A new El Niño-Southern Oscillation forecasting tool based on Southern Oscillation Index

C. A. Varotsos and C. Tzanis

Climate Research Group, Division of Environmental Physics and Meteorology, Faculty of Physics, University of Athens, University Campus Bldg. Phys. V, Athens 15784, Greece

Received: 20 January 2012 – Accepted: 29 June 2012 – Published: 16 July 2012

Correspondence to: C. A. Varotsos (covar@phys.uoa.gr)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

An exploration of the temporal evolution of the El Niño Southern Oscillation (ENSO) during January 1876–November 2011 by means of a new time domain called *natural time* reveals that the major ENSO events provide precursory signals that are maximized in a time window of almost two years. This finding improves the accuracy of the short-term prediction models of the ENSO extreme events, preventing thus from its disastrous impacts in advance.

1 Introduction

The term El Niño/La Niña refers to the process of the irregular occurrence of the oceanic event of the extensive warming/cooling of the Central and Eastern Tropical Pacific. El Niño/La Niña leads to a major shift in weather patterns across the Pacific, like the increased convection (altering the Walker circulation) or cloudiness in the Central Tropical Pacific Ocean, weaker/stronger than normal trade winds across the Pacific Ocean.

A well-established link between the atmosphere and the oceanic El Niño/La Niña phenomenon is the seesaw back and forth in surface air pressure between the Eastern and the Western South Pacific which is known as Southern Oscillation. The strength of this oscillating atmospheric bridge is measured by the Southern Oscillation Index (SOI), which is computed from the monthly surface air pressure difference between Tahiti (17°40′ S, 149°25′ W) and Darwin (12°27′ S, 130°50′ E).

It is widely recognized that El Niño episodes are associated with negative values of the SOI, while La Niña episodes with positive values of the SOI, consisting thus a composite oceanic-atmospheric phenomenon, notably the El Niño/La Niña-Southern Oscillation (ENSO) phenomenon. In this context, Wunsch (1999) stated that at least some of El Niño phenomenon is an oceanic response to purely stochastic atmospheric forcing.
The global concern for ENSO’s prediction stems from the fact that ENSO has climatological impacts in regions far removed from the Tropical Pacific (e.g. teleconnections) and may be linked to extreme weather conditions (e.g. floods and droughts), changes in the incidence of epidemic diseases (e.g. malaria), severe coral bleaching, civil conflicts etc (Klein et al., 1999; James et al., 2003; Eckhardt et al., 2004; Marshall and Schuttenberg, 2006; Hsiang et al., 2011; Schiermeier, 2011). That is why ENSO has been shown to be of great use to climatologists as a climate index (Stenseth et al., 2003). For instance, the effect of 2010–2011 La Niña (the strongest ever observed) on Eastern Australia was devastating (i.e. Australia’s second wettest year on record).

In general, in the climate predictions we try to find out the procedures that would lead to the predictions that are most accurate (for quality for an individual case e.g., a single location for one time), with the highest possible skill (for quality over a collection of cases e.g., either many locations for a single time, many times for a single location, or, more typically, for many cases over both time and space). In this context Barnston et al. (2005) stressed that the best prediction may not be accurate with respect to the observed result in individual cases, or even skillful on average over many such cases, since the signal may be disturbed by the climate noise.

More recently, seasonal forecasts of 3-month-average surface temperature or precipitation have been clearly demonstrated to have skill in particular seasons, regions, and circumstances resulting in a better quantification of the climate effects of the ENSO phenomenon (Livezey, 1990; Kumar et al., 2000; Shukla et al., 2000; Graham et al., 2000). Along these lines the Seasonal Diagnostics Consortium activities lead to greater acceptance in the use of dynamical methods to better understand and predict climate variations on seasonal-to-decadal time scales.

