NOTES AND CORRESPONDENCE

Spectral Estimation from Time Series Models with Relevance to the Southern Oscillation*

PAO-SHIN CHU

Department of Meteorology, University of Hawaii, Honolulu, Hawaii

RICHARD W. KATZ

*Environmental and Societal Impacts Group, National Center for Atmospheric Research,** Boulder, Colorado*

14 April 1988 and 23 August 1988

ABSTRACT

The theoretical spectra of certain parametric time series models with relevance to the Southern Oscillation (SO) are determined and compared with those based on a frequency-domain approach. Consistent spectral estimates are found for the two models selected in our earlier studies of the SO. All these results yield larger power at low frequencies and a dominant peak around 3–4 yr. Some reasons are offered for the slightly different behavior of the spectra as derived from the time-domain and frequency-domain approaches.

For the sake of comparison, the spectra of other simpler time series models are also calculated. While larger power is found at low frequencies, no spectral peak exists in these simpler models. Some implications of the quasi-periodic behavior found in the more complex models (i.e., an intermediate peak in the spectrum) are discussed in the context of the persistence and forecasting of the SO.

1. Introduction

The Southern Oscillation (SO) is the most dominant mode in short-term climate variation over the globe, particularly in the tropics and midlatitudes, accounting for a significant portion of the variance in the global climate system. The state of the SO is generally described by an index (SOI), which is the normalized sea level pressure difference between Tahiti and Darwin. Based on the Fourier transform of the autocovariance or autocorrelation function (hereafter referred to as the frequency-domain approach), many studies have found a broad range of large power in the spectrum of the SOI from 2–10 yr, with a peak around 3–5 yr (e.g., Chen 1982). Similar spectral properties are also revealed in the Darwin pressure series alone (Trenberth 1976; Trenberth and Shea 1987). This quasi-periodic behavior is also found in the sea surface temperature over the tropical eastern Pacific (e.g., Rasmusson and Carpenter 1982) and is sometimes referred to as the El Niño–Southern Oscillation (ENSO) "signal."

Recently, Chu and Katz (1985; hereafter referred to as CK1) showed that fluctuations in the SO can be adequately represented by parametric time series models. For example, a third-order autoregressive [AR(3)] model is representative of the seasonal SO fluctuations. In brief, on the seasonal time scale, the current state of the SO can be expressed in terms of its immediate past three states and a current white noise input. Now the question arises as to what sort of theoretical spectrum a parametric time series model can produce. In particular, what would be the preferred spectral characteristics as determined from the models such as selected by CK1? Will their behavior be similar or different from those calculated from the frequency-domain approach? As noted by Jolliffe (1983) and Koscielny and Duchon (1984) in modeling other climate time series, a low-order autoregressive (AR) process can produce time series with quasi-periodicities (i.e., intermediate peaks in the spectrum), thus suggesting that the models selected in CK1 might well exhibit cyclic behavior consistent with that observed for the SOI. However, in practice, this consistency is not guaranteed, since the shape of the spectrum depends on the specific parameter values of the model.

Given this complexity, the purpose of this note is to assess whether the well-known quasi-periodic nature of the SO also could be reproduced from a time-domain approach using models such as those found in CK1.

* Contribution No. 87-19 of the Department of Meteorology, University of Hawaii.

** The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: Prof. Pao-Shin Chu, Dept. of Meteorology, University of Hawaii at Manoa, 2525 Correa Road, Honolulu, HI 96822.

This approach provides another criterion by which to assess the adequacy of the models in representing the temporal behavior of the SO, recalling that these models were originally selected on the basis of a different time-domain criterion, i.e., the Bayesian information criterion. The theoretical spectra of parametric time series models are formulated in section 2. Section 3 provides an application to the SOI, and section 4 contains the summary and discussion.

