551.513.7:551.465.7

Prediction of El Niño events using a physical model

By MARK A. CANE and STEPHEN E. ZEBIAK
Lamont-Doherty Geological Observatory of Columbia University, Palisades, NY 10964

SUMMARY

A theory for the genesis of El Niño events is developed. ENSO (El Niño/Southern Oscillation) is seen as an oscillation of the coupled ocean–atmosphere system, but only interactions taking place in the tropical Pacific are vital to its existence. As envisioned by Bjerknes, positive feedbacks between atmosphere and ocean along the equator maintain both the warm El Niño phase and the cold, non-El Niño phase of the ENSO cycle. It is hypothesized that changes in the zonal mean heat content of the equatorial ocean are essential for the transition from one phase to another. These ideas are illustrated with a simple, numerical model of the Pacific ocean–atmosphere system. Implications for prediction are discussed.

Glendower: I can summon spirits from the vasty deep.
Hotspur: Why so can I or any man; but will they come when you do call for them?

King Henry IV – Part I
Act III, Scene 1

1. INTRODUCTION

There are many sorts of prediction schemes and many ways of classifying them. This is true for picking horses, forecasting the weather, or predictions of climatic variations such as El Niño. Subjective schemes may be little better than hunches but the term 'subjective' need not be perjorative. Such methods may involve the ingenious determination of critical indices, or an intelligent use of analogues, or the sophisticated skills of an experienced weather forecaster.

Some objective procedures differ little in outline from corresponding subjective ones. Statistical procedures often consist of using key indices as predictors. The choice of indices may rest on a more or less subjective assessment of the important factors in the problem at hand, but such procedures can be quite sophisticated. In contrast to their subjective counterparts they typically return quantitative estimates of probabilities and amplitudes.

We are aware of three rather different objective schemes which have lately been used to predict El Niño and the Southern Oscillation (ENSO). Pazan and White (1986) identified a pattern of oceanic heat content anomaly in the western tropical Pacific which appears to be a precursor to El Niño events. White *et al.* (1987) use complex empirical orthogonal function analysis to identify the crucial pattern in the output of a linear shallow-water model driven by observed winds. (The model is that of Busalacchi and O'Brien (1981); the wind fields are processed at Florida State University using the procedure described in Goldenberg and O'Brien (1981).)

Inoue and O'Brien (1984) report a statistical scheme where the predictor is determined by running a dynamical ocean model forced by realistic winds. (The model and wind fields are those mentioned above.) The observed monthly mean winds are used up to the forecast time; the last month's winds are held constant for two additional months. The predictor used is the model thermocline depth anomaly averaged over the eastern equatorial Pacific. If its value exceeds a predetermined threshold, which varies with the month of the year, then an El Niño is considered imminent.

Barnett (1984) has developed a purely statistical technique which uses area-averaged wind anomalies in the western equatorial Pacific as predictors of eastern Pacific sea surface temperature (s.s.t.) anomalies. This procedure has significant skill at a lead time of a few months; recently, Barnett (private communication) has developed a method based on sea level pressure anomaly patterns which appears to have skill at longer lead times.

At the present time (August 1986), all the methods discussed above are predicting a warm event—an El Niño—for the end of the calendar year 1986. The Barnett (1984) and the Inoue and O'Brien (1984) procedures are conservative in that they do not predict an event unless it is highly likely. Their lead time is short; the former first predicted an event after July, the latter after May. The reliability of the White *et al.* (1987) procedure is not so well established, but this scheme would have allowed a prediction much further ahead—by the end of 1985.

The final entry among objective forecasting methods is the numerical integration of equations describing the dynamics and thermodynamics of the relevant physical system. This has become the dominant approach to weather prediction. In the case of El Niño it is the coupled ocean–atmosphere system for (at least) the tropical Pacific region which must be simulated. Though El Niño may be defined in purely oceanographic terms, it is not possible to predict for the ocean alone. The ocean circulation is *forced* by the atmosphere, and its future state will depend on the future state of the atmosphere. (This is especially true of the tropical oceans, but with the exception of some mesoscale features which are predominantly a consequence of internal ocean dynamics, it applies to virtually all the interesting aspects of the global ocean.) To predict El Niño one must predict both atmosphere and ocean.

This paper reports on efforts to predict El Niño by numerical integration of a physical model of the coupled ocean–atmosphere system. As far as we know, at present it is the only El Niño prediction scheme of this type. We do not wish to claim that the current version of our dynamical method gives better results than the statistically based methods outlined above. However, the example of numerical weather prediction encourages the expectation that predictions based on a physical model will ultimately prove superior. *Inter alia*, this approach creates an experimental apparatus for studying the physics of the ENSO cycle.

The first scheme developed for numerical weather prediction is the justly celebrated work of Richardson (1922). The forecasting model he describes is comparable in complexity to the atmospheric general circulation models (GCMs) in use today for the study of climate. The lack of sufficient computing power at the time would have been enough to defeat any attempt to make his scheme operational. It is well to keep in mind, however, that even if modern computers had been available, his brute force, GCM-like approach would have floundered because of an inadequate theoretical understanding of relevant atmospheric dynamics (Phillips 1970; Cane 1986b).

The first successful numerical weather forecasts, those of Charney *et al.* (1950), followed upon the major advances in dynamical meteorology associated with the development of quasi-geostrophic theory by Charney (1948) and others. This calculation was recognized as a landmark (by Richardson among others; see Ashford (1985), p. 243), though the results were not so striking as to make it a great popular success. (This was Richardson's opinion; *ibid.*) Charney *et al.* avoided the problems which plagued Richardson by resorting to a highly simplified description of atmospheric dynamics, the equivalent barotropic model. Several decades passed before it was possible to make successful numerical weather forecasts with a model as complex as Richardson's.

We regard our ENSO model as being at about the level of simplification of the equivalent barotropic model. A goal of the remainder of this paper is to persuade the reader that ENSO forecasting by a physical model has progressed to at least the late

1950s by the standards of numerical weather prediction. One would expect that the general progress in oceanography and meteorology in the past decades will provide a basis for rapid advances in ENSO prediction.

The plan of the remainder of the paper is as follows. The coupled ocean–atmosphere model is briefly described together with a few results. Our ideas on the mechanism of the ENSO cycle are then presented followed by a discussion of their implications for the forecasting of El Niño. We then describe the results of experimental forecasts of past years as well as the model forecast for 1986. We conclude with some speculations on the predictability of ENSO and the prospects for improvement in forecasting procedure.

2. Model description

The components of our coupled model were developed and tested independently. The atmospheric component, described by Zebiak (1986), was shown to reproduce the major features of the equatorial wind anomaly field when forced by observed ENSO s.s.t. anomalies (also see Weare 1986). The oceanic component, described in Zebiak and Cane (1987a; also see Zebiak 1984; Cane 1986a) was shown to simulate the overall evolution of s.s.t. anomalies during ENSO, when forced by observed tropical wind anomalies. The coupled model differs from others (see the review by McCreary (1985)) primarily in its treatment of the thermodynamics in the atmosphere and ocean, especially through the inclusion of a moisture feedback process in the atmosphere and a simple, but thermodynamically active, surface layer in the ocean.

Only a summary account of the model components will be given here. Both describe perturbations about the mean climatological state, with the climatology specified from observations. The Climate Analysis Center data set (see Rasmusson and Carpenter 1982) was used for this purpose. For completeness the full governing equations are given in an appendix.

(*a*) *Atmosphere.* If s.s.t. anomalies characteristic of El Niño are given, then the principal changes in the tropical circulation may be calculated. This has been amply demonstrated by simulations with atmospheric GCMs (e.g. Shukla and Wallace 1983; Lau 1985). Observations show that the tropical anomalies have a simple vertical structure with a universal form, namely, a reversal of polarity between the lower and upper troposphere (e.g. regions of low-level convergence lie below regions of upper-level divergence). Linear dynamical models with a single degree of freedom in the vertical have proved remarkably adept at reproducing the horizontal structure of the atmosphere (Matsuno 1966; Gill 1980) though the physical interpretation of these models is uncertain (Geisler and Stevens 1982; Zebiak 1982). Our model dynamics is of this type: i.e. steady-state, linear shallow-water equations on an equatorial beta plane. Linear dissipation in the form of Rayleigh friction and Newtonian cooling is used.

The circulation is forced by a heating anomaly distribution which depends partly on local evaporation anomalies (parametrized in terms of local s.s.t. anomalies), and partly in the low-level moisture convergence (parametrized in terms of the surface wind convergence). Several observational studies (e.g. Cornejo-Garrido and Stone 1977; Ramage 1977) as well as GCM calculations have demonstrated the important contribution of moisture convergence to the overall tropical heat balance.

The convergence feedback is incorporated into the model using an iterative procedure in which the heating at each iteration depends on the convergence field from the previous iteration. The scheme is analysed in detail in Zebiak (1986). The feedback is nonlinear because the moisture-related heating is operative only when the total wind field is convergent, and this depends not only on the calculated anomalous convergence, but also the specified *mean* convergence (see appendix). The feedback focuses the atmospheric response to s.s.t. anomalies into or near the regions of mean convergence; in particular, the Intertropical Convergence Zone (ITCZ) and the South Pacific Convergence Zone (SPCZ). Such a focusing is conspicuous in the observed wind anomalies during ENSO (see Rasmusson and Carpenter 1982).

(*b*) *Ocean*. The model ocean basin is rectangular, and extends from 124°E to 80°W, and from 29°N to 29°S. The dynamics of the model begin with the linear reduced-gravity model that is so successful in simulating thermocline depth anomalies and sea level changes during El Niño events (Cane (1986a) reviews this work). Such models produce only depth-averaged baroclinic currents, but the ocean surface current is usually dominated by the frictional (Ekman) component. Therefore, a shallow frictional layer of constant depth is added to simulate the surface intensification of wind-driven currents. The dynamics of this layer is also kept linear, but only by using Rayleigh friction to stand in for nonlinear influences at the equator (see Zebiak and Cane (1987b) for a discussion). Upwelling velocity is computed as the divergence of the surface layer transport. Inclusion of a surface layer allows a strong response to local winds; models which omit it will understate upwelling effects. Mean currents are generated by spinning up the model with monthly mean climatological winds. These 'climatological' currents are then used in the anomaly calculation.

The thermodynamics describe the evolution of temperature *anomalies* in the model surface layer. The governing equation is complete, including three-dimensional temperature advection by both the specified mean currents and the calculated anomalous currents. Although there may be local exceptions, the preponderant evidence is that surface heating does not contribute to the El Niño warming (Bjerknes 1969, 1972; Ramage and Hori 1981; Weare 1983). Instead the data indicate an inverse relation between s.s.t. and heat flux into the ocean because of increased evaporation. In the model, surface heat flux anomaly is taken to be proportional to the local s.s.t. anomaly, acting always to adjust the temperature field toward the climatological mean state. This (monthly) mean temperature structure is specified from observations. The thermodynamic equation thus has the following form (where barred quantities represent mean fields and unbarred quantities represent anomalies):

$$\frac{\partial T}{\partial t} = -\bar{\mathbf{u}}_1 \cdot \nabla T - \mathbf{u}_1 \cdot \nabla(\bar{T} + T) - \{M(\bar{w}_s + w_s) - M(\bar{w}_s)\}\bar{T}_z - M(\bar{w}_s + w_s)T_z - \alpha_s T \tag{1}$$

where $\mathbf{u}_1$ and w_s represent horizontal surface currents and upwelling, respectively, and the function $M(x)$ is defined by

$$M(x) = \begin{cases} 0, & x \leqslant 0 \\ x, & x > 0. \end{cases}$$

This function accounts for the fact that surface temperature is affected by vertical advection only in the presence of upwelling. The anomalous temperature gradient, T_z, is defined by

$$T_z = (T - T_e)/H_1, \tag{2}$$

where H_1 is the surface layer depth, and the T_e measures temperature anomalies entrained into the surface layer. T_e is a linear combination of T and T_d $(h, \bar{h})$, a function expressing the subsurface temperature anomaly as a function of the mean ($\bar{h}$) and anomalous (h) thermocline depths. The precise forms are given in the appendix, Eqs. (A12) and (A13). The variable h is obtained from the model dynamics. Details are provided in Zebiak and Cane (1987b), where it is also shown that the full complexity of (1) is required to give realistic results.