Undoubtedly the reliable prediction of ENSO major events (via the hydrodynamic coupled ocean-atmosphere and statistical models) is limited and thus its improved forecasting is still among the open geophysical problems (e.g. Balmaseda et al., 1994; Stone et al., 1996). It is due to the fact that ENSO is a non-stationary, nonlinear quasi-periodic interannual variation in global atmospheric and oceanic circulation patterns,
occurring at irregular intervals of 3–7 yr (roughly every five years), which, as such, is not yet fully understood (e.g. Ausloos and Ivanova, 2001, 2003). Of course, the principal question that arises is whether this climate anomaly is predictable and at what lead times. In this context, a number of coupled models have been employed for understanding and prediction of ENSO phenomenon (see the review articles in the special issue of J. Geophys. Res. on the Tropical Ocean Global Atmosphere – TOGA – programme, and reviews by Bigg, 1990; McCreary and Anderson, 1991; Neelin et al., 1994; Palmer and Anderson, 1994; Latif et al., 1998). These models justly assume that the atmosphere and ocean are in equilibrium, but exclude the intraseasonal oscillation and many extratropical teleconnections. An assessment of various models predictions are presented and discussed in Anderson and Davey (1998) assuming that every ENSO event is an individual and its timing, duration and magnitude can vary considerably, affecting its impact.

In this paper we use the SOI as a measure of the status of ENSO in order to describe an advanced ENSO forecasting tool that provides an alarm of SOI major events coming from the data of previous two years or so for regions worldwide. Our analysis is based not in the conventional time domain but in a new not continuous time domain termed natural time (Varotsos et al., 2002), in order to detect the novel dynamical features hidden behind the coupled ocean-atmosphere system which will enable us to predict and impending major events in ENSO.

This new analytical tool (natural time analysis – NTA), which extracts from a given time series the maximum information possible (Abe et al., 2005), has been implemented in various fields, including Biology, Earth Sciences and Physics (for a review see Varotsos et al., 2011). The first study in climate physics that made use of this modern method was very recently published (Varotsos and Tzanis, 2012) and focused on the investigation of the dynamical evolution of the ozone hole complex system over Antarctica, where significant precursory signals were clearly identified before the unprecedented event of the major, sudden stratospheric warming and the subsequent break-up of the Antarctic ozone hole into two holes in September 2002.
2 Data

For the purposes of the present study the SOI monthly averaged data is analyzed for the time interval January 1876–November 2011. This data set was obtained from the Long Paddock site and it is entitled: MonthlySOIPhase1887–1989Base. The SOI used here is calculated by employing the Troup’s formula (Troup, 1965; Power and Kociuba, 2011), notably:

\[ 10 \times \frac{\text{PA (Tahiti)} - \text{PA (Darwin)}}{\text{SDD}} \]  

(1)

where PA stands for the pressure anomaly that is defined as the deviation of the monthly mean sea level (air) pressure (i.e. \( P \) (Tahiti) – \( P \) (Darwin)) from the long-term average of \( P \) (Tahiti) – \( P \) (Darwin) for the calendar month (based on the period 1887–1989) and SDD is the standard deviation of the difference \( P \) (Tahiti) – \( P \) (Darwin) for the same calendar month (1887–1989 base period). It should be noted that Troup’s monthly SOI from 1876 onwards is derived from normalized Tahiti minus Darwin mean sea level pressure.

As mentioned in the Introduction SOI demonstrates the strength of the seesaw back and forth in surface air pressure between the Eastern and the Western South Pacific. The latter location is the most extensive region in the world of water warmer than 28°C and is considered a major source of atmospheric heating, which drives large-scale convective circulation patterns (e.g. Kahya and Dracup, 1993). Therefore, SOI may be justly regarded as a substantial component of the energy exchange between the coupled ocean/atmosphere systems.

3 Method and analysis

We analyse the SOI time series described in the previous section by employing the natural time analysis (NTA) introduced by Varotsos et al. (2002) and presented briefly in the following steps:
Firstly, we convert the original SOI time series to a new one disregarding the time of occurrences but keeping the temporal sequence of the events. To this end, for each SOI value \( Q_k \) we calculate the ratio \( \chi_k \) of the order of its occurrence \( k \) with the total number \( N \) of SOI values, notably:

\[
\chi_k \equiv \frac{\text{Order of occurrence of an event } (k)}{\text{Total number of the events } (N)} \quad (2)
\]

In this way, we are introducing a new time series which consists of the pairs \((\chi_k, Q_k)\). The new parameter \( \chi_k \) which replaced the conventional time \( t \) bears the name “natural time” and varies between zero and unity (being not continuous in contrast to the conventional time that is assumed continuous). It plays the role of the indicator of the occurrence of the \( k \)-th event proceeding with the occurrence of each consecutive \( k \)-th event with intensity \( Q_k \). Additionally, since in natural time \( Q_k \) corresponds to an intensity (i.e. a positive value) we consider as \( Q_k = \text{SOI}_k + \text{min(SOI)} \), where \( \text{min(SOI)} \) is the minimum value that SOI has reached during the period of our study.