2. Theoretical spectrum

a. General form

A stochastic process generated from a linear filter $\psi(B)$ with input a_t can be represented as the infinite-order moving average [i.e., MA(∞)] process:

$$X_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots = \sum_{j=0}^{\infty} \psi_j a_{t-j} = \psi(B) a_t \quad (1)$$

with $\psi_0 = 1$. The $\psi(B)$ is a polynomial and B is the backward shift operator, defined as $B^d a_t = a_{t-d}$. The input a_t is a white noise process, implying a lack of memory of prior states for this input process. The spectrum of this linear process is (Box and Jenkins 1976, p. 50):

$$p(f) = 2\sigma_a^2 |\psi(e^{-i2\pi f})|^2 \quad 0 \leq f \leq \frac{1}{2}, \quad (2)$$

where σ_a^2 is the variance of the white noise process. It is sometimes convenient to consider the spectral density function

$$s(f) = \frac{p(f)}{\sigma^2}, \quad (3)$$

where σ^2 is the variance of the X_t process. This form of normalization makes the area under the curve from $f = 0$ to $f = 0.5$ (i.e., the Nyquist frequency) equal to unity.

b. Autoregressive models

For a pth-order autoregressive [i.e., AR(p)] process,

$$\psi(B) = \phi^{-1}(B) \quad (4)$$

where $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 \cdots - \phi_p B^p$. Using (2) and (4), the generalized form of the theoretical spectrum of an AR(p) process can be expressed as

$$p(f) = \frac{2\sigma_a^2}{|1 - \phi_1 e^{-i2\pi f} - \phi_2 e^{-i4\pi f} - \cdots - \phi_p e^{-i2\pi p f}|^2}. \quad (5)$$

In meteorology and physical oceanography, an AR(1) (Markov or red-noise) process or an AR(2) process often can provide a good approximation to the data (e.g., Leith 1973; Shukla and Mo 1983; Moene 1986). Expressions for the power spectrum of an AR(1) or AR(2) process can be easily derived from (5) and are given by Jenkins and Watts (1968) or by Box and Jenkins (1976, p. 58, p. 62). In modeling the seasonal SOI, an AR(3) process was found to be optimal (CK1). For this particular process, (5) reduces to

$$p(f) = \frac{2\sigma_a^2}{1 + \phi_1^2 + \phi_2^2 + \phi_3^2 - 2(\phi_1 - \phi_1\phi_2 - \phi_2\phi_3)\cos 2\pi f - 2(\phi_2 - \phi_1\phi_3)\cos 4\pi f - 2\phi_3 \cos 6\pi f}. \quad (6)$$

This equation thus provides an explicit representation of the spectrum in terms of the individual parameters of the AR(3) process.

c. Autoregressive-moving average models

From (2), the generalized form of the theoretical spectrum of an autoregressive-moving average [ARMA(p, q)] process is

$$p(f) = 2\sigma_a^2 \frac{|1 - \theta_1 e^{-i2\pi f} - \cdots - \theta_q e^{-i2\pi q f}|^2}{|1 - \phi_1 e^{-i2\pi f} - \cdots - \phi_p e^{-i2\pi p f}|^2}. \quad (7)$$

Here ϕ's and θ's are the autoregressive (AR) and moving-average (MA) parameters, respectively; and p and q denote the order of the AR and MA terms, respectively. The simplest form of a mixed model is the ARMA(1, 1) process, but this model does not adequately represent the monthly SOI. Instead, an ARMA(1, 7; 1) process was found by CK1 to represent the monthly data more accurately. This process is defined as

$$X_t = \phi_1 X_{t-1} + \phi_7 X_{t-7} + a_t - \theta_1 a_{t-1}. \quad (8)$$

From (7), the spectrum of this process is given by

$$p(f) = 2\sigma_a^2 \frac{1 + \theta_1^2 - 2\theta_1 \cos 2\pi f}{1 + \phi_1^2 + \phi_7^2 - 2\phi_1 \cos 2\pi f - 2\phi_7 \cos 14\pi f + 2\phi_1\phi_7 \cos 12\pi f}. \quad (9)$$