(*c*) *Coupling*. In the component models, the ocean affects the atmosphere exclusively through the s.s.t. field, and the atmosphere affects the ocean through the surface wind stress alone. Though it was reasonable to consider a steady-state atmosphere when prescribing seasonal mean s.s.t. forcing (as in Zebiak (1986)) this must be reconsidered when coupling to a time-dependent ocean in which the s.s.t. field can change on the timescale of a few days. For boundary forcing variations on this timescale, the atmospheric response due to moisture convergence feedback cannot reasonably be assumed to be in continuous equilibrium, since the transport time between regions of moisture input and corresponding latent heat release during ENSO can easily approach one month. This is an important consideration for the coupled model. If the atmospheric response is assumed to be in continuous equilibrium, then the *change* in the wind field between successive s.s.t. time steps will be overestimated. As a result, more rapid changes in s.s.t. will occur because of local wind effects. This induces yet larger changes in the atmosphere, and the combined interaction favours an artificially rapid development of anomalies, particularly at small scales where the atmospheric convergence feedback is most efficient.

To circumvent this, a procedure was adopted which effectively gives the steady-state response as before for timescales of one month or more, but restricts the feedback for shorter timescales. This is accomplished by altering the criterion governing the number of convergence feedback iterations that are performed (see Zebiak 1984). Other methods could be used to produce a similar result. For example a spatial smoother could be applied after each iteration, or time dependence could be added explicitly to the atmosphere model. The present method was chosen because it requires less computation.

3. Coupled model results

A numerical experiment with the coupled model was initiated with an imposed $2\,\mathrm{m\,s^{-1}}$ westerly wind anomaly of four-months duration beginning in December of the year designated -1. There was no external forcing thereafter: aside from the model physics, evolution of anomalies in s.s.t., winds, etc. depends only on this initial condition and on the monthly mean climatological fields specified in the component ocean and atmosphere models. Furthermore, because of the damping in the model the initial conditions are largely forgotten within a decade.

A 90-year time series of model s.s.t. anomaly averaged over the eastern equatorial Pacific is shown in Fig. 1. There are peaks of varying amplitude occurring at irregular intervals but typically three to four years apart. They tend to be phase-locked to the annual cycle, with major events reaching maximum amplitude at the end of the calendar year and decaying rapidly thereafter. All these features are characteristic of observed El Niño events (e.g. Cane 1983). The amplitude of model events is similar to that of observed ones, though the model did not produce anything as extraordinary as the 1982/83 event. The model is somewhat more regular than nature; the high frequency fluctuations present only in the real atmosphere and ocean may account for the broader natural spectrum.

Figure 2 depicts the evolution of s.s.t. during the El Niño event of model year 31. In December of the preceding year there was no discernible anomaly; by March of year 31 there is a small but systematic warming in the eastern Pacific; by December the

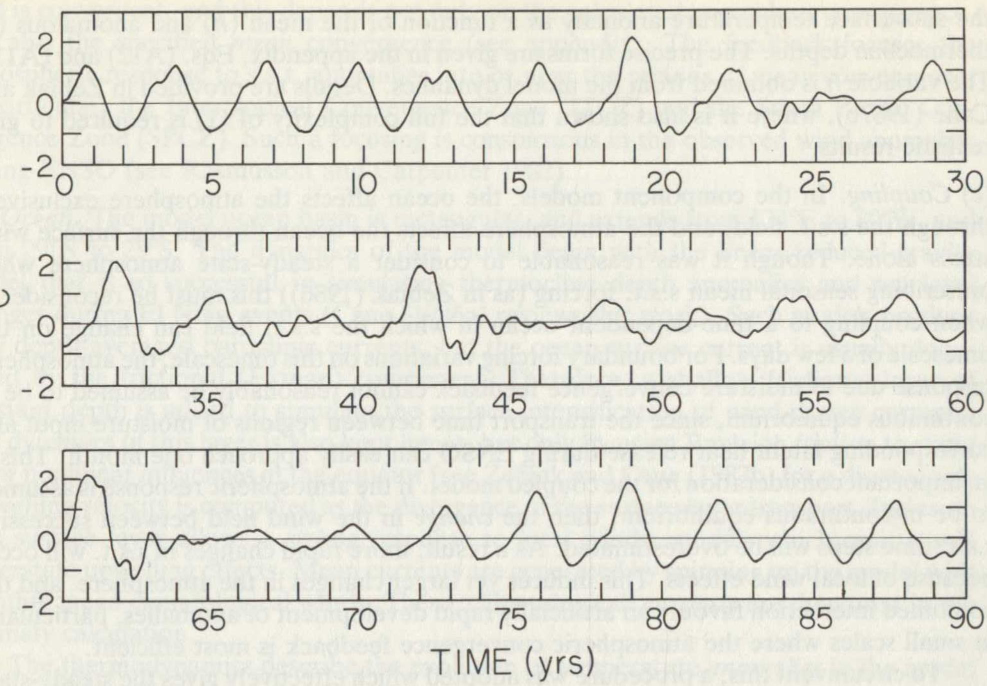


Figure 1. Sea surface temperature anomalies averaged over the eastern equatorial Pacific region NINO3 (5°N–5°S, 90°W–150°W) for 90 years of coupled model integration. (From Cane and Zebiak 1985).

anomaly extends to the dateline, with a maximum at about 135°W. The model patterns and amplitudes are fairly realistic except near the South American coast, where the model's coarse resolution precludes an accurate simulation of coastal upwelling processes.

Figure 3 shows the evolution of zonal wind along the equator. The most prominent feature is the band of westerly anomalies in the central Pacific, which bears a resemblance to the typical ENSO anomaly (Cane 1983). However, the model omits the commonly observed initial anomalies in the vicinity of the dateline and so fails to reproduce the observed eastward progression of the westerly anomaly patch. The ocean model generates

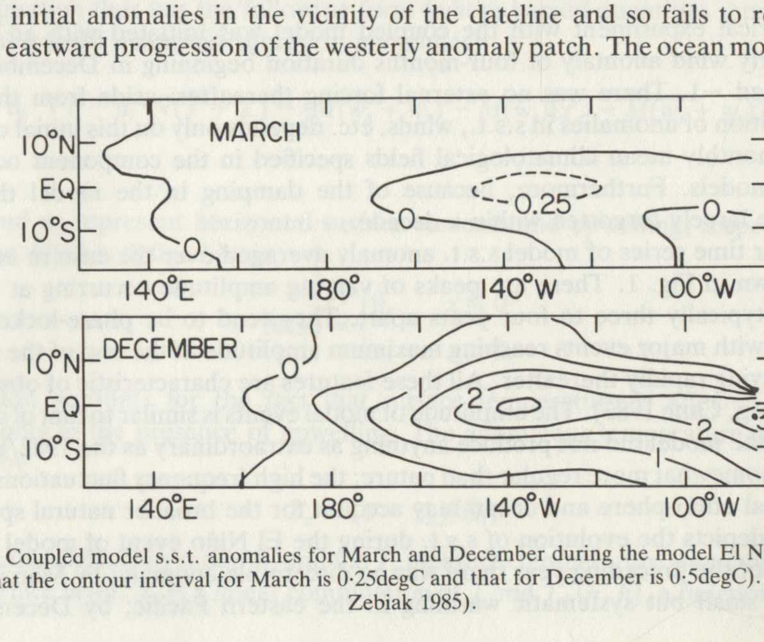


Figure 2. Coupled model s.s.t. anomalies for March and December during the model El Nino event in year 31 (note that the contour interval for March is 0·25degC and that for December is 0·5degC). (From Cane and Zebiak 1985).

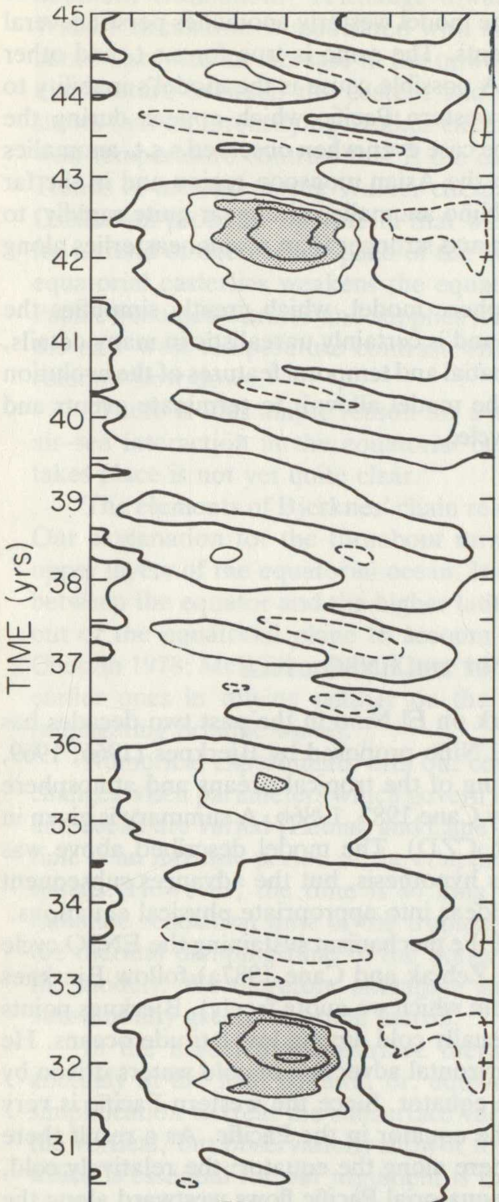


Figure 3. Time–longitude sections for years 30–45 of the coupled model integration showing the forcing for the gravest mode oceanic Kelvin wave, a measure of zonal wind anomalies along the equator. Positive (westerly) anomalies are indicated with solid lines, and negative (easterly) anomalies are indicated with dashed lines. Large westerly anomalies (0·15 dyn cm^{-2}) are stippled.

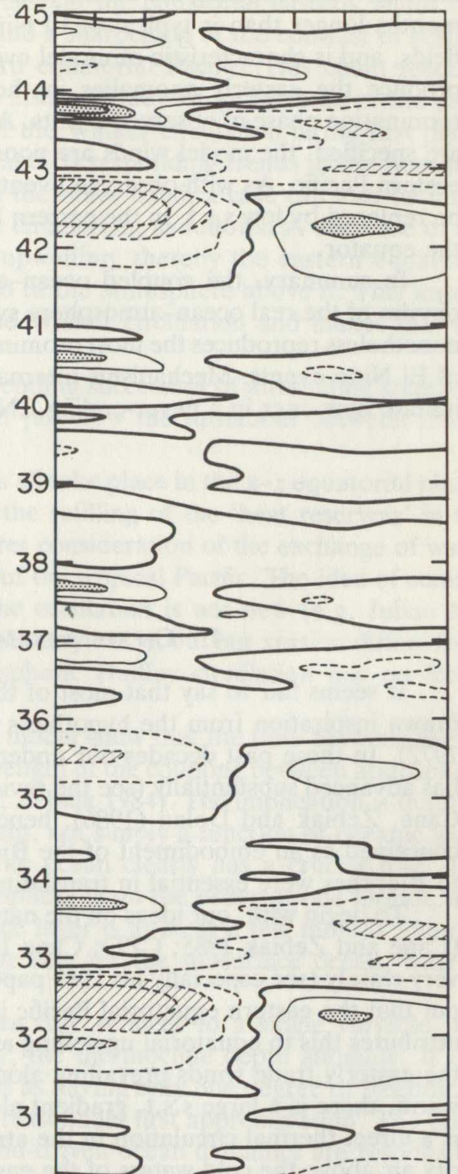


Figure 4. Model thermocline depth anomaly at the equator, for years 30–45 of the 90-year integration. Positive anomalies are indicated with solid lines, and negative anomalies are indicated with dashed lines. The contour interval is 10 m. Anomalies greater than +20 m are stippled; anomalies greater than −20 m are hatched.

too little variability in that region, so the model atmosphere has no s.s.t. anomaly to respond to. After this initial phase the spatial and temporal patterns are generally realistic until the year following the event, year 32. The model westerly anomalies persist several months longer than is typical of El Niño events. The same is true for s.s.t. and other fields, and is characteristic of model events. A possible cause is the model's inability to produce the easterly anomalies in the far western Pacific which appear during the termination phase of observed events. As is the case even when observed s.s.t. anomalies are specified, the model winds are poorest in the Asian monsoon region and in the far eastern Pacific. As with observed events, El Niño anomalies disappear quite rapidly, to be replaced by low s.s.t. in the eastern Pacific and stronger than normal easterlies along the equator.