To briefly describe the methodology of NTA let us suppose an excerpt of the time series of SOI maxima events shown in Fig. 1a. The new time series which depicts the evolution of the pair \((\chi_k, Q_k)\) or in other words, the original time series conversed in the natural time domain is shown in Fig. 1b.

In the case of the “uniform” distribution (e.g. when the system is in a stationary state emitting uncorrelated bursts of energy) the \( Q_k \) are positive independent and identically distributed random variables. Then as \( N \to \infty \), the probability value \( p(\chi) \) of the probability variable \( \chi \) tends to \( p(\chi) = 1 \) leading to an average value of natural time:

\[
\langle \chi \rangle = \int_0^1 \chi p(\chi) d\chi = \frac{1}{2} \quad (3)
\]

Secondly, we calculate the entropy \( S \) of SOI values in the natural time domain as follows:
By definition: \( S \equiv \langle \chi \ln \chi \rangle - \langle \chi \rangle \ln \langle \chi \rangle , \)

where: \( \langle \chi \rangle = \sum_{k=1}^{N} p_k \chi_k , \langle \chi \ln \chi \rangle = \sum_{k=1}^{N} p_k \chi_k \ln \chi_k , \)

with \( p_k = Q_k / \sum_{n=1}^{N} Q_n , \)

and \( (p_k) \) is a normalized intensity of the \( k \)-th event (i.e. the “probability” to observe it).

Thus, \( S = \sum_{k=1}^{N} p_k \chi_k \ln \chi_k - \left( \sum_{k=1}^{N} p_k \chi_k \right) \ln \left( \sum_{m=1}^{N} p_m \chi_m \right) \) \( (4) \)

This calculation of the entropy is made at a window of SOI events (monthly mean values) of length \( (i) \) which is sliding, each time by one month, running thus the whole time series of SOI. We next repeat this calculation at longer and longer windows with \( 3 \leq i \leq 84 \) months. We have selected an upper limit of 84 months because it is the closest time to the maximum ENSO period (7 yr).

Thirdly, we perform the same entropy calculations as above, but this time considering the time reversal, e.g., in the excerpt of Fig. 1 the last event is now read as the first one, the last but one event as the second one, etc. We label this entropy obtained upon considering the time reversal, by \( (S_\)\). It was found (Varotsos et al., 2005) that in general \( S_\) is different from \( S \), and hence the entropy in natural time shows the breaking of the time reversal symmetry, thus revealing the profound importance of considering the (true) time arrow in classifying similar looking signals of different dynamics.

Finally, we calculate the difference \( \Delta S_i = S_i - (S_\)_i \) for each \( (S) \) and \( (S_\)\) at a scale \( (i) \) (= number of successive events, i.e. the number of months).

For instance, Fig. 2 shows the entropy change in natural time under time reversal \( \Delta S_i \) for the window length \( i = 36 \) months (left scale) sliding each time by one month through the whole time series (January 1876–November 2011) of SOI (right scale).

From physical point of view, it has been shown that an increasing trend in a time series corresponds to negative \( \Delta S \) and vice-versa (see p. 163 of Varotsos et al., 2011).

It should be noted here that natural time has been applied to intermittent time series as in the case of earthquakes, where individual events can be easily identified.
In addition, natural time has been also applied to time series analysis, e.g., see pp. 126–128 of Varotsos et al. (2011). In the latter frame, the values of the time series itself (after adding an appropriate constant so that they become positive) constitute the quantity $Q_k$; in the present case the corresponding mean values per month constitute the time series. It has been proven (Varotsos et al., 2005) that $S$ and $S_-$ exhibit stability or experimental robustness. In particular, this proof was made in terms of an early suggestion by Lesche (Lesche stability criterion), which states that an entropic measure (like the one we used here) is stable if its change upon an arbitrarily small deformation on the distribution $p_k$ (representing fluctuation of experimental data) remains small (see pp. 161, 162 of Varotsos et al., 2011).