3. Spectrum of SOI

a. Frequency-domain approach

As a first step of the analysis, a smoothed estimate of the spectral density function from a frequency-domain approach is obtained using the following formula:

$$\hat{s}(f) = 2\,[1 + 2 \sum_{k=1}^{m-1} r(k)W(k)\cos(2\pi f k\Delta t)]\,\Delta t, \tag{10}$$

where m, $r(k)$, $W(k)$, and Δt are the maximum lag (i.e., truncation point), sample autocorrelation function, lag window, and time interval between observations (i.e., 1 month or 1 season), respectively. We employ the Parzen lag window (e.g., Jenkins and Watts 1968, p. 244) i.e.,

$$W(k) = \begin{cases} 1 - 6\left(\dfrac{k}{m}\right)^2 + 6\left(\dfrac{|k|}{m}\right)^3, & |k| \leq \dfrac{m}{2} \\ 2\left(1 - \dfrac{|k|}{m}\right)^3, & \dfrac{m}{2} < |k| \leq m \\ 0, & |k| > m. \end{cases} \tag{11}$$

This window is used commonly to smooth the sample spectrum of atmospheric time series.

The intelligent choice of the truncation point, m, is a rather difficult task in spectral analysis and certainly lacks a clear-cut criterion. Empirical rules indicate that an appropriate value of m should be close to $n/6$ when $100 < n < 200$, where n represents the total number of observations used in the spectral analysis (Chatfield 1975). As n increases above 200, an appropriate value of m should be correspondingly smaller than $n/6$. In particular, for $1000 < n < 2000$, a value of m should be less than $n/10$. Keeping these rules in mind, we choose m as 30 for the seasonal case ($n = 194$) and 60 for the monthly case ($n = 584$).

The sea level pressure observations at Tahiti and Darwin were obtained from the *Monthly Climatic Data for the World and the World Weather Records*. As in CK1, the pressure series at each station is normalized first and the Tahiti-minus-Darwin difference forms the monthly or seasonal SOI. As shown in Fig. 1, the preferred period of the seasonal SOI based on the frequency-domain approach is at time scales of more than 8 seasons; a large peak is present near 20 seasons (5 yr). Analyzing Tahiti and Darwin pressure series individually, Trenberth (1976) noted that a spectral peak near 20 seasons is present for both stations. His peak is statistically significant at about the 5% level relative to red noise [i.e., an AR(1) process]. More recently, Trenberth and Shea (1987) found a spectral peak centered around 60 months (or 20 seasons) for a detrended Darwin pressure time series. This quasi-periodicity is considered to be one of the major features of the ENSO phenomenon.

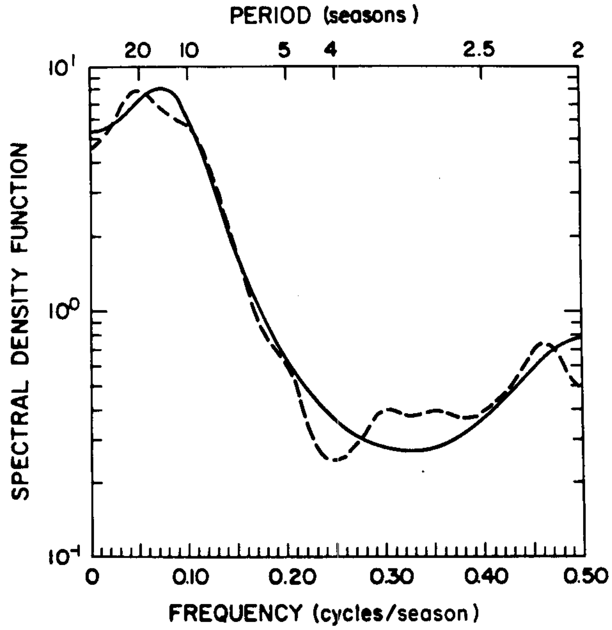

FIG. 1. Spectral density function of the seasonal SOI from the frequency-domain approach (broken curve) and time-domain approach (solid curve). Analysis period is from spring 1935 to summer 1983. Parzen lag window is used with a maximum lag of 30 seasons (broken curve). Note that the scale on the ordinate is logarithmic, whereas the scale on the abscissa is linear.

b. Time-domain approach

As mentioned earlier, ARMA processes may possess theoretical spectra with intermediate peaks under certain conditions. For example, the spectrum of an AR(2) process contains a peak or trough if

$$|\phi_1(1 - \phi_2)| < |4\phi_2| \tag{12}$$

(Jenkins and Watts 1968, p. 229). The spectrum contains a peak if, in addition to (12), $\phi_2 < 0$.

