HOW PREDICTABLE IS EL NIÑO?

BY A. V. FEDOROV, S. L. HARPER, S. G. PHILANDER, B. WINTER, AND A. WITTENBERG

The interplay between the Southern Oscillation and atmospheric noise is the likely reason why each El Niño is distinct, and so difficult to predict.

A recent analysis of 12 statistical and dynamical models used for El Niño predictions finds that at the long (1–2 yr) and even medium (6–11 months) ranges there were “no models that provided useful and skillful forecasts for the entirety of the 1997/98 El Niño” (Landsea and Knaff 2000). Most of the models were wrong in predicting the timing of the onset and/or demise of El Niño, and unable to predict the full duration and even one-half of the actual amplitude of the event. An earlier study by Barnston et al. (1999) reached similarly discouraging conclusions. Regarding El Niño of 2002, the forecasts again cover a very broad spectrum of possibilities (Kirtman 2002). Why is it so difficult to predict El Niño? How predictable is this phenomenon?

There is agreement that the continual Southern Oscillation—El Niño in its warm phase—depends on ocean–atmosphere interactions that amount to a positive feedback. (The winds induce sea surface temperature patterns that in turn affect the winds.) The intensity of the feedbacks is a matter of considerable debate. Figure 1 shows a range of possibilities, from a strongly damped oscillation (Fig. 1a), to a self-sustaining and approximately linear oscillation (Fig. 1b) to an oscillation that is so unstable that nonlinearities introduce irregularities (Fig. 1c). In the latter case, instabilities amplify errors in the initial conditions and thus limit predictability. If reality were to correspond to this parameter range then the predictability of El Niño would be similar to the predictability of weather. An appropriate approach is then an ensemble of calculations, each one starting from a slightly different initial condition. Cane et al. (1986), Zebiak and Cane (1987), Zebiak (1989), Tziperman et al. (1995), and others explore this range of parameters.

Figure 1a shows a very different range of parameters, that of a highly damped oscillation. Now random atmospheric disturbances are always responsible for initiating developments. Predictability is once again very limited, not because of errors in initial conditions as in the case of weather, but rather because of the random disturbances. (Penland and Sardeshmukh 1995; Penland 1996; Thompson and Battisti 2000, 2001, explore this parameter range.) To insist that westerly wind bursts, which occur at ran-
FIG. I. Evolution of the SST along the equator in response to an initial burst of westerly winds, in the coupled model of Neelin (1990). The strength of the ocean-atmosphere coupling increases from (a) to (b) to (c). Temperatures exceed 30°C in shaded regions.

dom times, are always responsible for the initiation of El Niño, is to claim that the Southern Oscillation is strongly damped, as in Fig. 1a.

Different models yield different forecasts (e.g., for El Niño of 1997 and 2002) because El Niño is damped in some of the models, and is highly unstable in others. (For a comprehensive list of references to studies earlier than 1998 see the special issue of the Journal of Geophysical Research, June 1998, Vol. 103, no. C7, which is devoted to a series of articles that review different aspects of El Niño.) To make progress we have to determine which of the panels in Fig. 1 corresponds to reality, and we then have to develop models that capture that particular parameter range. At present the realistic range of parameters is a topic of considerable debate (e.g., Latif et al. 1998; Chang et al. 1996) but there is growing evidence that conditions in the Pacific, for the past few decades, correspond to neither of the extremes shown in Figs. 1a,c. If Fig. 1a were appropriate so that El Niño appears in response to westerly wind bursts, then it is difficult to explain why westerly wind bursts excite El Niño on some occasions but not others, and why El Niño is associated with a distinctive timescale of 5 yr since 1982.

If, on the other hand, Fig. 1c is appropriate, then the distinctive timescale is explained, but the role of westerly wind bursts is problematic. Each of these extreme cases explains some, but not all of the features that are observed. It is therefore reasonable to explore whether a compromise between the extremes is feasible, whether the appropriate panel in Fig. 1 is somewhere between Figs. 1a and 1b. A growing body of evidence that supports this possibility comes from a variety of modeling studies. Some concern stability analyses of ocean–atmosphere interactions; others force coupled ocean–atmosphere models with the observed “atmospheric noise” to determine under what conditions the spectrum of the observed Southern Oscillation can be reproduced (Neelin et al. 1998; Kirtman and Schopf 1998; Thompson and Battisti 2000, 2001; Fedorov and Philander 2000, 2001; Chang et al. 1996; Eckert and Latif 1997; Blanke et al. 1997; Roulston and Neelin 2000).