4 Discussion and results

As explained above the calculation of the entropy change in the natural time domain is made at windows lengths which vary in $3 \leq i \leq 84$ months.

In an effort to exploit $\Delta S_i$ dynamics to reveal the forecasting power of the upcoming SOI events the receiver operating characteristics (ROC) tool has been employed (Fawcett, 2006). This tool has already been used in other fields for prediction purposes (e.g. Garber et al., 2009; Sarlis et al., 2011) assessing the skill of various prediction tools. In this regard ROC graph obtained when using $\Delta S_i$ for various values of ($i$), e.g. $i = 6, 12, 20, 24, 48$ as a binary predictor for the value of SOI in the next month is depicted in Fig. 3.

When $\Delta S_i$ is larger or equal than a threshold $\Delta S$, a time increased probability (TIP) is set on. If the monthly mean value of SOI in the next month is smaller or equal than a target value $T$, the prediction is true. In a ROC, the True Positive Rate (hit rate) versus the False Positive Rate (false alarm rate) is plotted. Hit rate is simply the percentage of large negative monthly mean values of SOI that have been successfully predicted using the data of the previous months in calculating $\Delta S_i$, where as false alarm rate is the ratio of false alarms over the total number of cases in which SOI was larger than $T$. 
By varying $\Delta S$, and keeping the target value $T$ constant we obtain curves like the ones shown in Fig. 3 for $T = -5$ or $T = -15$. A random predictor leads to a ROC graph along the diagonal as shown by the black line.

From the results drawn from the ROC analysis described above is evident that the best hit rate (highest skill) is obtained from $\Delta S_{20}$ to $\Delta S_{24}$ (i.e., for SOI data almost two years before). In this context, Fig. 4 shows the entropy change in natural time under time reversal $\Delta S_i$ for the window length $i = 20$ months (left scale) sliding each time by one month through the whole time series (January 1876–November 2011) of SOI (right scale).

In the following, a trial of the capability to forecast an SOI event via the proposed NTA technique is attempted, keeping in mind that although SOI is commonly used to measure SO, Trenberth and Hoar (1997) regard it as unreliable prior to 1935.

Let's focus first on feasibility to predict the strongest El Niño episode in the last century, in 1982–1983, which had a unique pattern in terms of the warming and the time of onset.

Figure 5 illustrates the value of $\Delta S_{20}$ (red points – left scale) along with SOI (blue points – right scale). The black line represents when the time increased probability (TIP) is on. When $\Delta S_{20}$ exceeds the value $\Delta S$ corresponding to False Positive Rate (false alarm rate) equal to 50% for $T = -5$ the TIP is set on for the next month.

The results obtained from the NTA focusing on the strong ENSO event of 1997–1998 are displayed in Fig. 6. From the time march of the $\Delta S_{20}$ shown in this figure, it is evident that thin strong ENSO event could be predicted in advance via the NTA tool.

It is worthwhile noting that similar successful results to those shown in Figs. 5, 6 were also obtained by testing NTA tool for the prediction of the 2009–2010 strong El Niño event. Also, similar results to those discussed above were also obtained for window lengths $20 < i \leq 24$.

A plausible mechanism which dictates the time window of around two years could be the Quasi-Biennial Oscillation (QBO) in the zonal wind of the tropical stratosphere, which drives the mean meridional circulation inducing a warm or cold anomaly during
its descending zonal mean westerly or easterly shear, respectively. In other words, QBO modulates deep convection in the upper tropical troposphere (Huesmann and Hitchman, 2001). However, the tropical convection is also influenced by the ENSO signal, which produces longitudinal shifts (e.g. Curtis and Adler, 2000), giving rise thus to a robust link between the QBO and ENSO signals. The further investigation of this distinct link is the main objective of an upcoming paper. Finally it should be pointed out that the use of various window lengths in the present study is reminiscent of various window lengths used in many published tools in the literature (e.g. Lenton et al., 2012), like the “optimal climate normals” approach, of Huang et al. (1996). The latter however is based on the construction of a new time series of backward-looking averages, but the method presented here is substantially different, in both the time scale employed and the entire method applied and explained above.