In summary, the coupled ocean–atmosphere model, which greatly simplifies the physics of the real ocean–atmosphere system and is certainly unrealistic in many details, nonetheless reproduces the most prominent spatial and temporal features of the evolution of El Niño events. Mechanisms internal to the model allow it to terminate events and initiate new ones in a never-ending ENSO cycle.

4. On the mechanism of the ENSO cycle

It seems fair to say that most of the work on El Niño in the past two decades has drawn inspiration from the hypothesis for El Niño proposed by Bjerknes (1966, 1969, 1972). In these past decades our understanding of the tropical oceans and atmosphere has advanced substantially (see the reviews by Cane 1983, 1986b. A summary is given in Cane, Zebiak and Dolan (1986), henceforth CZD). The model described above was conceived as an embodiment of the Bjerknes hypothesis, but the advances subsequent to Bjerknes were essential in translating his ideas into appropriate physical equations.

To begin with, our ideas on the nature of the mechanism sustaining the ENSO cycle (Cane and Zebiak 1985; CZD; Cane 1986b; Zebiak and Cane 1987a) follow Bjerknes very closely (see especially his 1969 paper, from which we quote freely). Bjerknes points out that the eastern equatorial Pacific is unusually cold among low-latitude oceans. He attributes this to equatorial upwelling and horizontal advection of cold waters driven by the easterly trade winds prevailing along the equator. Since the western Pacific is very warm, there is a large s.s.t. gradient along the equator in the Pacific. As a result there is a direct thermal circulation in the atmosphere along the equator: the relatively cold, dry air above the cold waters of the eastern equatorial Pacific flows westward along the surface toward the warm west Pacific. "There, after having been heated and supplied with moisture from the warm waters, the equatorial air can take part in large-scale, moist-adiabatic ascent." Some of the ascending air joins the poleward flow at upper levels associated with the Hadley circulation, and some returns to the east to sink over the eastern Pacific. There is a zonal surface pressure gradient associated with this equatorial circulation cell, high in the east and low in the west.

Bjerknes named this the 'Walker circulation', because he felt that fluctuations in this circulation initiated pulses in Walker's Southern Oscillation. It can have such global consequences because "it operates a large tapping of potential energy by combining the large-scale rise of moist air and descent of colder dry air."

The Walker circulation is the link between eastern Pacific s.s.t. anomalies and the Southern Oscillation: "A change toward a steeper pressure slope at the base of the Walker circulation is associated with an increase in the equatorial easterly winds and hence also with an increase in the upwelling and a sharpening of the contrast of surface temperature between the eastern and western equatorial Pacific. This chain reaction shows that an intensifying Walker circulation also provides for an increase of the east–west temperature contrast that is the cause of the Walker circulation in the first place. Trends of increase in the Walker circulation and corresponding trends in the Southern Oscillation probably operate in that way. On the other hand, a case can also be made for a trend of decreasing speed of the Walker circulation, as follows. A decrease of the equatorial easterlies weakens the equatorial upwelling, thereby the eastern equatorial Pacific becomes warmer and supplies heat also to the atmosphere above it. This lessens the east–west temperature contrast within the Walker circulation and makes that circulation slow down."

"There is thus ample reason for a never-ending succession of alternating trends by air–sea interaction in the equatorial belt, but just how the turnabout between trends takes place is not yet quite clear."

The elements of Bjerknes' chain reactions all take place in the x–z equatorial plane. Our explanation for the turnabout turns on the refilling of the 'heat reservoir' in the upper layers of the equatorial ocean. It requires consideration of the exchange of water between the equator and the higher latitudes of the tropical Pacific. The idea of coming out of the equatorial plane to account for the oscillation is not new (e.g. Julian and Chervin 1978; McWilliams and Gent 1978; McCreary 1983), but our version differs from earlier ones in relying neither on the atmospheric Hadley circulation nor on freely propagating oceanic waves.

Numerical experiments with our coupled model show that the time between events changes when parameters which govern the strength of the coupling between atmosphere and ocean are varied (Zebiak and Cane 1987a; Zebiak 1984). The implication is that this time is an intrinsic period of the *coupled* system, not simply a function of oceanic wave speeds. However, the time is so long that the ocean clearly has a crucial role. The radiative relaxation time of the tropical atmosphere is on the order of one month, and the thermal damping time of the ocean surface layer is at most a few months (four in the model). We therefore anticipate an active role for the subsurface ocean, with its substantially greater thermal inertia.

In the model the subsurface thermal structure reduces to a single variable, the anomaly in the heat content, or, equivalently, the thermocline depth anomaly. It is a simplification to treat the subsurface variations as having only one degree of freedom in the vertical, but observations show it to be a reasonable first approximation. The point which is essential for our argument is that wind-driven ocean dynamics are responsible for the subsurface changes. The subsurface ocean affects the atmosphere, and hence the feedback process, through the temperature of the water which upwells into the surface layer. Upwelling is an important influence on s.s.t. at the South American coast and along the equator, especially in the eastern Pacific. As Bjerknes noted, this is the region where the s.s.t. variations are large. The model suggests that the warming of upwelled waters associated with the depression of the thermocline is the dominant cause of the warm anomalies there during the onset stage of an El Niño event, a result which is consistent with what data there are. It is certainly the fastest mechanism for converting a remote influence into an s.s.t. anomaly.

Our thesis is that an El Niño event cannot start until the heat content in the equatorial band is high enough to sustain the Bjerknes 'chain reaction'. The prevalence of 30–60-

day oscillations in the western Pacific (e.g. Lau and Chan 1985) ensures that episodes of westerly wind anomalies are frequent. They always generate oceanic Kelvin waves which tend to push down the thermocline in the east. However, if the thermocline is initially too shallow then the resulting increase in s.s.t. will not alter the east–west temperature contrast enough to influence the atmosphere. Hence the initial wind anomaly is not reinforced and no event takes place.

Now consider the variations of the equatorial thermocline through an ENSO cycle. As the peak of an El Niño event is approached water moves not only west to east, but also poleward, emptying the reservoir of warm water at the equator. As a result the thermocline has already begun to shallow at the time when the wind and s.s.t. anomalies are reaching their peak. This phase relation may be seen in the model by comparing Fig. 4 with Fig. 3. For example, at the beginning of year 31 the thermocline along the equator is already anomalously deep although the wind anomalies are still predominantly easterly; by the autumn of year 32 the equatorial thermocline is anomalously shallow at all longitudes although the wind anomalies are still somewhat westerly. Figure 5, which is based on tide gauge observations in the equatorial Pacific, shows that the heat content begins to fall before the surface anomaly maximum in December 1982–January 1983 (cf. the 1976/77 event).

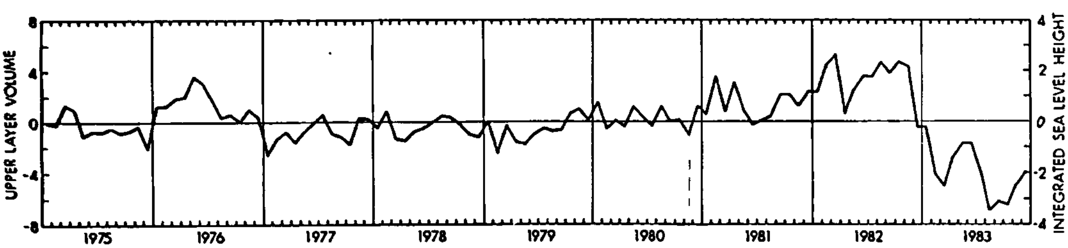


Figure 5. Upper layer volume of the tropical Pacific ($10^{14}m^3$) from 1975 to 1983 relative to its mean value ($70\times10^{14}m^3$). The volume is estimated on the basis of tide gauge data for the region from 15°N to 15°S. (From Wyrtki 1985).

The decline in heat content is truly the beginning of the end of the El Niño event. The absence of the subsurface warm anomaly means that the s.s.t. anomaly can no longer be maintained. As the temperature drops in the eastern Pacific the westerly wind anomaly loses its *raison d'etre*. In the rapid collapse of the El Niño the ocean overshoots its mean state, leaving the eastern ocean very cold, the equatorial easterlies stronger than normal, and the heat content below its mean value. The transition back to an El Niño state cannot take place until enough warm water flows back from higher latitudes to refill the equatorial heat reservoir.

Additional support for the importance of the mean zonal heat content may be drawn from the numerical experiments illustrated in Figs. 6 and 7. For the experiments in Fig. 6 the model was simplified by allowing advection of temperature by only the averaged currents of the upper ocean (i.e. the reduced gravity model currents of (A4)–(A7)): in effect the surface layer has been eliminated. Also, the mean conditions were held fixed at their July values. The model now oscillates regularly. In these experiments the relation between the deep temperature T_d and the thermocline depth anomaly h was altered by, for example (Fig. 6(c)), replacing $T_d(h)$ by $T_d(h - 2h^*)$, where h^* is the mean thermocline depth anomaly averaged along the equator. This is tantamount to tripling the rate at which the equatorial heat reservoir is refilled, and has the effect of decreasing the period between events. In general such changes (cf. Fig. 6) have a roughly linear effect on the period, with the experiment in which the heat content is held fixed showing no oscillation at all.

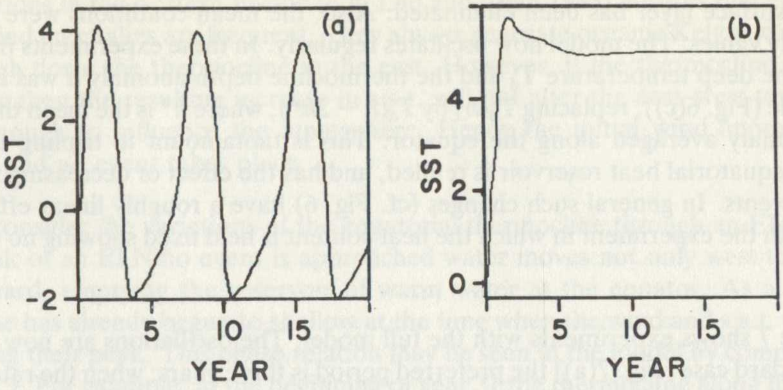
Figure 7 shows experiments with the full model. The oscillations are now irregular. In the standard case (Fig. 7(a)) the preferred period is three years; when the rate at which the equatorial heat content refills is halved (Fig. 7(b); $T_d(h)$ replaced by $T_d(h - 0{\cdot}5h^*)$) the preferred period roughly doubles to 5 or 6 years; when the rate is tripled (Fig. 7(c)) it decreases to two years. In these cases with the full annual cycle the periods are quantized in integral numbers of years, but the changes have the same sense as in Fig. 6.