These results suggest that a useful analogy for the Southern Oscillation (but one of limited validity) is a slightly damped, swinging pendulum sustained by modest blows at random times. In the absence of noise, El Niño would be perfectly predictable because the Southern Oscillation would be perfectly periodic while its amplitude slowly attenuates. Noise sustains the oscillation and makes it irregular. Although predictability is now relatively insensitive to errors in initial conditions—the instabilities are too weak to amplify those errors—the initial conditions do matter because they describe the phase of the Southern Os-
oscillation and strongly influence the impact of random disturbances. For example, a burst of westerly winds when the oscillation enters its El Niño phase is very different from the impact of the same winds when the oscillation enters its La Niña phase. A westerly wind burst during La Niña can diminish the amplitude of the subsequent El Niño; one that occurs as El Niño is developing can accelerate that development and amplify the event; one after the peak of El Niño will merely prolong its duration (Fedorov 2002).

Thus, predictions for a specific period depend critically on information about the phase of the Southern Oscillation at the start of that period. This means that developments at a certain time depend on two sets of phenomena with very different timescales. The Southern Oscillation, a natural mode of the coupled ocean–atmosphere, has a period of several years. To determine its phase, the available data have to be low-pass filtered. In the next section we introduce phase diagrams, based on the energetics of the Southern Oscillation, that effectively provide the required information. Next it is necessary to turn to the short timescales, of days and weeks, of the random disturbances. (Various aspects of the connection between interannual and intraseasonal variability have been explored in studies such as those of Lau 1986; Lau and Chan 1988; Lau and Shen 1988; Hendon et al. 1998; Vecchi and Harrison 2000; Kessler and Kleeman 2000; Kessler 2001; Zhang et al. 2001; Fedorov 2002; and many others). The short-term response to these disturbances can correspond to nonnormal modes whose structure depends on the initial state and on the structure of the disturbance (Moore and Kleeman 1997, 1999a,b). Here in the section titled “Probabilistic Prediction of El Niño,” we assume that we have information about the statistics of the random atmospheric disturbances. We then make an ensemble of runs that differ, not in initial conditions as in the case of weather forecasts, but in the atmospheric noise that is superimposed as the calculations proceed. Each member of the ensemble has a different realization of the noise. The goal is not to make an actual prediction, but to introduce and demonstrate methods and tools for forecasts.

One of our main results is that, although Landsea and Knaff (2000) and Barnson et al. (1999) assessed the predictions of El Niño of 1997/98 fairly, the real problem with those predictions was the manner in which they were presented. Most were presented as deterministic forecasts that described only one of several possibilities. We argue that the exceptionally large amplitude of El Niño in 1997/98 could not have been anticipated far in advance because it depended on the unusual occurrence of a succession of westerly wind bursts (McPhaden and Yu 1999; Lagerloef et al. 1999; Perigaud and Cassou 2000; Vecchi and Harrison 2000; Villard et al. 2001; Boulanger et al. 2001). To cope with such possibilities, it is best to present probabilistic forecasts. This paper explores the prediction of a specific episode in terms of a probability distribution function, by means of an ensemble of runs.

