MAUSAM, **53**, 2 (April 2002), 165-176

551.58

Prediction of southern oscillation index using spectral components

R. P. KANE

Instituto Nacional de Pesquisas Espacias – INPE

C.P. 515, 12201-970 –São José dos Campos, SP, Brazil

(*Received 21 March 2001, Modified 7 November 2001*)

सार - जिस स्थान पर दोलन की आवृतिताओं का पता एम.ई.एम. (अधिकतम एंट्रापी प्रणाली) द्वारा चला है वहाँ पर एम. आर. ए.(बहुसमाश्रयण विश्लेषण) में इनका उपयोग इनके आयामों और अवस्थाओं का आकलन करने के लिए एम.ई.एम - एम.आर.ए. की सरल पद्धति द्वारा एस.ओ.आई (दक्षिणी दोलन सूचकांक, ताहिती माइनस डार्विन समुद्र सतह वायुमंडलीय दाब अंतर) की समय श्रृंखलाओं का स्पैक्ट्रल विश्लेषण करने के लिए किया गया है। इनमें से सबसे अधिक सुस्पष्ट तीन अथवा चार समय श्रृंखलाओं का उपयोग पुनः संरचना एवं पूर्वानुमान के लिए किया गया है। पराश्रयी आँकड़ों के रूप में 1935-80 की अवधि के आँकड़ों का उपयोग करने पर यह पता चला है कि एस.ओ.आई. के पुनः संरचित मान प्रेक्षित मानों के अनुरूप थे तथा इस समय अधिकांश एल नीनो (एस.ओ.आई. न्यूनतम) तथा ला नीनो (एस.ओ.आई. अधिकतम) सही स्थान पर पाए गए हैं। किन्तु निरपेक्ष आँकड़ों (1980 से आगे) के साथ अनुरूपता नगण्य पाई गई है। पहले के 1945-80, 1955-80, 1965-80 आंकड़ों की पराश्रयी आँकडों के रूप में गणना नहीं करने पर 1980 से आगे के आंकड़ों की अनुरूपता पुनः नगण्य पाई गई है। जब 1980 से आगे के आँकड़ों को ही पराश्रयी आँकडों के रूप में उपयोग किया गया तो उस समय ये बेहतर ढंग से अनुरूप पाए गए हैं जिससे यह पता चलता है कि समय के साथ स्पैक्ट्रल अभिलक्षणों में पर्याप्त परिवर्तन आया है तथा भविष्य के पूर्वानुमानों के लिए ये आँकड़ें अधिक सही हैं। 1997 के एल नीनो के संबंध में 1985 से आगें के वर्षो के आँकड़ों में ही विवरण दिया गया है। किन्तु 1990 से लेकर आगे के वर्षो के आँकड़ों में 3.5 वर्षों की केवल एक घटना (1991) की केवल एक लहर के बारे में ही विवरण दिया गया है जिससे 1997 और 1994 की घटनाएँ स्पष्ट होती है। 1989-95 में 3 संयुक्त और तीव्र गति से घटी घटनाओं में से डा. गिल के यू.एल.सी.ए. का समूह एस.ओ.आई विश्लेषण के लिए एस.एस.ए. (सिंगुलर स्पैक्ट्रल विश्लेषण) एम.ई.एम. का संयुक्त रूप से उपयोग कर रहा है। 1980 के दशक में ये बेहतर रूप से अनुरूप पाए गए हैं किन्तु 1989-95 की संरचनाओं के बारे में विवरण नहीं दिया गया है। हाल ही के वर्षों में इनके एस.एस.ए. से फिल्टर किए गए एस.ओ.आई. ~3.5 वर्षों के साधारण सिनुसाइड (पूर्वानुमान के लिए प्रयुक्त होने वाले) हैं। इसके द्वारा 1997 की एल नीनो घटना का चरम अवस्था में ही पूर्वानुमान किया गया था जिसके लिए फरवरी 1997 तक के आँकड़ों का उपयोग किया गया था। लेकिन इसके पश्चात मार्च 1997 की अकस्मात घटना घटने तथा जून 1998 में अकस्मात समाप्त होने का पूर्वानुमान नहीं किया जा सका था।

3.5 वर्षों की लहर का ही उपयोग करते हुए वर्ष 2000-2001 के एल नीनों के बनने की संभावना प्रकट की गई थी। तथापि बहुत अधिक दिनों तक बनी रहने वाली क्रियाशील प्रतीत हो रही ला नीना घटना के संबंध में अभी तक (2001 सितंबर तक) एल नीनो जैसी परिस्थितियों का कोई संकेत नहीं मिल पाया है।

ABSTRACT. The time series of SOI (Southern Oscillation Index, Tahiti minus Darwin sea-level atmospheric pressure difference) was spectrally analysed by a simple method MEM-MRA, where periodicities are detected by MEM (Maximum Entropy Method) and used in MRA (Multiple Regression Analysis) to get the estimates of their amplitudes and phases. From these, the three or four most prominent ones were used for reconstruction and prediction. Using data for 1935-80 as dependent data, the reconstructed values of SOI matched well with observed values and most of the El Niños (SOI minima) and La Niñas (SOI maxima) were located correctly. But for the independent data (1980 onwards), the matching was poor. Omitting earlier data, 1945- 80, 1955-80, 1965-80 as dependent data again gave poor matching for 1980 onwards. When data for 1980 onwards only were used as dependent data, the matching was better, indicating that the spectral characteristics have changed considerably with time and recent data were more appropriate for further predictions. The 1997 El Niño was reproduced only in data for 1985 onwards. For 1990 onwards, only a single wave of 3.5 years was appropriate and explained the 1997 and 1994 events but only one (1991) of the 3 complex and quick events that occurred during 1989-95. The UCLA group of Dr. Ghil has been using the SSA (Singular Spectrum Analysis)-MEM combination for

SOI analysis. For the 1980s, they got very good matching, but the 1989-95 structures were not reproduced. For recent years, their SSA-filtered SOI (used for prediction) is a simple sinusoid of ~3.5 years. It predicted the El Niño of 1997 only at its peak and even after using data up to February 1997, the abrupt commencement of the event in March 1997 and its abrupt end in June 1998 could not be predicted.

Using only a 3.5 years wave, an El Niño was expected for 2000-2001. However, a very long-lasting La Niña seems to be operative and there are no indications as yet (September of 2001) of any El Niño like conditions.

Key words – Southern oscillation index, SOI, Singular spectrum analysis, SSA, Maximum entropy method, MEM, El-Niño.

1. Introduction

Spectral components are important characteristics of a time series and can be used for predictions. In earlier days, Fourier analysis was copiously used, where periodicities were fixed, namely the fundamental period *T* (the data length) and its simple fractions *T*/2, *T*/3 *etc*. and the amplitudes and phases of the same were estimated. Soon followed the lag-auto-correlation methods BT (Blackman and Tukey, 1958), where the power in only certain frequencies related to the lag m could be evaluated, namely, $f = r/2\text{m}$, where r = 1, 2, 3 etc. up to m. The lag m is recommended to be ~25% of the data length. Thus, for a series of 100 data points, m=25 and the periodicities $T(= 1/f = 2\text{m}/r)$ that could be studied would be 50.0, 25.0, 16.6, 12.5, 10.0 etc. Smoothing for intermediate frequencies and the confidence levels could be obtained by the procedure outlined in Jenkins and Watts (1968). Burg (1967) developed MEM (Maximum Entropy Method) where it is assumed that the time series can be estimated well by an auto-regressive (AR) process of some order (see review by Ulrych and Bishop, 1975). The order, called LPEF (Length of the Prediction Error Filter) decides the resolution. At small LPEF, nearby frequencies are not resolved. At larger LPEF, resolution is good but spurious peaks may appear that can be confused with physically valid ones. Hence, only small LPEF is recommended, at the cost of resolution. The main advantage of MEM is its high resolution, much higher than that of BT.