The outcome of a different set of experiments exploring the role of heat content changes is summarized in Fig. 8. A coupled run, in which an El Niño was about to occur (heavy curve), was restarted four times with $T_d(h)$ replaced by $T_d(h - d)$ with $d = 1$, 2, 3 and 4 m (light curves, top to bottom). The effect is equivalent to reducing the heat content anomaly uniformly by d metres of warm water. The El Niño occurs more or less as before in the $d = 1$ m case, but raising the thermocline by another metre makes the difference between an El Niño and a mild warming. Raising it further still eliminates the warming altogether.

That the wind and thermocline anomalies are not precisely in phase appears to be necessary for the oscillation to occur. If they were in phase the positive feedback would simply increase the amplitude of the anomaly forever. The linear model of Philander *et al.* (1984) exhibits such growing, non-oscillatory behaviour. A more realistic model would warm until the coupled system's equilibration mechanisms took over and locked the system in its maximum warm state.

If one considers how the wind anomaly is generated, it is plausible that the wind would lag the thermocline. A change in thermocline depth changes the temperature rapidly at the equator, but it takes some time for advection to spread the anomaly north and south to a scale which will fully influence the atmosphere.

It is a consequence of linear equatorial ocean dynamics that the changes in the equatorial thermocline will lead those in the wind. The slope along the equator is closely in phase with the wind (e.g. Philander 1979; Cane and Sarachik 1981), but the zonal mean thermocline depth and the depth at the east lead. Cane and Sarachik have shown this to be characteristic of the equatorial ocean's response to low frequency periodic wind forcing and thus to all low frequency forcing. (Here 'low frequency' means timescales comparable with or greater than the time for a Kelvin wave to cross the ocean; this is about 70 days for the Pacific.) The reason for the lead may be better understood by considering the response of the ocean to a wind (a westerly wind, for example) suddenly switched on over a resting ocean at $t = 0$ (Cane and Sarachik 1977).



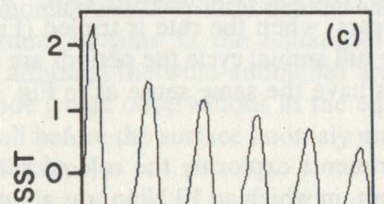
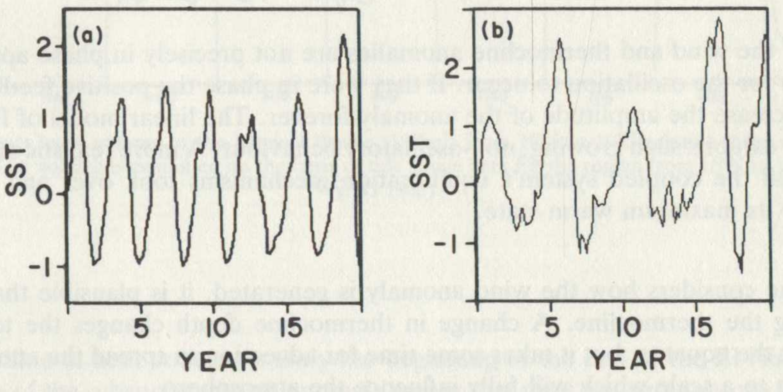


Figure 6. NINO3 s.s.t. from model runs with the surface layer removed and perpetual July conditions. (a) Standard case. (b) Equatorial heat content held constant. (c) Changes in equatorial heat content accelerated by a factor of three (see text).



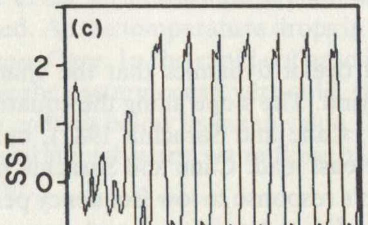


Figure 7. As Fig. 6. but with the complete model. (a) Standard case. (b) Changes in equatorial heat content slowed by a factor of two. (c) Changes in equatorial heat content increased by a factor of three (see text).

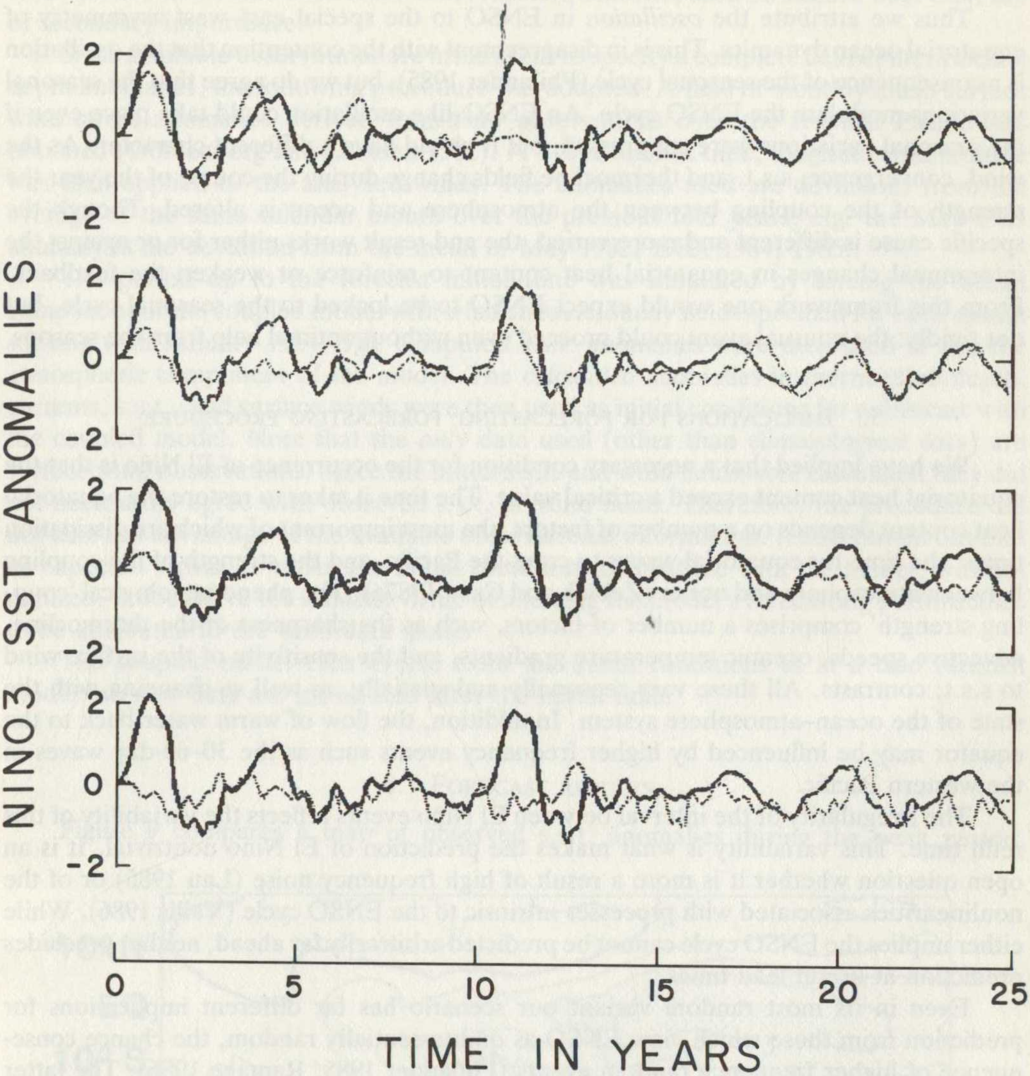


Figure 8. NINO3 s.s.t. when the equatorial heat content is altered as described in the text. The dark curve is the standard case. For the light curves, the thermocline was artificially raised by 1, 2, 3 and 4 m from top to bottom.

Initially, the westerly winds induce an equatorward Ekman flux, resulting in an increase in the equatorial heat content. This begins to change as the influence of the boundaries makes itself felt. Kelvin waves leaving the western side and Rossby waves from the eastern side set up a pressure gradient across the ocean, and the accompanying poleward geostrophic flow counteracts the equatorward mass flux. The thermocline rises in the west and sinks in the east. The 'hole' at the west is dug by Kelvin waves and so is equatorially confined, while at the east the Rossby waves spread the extra mass over a wide latitude band. The net result is to decrease the heat content near the equator and export mass to higher latitudes.

Thus we attribute the *oscillation* in ENSO to the special east–west asymmetry of equatorial ocean dynamics. This is in disagreement with the contention that the oscillation is a consequence of the seasonal cycle (Philander 1985), but we do agree that the seasonal variations modulate the ENSO cycle. An ENSO-like oscillation could take place even if the seasonal variations were suppressed, but it would have a different character. As the wind, convergence, s.s.t. and thermocline fields change during the course of the year the strength of the coupling between the atmosphere and ocean is altered. Though the specific cause is different and more varied, the end result works either for or against the interannual changes in equatorial heat content to reinforce or weaken the feedback. From this framework one would expect ENSO to be locked to the seasonal cycle, but not rigidly: the unusual event could proceed even without optimal help from the seasons.

5. Implications for forecasting: forecasting procedure

We have implied that a necessary condition for the occurrence of El Niño is that the equatorial heat content exceed a critical value. The time it takes to restore the equatorial heat content depends on a number of factors, the most important of which are dissipation times, the time for equatorial waves to cross the Pacific, and the strength of the coupling between atmosphere and ocean (Zebiak and Cane 1987a). The phenomenological 'coupling strength' comprises a number of factors, such as the sharpness of the thermocline, advective speeds, oceanic temperature gradients, and the sensitivity of the surface wind to s.s.t. contrasts. All these vary seasonally and spatially, as well as changing with the state of the ocean–atmosphere system. In addition, the flow of warm water back to the equator may be influenced by higher frequency events such as the 30–60-day waves in the western Pacific.

The irregularity of the interval between El Niño events reflects the variability of this refill time. This variability is what makes the prediction of El Niño nontrivial. It is an open question whether it is more a result of high frequency noise (Lau 1985) or of the nonlinearities associated with processes intrinsic to the ENSO cycle (Vallis 1986). While either implies the ENSO cycle cannot be predicted arbitrarily far ahead, neither precludes prediction at useful lead times.

Even in its most random variant our scenario has far different implications for prediction from those which view ENSO as quintessentially random, the chance consequence of higher frequency random events (Philander 1985; Ramage 1986). The latter offer no hope of forecasting ENSO until its high frequency causes can be predicted, an unlikely prospect. In the former El Niño is the outcome of a deterministic sequence, albeit one which can be distorted by chance occurrences. The prognosis for prediction is far more optimistic.

In our account of ENSO all the essential interactions take place in the tropical Pacific region. Consequently, while information from other parts of the globe might improve the results, it is possible that predictions based solely on data from this region could be successful. There are practical advantages in limiting the region which must be covered, though it is unfortunate that the critical region is one which is so poorly observed. It is also unfortunate that data for the most important variable, the field of thermocline depth anomalies, are so sparse. Of the other two most important variables, s.s.t. is known to no better than 0·5 degC (Reynolds, private communication), and the surface winds to no better than $2\,\mathrm{m\,s^{-1}}$ (Halpern and Harrison 1982). For both variables, the errors are generally higher than the quoted numbers. *In situ* wind measurements are few and are concentrated in a handful of shipping lanes. Our ENSO scenario does have the fortunate consequence that some even more inaccessible quantities, such as surface heat flux, are of secondary importance.

