s spectra of the bell-shape type. The coherence function estimate obtained here behaves in the manner predicted by Kleeman [12]. Our experimental results provide some verification for those theoretical models as well as a detailed description of other statistical properties of the ENSO system at climatic time scales in both time and frequency domains.

REFERENCES

- [1] M. Latif, S. Keenlyside. El Niño/Southern Oscillation response to global warming, PNAS, Vol. 106, No.49, 20578-20583, 2009.
- [2] H.A. Dijkstra. Non-Linear Physical Oceanography, Kluwer

- Academic Publishers, Dordrecht, 2010.
- [3] SOI data, Online available from <http://www.bom.gov.au/climate/current/soihtml.shtml>
 - [4] HadCRUT3, Online available from <http://www.cru.uea.ac.uk/cru/data/temperature/>
 - [5] G. E. P. Box, G. W. Jenkins, G. C. Reinsel. Time Series Analysis, Forecasting and Control, Third Edition. Prentice-Hall, New Jersey, 1994.
 - [6] V. Privalsky. Parametric model; order selection, in Probability and Mathematical Statistics Encyclopedia. Yu. V. Prokhorov, ed. The Great Russian Encyclopedia Publishers, Moscow, Russia, 442-443, 1999 (in Russian).
 - [7] J. S. Bendat, A. G. Piersol. Random Data. Analysis and Measurement Procedures, Fourth Edition. Wiley and Sons, NJ, 2010.
 - [8] V. Privalsky, I. Protsenko, G. Fogel. The sampling variability of autoregressive spectral estimates for two-variate hydrometeorological processes, Proc. 1st World Congress of the Bernoulli Society on Mathematical Stat. Theory and Applications, Vol. 2. VNU Science Press, Utrecht, 651-654, 1987.
 - [9] W. Kelley, A. Peterson. Difference Equations. An Introduction with Applications, Academic Press, MA, 2000.
 - [10] F.-F. Jin. An equatorial ocean recharge paradigm. Part I: Conceptual model, J. Atmos. Sci., Vol. 54, No.7, 811-829, 1997.
 - [11] F.-F. Jin. An equatorial ocean recharge paradigm. Part II: A stripped-down coupled model, J. Atmos. Sci., Vol. 54, No.7, 830-847, 1997.
 - [12] R. Kleeman. Spectral analysis of multidimensional stochastic geophysical models with an application to decadal ENSO variability, J. Atmos. Sci., Vol. 68, No.1, 13-25, 2011.
 - [13] A. F. Z, Levin, F.-F. Jin. Noise-induced instability in the ENSO recharge oscillator, J. Atmos. Sci., Vol. 67, No.2, 529-542, 2010.
 - [14] H. A. Dijkstra, G. Burgers. Fluid dynamics of El Niño variability, Annu. Review Fluid Mech., Vol. 34, 531-538, 2002.
 - [15] A. M. Yaglom. An Introduction to the Theory of Stationary Random Functions, New York: Dover Publications, 1962.
 - [16] B. Blanke, J. D. Neelin, D. Gutzler. Estimating the effects of stochastic wind stress forcing on ENSO irregularity, J. Climate, Vol. 10, No.7, 1473-1486, 1997.
 - [17] J.-S. Von Storch, B. Wang, Z. Wang. The ENSO-spectrum: a result of deterministic chaos or stochastic forcing? Dynamics of Atmospheric and Oceanic Circulation and Climate, China Meteorological Press, Beijing, 579-600, 2001.
 - [18] J. Bjerknes. Atmospheric teleconnections from the Equatorial Pacific, Monthly Weather Review, Vol. 97, No.3, 163-172, 1969.
 - [19] S. G. Philander. El Niño, La Niña, and the Southern Oscillation, Academic Press, Inc., CA, 1989.
 - [20] V. Privalsky, D. Jensen. Assessment of the influence of ENSO on annual global air temperatures, Dynamics of Atmospheres and Oceans, Vol. 22, No.3, 161-178, 1995.
 - [21] K. Hasselmann. Stochastic climate models, Part I. Theory. Tellus, Vol. 28, No.6, 473-485, 1976.
 - [22] V. Privalsky. Estimating spectral densities of large-scale processes, Izv. Acad. Sci. USSR, Atmospheric and Oceanic Physics, Vol. 12, No.9, 598-600, 1976. (Published by American Geophysical Union and American Meteorological Society.)
 - [23] V. Privalsky. Statistical predictability of large-scale atmospheric processes. Izv. Acad. Sci. USSR, Atmospheric and Oceanic Physics, Vol. 13, No.4, 261-265, 1977. (Published by American Geophysical Union and American Meteorological Society.)
 - [24] C. Frankignoul, K. Hasselmann. Stochastic climate models, Part II. Application to sea-surface temperature anomalies and thermocline variability, Tellus, Vol. 29, No.4, 289-305, 1977.