Penland *et al*. (1991) noticed that the spectral resolution and statistical significance of the MEM analysis at small LPEF can be improved by subjecting the data to an adaptive filter, which consists of projecting the data onto the leading temporal empirical orthogonal function obtained from SSA (Singular Spectrum Analysis) (Vautard and Ghil, 1989), which is a univariate application of PCA (Principal Component Analysis, *e.g*., Preisendorfer, 1988) in the time domain. Penland *et al*. (1991) illustrated the utility of the SSA-MEM combination by analyaing a synthetic time series and a time series of atmospheric angular momentum. Keppenne and Ghil (1992) (henceforth referred to as KG) used SSA-MEM for prediction of SOI (Southern Oscillation Index), which is intimately related to the oceanic El Niño phenomenon.

Whereas the elegance of the SSA-MEM method is well established, a simpler method initiated by Kane (1977) does not seem to have attracted any attention. There, only the best property of MEM, namely the high spectral resolution, is used to locate possible peaks, and MEM is used no more thereafter. Instead, the prominent-most peaks T_k (about a dozen) are used in MRA (Multiple Regression Analysis, Bevington, 1969), and the amplitudes and phases determined and the significant ones only used for prediction. In the present communication, prediction results for SOI by this simple MEM-MRA method are presented and compared with the results obtained by KG for SOI by the SSA-MEM. In particular, results for the 1997-98 El Niño are compared in detail.

2. Methodology

As mentioned above, MEM is used only for locating the peaks. These peaks T_k (about a dozen) are used in the expression :

$$f(t) = A_0 + \sum_{k=1}^{n} \left[a_k \mathrm{Sin}(2\pi t / T_k) + b_k \mathrm{Cos}(2\pi t / T_k) \right] + E$$

$$= A_0 + \sum_{k=1}^{n} \mathrm{Sin}(2\pi t / T_k + \phi_k) + E \qquad (1)$$

where $f(t)$ is the observed time series and E the error factor. A Multiple Regression Analysis (MRA, Bevington 1969) was then carried out to obtain the best estimates of A_0, (a_k, b_k) and their standard errors, by a least-square fit. From these, the amplitudes r_k and their standard error a (common for all r_k in this methodology, which assumes white noise) can be calculated and any r_k exceeding 2σ would be significant at a 95% (a priori) confidence level. Note that MESA and MRA are two independent

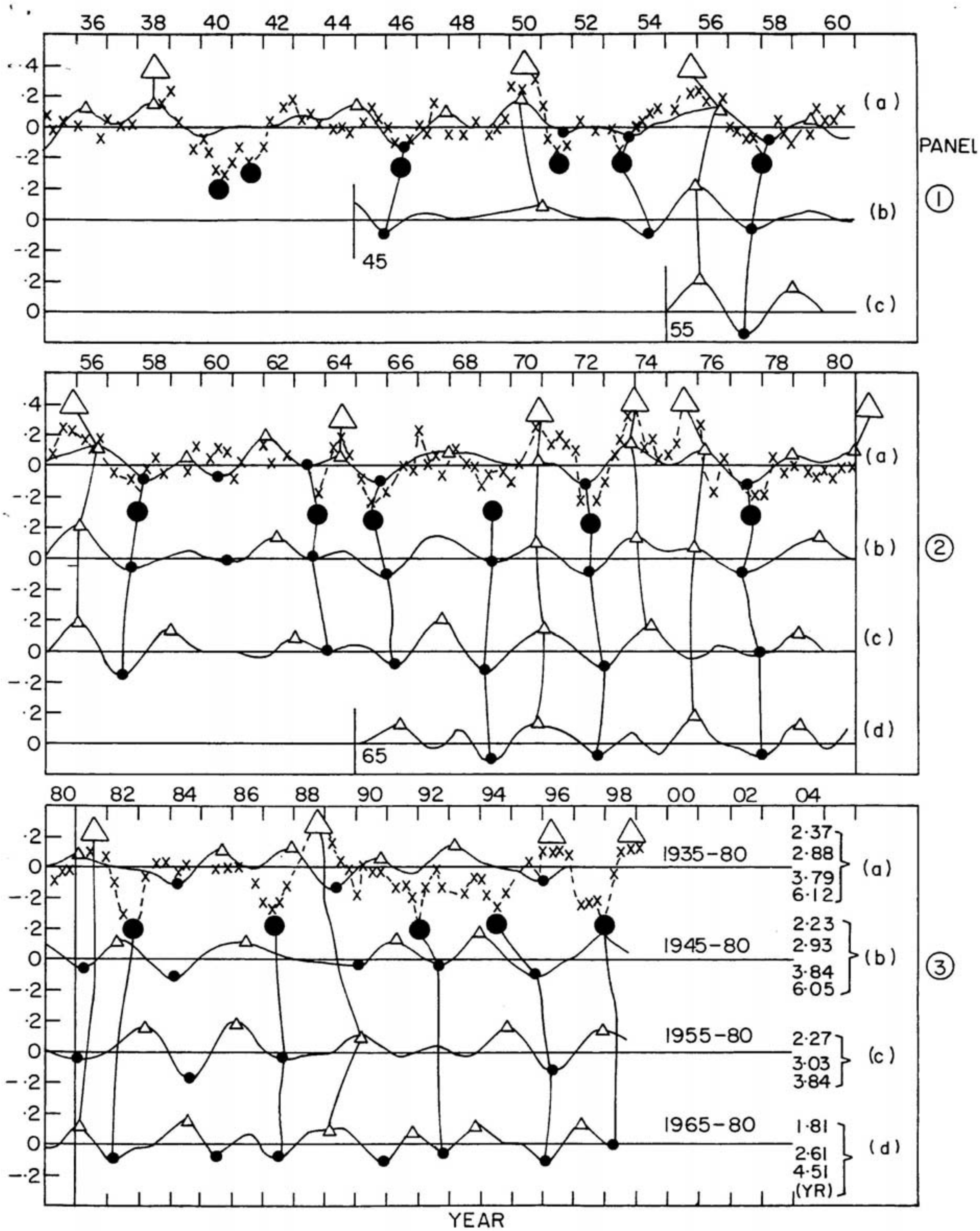
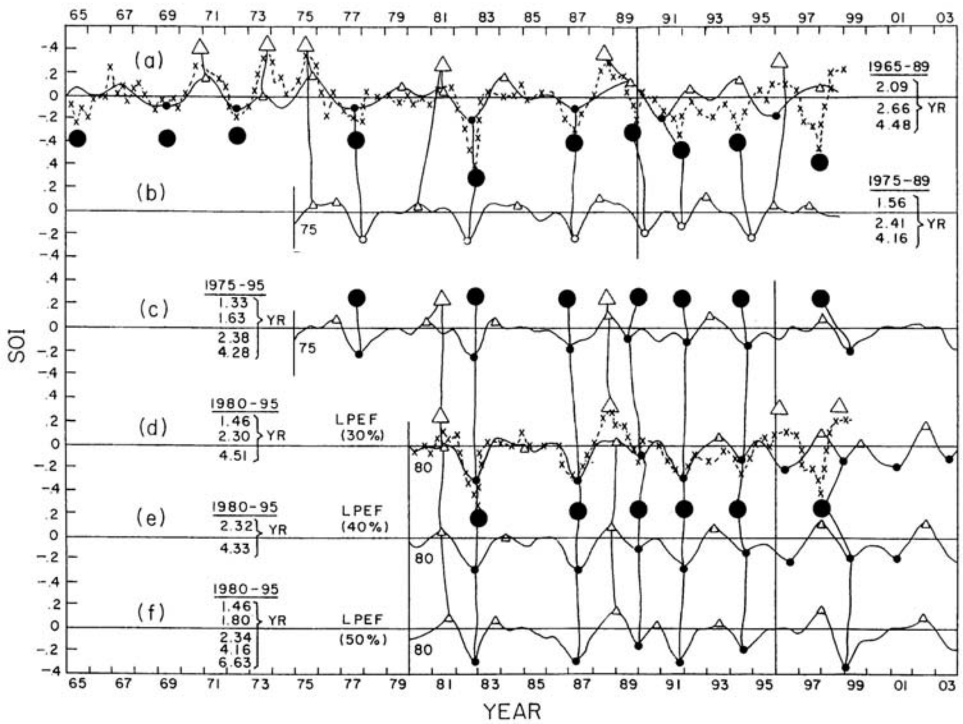