1) SEASONAL SOI

From CK1, the three parameter estimates and the white noise variance for an AR(3) model of the seasonal SOI data are $\hat{\phi}_1 = 0.6885$, $\hat{\phi}_2 = 0.2460$, $\hat{\phi}_3 = -0.3497$, and $\hat{\sigma}_a^2 = 1.505$. Given these values, it is possible to estimate the theoretical spectrum of an AR(3) process by means of (6). Also shown in Fig. 1 is the spectrum of this AR(3) process. A large portion of power is found in the lower frequency end, with a peak somewhere near 4 yr (14.3 seasons). Moreover, both spectra are marked by an increase in energy in the high-frequency tail. Thus, the spectral estimates from the time-domain and frequency-domain approaches are reasonably consistent.

Nevertheless, one notices that the spectral peak as determined from the time-domain approach does not exactly match the peak from the frequency-domain approach. One reason for this difference is that the selected time series model in CK1 can only be regarded as representative of the general behavior of the SO fluctuations, but not representative of its detail, since to achieve parsimony and maximize predictability (Chu and Katz 1987; hereafter referred to as CK2) it involves only three parameters. For instance, the theoretical autocorrelation function of the selected model behaves like the sample autocorrelation function of the SOI, but only up to about lag-9 (see Fig. 6 in CK1). In deriving the spectral estimates from the frequency-domain approach, lags up to 30 seasons are used [see (10)–(11)]. Although the autocorrelations after lag-9 are small, their cumulative contributions to the estimated spectrum (10) will make a difference in such a way that the energy distribution for the frequency- and time-domain approaches will be slightly different. In particular, the spectral density function estimated by the time-domain approach will necessarily be smoother in appearance than the same function estimated by the frequency-domain approach. Another reason for the difference between these two approaches is simply the uncertainty in point estimates of the sample spectrum, because the sample size is finite.

Now the question becomes whether a lower-order model than AR(3) also produces a quasi-periodicity such as the one shown in Fig. 1. The autocorrelation function of an AR(1) process always decays exponentially to zero, implying no intermediate spectral peak. For the case of an AR(2) process, different parameter estimates (i.e., different values of $\hat{\phi}_1$ and $\hat{\phi}_2$) give rise to various forms of spectra, either with a single intermediate peak or trough or with no intermediate peak or trough (Jenkins and Watts 1968, p. 229). For the AR(2) model of the seasonal SOI, since $\hat{\phi}_1 = 0.670$ and $\hat{\phi}_2 = 0.021$, the inequality (12) is not satisfied. Hence, the spectral density function has no intermediate peak. Keeping the estimate of ϕ_1 fixed, the estimate of ϕ_2 would need to be less than -0.20 or more than three standard errors less than its actual value to produce an intermediate peak in the spectrum.

2) MONTHLY SOI

Given the parameter estimates and the white noise variance for the ARMA(1, 7; 1) model of the monthly SOI as given in CK1, we can calculate the spectrum of this model directly by means of (9). The theoretical spectral density function is shown in Fig. 2, with a preferred peak at approximately 40 months. This result is consistent with the estimated spectral density function also shown in Fig. 2, in that a spectral peak occurs near 50 months. The difference between these two spectra, again, is in part due to the fact that the selected ARMA(1, 7; 1) model only mimics the general behavior of the monthly SO fluctuations. The uncertainty in point estimates from the frequency-domain approach also contributes to this difference. In an attempt to see whether a model simpler than the ARMA(1, 7; 1) can produce quasi-periodicities that the SOI has been claimed to possess, the spectral density function of the simpler ARMA(1, 1) model of the monthly SOI has been computed (not shown). While large power is found at low frequencies for this ARMA(1, 1) model, it is clear that there is no intermediate peak at around 40 or 50 months in its spectral density function.