THE ENERGETICS OF THE OCEAN–ATMOSPHERE INTERACTIONS. Consider an undamped pendulum swinging back and forth in a perfectly periodic, sinusoidal manner. Such a simple system can be described in terms of potential energy $E$ and its time derivative $dP/dt$. In a phase diagram that has as its axes nondimensionalized $dP/dt$ and $P$, the motion of the pendulum corresponds to a perfect circle. In the presence of dissipation, the circle becomes a gradual spiral into the origin. The coupled ocean–atmosphere system is of course far more complex than a simple pendulum. Nonetheless, under certain conditions—that of Fig. 1b—that system is capable of perfectly periodic oscillations analogous, to some degree, to the motion of such a pendulum. The analogy is based on the energetics of the Southern Oscillation, which amount to the approximate equation

$$dE/dt = W,$$  

(1)

which relates $E$, the perturbation available potential energy of the tropical Pacific Ocean, and $W$, the work done on that ocean by the winds, per unit time. 1 Sea surface temperatures in the eastern tropical Pacific are highly anticorrelated with $E$ (which measures the slope of the thermocline) so that negative values of $E$ correspond to El Niño conditions, and positive values to La Niña conditions. [For a more detailed discussion of the energetics, and for definitions of the different variables see Goddard and Philander (2000) and the appendix]. Because the Southern Oscillation, in reality, is irregular there should be departures from Eq. (1), departures that ought to shed light on the processes that cause the irregularities. This means that a study of the energetics can help us resolve the debate about which panel in Fig. 1 corresponds to reality, and can therefore shed light on the predictability of El Niño. It is of interest to note that the latent heat lost by the ocean to the atmosphere, although of criti-
cal importance to the atmosphere, is of secondary importance to the energetics of the ocean. Variations in the kinetic energy of the oceanic currents are similarly of minor importance. If these factors were truly negligible and Eq. (1) were precise then variations in $W$ should lead those in $E$ by a quarter period—on the order of a year.

The departures from the balance in Eq. (1) can be quantified by calculating the correlation between $E$ and $W$. Ideally the calculations would be based on oceanic measurements assimilated into a realistic oceanic GCM, the way atmospheric datasets are generated for analyses of the dynamics and energetics of the atmosphere. Here we use data from a realistic ocean general circulation model similar to the one used by Goddard and Philander (2000). The winds forcing the model are obtained by matching the Comprehensive Ocean–Atmosphere Data Set (COADS) wind data (before 1993) and modified winds from the National Centers for Environmental Prediction (NCEP) re-analyses for the years since 1993. The domain for calculations of $E$ and $W$ is 15°S–15°N and 130°E–80°W. The data obtained from this model indicate that the correlation between $E$ and $W$, in the interannual frequency band, has a maximum value of 0.8 for $E$ lagging $W$ by about 8 months. This lag is clearly evident in Fig. 2 and implies that, by calculating and monitoring wind power, El Niño can be anticipated about 8 months in advance.

The imperfect correlation between $W$ and $E$ is one measure of the degree to which the energetics of the observed Southern Oscillation depart from Eq. (1). Another measure is available in Fig. 3, which shows variations in $E$ and $W$ relative to their time-averaged values for the period 1980–2000. Striking features of Fig. 3 are the spirals, all consistently anticlockwise. If we start from the years of an intense El Niño, that of 1982 or 1997 say, then the spirals are convergent, implying a damped oscillation. The decay time of the Southern Oscillation is estimated to be on the order of 10 yr. On a few striking occasions, the spirals in Fig. 3 uncoil—leading up to the very intense El Niño episodes of 1982 and 1997, for example. Westerly wind bursts—the red portions of the curves indicate their timing—are seen to contribute significantly to those developments. On other occasions the bursts have little or no effect. [Note that westerly wind bursts that lead to ocean–atmosphere interactions and hence contribute to the development of El Niño are difficult to separate from the wind relaxation associated with El Niño, see Fedorov (2002)]. The causes of westerly wind bursts, and their relation to the Madden–Julian Oscillation (e.g., Madden and Julian 1994; Slingo et al. 1999; Hendon et al. 1999; Vecchi and Harrison 2000; Kessler 2001) or to northerly surges from the extratropics (e.g., Yu and Reinecker 1998; Yu et al. 2002) are still being debated.

A striking feature of Fig. 3 is how orderly the trajectories are. If we were dealing with a high unstable phenomenon, similar to weather, then we could expect chaotic trajectories. At the other extreme, if we were dealing with a highly damped system, then the trajectory would spend most of its time at the origin and would venture away briefly during El Niño, returning promptly to the origin. Thus, Fig. 3 indicates that the appropriate parameter range for the observed Southern Oscillation is neither Fig. 1a nor Fig. 1c. It appears that the oscillation is slightly damped and is sustained by modest disturbances (i.e., it is between Figs. 1a and 1b).