Fig. 1. Plot of 3-monthly values (4 values per year) of SOI (Southern Oscillation Index T-D) for 1935-60 in Panel 1, 1955-80 in Panel 2, and 1980 onwards in Panel 3. In each Panel, the top plot (a) shows the original 3-monthly values (crosses and dashes) and the big triangles above it mark observed SOI maxima (La Niños) and the big full circles below it mark observed SOI minima (El Niños). The superposed full curve in Panels 1 and 2 shows the reconstruction by using 4 periodicities obtained by analysis of dependent data for 1935-80. In Panel 3, it represents prediction. Plots (b), (c), (d) show data analysis for 1945-80, 1955-80, 1965-80, respectively



Figs. 2(a-f). Same as Fig. 1, for dependent data for 1965 onwards, (a) 1965-89, (b) 1975-89, (c) 1975-95, (d, e, f) 1980-95 with LPEF= 30%, 40%, 50% respectively

procedures, in no way related to each other. For prediction, only highly significant peaks were chosen and if these were too many, only the 4 most prominent were selected.

3. Data

Southern Oscillation Index is represented by the sea-level atmospheric pressure difference (T-D) between two locations T (Tahiti, 18° S, 150° W) and D (Darwin, 12° S, 131° E). If their seasonal variation is eliminated, the monthly difference (T-D) shows considerable interannual variability and (T-D) minima generally occur in years of El Niño events, though sometimes some dephasing is seen (Deser and Wallace, 1987). Data for (T-D) were obtained from Parker (1983) and updated from "Monthly Climatic Data for the World". However, these data are also available in the Archives of the web site of Climate Prediction Center, NOAA, for 1950 onwards.

The monthly values of (T-D) are often erratic. Some smoothing is necessary. KG used 5-monthly running means. However, El Niño events last for several months and are seen clearly even in seasonal values. Hence, only 3-monthly values (JFM, AMJ, JAS, OND, and four values per year) were used for spectral analysis. The LPEF used for MEM was usually 30%, but in some cases, LPEF=40% and 50% were also used.

4. Predictions by the simple MEM-MRA combination

Reliable data for (T-D) were available from 1935 onwards. As dependent data, only values for 1935-80 were used. Fig. 1 shows the results for 1935-60 in Panel 1, 1955-80 in Panel 2 and 1980 onwards in Panel 3. In each panel, the top plot (a) shows the original seasonal values as crosses and dashes. The big triangles above these plots indicate observed (T-D) *maxima* and represent La Niña events. The big full circles below these plots indicate observed (T-D) *minima* and represent El Niño events. In Panel 1(a) and 2(a), the superposed solid curve represents the reconstruction of (T-D), by using 4 significant periodicities 2.37, 2.88, 3.79, 6.12 years, obtained in the MEM analysis of the data for 1935-80. In Panel 3(a), the solid curve represents prediction, using the same periodicities. On the solid curve itself, maxima are marked by small triangles and minima by small dots. As can be seen, in Panel 1(a) and 2(a), the small circles occurred almost simultaneously with the big circles (shown by connecting lines) and the small triangles occurred almost simultaneously with the big triangles, indicating that the reconstructed solid plot reproduced the La Niña (triangles) and El Niño (circles) events satisfactorily. However, the real test is in the independent data for 1980 onwards in Panel 3(a). Here, the matching is not at all satisfactory, with small circles having considerable lags or leads with the big circles and small triangles having considerable lags or leads with the big triangles. Thus, the predictions failed.

In Panel 1(b) and Panel 2(b), ten years of earlier data were omitted and the data used for MEM were for 1945-80. The reconstructed (solid plot) was obtained by using periodicities 2.23, 2.93, 3.84, 6.05 years (almost the same as for 1935-80) and the maxima and minima show good match with the big circles and triangles; but in Panel 3(b), the predictions do not show a good match. In plots (c), the data used for MEM were for 1955-80 (periodicities 2.27, 3.03, 3.84 years, almost the same as for 1935-80 and 1945-80) and in plots (d), the data used for MEM were for 1965-80 (periodicities 1.81, 2.61, 4.51 years, different from those for plots (a, b, c). Whereas the reconstructions are satisfactory in the dependent data up to 1980, the independent data after 1980 do not show good match. Thus, all the predictions are failures. However, the failure is less (matching in the data for 1980 onwards is better) for Panel 3(d), indicating that use of earlier data may not be desirable. The fact that the periodicities used in (d) are different from those used in (a, b, c), indicates a change in spectral characteristics after about 1965.

In Fig. 2, only data for 1965 onwards are used for MEM-MRA analysis. In Fig. 2(a), data for 1965-89 are used as dependent data, with periodicities 2.09, 2.66, 4.48 years [similar to Fig. 1(d)]. The reconstructed (solid) plot matches well with the original values (crosses and dashes) up to 1987, with the prominent El Niños of 1965, 1969, 1972, 1977, 1982, 1987 well reproduced. But the events after 1987 are not at all reproduced well. In Fig. 2(b), another 10 years from the earlier data are omitted and only data for 1975-89 are used, with periodicities 1.56, 2.41, 4.16 years, slightly different from those in Fig. 2(a). The match in the independent data is now better, with El Niños of 1989, 1991, and 1994 well reproduced; but the El Niño of 1997 is not reproduced.

In Fig. 2(c), dependent data used are for 1975-95, with periodicities 1.33, 1.63, 2.38, 4.28 years. It is interesting to note that lower periodicities are now needed (less than 2 years) to get a good match, which is good up to 1994, but the Ei Niño of 1997 is still not reproduced.

In Fig. 2(d), data used for MEM are only for 1980-95, with periodicities 1.46, 2.30, 4.51 years. The match up to 1995 is good, but the 1997 El Niño is still not reproduced, indicating that the 1997 event probably may not belong to the recent spectral sequence.

So far, only a LPEF=30% was used. Since data lengths were becoming smaller and smaller, a larger LPEF could have been more appropriate. Fig. 2(e) used data for 1980-95 with LPEF=40% and periodicities 2.32, 4.33 years only. The 1997 El Niño is still not reproduced. In Fig. 2(f), LPEF=50% was used, with many more periodicities, namely, 1.46, 1.80, 2.34, 4.16, 6.63 years. The 1997 El Niño is still not reproduced.

In Fig. 3(a), the 1997 event was included in the dependent data. MEM for 1980-98, with periodicities 2.34 and 4.31 years and LPEF=30% reproduced all earlier events, but the 1997 El Niño was not reproduced (even in the dependent data), indicating that this event probably occurred so to say "out of turn". Instead, an El Niño was predicted for 1999, probably a wrong prediction. In Fig. 3(b), same data were used with LPEF=40% and many more periodicities were used, namely, 1.48, 1.87, 2.35, 4.00, 7.06 years. For the first time, the 1997 El Niño was reproduced, but the additional El Niño for 1999 persisted.