Since available observations are insufficient to specify a complete field of thermocline depth anomalies, the following procedure was adopted. A field of monthly mean surface wind stress anomalies derived from ship observations over the tropical Pacific was obtained (Goldenberg and O'Brien 1981). A 1–2–1 filter in time, longitude and latitude was then applied to the analysed winds. The anomalies used are deviations from the average of the same calendar month over the previous four years; e.g. the May 1985 anomaly is the deviation from the mean of May 1982, 1983, 1984, 1985.

The period up to the forecast initial time was simulated by forcing the ocean component of the coupled model with wind stress anomaly fields specified for each month starting with January 1964. The computed s.s.t. anomalies were then used to run the atmospheric component of the model. The *calculated* anomalies in thermocline depth, currents, s.s.t., and surface winds were then used as initial conditions for a forecast with the coupled model. Note that the *only* data used (other than climatological data) are surface wind observations. Since the initial s.s.t. and wind fields were calculated they did not necessarily agree with observed s.s.t. or wind fields. Therefore, the procedure did not take full advantage of the available observational information. It also turned out that its expected virtue, to make the initial conditions compatible with the model, was not realized. It does have the didactic virtue of allowing the model's forecasting performance to be attributed to the wind data alone.

The coupled model runs evolve from the initial conditions as in a true forecast model: no new data are introduced after the initial time.

6. Forecast results

Figure 9 compares a map of observed s.s.t. anomalies during the peak period,

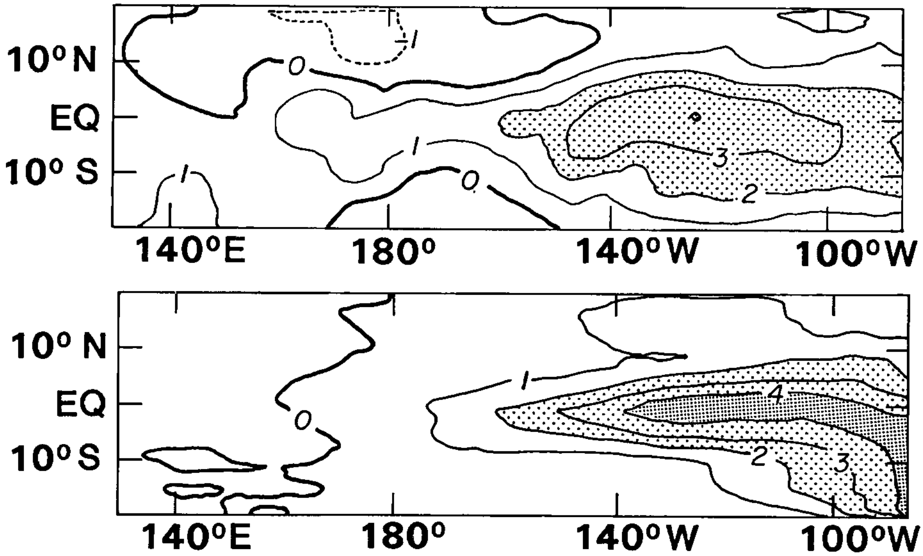


Figure 9. Sea surface temperature anomalies (degC) in January 1983. Top: observed, based on the analysis of the Climate Analysis Center (CAC) of NOAA. Bottom: predicted by the model forecast initiated in January 1981, two years earlier. (From CZD)

January 1983, with that from the forecast initiated in January 1981. The result is typical of the more successful forecasts. As in earlier work (Zebiak and Cane 1987a), including cases with the same ocean model driven by observed winds (Zebiak and Cane 1987b), the gross character of the eastern Pacific El Niño s.s.t. anomaly is reproduced though its meridional and westward extent is understated.

Forecasts were initiated in the periods preceding each of the El Niño events which have occurred since 1970; i.e. 1972, 1976 and 1982. An additional set was made for the 'non-event' of 1979. None was attempted in the years before 1970 because the wind analyses available for these earlier years are of distinctly lower quality. In each period there are six forecasts spaced three months apart with the sequence ending in January of the nominal year. For example, for 1972 forecasts were initiated from October 1970, January, April, July, October 1971, and January 1972.

The forecast results will be summarized in terms of the s.s.t. anomaly averaged over the eastern equatorial Pacific area NINO3 (Fig. 10). The NINO3 index was devised by the Climate Analysis Center of NOAA (National Oceanic and Atmospheric Administration) because a warming in this region strongly influences the global atmosphere (Rasmusson and Wallace 1983). It is probably the best single indicator of an ENSO episode likely to impact global climate.

Figure 10 shows that the model generally succeeds in predicting warmings in those

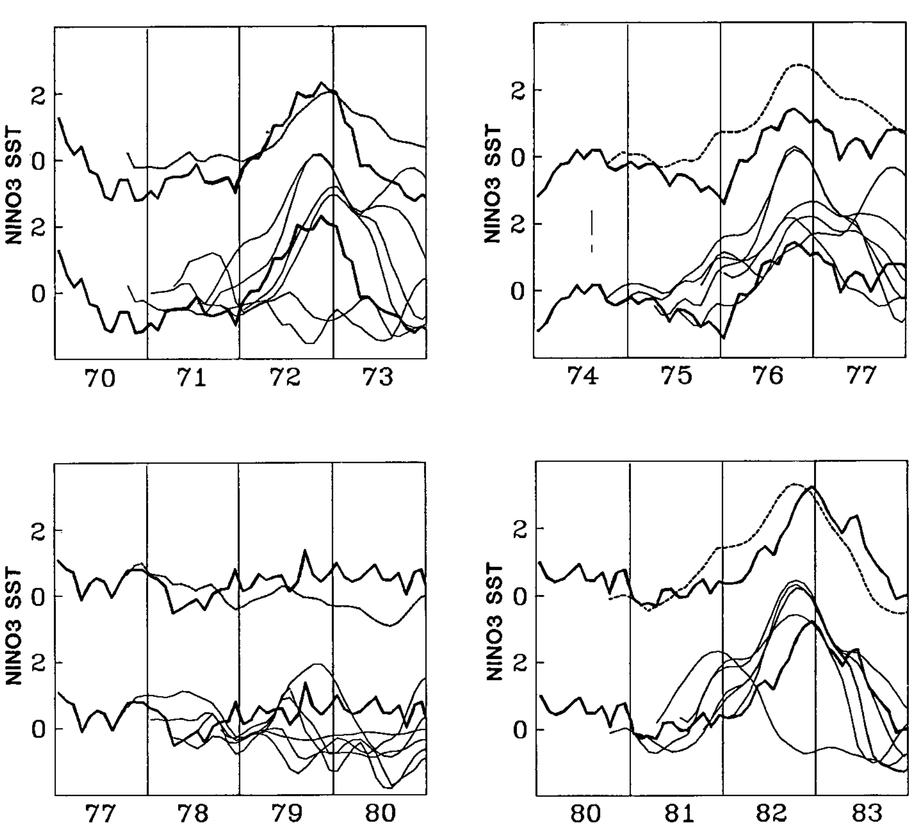


Figure 10. Sea surface temperature anomalies (degC) averaged over the eastern equatorial Pacific NINO3 region (90°W to 150°W, 5°S to 5°N) for four periods centred on 1972, 1976, 1979 and 1982. The heavy curves are the observed values as reported by CAC/NOAA. The light curves in each panel are from six forecasts initiated at three-month intervals from the October two years ahead to the January of the nominal year. The dashed curve is the average of the six forecasts (the consensus forecast). (From CZD)

years when they were observed to occur and in predicting their absence when they did not occur. As with the observed events, the peaks in the NINO3 index are always forecast to take place at the end of the calendar year. The individual forecasts tend to give too strong an event, especially for the 1976 El Niño. The rise in temperature during 1976 is well predicted, but five of the six forecasts begin the year at least 2 degC too warm. They entirely missed the 1975 cooling trend, which culminated in the lowest NINO3 s.s.t. in the entire data record. This failure is present even when the model was forced by observed winds, as in the run creating initial conditions (i.e. the initial s.s.t. for the January 1976 forecast in Fig. 10). On the other hand, the model never predicts an El Niño for 1975, though at the time observers saw enough precursors of an event early in the year to call an El Niño Watch.

We wish to establish that the model's ability to predict an El Niño event is significantly better than chance. Because the distribution of states in the tropical Pacific is strongly bimodal (e.g. Cane 1983), it is straightforward to distinguish El Niño events from non-events. In terms of NINO3 s.s.t. we will say there is an El Niño if there is a rise of at least 2 degC within 12 months leading to an anomaly of at least +1 degC for at least 3 months. These criteria are derived from the observed events but are somewhat arbitrary; the set of forecasts would have similar scores for any other criteria able to separate observed events from non-events.

According to these criteria, 19 of the 24 forecasts are correct. (The forecasts would be judged to be less successful if more detailed agreement with observation were demanded.) January 1971, April 1971, October 1975 and April 1981 fail to predict the upcoming event while April 1978 falsely calls for an El Niño in 1979. Of the five failures, three are from April starts. In the spring the intertropical convergence zone is closest to the equator and the s.s.t. in the east is highest. As a result the *local* coupling between model atmosphere and ocean is strong and noise in the initial state often amplifies, leading to poor results. Nature is also least predictable in the spring: it is the time when the Southern Oscillation index shows the smallest correlation with its future value (Wright 1985).

Meteorological forecasts are typically evaluated relative to climatology or persistence, but either inflates the skill for infrequent events like El Niño. Table 1 offers another

TABLE 1. SUMMARY OF FORECAST RESULTS FOR PAST YEARS

Start month	Number right	Number of cases	Chance by coin flips
One-year forecasts			
Oct.	7	8	0·04
Jan.	8	8	0·004
Apr.	1	4	0·94
Jul.	4	4	0·06
Total	20	24	0·0008
Two-year forecasts			
Oct.	3	4	0·31
Jan.	3	4	0·31
Total	6	8	0·14

Results correspond to the 24 forecasts of Fig. 9. For January, 'one-year forecasts' refers to the same calendar year; for all other start months, to the next calendar year. The last column shows the probability of correctly predicting at least as many cases as the model did, by flipping a coin.

standard. It may seem artificially favourable to compare against a guessing strategy which makes no apparent use of the known 3- to 4-year quasi-periodicity of the ENSO cycle. However, we have forecast only for periods three or four years apart when events are seemingly due. Hence the occurrence or non-occurrence of an event in any of the years in our sample is roughly equally likely, so that flipping a coin is a reasonable guessing strategy.

In order to establish a rigorous confidence level the number of *independent* forecasts must be known. A precise determination of independence would require many more model runs, but estimates adequate for present purposes can be made readily. Treating forecasts three months apart as independent derives from the observation that qualitative behaviour (such as incorrect forecasts or maintaining high s.s.t. in the year after an event) rarely persists longer. Doing so probably overstates the confidence level. At the other extreme, a very conservative estimate is obtained by assuming the forecasts within each set are highly correlated. Then there are only four independent forecasts, each given by the majority of the six within each set. Since all four are correct there is one chance in 16 of doing as well by coin flipping. Our most plausible estimate is based on the fact that while indices of the Southern Oscillation can be highly correlated for many months, months preceding April of a given year correlate poorly with months following it (Wright 1985). This motivates treating the (majority of) the first October, January and April of a set as one forecast, independent of a second forecast derived from the subsequent July, October and January. There are now eight independent forecasts, the last seven of which are correct. The probability of guessing so successfully is about 0·04.

Another measure of model performance is obtained by considering events and non-events separately. Altogether, 14 of 18 occurrences of El Niño and 5 of 6 non-occurrences are forecast correctly. The joint probability of getting at least that many of each type right by an optimal random strategy is 0·004. (The optimal random strategy assumes *a priori* knowledge of the number of events and non-events in the sample, as well as the score to beat.) The corresponding joint probabilities for the 'conservative' and 'most plausible' cases are 0·06 and 0·04, respectively.