As mentioned before, from physical point of view an increasing trend in a time series corresponds to negative $\Delta S$ and vice-versa, thus in principle the predictive power of $\Delta S_{20}$ for the value of the next month could thought as containing the persistence of the SOI time series. But, $\Delta S$ captures additional properties of the dynamics. This can be seen, for example, in Fig. 7, where we compare three predictors for the future value of SOI after a lag $l = 24$ months, i.e., SOI$(k + 24)$. In this example, we clarify that SOI$(k + 24)$ and SOI$(k)$ are anticorrelated. These predictors are: $-\Delta S_{40}(k)$ (red, $T = -5$, and blue, $T = -15$, lines with points), $-\Delta S_{48}(k)$ (red, $T = -5$, and blue, $T = -15$, lines without points) and the value of $-\text{SOI}(k)$ (magenta, $T = -5$, and cyan, $T = -15$, lines with points). We observe that $-\Delta S_{40}(k)$ performs much better than $-\text{SOI}(k)$ for false positive rates in the range 0.1 to 0.9. The above result confirms that the present method through $\Delta S$, beyond its aforementioned successes, enables prediction well in advance.

An alternative way of understanding intuitively the aforementioned findings is the following: let us consider the simple view that $S$ could be thought as a measure of the “disorder” (in a number of successive events belonging to the natural time window length considered into the calculation). Upon approaching a phase change (or critical
point), the difference $\Delta S$ between the “disorder” looking in the (immediate) future, i.e., $S$, and that in the (immediate) past, i.e., $S_-$, becomes of profound importance when compared to the corresponding difference under more or less “normal” conditions; this reflects that the amplitude of $\Delta S$ maximizes (see p. 423 of Varotsos et al., 2011).

5 Conclusions

The analysis of the SOI time series by using the entropy defined in the natural time domain allows for detection of those characteristics of the dynamics of the complex ocean-atmosphere system that could be employed for the prediction of the ENSO events. We have shown that the calculation of the entropy change under time reversal of the SOI throughout the time period January 1876–November 2011 in the natural time domain at windows lengths which vary in $3 \leq i \leq 84$ months provides the strongest precursory signal at the window lengths of around two years. This means that the calculation of the entropy change of the SOI during the two previous years, approximately, allows warning one month before the advent of an ENSO event.

Investigation of the special characteristics of the entropy change of the SOI on the natural time domain revealed the existence of strong precursory signals in advance the occurrence of the major ENSO events of 1982–1983 and 1997–1998.

References

A new El Niño-Southern Oscillation forecasting tool

C. A. Varotsos and C. Tzanis


A new El Niño-Southern Oscillation forecasting tool

C. A. Varotsos and C. Tzanis


Fig. 1. An excerpt of the time series of SOI maxima (a) in conventional time and (b) in the natural time $\chi$ (evolution of the pair $(\chi_k, Q_k)$). $Q_k$ quantifies the intensity of the $k$-th event.
Fig. 2. The time march of the entropy change $\Delta S_{36}$ in natural time for the window length $i = 36$ months (red points – left scale) along with SOI (blue points – right scale).
Fig. 3. The hit rate versus false alarm rate by varying $\Delta S_i$ and keeping the target value $T$ constant.
Fig. 4. The time march of the entropy change $\Delta S_{20}$ in natural time for the window length $i = 20$ months (red points – left scale) along with SOI (blue points – right scale).
Fig. 5. The time increased probability is set on when $\Delta S_{20}$ (red line) exceeds the threshold $\Delta S$ value predicting thus the major ENSO event of the last century in 1982–1983.
Fig. 6. As in Fig. 5, but for the NTA results for the successful prediction of ENSO event of 1997–1998 (one of the strongest in the 20th century).
Fig. 7. Comparison of three predictors for the future value of SOI after a lag $l = 24$ months, i.e., SOI($k + 24$). The three predictors are: $-\Delta S_{40}(k)$ (red, $T = -5$, and blue, $T = -15$, lines with points), $-\Delta S_{48}(k)$ (red, $T = -5$, and blue, $T = -15$, lines without points) and the value of $-\text{SOI}(k)$ (magenta, $T = -5$, and cyan, $T = -15$, lines with points) and lead to the depicted ROCs.