Given how important westerly wind bursts and other atmospheric disturbances are to the development of El Niño, how predictable is that phenomenon? The phase diagrams of Fig. 3, divided into four quadrants (I, II, III, IV), shed light on this question. The trajectories in Fig. 3 are always counterclockwise, and normally pass through all four quadrants to complete a cycle. El Niño conditions, because they correspond to the lower parts of quadrants I and II, are least predictable.
likely when the oceanic state is in quadrant III; at such times we can confidently state that El Niño will not occur. Conditions are most uncertain, and the development of El Niño is most dependent on westerly wind bursts, in quadrant IV. (Some coupled general circulation models that simulate a Southern Oscillation predict a new El Niño event every time the system is in quadrant IV.) Since a complete cycle takes a fixed time, the longer the system stays in quadrant IV, the more probable El Niño becomes.

When the system spends considerable time close to the origin in Fig. 3—that was the case in the early 1990s—then the predictability of El Niño (or of any future developments) is very low. On the other hand, when the trajectory is far from the origin in quadrant I, then El Niño conditions are certain and after a while inevitably will give way to La Niña conditions as the trajectory moves through quadrants II and III. A public alert of a developing El Niño was sounded (in the press and on television) in June 1997. At that time the trajectory was deep in quadrant I, so that developments over the next several months could be anticipated with reasonable certainty.

The impact of a wind burst on El Niño depends not only on its timing and duration, but also on its intensity and spatial structure, which may cause subsequent developments to correspond to the evolution of a “nonnormal mode” (associated with a fast transient growth or singular vectors) rather than the gravest normal mode (e.g., Penland and Sardeshmukh 1995; Penland 1996; Xue et al. 1997a,b; Moore and Kleeman 1997, 1999a,b). It is plausible that the eastward propagation of thermocline and wind disturbances during El Niño of 1982 and 1997 can be interpreted as the rapid nonnormal development over the brief period of approximately a year, in contrast to the much longer period of 5 yr for a complete oscillation (the normal mode). This is consistent with the results of Thompson and Battisti (2000, 2001) who use a linearized version of the Cane–Zebiak model with stochastic forcing and argue that the greatest part of the variability in the model comes from perturbations that project onto the first singular vector, and then grow rapidly.

In Fig. 3, the seasonal cycle appears as the movements back and forth between quadrants III and IV. [The seasonal cycle involves only modest movements

---

**Fig. 3.** (a) The E–W phase diagrams for the period 1979–98. Alternate years are shown in solid and dashed lines, respectively. The red portions of the plot indicate the occurrence of westerly wind bursts. (The intensities and spatial structure of bursts vary considerably.) Here I, II, III, IV indicate different quadrants. (b) The E–W phase diagram for the period 1997–2000. Alternate years are shown in solid and dashed lines, respectively. The phase trajectory starts in Apr of 1997 and ends in Jan 2001. The red-marked portions of the plot indicate strong activity of the westerly wind bursts. Here I, II, III, IV indicate different quadrants. The small cycles on the upper half-plane are associated with the anomalously strong seasonal fluctuations that may lead to modest changes in potential energy, but larger changes in wind work.
of the thermocline, in comparison with El Niño, so that \( E \) does not change significantly. The interaction between the annual cycle and the Southern Oscillation appears to be another cause of El Niño irregularity, see Chang et al. (1994); Tziperman et al. (1995); Jin et al. (1996). The seasonal cycle is prominent in quadrants III and IV, which we associate with La Niña, because the shallow thermocline during La Niña favors a large amplitude for the seasonal cycle (Gu and Philander 1995). El Niño starts to develop in quadrant IV and therefore tends to do so only at a certain phase of the annual cycle, giving the impression of phase locking to the annual cycle. For the purpose of predicting El Niño it would be useful to filter out the seasonal cycle completely. This, however, is problematic because of the skewness of the Southern Oscillation. For example, the warm conditions of El Niño persisted for about a year in 1982 and in 1997, and were then followed by almost 4 yr of cold La Niña conditions before El Niño returned. (The nonlinear mechanisms that give rise to this skewness have not yet been identified.) If we were to minimize the seasonal cycle by averaging over a year, then we lose the ability to make short-term forecasts. During 2001 conditions in the tropical Pacific moved from quadrant IV toward III and back toward IV as part of the seasonal cycle. The stage was set for El Niño to develop during 2002, but predicting the intensity of the event is difficult because it is not known whether or not westerly wind bursts will matter.