Forecasts started from the same calendar month in different years differ only in their initial conditions. Their success attests to the thesis that the future evolution of ENSO is implicit in the initial conditions. Consider, for example, the eight one-year forecasts initiated in January. Even if one knew in advance that there were three El Niño events and five non-events, the chance of matching the model performance by guessing all eight correctly is only 1 in 56.

Rather than considering forecasts at 3-month intervals, we now examine a more useful forecasting strategy. A forecast is initiated in each of the six successive months from August to January to obtain the prediction of the ensemble one and two years ahead. Table 2 summarizes the results of doing this from every year for which we have data (1970–1984) with the exception of the three El Niño years in the sample. A few examples are shown in Fig. 11. There are 12 sets of one-year forecasts. In nine of these all six months agree so the interpretation of the model prediction is unambiguous. In all nine the prediction is correct. By the rather stringent standard which scores the other three as wrong, the probability of doing as well or better by chance is 0·07.

It would be more accurate to judge that the three provide unclear guidance. In two of the cases, where there was no event in the following year, half of the months forecast correctly and other half did not. The remaining case is the set from August 1975 through January 1976. As can be seen from Fig. 11 the model actually does a good job of predicting the odd event to come. However, four times out of six our temperature rise criterion is not met because the model starts out too warm at the beginning of 1976. One

TABLE 2. SUMMARY OF THE SETS OF FORECASTS INITIATED IN THE SIX SUCCESSIVE MONTHS FROM AUGUST THROUGH JANUARY OF EACH YEAR FROM 1970 TO 1984 (EXCLUDING EL NIÑO YEARS)

Forecast from	Number correct (out of six) One-year	Two-year	Forecast from	Number correct (out of six) One-year	Two-year
1970	6	5*	1978	6	6
1971	6*	1	1979	6	2
1973	6	3	1980	6	6*
1974	6	6*	1981	6*	6
1975	2*	5	1983	3	5
1977	6	6	1984	3	?

Years of El Niño events are starred. In the first line of the table, for example, the forecasts are from August 1970 through January 1971; the one-year forecasts are for 1971 and the two-year forecasts are for 1972.

might quibble with our scoring system, but it undoubtedly would have been difficult to interpret such a forecast at the time.

As might be expected, the two-year forecasts are not as successful: in five cases all six individual forecasts are correct; in three cases five of six are correct; in the remaining

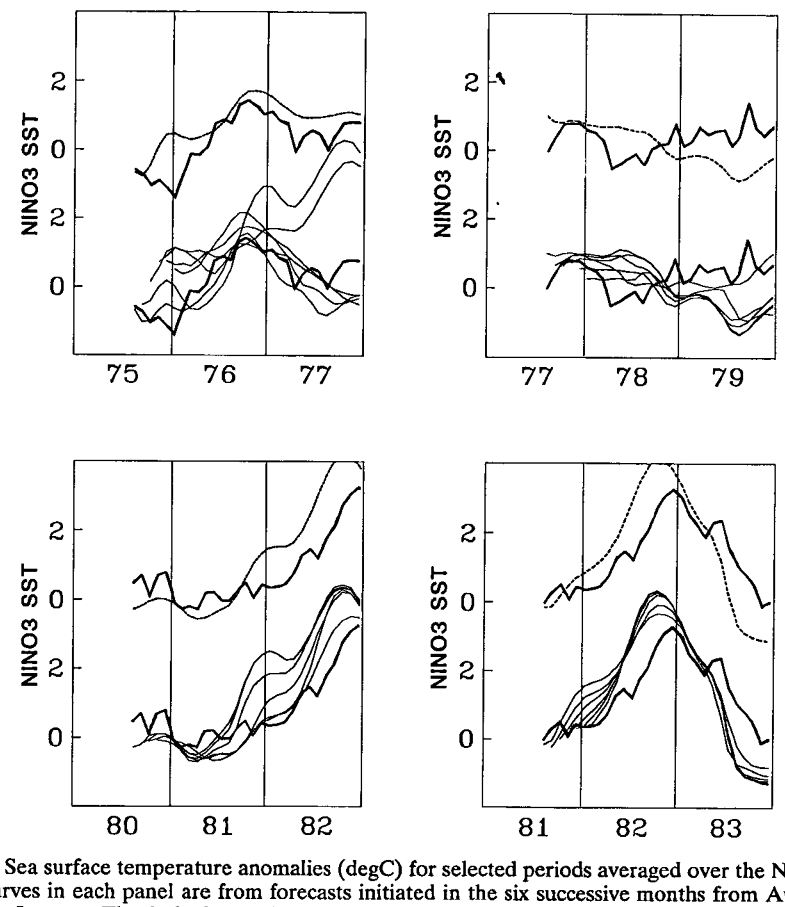


Figure 11. Sea surface temperature anomalies (degC) for selected periods averaged over the NINO3 region. The light curves in each panel are from forecasts initiated in the six successive months from August through the following January. The dashed curve is the average of the six (the consensus forecast) and the heavy curves are the observed values. (From CZD)

cases 1, 2 or 3 are correct. Counting 8 of the 11 as successes gives an 89% confidence level. The most consistent failure is the 1971 set, where all predicted the 1972 El Niño but five of the six had the warming persist through 1973. Four of the 1979 forecasts predicted an event in 1981, a year early. The 1983 set is the most ambiguous; note that for the years since the outsized 1982 event both the one-year and two-year forecasts have been inconsistent.

We offer one final example: the prediction for 1986. At the time of writing, September 1986, this forecast cannot yet be evaluated. Data for August (Kousky, private communication) indicate the trades have weakened and the eastern equatorial Pacific is anomalously warm, but the amplitudes are small. Perhaps an El Niño will follow, but perhaps it will never develop beyond a mild warming. It is surprising to be so late in the year without a definite reading as to whether or not an El Niño will occur.

Figure 12 shows the forecasts initiated in August 1985 through January 1986. All six predict an El Niño. The amplitude of the consensus forecast suggests it will be a moderate event, not as strong as the events of 1982, 1972 or 1957. However, the only other moderate event in our study was the peculiar 1976 El Niño; experience with the moderate event of 1965 and the even weaker 1969 one might help in interpreting the present case. For the three events in our sample (1972, 1976 and 1982) the consensus overstated the maximum NINO3 s.s.t. anomaly amplitude by 30% (more precisely, 28%, 36% and 28%, respectively). Rescaling the 1986 forecast accordingly indicates an amplitude of 1·9 degC, well below the 3·2 degC of 1982 or the 2·5 degC of 1972, but above the 1·4 degC of 1976.

How much confidence should be placed in the prediction of a 1986 El Niño? Though all of the 1985 forecasts predict an event for 1986, none of the two-year forecasts from 1984 predict it. This is unprecedented among the limited set of cases we have studied. Also, both the 1983 and 1984 one-year forecasts lack the consistency which is typical of earlier years. On the other hand, and in our view more importantly, no previous case where all six one-year forecasts agree was incorrect. Moreover, except for June 1986, all forecasts from March 1985 to July 1986 give indications of a warm event at the end of 1986.

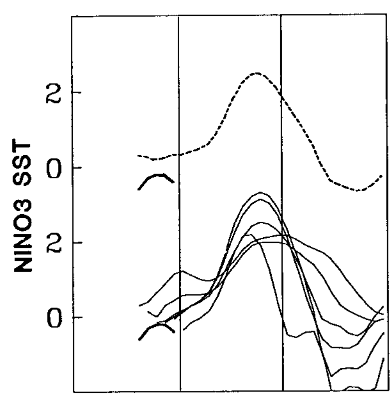


Figure 12. Forecast for 1986: sea surface temperature anomalies (degC) in the NINO3 region. The light curves are from the forecasts initiated in the six successive months from August 1985 through January 1986. The dashed curve is the average of the six and the heavy curves are the observed values. (From CZD)

7. Implications of the forecasts for the nature of ENSO

The forecasts may be regarded as a set of numerical experiments bearing on the nature of ENSO. The model appears to be able to simulate the major features of observed El Niño events, and largely to avoid predicting events when it is inappropriate. Though hardly irrefutable evidence, the model's ability to generate observed El Niño events from apparently innocuous, 'normal' initial conditions adds some empirical support for the contention that the cycles of ocean–atmosphere variations occurring in the model work in the same way as nature's ENSO cycle.

The degree of forecasting skill obtained despite the crudeness of the model is telling. It suggests that the mechanism responsible for the generation of El Niño events and, by extension, the entire ENSO cycle, is large scale, robust and simple: if it were complex, delicate or dependent on small-scale details this model could not succeed. Neither the ocean nor the atmosphere is the prime mover in the ENSO cycle. It is essential for ENSO that the two-way coupling between ocean and atmosphere is active in both the El Niño and the non-El-Niño phase of the cycle. The only region where the model does a creditable job of simulating the requisite interactions is near the equator. Its success with ENSO is a confirmation of Bjerknes' emphasis on the interactions which take place in the equatorial plane.

Our addition to the basic Bjerknes hypothesis required the model to include the dynamics of the upper layers of the tropical ocean. The tropical Pacific region remains the locus for all of the physics responsible for the existence of the ENSO cycle. This is not to deny that events in the Indian Ocean sector, in mid-latitudes, etc., may influence the evolution of an El Niño event. Moreover, the evidence is compelling that the impact of ENSO is global. However, we do mean to assert that the essence of the oscillation takes place in the tropical Pacific.

Though able to generate the ENSO cycle and to forecast particular El Niño events, the model is unable to generate the 30–60-day waves. The clear implication is that these waves have no more than a minor role in ENSO. Observations seem to carry the same implication: these waves are virtually ubiquitous, and there appears to be nothing distinctive about their behaviour prior to an El Niño. We believe they are properly viewed as a rather powerful source of noise on the larger scale, lower frequency ENSO cycle. As such they have important consequences for prediction; this is discussed further below.

We have characterized El Niño as a phase of a recurring cycle, ENSO. The deterministic evolution of the cycle returns the ocean–atmosphere system to a state where small perturbations can grow into an El Niño. Our view is that a necessary condition for such an instability is that the equatorial heat content exceed a threshold value. Once El Niño is ready to happen it can be initiated by any appropriate perturbation. The 30–60-day waves are the most likely candidate, but were they absent, something else would play this role. This description is in the spirit of the classical paradigm for hydrodynamic instability: it is assumed that a suitable perturbation will be available to grow; the crucial question is whether the necessary and sufficient conditions for instability are met.

In this scenario, the interval between events is the time for the equatorial heat reservoir to refill. In our view, the refilling does not result from the action of free equatorial waves alone. As with the events themselves, it is an interactive process, and the time it takes depends on the strength of the coupling between atmosphere and ocean. *A priori*, it seemed conceivable to us that this refill time might be determined by random events, so a scenario assigning to them a dominant role could not be ruled out. The demonstrated ability of our model to forecast at long lead times suggests that the process is largely deterministic, though, as discussed below, occurrences external to the basic ENSO cycle may well exert an important influence.

A concise statement is that ENSO is a relaxation oscillation of the coupled system: the slow build-up to the necessary condition for instability during the cold phase presages the rapid warming during the event itself; an even more rapid return to cold conditions follows. In the model the mean climatic state is unstable, so a relaxation to mean conditions will lead into the next El Niño event. In both the model and nature the interval between events is irregular; as noted above, while this could be the result of random fluctuations, it could also be a consequence of the deterministic physics intrinsic to ENSO.

If ENSO is a chaotic deterministic system then the predictability is inherently limited. There is a bifurcation surface dividing states of the ocean–atmosphere system which will soon culminate in an El Niño from those which will remain near normal. On one side, far from this surface an El Niño is sure to occur; far away on the other side its absence is equally certain. Near the surface, where a small change in present position can result in a large difference in the future path of the system, the outcome is uncertain.