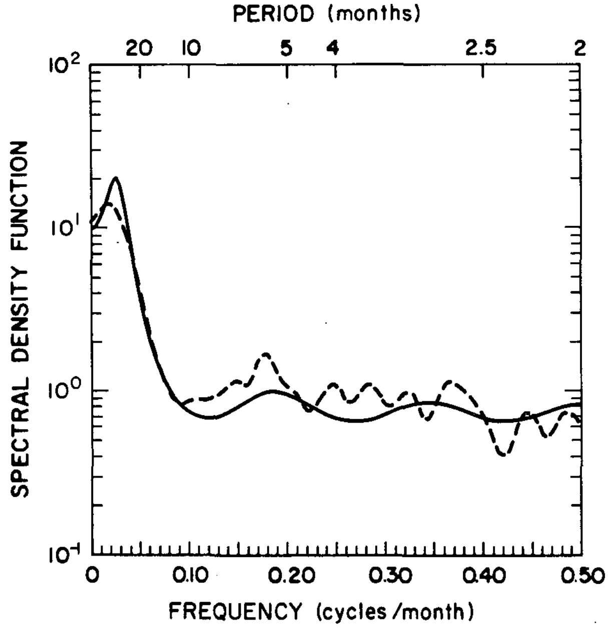

FIG. 2. Spectral density function of the monthly SOI from the frequency-domain approach (broken curve) and time-domain approach (solid curve). Analysis period is from January 1935 to August 1983. Parzen lag window is used with a maximum lag of 60 months (broken curve).

4. Summary and discussion

The theoretical spectra of time series models which are representative of the monthly and seasonal SO fluctuations have been determined. Results indicate that the two models selected by CK1 and CK2 produce a dominant spectral peak in the range between 3 and 4 yr and small power at high frequencies, features consistent with those obtained from the frequency-domain approach in this paper and previously by other researchers. For the seasonal index, the spectra of red-noise and AR(2) models are marked by large power at low frequencies but no intermediate peak. This same feature is also true for the monthly index when an ARMA(1, 1) model is assumed.

The two stochastic models selected in CK1 and CK2 have some important implications. For instance, since the AR(3) model involves three preceding seasonal

SOI values and a random noise input, it implies that only short-term memory (in the sense of conditional independence) is sufficient to describe the current behavior of the SO fluctuations. As such, other complex physical mechanisms such as extratropical atmospheric forcing may become of secondary importance once the ENSO has started to swing. This interpretation also appears plausible when applied to an independent forecasting experiment as shown in CK1 and CK2, in which the forecast skill is mainly derived from relatively short-term "persistence" in the SOI. This persistence, however, is determined by its past three seasonal values and the associated parameters which, loosely speaking, can be regarded as specifying the decay rates of anomalies. It should be noted that our definition of persistence is somewhat different from the more common usage in which only the most recent value is employed.

But what causes the persistence in the SO? The atmosphere alone does not intrinsically have such a relatively long time scale. The tropical Pacific Ocean, though, may exhibit persistence on such a time scale, but its anomalous heat content is intimately related to the atmospheric forcings (e.g., Wyrtki 1975). Thus, the persistence observed in the SO is possibly a manifestation of positive feedback between ocean and atmosphere, since the SO is closely connected to sea surface temperature anomalies in the tropical Pacific.

This feedback may involve surface westerly wind bursts in the equatorial western or central Pacific, as observed in the early phase of most ENSO events. Through dynamic processes, these bursts result in warm upper ocean waters located downstream from the wind forcing region (e.g., McCreary 1976; Luther et al. 1983). Wright (1985) also stressed the importance of positive feedback in the persistence of the SO, but his interpretation is based on cloudiness and sea surface temperature relationships in the eastern tropical Pacific.