Fig. 3(c) used data for 1985-98, with LPEF=30% and periodicities 2.44, 8.54 years, while Fig. 3(d) used the same data with LPEF=50% and periodicities 1.91, 2.35, 3.67, 7.44 years. In both the cases, the 1997 El Niño was reproduced, but the (false?) 1999 El Niño persisted.

Fig. 3(e) used a very short data set 1990-98, with LPEF=30%. With only one periodicity, namely 3.4 years, the 1994 and 1997 El Niños were reproduced, but the earlier 1990 and 1992 El Niños were mingled into one event of 1991.

The analysis described above seems to indicate that:

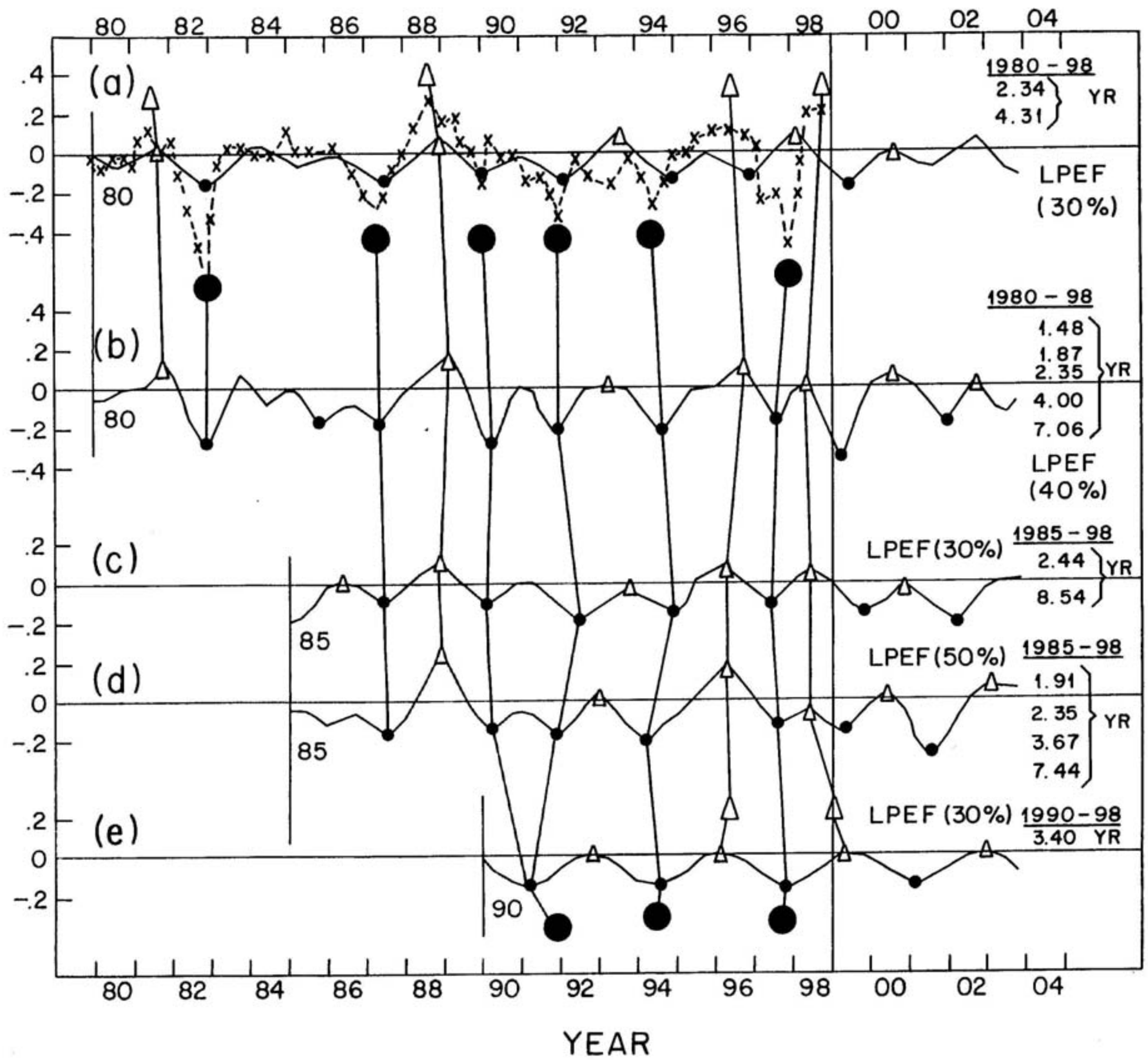
(*i*) For predictions of recent events, data before about 1965 should not be used.

(*ii*) Even after 1965, more recent events are predicted better by using more recent data only.

(*iii*) The 1997 El Niño prediction is somewhat illusive. The event seems to fit only a recent 3.4 years wave.

5. Predictions by KG by the SSA-MEM

KG used deseasoned standardized SOI data (T-D) for 1941 onwards. They used SSA to compute a data-adaptive prefilter by retaining only the leading, statistically significant temporal principal components (T-PCs) of the SOI time series. For 1940-90, the SSA-filtered SOI and the 5-month running means of observed SOI matched very well and most of the well known El Niños and La Niñas in this interval were reproduced. Using data up to February 1992, KG predicted a La Niña



Figs. 3(a-e). Same as Fig. 1, for dependent data for 1980 onwards, (a, b) 1980-98, with LPEF= 30%, 40%, (c, d) 1985-98, with LPEF=30%, 50%, (e) 1990-98, LPEF=30%

event for the 1993-94 winter. As can be seen from our Fig. 3(a), the SOI variations during 1990-95 were somewhat complicated. The group of Dr. Michael Ghil and his colleagues at the Department of Atmospheric Sciences and the Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California, USA have been updating their SOI predictions and reporting the results in the Experimental Long-Lead Forecast Bulletins (henceforth referred to as ELLFB) in quarterly issues. In the March 1996 (Vol. 5, No. 1) issue, Keppenne and Lall (1996) introduced several improvements in the linear prediction schemes used earlier to issue the SSA-MEM predictions of SOI. However, for the 60-month lead SOI forecasts issued at intervals of 24 months between August 1981 and August 1993 and further up to November 1995, they remarked that whereas the 1982-83 and 1986-87 El Niño events could have been forecasted at leads of several years, and the forecast skill of the 1985 and 1988 La Niñas was also impressive, the skill was much lower in the early 1990s, where all their forecasts missed the doubly recurring mild El Niño events. Since then, their group has been giving forecasts not only for SOI but also for the area-averaged SST (sea surface temperature) in the Niño 3 region (5° N-5° S, 150° W-90° W). In the March 1999 issue (Vol. 8, No.1) of ELLFB, Saunders *et al*. (1999) gave the 12-month lead forecasts for Niño 3 SST, which are