**PROBABILISTIC PREDICTION OF EL NIÑO.**

If the Southern Oscillation is indeed slightly damped and is sustained by random disturbances then predictions ought to be probabilistic, on the basis of an ensemble of forecasts. Such an approach has been explored by means of coupled general circulation models. The members of the ensemble differ in the initial conditions from which the forecasts start. They also differ from each other in another respect: the models generate their own atmospheric noise, which changes from one ensemble member to the next. At present the results are of very limited usefulness because none of the models reproduces a realistic Southern Oscillation (Latif et al. 2001), and because the atmospheric noise generated by these models is not very realistic.

The difficulties with the coupled GCMs just mentioned can be bypassed by using simpler coupled models that are tuned to be in the appropriate parameter range, and by adding to their noise-free atmospheric components the observed noise. This can be done by subtracting, from a standard wind dataset—that from The Florida State University (Stricherz et al. 1997), for example—the deterministic part of the wind field that is correlated with sea surface temperature changes in the tropical Pacific. A shortcoming of this approach is that it ignores that a certain part of the random wind fluctuations merits special attention. We are referring to the westerly wind bursts that sporadically appear near the date line in the equatorial Pacific and that are capable of having an exceptionally strong influence on subsequent developments. That may have happened in 1997, for example. To take the special importance of these winds into account we took zonal wind data for the period 1980–2000 for the location around 160°W on the equator and generated atmospheric “noise,” with similar statistics, in the following manner. The spatial structure of the winds is assumed to be a Gaussian shape, centered on the equator, with \( e \)-folding half-widths of 20° longitude and 7° latitude. The winds are multiplied by a time-dependent factor that yields an evolving red noise process with a mean of zero, a standard deviation of 2 m s\(^{-1}\), and a decorrelation time of 1 week. (This noise serves for conceptual purposes only. For actual forecasts, the statistics should include seasonal dependence, the dependence on the phase of the Southern Oscillation, etc.)

Given different realizations of the atmospheric noise, we then make an ensemble of predictions with a coupled ocean–atmosphere model, with all the calculations starting from the same initial conditions, each with its own superimposed atmospheric noise. [Eckert and Latif (1997) use a similar approach to study stochastic properties of a hybrid coupled model of the tropical ocean-atmosphere, while Moore and Kleeman (1998) investigate the relation between the skill of an ensemble of ENSO forecasts and the ensemble spread in an intermediate coupled model.] A forecast of the type shown in Fig. 4 can then be prepared on the basis of such an ensemble of forecasts—200 for this example (for details see Wittenberg 2002). The coupled model used to obtain the results of Fig. 4 is a modified version of the one developed by Zebiak and Cane (1987), configured in such a manner that the spontaneous Southern Oscillation is slightly damped and has a period of 4 yr. The model has been simplified by omitting the effects of the seasonal cycle, meridional asymmetry, and nonlinearities associated with atmospheric convection. This model has some unrealistic features but, for the purpose of this paper—exploring the type of forecasts appropriate for El Niño—the model has the enormous advantage of requiring very modest computer resources, thus permitting a large ensemble of predictions.
The predictions shown in Fig. 4 all start at the transition from La Niña to El Niño, when the temperature anomaly in the western equatorial Pacific is close to zero. In Fig. 4, the dashed line shows how sea surface temperature anomalies in the eastern equatorial Pacific—departures from the expected seasonal value—would evolve in the absence of the noise. The heavy solid line in Fig. 4 is one example of how noise can affect that evolution. The light contours show the fraction of the forecasts that indicate temperatures colder than a certain value. The colors in the figure, which in effect show how the forecasts are distributed, indicate how the probability distribution function for the temperature anomalies changes (widens) with time. Initially this function has a sharp peak (the red region) but it becomes relatively flat after approximately 18 months. That, presumably, is the limit of predictability for the chosen initial conditions, and for the chosen noise. Note that, at the peak of a strong El Niño event, predictability may be higher than indicated here; it will be lower when the values of $E$ and $W$ are close to zero.