The existence of a well-defined spectral peak as found by many researchers (see Figs. 1 and 2) suggests that the ENSO phenomenon is a recurrent one, with the interval between events ranging from 2 to 10 yr. Initially, this kind of quasi-periodic feature offered some hope for prediction until the next event. Nevertheless, no successful stochastic climate prediction has been made beyond the limits of the persistence of the SO (Wright 1985). In fact, based on the SOI time series itself, it is demonstrated in CK2 that the theoretical predictability of the SO almost drops to zero beyond 10 months or 3 seasons ahead. Although the spectral peak as derived from time series models is similar to that obtained from a frequency-domain approach, caution is warranted in the interpretation of the broad peak in the context of SO prediction.

Acknowledgments. Thanks are due to James Frederick and Wen Wang for assisting with computations and to Kevin Keen for comments on a draft version of this manuscript. Discussions with Norman Hui and Roger Lukas contributed to this work. Both of the reviewers and the editor helped to improve the final version. This research is partially supported by the National Oceanic and Atmospheric Administration under Contracts NA-80-RAH-00002 and NA-85-ABH-00032, and by the Water Resources Research Center of the University of Hawaii.

REFERENCES

Box, G. E. P., and G. M. Jenkins, 1976: Time Series Analysis: Forecasting and Control (rev.). Holden–Day, 575 pp.

Chatfield, C., 1975: The Analysis of Time Series: Theory and Practice. Chapman and Hall, 263 pp.

Chen, W. Y., 1982: Assessment of Southern Oscillation sea-level pressure indices. *Mon. Wea. Rev.,* **110,** 800–807.

Chu, P.-S., and R. W. Katz, 1985: Modeling and forecasting the Southern Oscillation: A time-domain approach. *Mon. Wea. Rev.,* **113,** 1876–1888.

——, and ——, 1987: Measures of predictability with applications to the Southern Oscillation. *Mon. Wea. Rev.,* **115,** 1542–1549.

Jenkins, G. M., and D. G. Watts, 1968: Spectral Analysis and Its Applications. Holden–Day, 523 pp.

Jolliffe, I. T., 1983: Quasi-periodic meteorological series and second-order autoregressive processes. *J. Climatol.,* **3,** 413–417.

Koscielny, A. J., and C. E. Duchon, 1984: Autoregressive modelling of the tropical stratosphere quasi-biennial oscillation. *J. Climatol.,* **4,** 347–363.

Leith, C. E., 1973: The standard error of time-average estimates of climatic means. *J. Appl. Meteor.,* **12,** 1066–1069.

Luther, D. S., D. E. Harrison and R. A. Knox, 1983: Zonal winds in the central equatorial Pacific and El Niño. *Science,* **222,** 327–330.

McCreary, J. P., 1976: Eastern tropical ocean response to changing wind systems with application to El Niño. *J. Phys. Oceanogr.,* **6,** 634–645.

Moene, A., 1986: Associations between North Atlantic sea surface temperature anomalies, latitude of the polar front zone, and precipitation over Northwest Europe. *Mon. Wea. Rev.,* **114,** 636–643.

Rasmusson, E. M., and T. H. Carpenter, 1982: Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Mon. Wea. Rev.,* **110,** 354–384.

Shukla, J., and K. C. Mo, 1983: Seasonal and geographical variation of blocking. *Mon. Wea. Rev.,* **111,** 388–402.

Trenberth, K. E., 1976: Spatial and temporal variations of the Southern Oscillation. *Quart. J. Roy. Meteor. Soc.,* **102,** 639–653.

——, and D. Shea, 1987: On the evolution of the Southern Oscillation. *Mon. Wea. Rev.,* **115,** 3078–3096.

Wright, P. B., 1985: The Southern Oscillation: An ocean-atmospheric feedback system? *Bull. Amer. Meteor. Soc.,* **66,** 398–412.

Wyrtki, K., 1975: El Niño–The dynamic response of the equatorial Pacific Ocean to atmospheric forcing. *J. Phys. Oceanogr.,* **5,** 572–584.