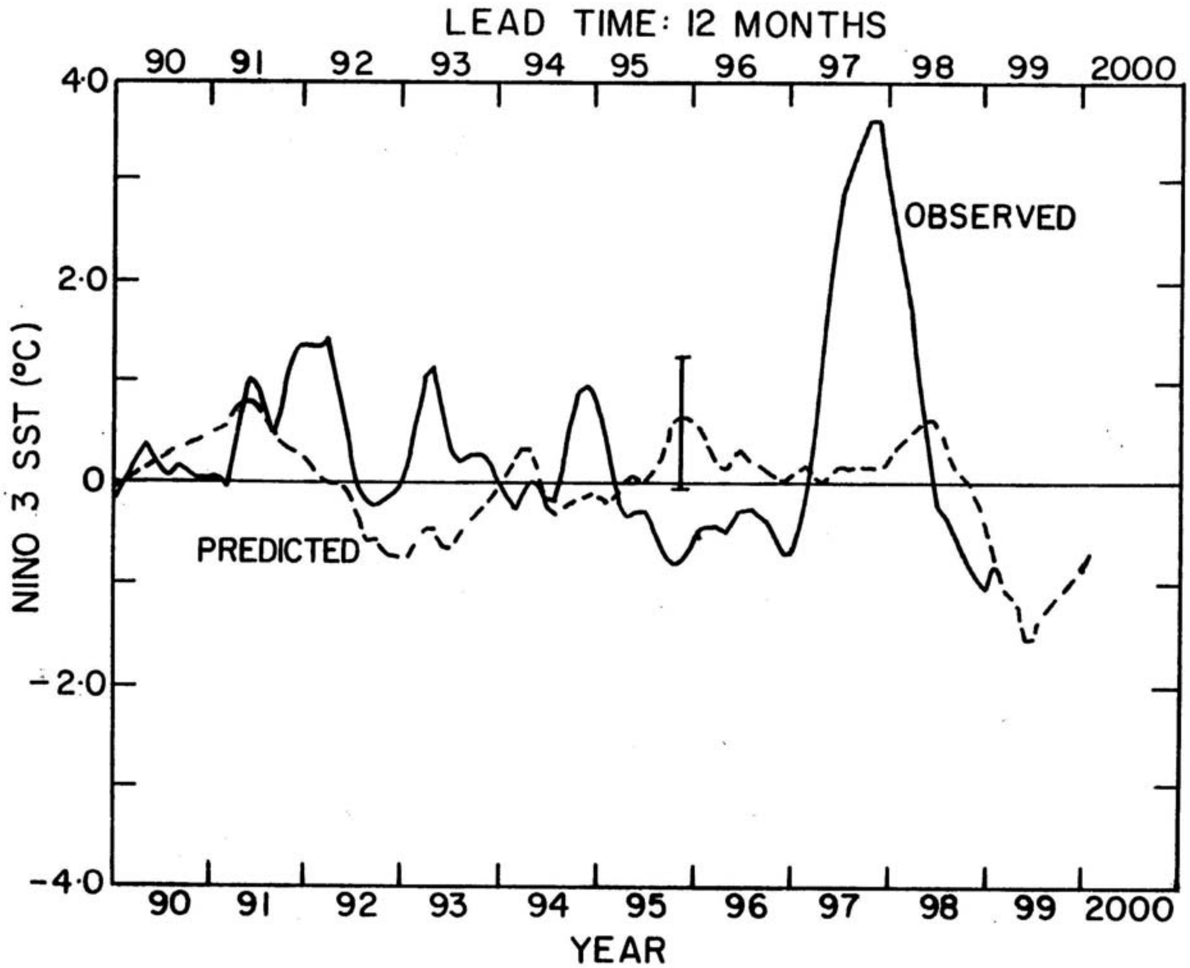


Fig. 4. The 12-month lead forecasts of the area-averaged Niño-3 SST (dashed line) using SSA-MEM (Saunders *et al*., 1999) and the observed Niño-3 SST (full line)

reproduced here in our Fig. 4. The solid plot is for observed values while the dashed plot is for the 12-month lead predictions. Judgement of fitness by visual inspection can be subjective. But the matching between the solid and dashed lines seems to be poor. The dashed line does not reproduce the complex structures of 1991-95, and for the El Niño of 1997, the dashed line indicates a commencement of a small size event by about August 1997, while the event started several months earlier (March 1997) and achieved gigantic proportions, unprecedented in known history.

Let us examine the predictions of the SOI for this event as reported by of Dr. Michael Ghil and his colleagues in various issues of ELLFB. Fig. 5 shows sequentially from October 1996 onwards (Fig. 5, plot 1), the evolution of observed SOI (5-month running means, thin line), the SSA-filtered SOI (thick line) and the 12-month lead predictions (dashed line), as extrapolated from the thick line values. In plot 1 (Jiang *et al.*, 1996), values up to October 1996 were used, and the predicted SOI values for November 1996 - October 1997, are slightly declining but all positive (above the zero line), indicating normal or La Niña type conditions. This prediction was obviously wrong; because a severe decrease in SOI started only a few months later (March 1997). In plot 2 (Jiang *et al*., 1997a), values up to January 1997 are used and the prediction is still of SOI values positive, up to January 1998, again a wrong prediction. In plot 3 (Jiang *et al*., 1997b), values up to April 1997 are used and a severe decrease was already observed, but the predicted SOI are still very small, though slightly negative. In plot 4 (Saunders *et al*., 1997), values up to July 1997 are used and thus, a considerable portion of the event has gone into prediction. This is true for plot 5 (Saunders *et al*., 1998a) as also for plot 6 (Saunders *et al*., 1998b)

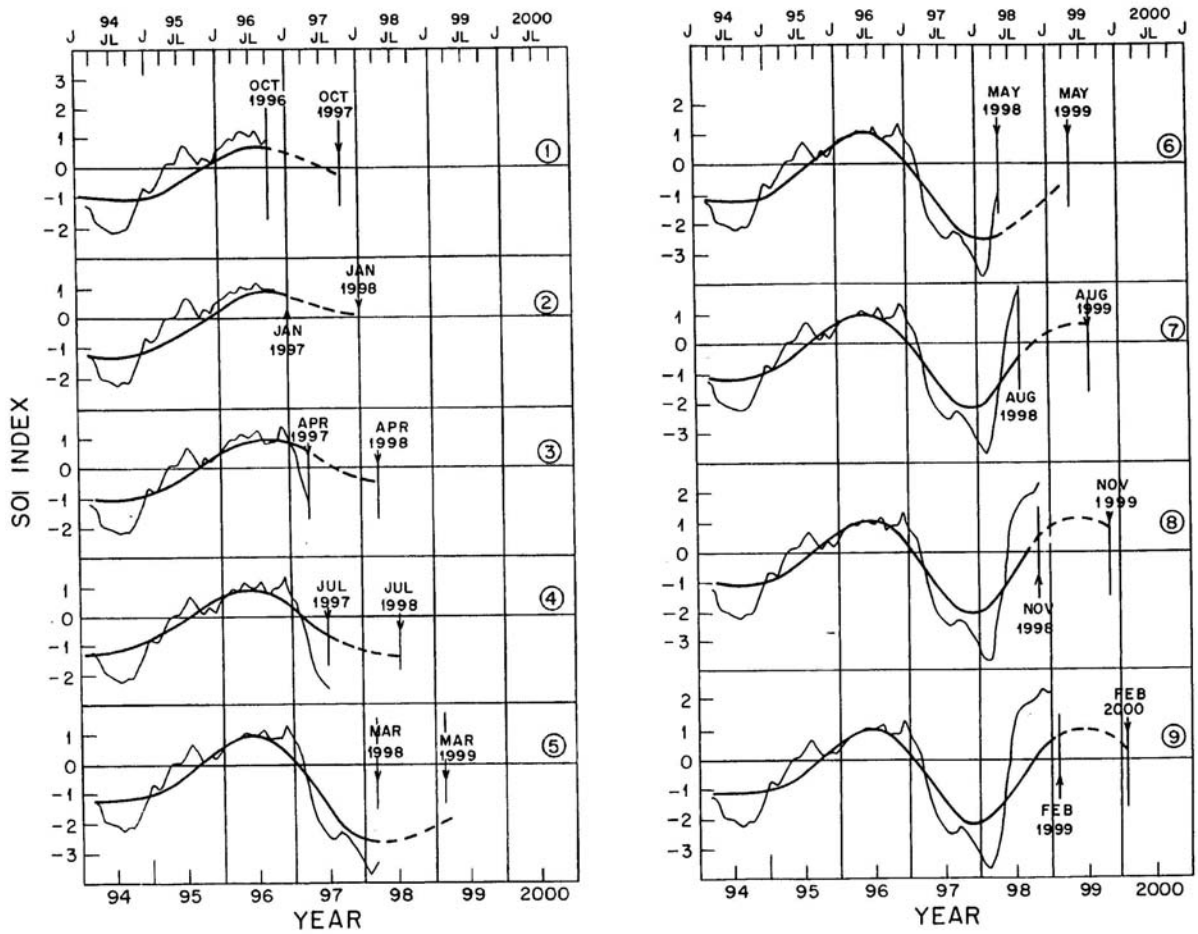
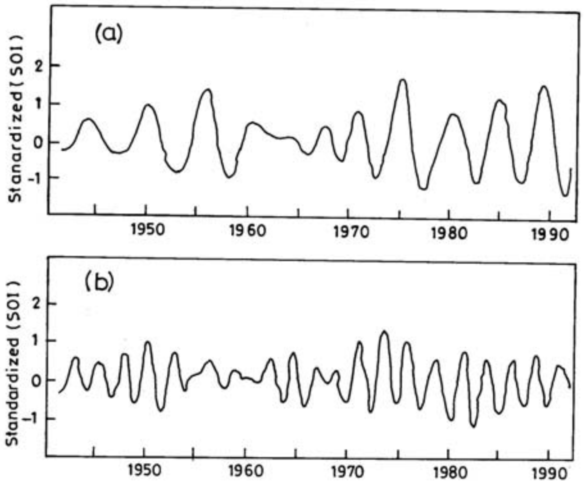