CONCLUSIONS. Thus far, attempts to forecast El Niño have not been very successful (Landsea and Knaff 2000; Barnston et al. 1999). However, the factors that cause the irregularity of the Southern Oscillation—random atmospheric disturbances whose influence depends on the phase of the oscillation—are such that the predictability of specific El Niño events is inevitably limited. That is especially true of the intensity of El Niño. For example, the occurrence of an event in 1997 was predictable on the basis of information about the phase of the Southern Oscillation, but the amplitude of the event could not have been anticipated because it depended on the appearance of several wind bursts in rapid succession. (This is a rare event that happened in 1982 and again in 1997.) Given that deterministic predictions are bound to be unreliable, it is preferable to make probabilistic forecasts on the basis of an ensemble of forecasts. The format of Fig. 4 is one way in which the results can be presented.

ACKNOWLEDGMENTS. We thank P. Chang, M. McPhaden, C. Landsea, S. Zebiak, and two anonymous reviewers for valuable comments. This research was supported by NOAA (under Contract NA16GP2246) and NASA (under JPL Contract 1229832).

APPENDIX: CALCULATION OF $E$ AND $W$.

The values of $E$ and $W$ can be calculated (Goddard and Philander 2000) as

$$E = -\frac{\rho}{2} \iiint (\partial_2^2 - \bar{\rho}^2) \hat{p}_z dxdydz, \quad (A1)$$

and

$$W = \iint (u \tau - \bar{u} \tau) dxdy, \quad (A2)$$

where $\rho = \rho(x,y,z,t)$ is the density field minus the reference density $\bar{\rho}(z)$ that corresponds to the static stability, $\hat{p}_z$ is the vertical gradient of the reference density, $u$ and $\tau$ are the zonal near-surface current and
the zonal wind stress, respectively. The bars indicate mean (climatological) values. The domain for integration in x and y is 15°S–15°N and 130°E–80°W. The depth range for integration in z in (Al) is 30–400 m. Note that as perturbation variables, E and W can be both positive and negative.

When linearized, Eqs. (A1) and (A2) reduce to

$$ E = -g \int \int \int \frac{\partial \rho'}{\partial z} dx dy dz, $$  \hspace{1cm} (A3)

and

$$ W = \int \int (\bar{\rho} \tau' + u' \tau) dx dy, $$  \hspace{1cm} (A4)

where the primes indicate perturbation values. Expressions (A3) and (A4) were used in calculating E and W for Figs. 2–3.

REFERENCES


Meteorology today is the beneficiary of the fundamental work in weather analysis and forecasting of Sverre Petterssen (1898–1974), a giant in the field and an international leader in meteorology during its formative era. In this lively and insightful autobiographical memoir, written just before his death, Petterssen shares intimate memories from his childhood in Norway, his education and service with the famous Bergen school of meteorology, and his extensive experiences in polar forecasting and as head of the meteorology department at MIT. The crisis of World War II comes alive in his passionate recounting of how forecasts were made for bombing raids and special operations, including the contentious forecasts for D-Day.

Sverre Petterssen's complete autobiographical memoir, published here for the first time in English, offers a fascinating view of a man, an era, and a science. Anyone interested in weather, World War II, the history of science, or Norwegian history will enjoy this book.

Weathering the Storm: Sverre Petterssen, the D-Day Forecast, and the Rise of Modern Meteorology, ISBN 1-878220-33-0, 326 pp., hardbound, $70 list/$50 member. To place an order, submit your prepaid orders to: Order Department, AMS, 45 Beacon Street, Boston, MA 02108-3693; call 617-227-2425 to order by phone using Visa, Mastercard, or American Express; or send e-mail to amsorder@ametsoc.org. Please make checks payable to the American Meteorological Society.