It is plausible that most of the time the system is far from this surface and its future course can be predicted for some time to come. Our forecast results suggest this is indeed the case and that the prediction time is typically several years or more.

Accurate prediction at those times when the system is close to a bifurcation point demands that its state be known with great precision. In practice the data may be too poor to allow this, and, since the prediction model is an imperfect representation of the physical system, it may mislocate the bifurcation surface in the phase space. Figure 13 suggests that the initial states shown in Figs. 14 and 15 are examples of states near a

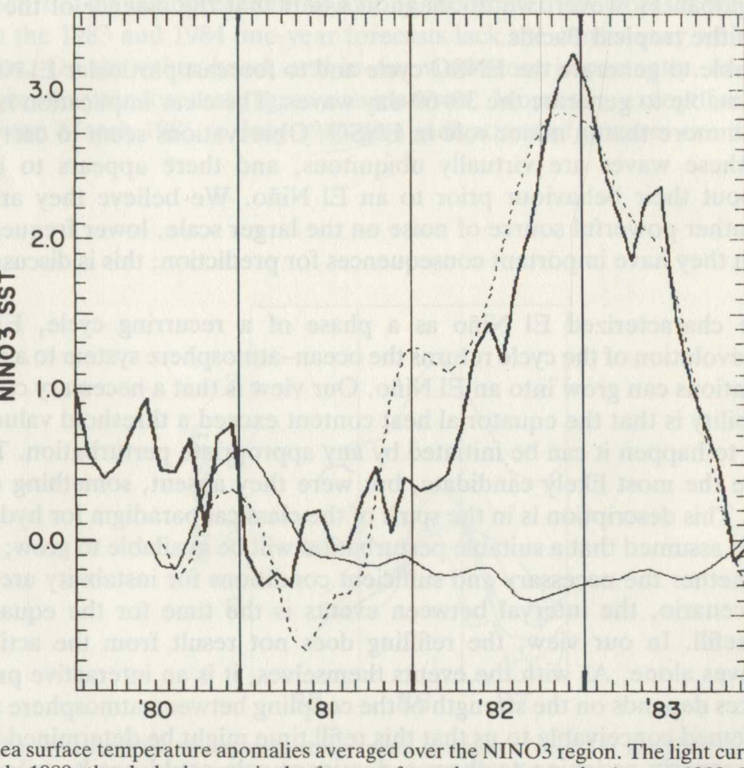


Figure 13. Sea surface temperature anomalies averaged over the NINO3 region. The light curve is the forecast initiated in July 1980; the dashed curve is the forecast from the next month, August 1980. The heavy curve shows the observed values.

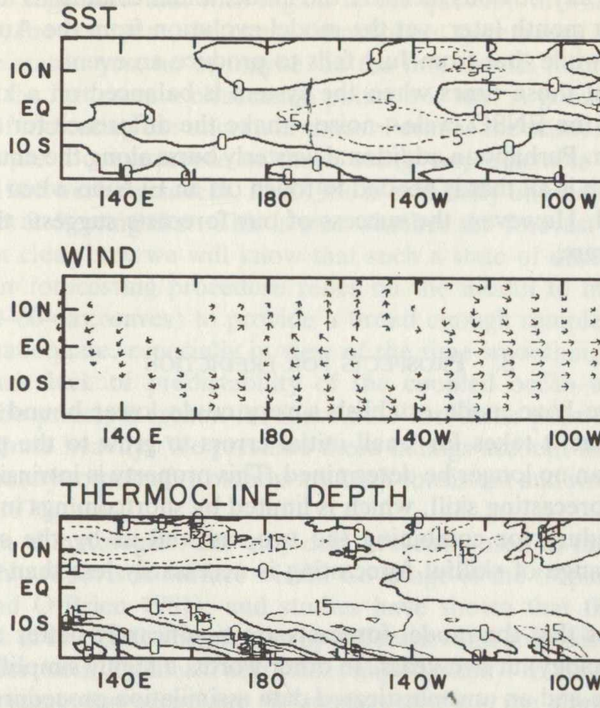


Figure 14. Initial conditions for the forecast from July 1980. Top: s.s.t. anomalies (units of 0·1degC). Middle: surface wind anomalies (the maximum is about $1\,\mathrm{m\,s^{-1}}$). Bottom: thermocline depth anomalies. Contour interval is 15 m.

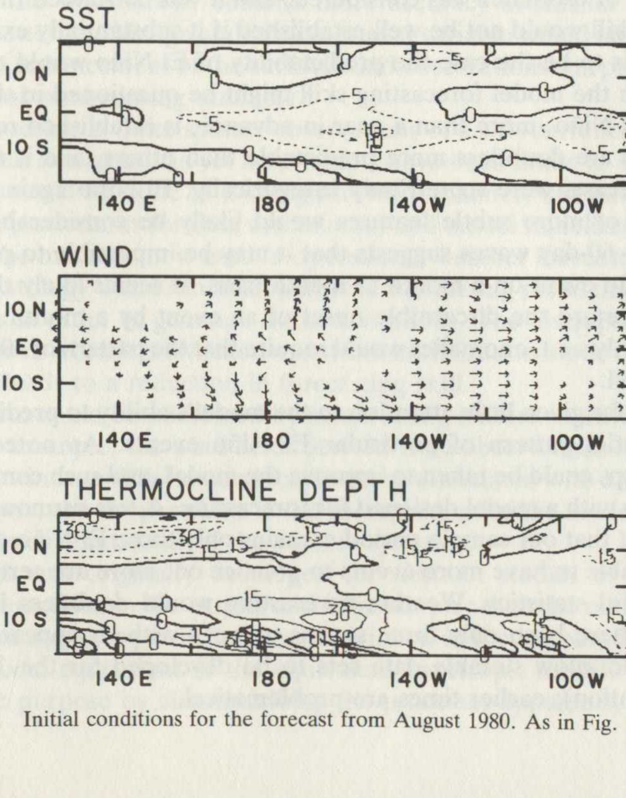


Figure 15. Initial conditions for the forecast from August 1980. As in Fig. 14 for July.

bifurcation point. By any obvious measure the model initial conditions for July 1980 are quite close to those a month later, yet the model evolution from the August start leads to the 1982 El Niño while that from July fails to produce an event.

It may be that at those times when the system is balanced on a knife-edge, then events extraneous to the ENSO cycle—noise—make the difference for the future state of the coupled system. Perhaps an additional westerly burst along the equator associated with a 30–60-day wave is all that is needed to touch off an El Niño when otherwise none would have occurred. However, the success of our forecasts suggests that if there are such times they are rare.

8. Prospects for prediction

The forecasts we have made establish a very crude lower bound on the limit of predictability, the time it takes for small initial errors to grow to the point where the system's behaviour can no longer be determined. This property is intrinsic to the system. It is different from forecasting skill, which is limited by shortcomings in the model, the data, and the procedure for combining the two, as well as by the system's limited predictability. The range of skillful forecasting is necessarily less than or equal to the limit of predictability.

Table 2 indicates that the model forecasts are significantly better than chance (or persistence or climatology) at two years. In other words, a highly simplified model using notoriously poor data and an unsophisticated data assimilation procedure has significant forecasting skill at a lead time of two years. We conclude that the predictability of ENSO is at least two years. It may well be considerably longer.

Our assertion that forecasting skill has been established depends on estimates of the number of independent forecasts, or, equivalently, the decorrelation time between forecasts. A range of estimates was considered, and it was concluded that this time is a year or less. Our skill would not be well established if it substantially exceeds one year. However, were this to be the case the predictability of El Niño would also exceed one year. Thus, though the model forecasting skill might be questioned in this manner, the predictability of El Niño, more than a year in advance, is established regardless.

Some features are doubtless more predictable than others, and it must be kept in mind that our forecasts were scored only categorically: El Niño against non-El-Niño. The predictability of more subtle features would likely be considerably shorter. The presence of the 30–60-day waves suggests that it may be impossible to predict the early stages of an El Niño event on a month to month basis. It seems likely that these waves could advance or retard the discernible onset of an event by a month. If so, accurate prediction of monthly s.s.t. anomalies would require that the pattern of 30–60-day activity be predicted as well.

We have thus far given little attention to the model's ability to predict details of the evolution and spatial pattern of particular El Niño events. As noted below, many straightforward steps could be taken to improve the model, and such comparisons would be more rewarding with a model designed for forecasting. A less surmountable difficulty arises from the fact that our sample period contains only three El Niño events. It would be extremely valuable to have more events to practice on; more are certainly needed to compute meaningful statistics. Weather forecasting would doubtless have developed more slowly had there been only three storms to work with. It appears that available observations should allow suitable data sets to be developed for the 1960s (O'Brien, private communication); earlier times are problematical.

Earlier, we suggested that with regard to the question of whether there will be an El Niño in, say, the next year, the ocean–atmosphere system may be thought of as being in one of three states: yes, no or maybe. In the first two its future evolution is readily predictable; in the last, near a bifurcation point, it is not. A good forecast model ought to be capable of indicating the uncertain times. Perhaps the scatter in the forecasts from 1983 and 1984 are a consequence of those being 'unpredictable' times. The forecasts from 1985 for 1986 were consistent, so if 1986 is especially unpredictable then the model has failed in not indicating that. This is true whether the forecast proves false or not—though it is not clear how we will know that such a state of uncertainty existed in the real world. Our forecasting procedure relies on the month to month variability (due primarily to 30–60-day waves) to provide a broad enough sample of initial conditions. This may be inadequate, especially in view of the time-smoothing applied to the data.

The intrinsic lack of predictability of the coupled ocean–atmosphere system is probably not the principal reason for inaccurate forecasts: poor data and model inadequacies contribute heavily. We presume these failings account for the cases where the forecasts from many consecutive months are both consistent and wrong (e.g. the forecasts for 1973 from 1971).

Apart from climatologies, the only data presently used in the model forecasts are derived from ship reports of surface winds. Coverage of the tropical Pacific is poor (see Goldenberg and O'Brien 1981), and studies have shown that the differences among model hindcasts of sea level when driven by different wind analyses are often comparable to the signal (Harrison, Kuklinski and Cane, unpublished). In view of this the success of the forecasts is somewhat surprising. It suggests that it is the large-scale features of the surface wind field—the only part one would expect to be correct—which are important for ENSO. This is consistent with our emphasis on the integrated heat content, a part of the oceanic response which integrates the surface wind stress field, allowing errors to cancel. (In contrast, the ridges and troughs in the dynamic topography, which depend on the curl of the wind stress, demand more accuracy at smaller scales.) It is to be expected that better knowledge of this most crucial variable would improve the forecasts, but a significant enhancement of the surface wind observations will probably have to wait for the advent of wind-sensing satellites.

Though coverage is quite sparse, observations of s.s.t., currents and ocean thermal structure are also available. As yet no attempts have been made to factor these into the initial fields. It could be done in a straightforward manner, but ultimately will require more elaborate procedures for data assimilation and model initialization, comparable in complexity to the methods now used in operational weather forecasting. The model has a different climatology when run in an unforced forecasting mode than when forced with observed winds in order to generate initial conditions. The discrepancy gives rise to an initialization shock when it is switched over from initialization mode to forecasting mode. This doubtless leads to a reduction in forecasting skill.

The problem could be addressed by constructing a more sophisticated initialization procedure. For example, the wind fields derived from observations have more structure than the model winds, a difference which could be eliminated with an additional analysis scheme. However, the deeper cause is the shortcomings of the model, its limited ability to achieve realistic simulations of the climate system.