Fig. 5. Observed 5-monthly means of SOI (thin full line) up to the month marked by the first vertical line in each panel (*e.g*., October 1996 in Panel 1), the SSA-MEM forecasts (thick smooth line) of the UCLA group of Dr. Ghil, (as given in the various issues of ELLFB) by using observed data up to that date, and predictions (dashed line) for the next 12 months up to the second vertical line in each panel (*e.g*., October 1997 in Panel 1). Successive panels 1, 2, 3 etc. are for roughly successive 3 months. Note that the forecast (thick line) is essentially a simple sinusoid of ~3.5 year period

and the predictions here are for negative values of SOI, which came true. However, these predictions look like predicting when the event already occurred. Also, in plot 6, the event had already started recouping, but the predictions are still for substantially negative SOI. In June 1998, SOI values had already turned positive, but the prediction was for negative values for the next 11 months, a wrong prediction. In plot 7 (Saunders *et al*., 1998c), plot 8 (Saunders *et al.*, 1998d) and plot 9 (Saunders *et al.*, 1999), observed values had already turned positive and were used into the prediction calculation, which was of positive SOI, correctly. Thus, the predictions have been crude, correct in the peak of the event but unable to detect the commencement or termination with accuracy of a few months.

The reason for this inability to detect the fine structure is easy to understand. As mentioned by Saunders *et al*. (1997), the skill of the SSA-MEM is based on the low–frequency, oscillatory components of the system, and so, forecasts tend to filter out short-lived from long-term anomalies. Earlier, KG mentioned that in contrast with standard spectral analysis in which basic functions are given a priori (*e.g*., the sines and cosines of Fourier analysis), in SSA they are determined from the data themselves to form an orthogonal basis that is optimal in the statistical sense. However, in practical terms, all that has happened for the SSA-filtered SOI series (Fig. 5, thick line) is a smooth wave of periodicity ~3.5 years, applicable for recent predictions. Even though Jiang *et al*. (1995) rnentioned quasi-quadrennial (QQ) and quasi-



Figs. 6(a&b). The reconstructed time series obtained by combining the variance associated with (a) low frequency (3-4 year) ENSO component and (b) high frequency (2-3 year) ENSO component, reproduced from KG (Keppenne and Ghil, 1992)

biennial (QB) components of the ENSO variability, only the QQ seems to be operative for recent years. With a 3.5 year wave, extrapolation backwards from 1997 would fit the 1994 event as well as the 1991 and the 1987 events. But the 1989 event would be missed. Obviously, QBOs are needed to explain short-spaced events. However, as seen in Fig. 3(d), addition of QBOs reproduces the 1987, 1989, 1991, 1994, 1997 events but adds one more event in 1999. Thus, there is no absolute solution, not in our simple MEM-MRA combination, much less in the SSA-MEM combination. The QQ component may be reasonably stable, but the QB component seems to be highly transient.

Transient nature of the QQ and QB components was known to KG. Their Fig. 2 is reproduced here in Fig. 6, where (a) shows the reconstructed SOI time series obtained by combining the variance associated with the low frequency part (QQ, ~4 years), while (b) shows the high frequency part (QB, 2-3 years). As can be seen, both show considerable change of characteristics with time. In (a) for QQ, the spacing was larger during 1940-60 (~5 years) as compared to 1970-90 (~3.5 years), while the pattern was obscure during 1960-70. In (b) for QB, the pattern was probably the same, except for 1955-62 when the pattern was obscure. For the period 1950 onwards, Kane (1998) divided the SOI data into 3 successive intervals of 15 years each (60 seasonal values) and the MEM-MRA analysis showed:

(*i*) During 1953-67, one strong peak at 4.41 years.

(*ii*) During 1967-81, two strong peaks, at 2.49 and 4.28 years.

(*iii*) During 1981-95, a small peak at 2.35 years, a strong peak at 4.36 years.

If the data were divided into two parts only of 22 years each (88 seasonal values), the analysis showed.

(*i*) During 1953-74, a small peak at 2.88 years, a strong peak at 3.73 years.

(*ii*) During 1974-95, small but significant peaks at 2.44, 3.34 years, a strong peak at 4.51 years.

Thus, the spectral characteristics are highly variable. Obviously, for recent predictions, the recent patterns would be more relevant (if at all) and the earlier patterns would certainly be irrelevant. Extending the data base used to compute the forecast backward from June 1945 to August 1881 as was done by Keppenne and Lall (1996) does not seem to be appropriate for recent predictions, but is certainly useful for studying the changes of the spectral characteristics of the SOI time series over long intervals.

Besides the method of KG, two more schemes of predicting SOI are reported in ELLFB as follows:

(a) Analogue (Non-linear) forecasts of the Southern Oscillation Index time series, based on the non-linear time series forecasting techniques of Sugihara and May (1990). Drosdowsky (1997) used data up to February 1997 and predicted positive values of SOI for March, April and May 1997. The prediction was obviously wrong, as SOI values became negative in March 1997 and more so in April and May.

(b) Application of the El Niño-Southern Oscillation CLImatology and PERsistence (CLIPER) forecasting scheme, based on the optimal combination of persistence, month-to-month trend of initial conditions and climatology, the ENSO-CLIPER (Knaff and Landsea, 1997). Knaff and Landsea (1998) forecasted for March-May 1997 a negative value for SST anomaly, obviously a wrong prediction as March SST values were already positive.

Thus, both these methods failed to predict the commencement of the El Niño in March 1997. Also, predictions about terminations and size estimates were wrong.

6. Indications for near future

The SSA-filtered data plot (thick line) of SOI in Fig. 5 shows a smooth sinusoid which had a minimum in early 1994, a maximum in mid-1996, a minimum in 1997 end, and a maximum in mid-1999, indicating a wave of periodicity ~3.5 years. As such, the present La Niña (positive SOI) phase should have been over by 1999 end and the next minimum should have been seen in 2000-2001. Fig. 3(e) also indicated the same, on the assumption that only a 3.5 years wave is operative. As it happens, the positive SOI phase (La Niña) is continuing for the last 30 or more months, and till September 2001, there is no indication that this phase is coming to an end. The SST values in the Niño 1+2 region are still near normal, with no indication of a possible increase (NOAA website, http://www.cpc.ncep.noaa.gov/data/indices/). Of course, this does not preclude the possibility that an El Niño may erupt suddenly in near future, as happened in March 1997.

7. Conclusions and discussion

A simple method where MEM (Maximum Entropy Method) was used only to detect possible periodicities, which were then used in MRA (Multiple Regression Analysis) to estimate the amplitudes and phases of these periodicities, was applied to SOI (Southern Oscillation Index, represented by the sea-level atmospheric pressure difference T-D between the locations Tahiti and Darwin). Deseasoned SOI data were used for MEM analysis. The following was noted:

(*i*) Using data for 1935-80 as dependent data and obtaining the 4 most significant periodicities, the reconstructed series reproduced most of the El Niño and La Niña events up to 1980 reasonably well. But the predictions for 1980 onwards were not satisfactory.

(*ii*) Omitting earlier data, the analysis was repeated for 1945-80 and 1955-80 as dependent data. The reconstructed series showed satisfactory matching, but matching after 1980 was not good. Using data for 1965-89 as dependent data did not show improvement, but data for 1975-89 showed satisfactory match for the 1989, 1991, 1993 events (but not for the 1997 event).

(*iii*) Various combinations were tried out, *e.g*., 1975-95, 1980-95, 1980-98, 1985-98, 1990-98, as dependent data. Only 1985-98 was able to reproduce the 1997 event; but a (false?) El Niño in 1999 was predicted. The data for 1990-98 reproduced the 1997 event correctly and involved only one periodicity of ~3.5 years. Thus, the spectral characteristics of the SOI series seem to have changed considerably in recent decades and only very recent data are useful for further prediction, if at all.

(*iv*) KG (Keppenne and Ghil, 1992) made a similar analysis using SSA (Singular Spectrum Analysis) in combination with MEM. Using data for 1940-90, the reconstructed series matched very well with the original values (5-month running means). However, for the early 1990s, the predictions did not come out very well. For the late 1990s, the SSA-filtered SOI was essentially a sinusoid of ~3.5 years and explained the 1994 and 1997 events correctly. However, the abrupt commencement of the 1997 El Niño in March 1997 and its abrupt end in June 1998 (total interval 16 months only) could not be foreseen with such a smooth wave of 3.5 years periodicity. For predicting abrupt changes, some high frequency components are necessary, but these may be very transient in nature.

In the last two decades, forecasting technology has advanced considerably. Besides the predictions based on extrapolation of spectral components, several schemes using various combinations of GCMs (General Circulation Models) are used and the predictions published regularly in ELLFBs (Experimental Long-Lead Forecast Bulletins) quarterly with a few months lead. The real test of these models came during the El Niño of 1997, the strongest in known history. Beginning about November 1996, several models predicted some warming in 1997. The forecasts were published in the December 1996 (Vol. 5, No. 4) issue of ELLFB. By and large, the predictions of an event occurring in 1997 came true, but the event turned out to be unexpectedly strong and the commencement and termination months were not predicted correctly. For example, Unger *et al*. (1996) had synthesized information from some of the predictive sources into a single objective estimate; but their consolidated forecast turned out to be mistimed. Instead of the predicted commencement of the warming only in late 1997, the El Niño started early (in March 1997). As remarked by Trenberth (1998), in digesting these predictions in real time, it was not possible to make a forecast until about April 1997 when the SST warming became clearly evident in the eastern tropical Pacific. Particularly deceptive was the forecast of the Cane-Zebiak model. This model had successfully forecast the 1986-1987 event (Cane *et al*., 1986; Zebiak and Cane, 1987). Hence their prediction of "normal by fall of 1997, warming thereafter" (Zebiak and Cane, 1996) substantially inhibited the issuing of a forecast for March 1997. Even in the September 1997 issue of ELLFB, Saunders *et al*. (1997) made the following remark, "the observed sharp rise of SST anomalies over the last few months does not foretell an El Niño of unprecedented magnitude; it is associated, in all likelihood, with a short- lived spike added to an ENSO event of typical magnitude". A preliminary assessment of some forecasts from February-March 1997 is given in Barnston (1997). However, the statistical/inverse models did not seem to show skill of value. For subsequent evolution, Trenberth (1998) points out that the full climate models, consisting of a full atmospheric general circulation model (GCM) coupled to a dynamical ocean GCM, were the most successful, for the first time. However, these models are not perfect and require corrections of various kinds. For example, the NCEP model (Ming *et al*., 1996) underestimated the magnitudes by at least a factor of 2. The

ECMWF model (Balmaseda *et al*., 1996) indicated a leveling off in May 1997, which did not come true.

In the Executive Summary for the La Niña Summit Report of a Workshop held in Boulder, Colorado from 15-17 July, 1998, Glantz (1998) mentions the key points about prediction of the El Niño 1997 as follows: (a) The performance in forecasting the onset of the 1997-98 El Niño was largely mediocre. (b) Dynamical models as yet do not outperform the statistical ones, with respect to forecasting El Niño. The situation about predictions based on spectral components also seems to be not fully satisfactory, especially for detecting finer characteristics like commencements and terminations of El Niños etc.

The El Niño of 1997 is followed by a La Niña event, which started in June 1998 and SOI values at present are still positive. Forecasts of different models are divergent. In the June 2001 (Vol. 10, No. 2) issue of ELLFB, the “Summary of Forecasts” states the following:

Dynamic Models

The consensus among the dynamic models is for a warming in the tropical eastern Pacific, but some models predict near normal SST through the middle of 2002, while some predict cold conditions in the middle of 2002.

Statistical Models

The statistical models are predicting a return to near normal in the eastern Pacific, but some models predict a modest El Niño for the boreal winter 2001-02 while some others predict near normal conditions. There are equal probabilities for “cool”, “normal” and “warm” in the boreal winter season 2001-02.

Thus, no matter what happens, some model or other will be able to say, “We said so”. In our simple MEM-MRA scheme, a 3.5 years periodicity envisaged a new El Niño in winter of 2000-2001, but that did not come true. And even in September 2001, there are no indications of an El Niño, though a sudden appearance of an El Niño in near future is not precluded.

The unsatisfactory performance of the present day prediction models is obviously bothering many workers. Pielke (1998) argues that though commonly weather is considered an initial value problem while climate is assumed to be a boundary value problem, the boundaries being the ocean surface and the land surface, these boundaries may show changes over the same time period as the atmosphere. In that case, the nonlinear feedbacks (two-way fluxes) between the air, land, and water may turn these boundaries into interfaces between interactive mediums and the two-way fluxes must be considered as part of the predictive system. Thus, climate prediction would be essentially an Initial Value Problem and the present-day model-based forecasts of future climate should be viewed as sensitivity analyses rather than as reliable predictions.

In the present analysis, the values of atmospheric pressure at Tahiti and Darwin in milibars were used. Some parts of the analysis were repeated by using the standardized Southern Oscillation Index given in the Archieves of the Climate Prediction Center, NOAA. The results were exactly the same qualitatively and only slightly different quantitatively (mainly in units).

Acknowledgements

Thanks are due to Dr. Don Garrett for El Niño and other data through the Web site of Climate Prediction Center, NOAA, NCEP, Washington. This work was partially supported by FNDCT, Brazil under contract FINEP-537/CT.

References

Angell, J. K., 1981, “Comparison of variations in atmospheric quantities with sea surface temperature variations in the equatorial eastern Pacific”, *Mon. Wea. Rev*., **109**, 230-243.

Barnston, A. G., 1997, “ENSO forecasts for 1998: Experimental long lead predictions. Proc. Twenty-second Annual Climate Diagnostics and Prediction Workshop”, Boulder, Colorado Oct 6-10, 1997, US Dept. Commerce, 6-9.

Bevington, P. R., 1969, “Data Reduction and Error Analysis for the Physical Sciences”, McGraw-Hill, New York, 164-176.

Blackman, R. B. and Tukey, J. W., 1958, “The Measurement of Power Spectral”, Dover, New York, p 190.

Burg, J. P., 1967, “Maximum Entropy Spectral Analysis”, paper presented at the 37th Meeting. Society of Exploration Geophysics, Oklahoma City, October.

Cane, M. A., Zebiak, S. E., and Dolan, S. C., 1986, “Experimental forecasts of El Nino”, *Nature,* **321**, 827-832.

COLA, Center for Ocean-Land-Atmosphere Studies, Calverton, Maryland, USA, Experimental Long-Lead Forecast Bulletins, March, June, September, December, 1998, 1999, Volumes 7, 8.

CPC, Climate Prediction Center, NOAA, National Centers for Environmental Predictions, Washington D. C., Experimental Long-Lead Forecast Bulletins, March, June, September, December, 1996, 1997, Volumes 5, 6.

Deser, C. and Wallace, J. M., 1987, “El Niño events and their relation to the Southern Oscillation: 1925-1986”, *J. Geophys. Res*., **92**, 14189-14196.

Drosdowsky, W., 1997, “Anologue (Non-linear) forecasts of the Southern Oscillation Index time series”, CPC Experimental Long-Lead Forecast Bulletin, March 1997, **6**, 1, 53-54.

Glantz, M. H., 1998, "A La Niña Summit: A review of the Causes and Consequences of Cold Events", Executive Summary of the workshop held 15-17 July 1998 in Boulder, Colorado, Environmental and Societal group, NCAR, Boulder.

Jenkins, G. M., and Watts, D. J., 1968, "Spectral Analysis and its applications", Holden- Day, San Francisco.

Jiang, N., Neelin, J. D., and Ghil, M., 1995, "Quasi-quadrennial and quasi-biennial variability in the equatorial Pacific", *Climate Dynamics*, **12**, 101-112.

Jiang, N., Ghil, M., and Neelin, J. D., 1996, "Forecasts of Niño 3 SST anomalies and SOI based on Singular Spectrum Analysis combined with the Maximum Entropy Method", CPC Experimental Long-Lead Forecast Bulletin, December 1996, **5**, 4, 50-52.

Jiang, N., Ghil, M., and Neelin, J. D., 1997a, "Forecasts of Niño 3 SST anomalies and SOI based on Singular Spectrum Analysis combined with the Maximum Entropy Method".

Jiang, N., Ghil, M., and Neelin, J. D., 1997b, "Forecasts of Ni*no* 3 SST anomalies and SOI based on Singular Spectrum Analysis combined with the Maximum Entropy Method", CPC Experimental Long-Lead Forecast BuIletin, June 1997, **6**, 2, 63-65.

Kane, R. P., 1977, "Power spectrum analysis of solar and geophysical parameters", *J. Geomag. Geoelect*, **29**, 471-495.

Kane, R. P., 1998, "Spectral comparison of ENSO and stratospheric zonal winds", *Int. J. Climatol*., **18**, 1195-1208.

Keppenne, C.L. and Ghil, M., 1992, "Adaptive filtering and prediction of the Southern Oscillation Index", *J. Geophys. Res*., **97**, 20449-20454.

Keppenne, C.L. and Lall, U., 1996, "Complex singular spectrum analysis and multivariate adaptive regression splines applied to forecasting the Southern Oscillation", CPC Experimental Long-Lead Forecast Bulletin, March 1996, Vol. 5, No. 1.

Knaff, J. A., and Landsea, C. W., 1997, "An El Niño-Southern Oscillation Climatology and Persistence (CLIPER) Forecasting Scheme", *Weather Forecasting*, **12**, 633-652.

Knaff, J. A., and Landsea, C. W., 1998, "Application of the El Niño-Southern Oscillation CLImatology and PERsistence (CLIPER) Forecasting Scheme", COLA Experimental Long-Lead Forecast Bulletin, June 1998, **7**, 2, 50-51.

Ming, J., Kumar, A. and Leetmaa, A., 1996, "Forecasts of tropical Pacific SST using a comprehensive coupled ocean-atmosphere dynamic model", CPC Experimental Long- lead Forecast Bulletin 5, No. 4, December 1996, 10-14.

Parker, D. E., 1983, "Documentation of a Southern Oscillation Index", *Meteor. Mag*., **112**, 184-188.

Penland, C., M. Ghil and Weickmann, K.M., 1991, "Adaptive filtering and maximum entropy spectra", with application to changes in atmospheric angular momentum, *J. Geophys. Res*., **96**, 22, 659-671.

Pielke Sr., R. A., 1998, "Climate prediction as an initial value problem", *Bull. Amer. Meteor. Soc*., **79**, 2743-2746.

Preisendorfer, R. W., 1988, "Principal Component Analysis in Meteorology and Oceanography", ed. C. D. Mobley, p425. Elsevier, New York.

Saunders, A., Ghil, M., and Neelin, J. D., 1997, "Forecasts of Niño 3 SST anomalies and SOI based on Singular Spectrum Analysis combined with the Maximum Entropy Method", CPC Experimental Long-Lead Forecast Bulletin, September1997, **6**, 3, 43-44.

Saunders, A., Ghil, M., and Neelin, J. D., 1998a, "Forecasts of Niño 3 SST anomalies and SOI based on Singular Spectrum Analysis combined with the Maximum Entropy Method", COLA Experimental Long-Lead Forecast Bulletin, March 1998, **7**, 1, 40-42.

Saunders, A., Ghil, M., and Neelin, J. D., 1998b, "Forecasts of Niño 3 SST anomalies and SOI based on Singular Spectrum Analysis combined with the Maximum Entropy Method", COLA Experimental Long-Lead Forecast Bulletin, June 1998, **7**, 2, 56-58.

Saunders, A., Ghil, M., and Neelin, J. D., 1998c, "Forecasts of Niño 3 SST anomalies and SOI based on Singular Spectrum Analysis combined with the Maximum Entropy Method", COLA Experimental Long-Lead Forecast Bulletin, September, l998, **7**, 3, 53-55.

Saunders, A., Ghil, M., and Neelin, J. D., 1998d, "Forecasts of Niño 3 SST anomalies and SOI based on Singular Spectrum Analysis combined with the Maximum Entropy Method", COLA Experimental Long-Lead Forecast Bulletin, December 1998, **7**, 4, 41-43.

Saunders, A., Ghil, M., and Neelin, J. D., 1999, "Forecasts of Niño 3 SST anomalies and SOI based on Singular Spectrum Analysis combined with the Maximum Entropy Method", COLA Experimental Long-Lead Forecast Bulletin, March 1999, **8**, 1, 39-41.

Sugihara, G. and May, R. M., 1990, "Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series", *Nature*, **344**, 734-741.

Trenberth, K. E., 1998, "Development and forecasts of the 1997-98 El Niño: CLIVAR Scientific Issues", *Exchanges*, **3**, 4-14.

Ulrych, T. J. and Bishop, T. N., 1975, "Maximum Entropy Spectral Analysis and autoregressive decomposition", *Rev. Geophys*., **13**, 183-200.

Unger, D., Barnston, A., Dool, H. V. D., and Kousky, V., 1996, "Consolidated forecasts of tropical Pacific SST in Niño 3.4 using two dynamical models and two statistical models", CPC Experimental Long-lead Forecast Bulletin 5, 4, December 1996, 68-70.

Zebiak, S. and Cane, M., 1987, "A model El Niño-Southern Oscillation", *Mon. Wea. Rev*., **115**, 2262-2278.

Zebiak, S. and Cane, M., 1996, "Forecasts of tropical SST using a simple coupled ocean-atmosphere dynamic model", CPC Experimental Long-lead Forecast Bulletin 5, No. 4, December 1996, 4-9.