The model was originally developed for abstract studies of large-scale ocean–atmosphere interactions in the tropics and is far simpler than the numerical models used operationally for weather prediction. As mentioned in the introduction, we regard it as the coupled model equivalent of the equivalent barotropic model. The simplifications serve a didactic purpose by circumscribing the processes apparently needed to account for ENSO, but when forecasting is the goal it is preferrable to achieve the verisimilitude of the most complex models (i.e. GCMs). Our model's forecasting performance must suffer from the lack of mid-latitude variability and the absence of the 30–60-day waves. The model is unable to reproduce episodes of strong cold anomalies (e.g. 1975) and overstates easterly wind anomalies. The success of the forecasts in the face of these flaws—and many others—must be taken as an indication of the robust, large-scale character of the ENSO phenomenon.

A particular concern is the model's understatement of the variability in the western equatorial Pacific. We noted earlier that this prevents it from reproducing the observed eastward propagation of westerly anomalies during the early stages of El Niño. The absence in the model of warm anomalies west of the dateline in early 1986 may have kept the atmospheric component from generating the stronger than normal easterlies observed during that period. (Others would characterize it as a time when the 30–60-day westerly bursts were deficient. The effect on the monthly mean would be the same, though the implied mechanism might be different.) Strong easterlies would retard the development of an El Niño. Even if the predicted event does occur in 1986, the model would seem to have developed its event too early.

It seems obvious that a better prediction model would result from the coupling of an atmosphere and an ocean GCM. Ocean GCMs for the tropics are the more recent development (see Philander 1987, this volume), and though still somewhat immature, they are certainly capable of bettering the performance of the oceanic component of our model. Furthermore, the additional computing requirements are well within the capacity of current supercomputers. Though an atmospheric GCM would unquestionably improve on our model, the computing requirements are problematic. Despite advances in computer power, it remains a formidable task to run an atmospheric GCM hundreds of times for several years at a time.

The strongest conclusion to be drawn from the results reviewed here bears on the predictability of the coupled ocean–atmosphere system. They indicate that El Niño is generally predictable at least a year in advance. Without going beyond the present state of the art, there is considerable room for improvement in the observing system, the data assimilation procedure, and the model itself. In view of this, it does not seem unduly optimistic to suggest that a concerted effort would result in a skillful forecasting capability within a few years. The strong role operational weather forecasting has played in advancing the science of meteorology is a hopeful example of what such an operational capability might do for our understanding of climate variations.

Acknowledgments

Our thanks to Jim O'Brien and David Legler of Florida State University for providing the surface wind analysis, and to Sean Dolan, Karen Streech and Lori Suckling for their help in preparing the manuscript and figures. This work was supported by grants NA-84-AA-D-0031 from the U.S. TOGA Office of NOAA, NA-84-RAD-05082 from EPOCS/NOAA, and NAGW-916 from NASA. Lamont-Doherty Geological Observatory Contribution 4187.

APPENDIX

Atmospheric variables are defined in the conventional manner: (x, y) are zonal and meridional distance respectively and f the Coriolis parameter.

$$\beta_o = \frac{df}{dy}(y = 0);$$

(u_a, v_a) is the velocity in the one-mode atmosphere model; p/ρ_o is the atmospheric dynamic pressure; ε is a phenomenological damping coefficient.

The governing equations for the atmosphere (at iteration n) are (see Zebiak 1986)

$$\varepsilon u_a^n - \beta_o y v_a^n = -(p^n/\rho_o)_x \tag{A1}$$

$$\varepsilon v_a^n + \beta_o y u_a^n = -(p^n/\rho_o)_y \tag{A2}$$

$$\varepsilon(p^n/\rho_o) + c_a^2[(u_a^n)_x + (v_a^n)_y] = -\dot{Q}_s - \dot{Q}_1^{n-1} \tag{A3}$$

$$\dot{Q}_s = \alpha \exp\{(\overline{T} - 30)/16{\cdot}7\}T \tag{A3a}$$

$$\dot{Q}_1^n = \beta\{M(\overline{c} + c^n) - M(\overline{c})\} \tag{A3b}$$

where

$$M(x) = \begin{cases} 0, & x \leqslant 0 \\ x, & x > 0. \end{cases} \tag{A3c}$$

In (A3a), $\overline{T}(x, y, t)$ is the prescribed monthly mean s.s.t. in °C, and T is the anomalous s.s.t. In (A3b), $\overline{c}(x, y, t)$ is the prescribed monthly mean surface wind convergence, and c^n is the anomalous convergence at iteration n, defined by

$$c^n \equiv -(u_a^n)_x - (v_a^n)_y. \tag{A3d}$$

In the ocean model:

$\mathbf{u}_1 = (u_1, v_1)$ is the velocity in the surface layer;
$\mathbf{u}_2 = (u_2, v_2)$ is the velocity in the second layer;
H_1 is the fixed depth of the surface layer;
H_2 is the mean depth of the second layer;
h is the variation in second layer depth;
$H = H_1 + H_2$;
$(\tau^{(x)}, \tau^{(y)})$ is the surface wind stress;
g' is reduced gravity $= g(\Delta\rho/\rho)$ where $(\Delta\rho/\rho)$ is the relative density difference between the upper ocean and the deep inert ocean;
r is a Rayleigh damping for the upper ocean;
r_s is a Rayleigh damping for the surface layer.

The governing equations for the ocean (see Zebiak and Cane 1987a) are

$$u_t - \beta_o y v = -g'h_x + \tau^{(x)}/\rho H - ru \tag{A4}$$

$$\beta_o y v = -g'h_y + \tau^{(y)}/\rho H - rv \tag{A5}$$

$$h_t + H(u_x + v_y) = -rh \tag{A6}$$

where

$$\mathbf{u} = H^{-1}(H_1\mathbf{u}_1 + H_2\mathbf{u}_2) \tag{A7}$$

and the equations governing $\mathbf{u}_s \equiv \mathbf{u}_1 - \mathbf{u}_2$ are

$$r_s u_s - \beta_0 y v_s = \tau^{(x)}/\rho H_1 \tag{A8}$$

$$r_s v_s + \beta_0 y u_s = \tau^{(y)}/\rho H_1. \tag{A9}$$

Equations (A4)–(A9) allow $\mathbf{u}_1$ to be determined. From this, the entrainment velocity is calculated:

$$w_s = H_1[(u_1)_x + (v_1)_y]. \tag{A10}$$

The temperature equation for the surface layer is, then,

$$\partial T/\partial t = -\mathbf{u}_1 \cdot \nabla(\overline{T} + T) - \overline{\mathbf{u}}_1 \cdot \nabla T - \{M(\overline{w}_s + w_s) - M(\overline{w})\}\overline{T}_z - \\ - M(\overline{w}_s + w_s)(T - T_e)/H_1 - \alpha_s T, \tag{A11}$$

where $\overline{\mathbf{u}}_1(x, y, t)$ and $\overline{w}_s(x, y, t)$ are the mean horizontal currents and upwelling, respectively; $T(x, y, t)$ is the prescribed mean s.s.t., and $\overline{T}_z(x)$ is the prescribed mean vertical temperature gradient. The entrainment temperature anomaly, T_e, is defined by

$$T_e = \gamma T_d + (1 - \gamma)T \tag{A12}$$

and T_d is given the form

$$T_d = \begin{cases} T_1\{\tanh(b_1(\overline{h} + h)) - \tanh(b_1\overline{h})\}, & h > 0 \\ T_2\{\tanh(b_2(\overline{h} - h)) - \tanh(b_2\overline{h})\}, & h < 0) \end{cases} \tag{A13}$$

where $\overline{h}(x)$ is the prescribed mean upper layer depth. Parameter values used for the coupled simulation are:

$\varepsilon = (2\text{ days})^{-1}$, $c_a = 60\,\text{m}\,\text{s}^{-1}$, $\alpha = 0{\cdot}031\,\text{m}^2\,\text{s}^{-3}\,\text{K}^{-1}$, $\beta = 1{\cdot}6{\times}10^4\text{m}^2\,\text{s}^{-2}$, $r = (2{\cdot}5\text{ years})^{-1}$, $c^2 \equiv g'H \doteq 2{\cdot}9\,\text{m}\,\text{s}^{-1}$, $H = 150\,\text{m}$, $H_1 = 50\,\text{m}$, $r_s = (2\text{ days})^{-1}$, $\alpha_s = (125\text{ days})^{-1}$, $\gamma = 0{\cdot}75$, $T_1 = 28\,°\text{C}$, $T_2 = -40\,°\text{C}$, $b_1 = (80\,\text{m})^{-1}$, $b_2 = (33\,\text{m})^{-1}$.

References

Ashford, O. M. 1985 *Prophet or Professor: The Life and Work of Lewis Fry Richardson.* Adam Hilger Ltd. Boston

Barnett, T. P. 1984 Prediction of the El Niño of 1982–83. *Mon. Weather Rev.*, **112,** 1403–1407

Bjerknes, J. 1966 A possible response of the atmospheric Hadley circulation to equatorial anomalies of ocean temperature. *Tellus*, **18,** 820–829

1969 Atmospheric teleconnections from the equatorial Pacific. *Mon. Weather Rev.*, **97,** 163–172

1972 Large-scale atmospheric response to the 1964–65 Pacific equatorial warming. *J. Phys. Oceanogr.*, **2,** 212–217

Busalacchi, A. J. and O'Brien, J. J. 1981 Interannual variability of the equatorial Pacific in the 1960's. *J. Geophys. Res.*, **86,** 10 901–10 907

Cane, M. A. 1983 Oceanographic events during El Niño. *Science*, **222,** 1189–1194

1986a El Niño. *Annual Review of Earth and Planetary Sciences*, **14,** 43–70

1986b 'Introduction to Ocean Modelling', in *Advances in Physical Oceanographic Numerical Modelling.* Editor J. J. O'Brien. Reidel Press, Dordrecht, Holland

Cane, M. A. and Sarachik, E. S. 1977 Forced baroclinic ocean motion. II: The equatorial bounded case. *J. Mar. Res.*, **35,** 395–432

1981 The response of a linear baroclinic equatorial ocean to periodic forcing. *ibid.*, **39,** 651–693

Cane, M. A. and Zebiak, S. E. 1985 A theory for El Niño and the Southern Oscillation. *Science*, **228**, 1085–1087

Cane, M. A., Zebiak, S. E. and Dolan, S. C. CZD 1986 Experimental forecasts of El Niño. *Nature*, **321**, 827–832

Charney, J. G. 1948 On the scale of atmospheric motions. *Geofys. Publ.*, **17**, 1–17

Charney, J. G., Fjortoft, R. and von Neumann, J. 1950 Numerical integration of the barotropic vorticity equation. *Tellus*, **2**, 237–254

Cornejo-Garrido, A. G. and Stone, P. H. 1977 On the heat balance of the Walker Circulation. *J. Atmos. Sci.*, **34**, 1155–1162

Geisler, J. E. and Stevens, D. E. 1982 On the vertical structure of damped steady circulation in the tropics. *Q. J. R. Meteorol. Soc.*, **108**, 87–93

Gill, A. E. 1980 Some simple solutions for heat-induced tropical circulation. *ibid.*, **106**, 447–462

Goldenberg, S. B. and O'Brien, J. J. 1981 Time and space variability of tropical Pacific wind stress. *Mon. Weather Rev.*, **109**, 1190–1207

Halpern, D. and Harrison, D. E. 1982 'Intercomparison of tropical Pacific mean November 1979 surface wind fields'. Report 82-1, Department of Meteorology and Physical Oceanography, Massachusetts Institute of Technology, Cambridge, U.S.A.

Inoue, M. and O'Brien, J. J. 1984 A forecasting model for the onset of a major El Niño. *Mon. Weather Rev.*, **112**, 2326–2337

Julian, P. R. and Chervin, R. M. 1978 A study of the Southern Oscillation and Walker Circulation phenomena. *ibid.*, **106**, 1433–1451

Lau, K-M. W. 1985 Modeling the seasonal dependence of the atmospheric response to observed El Niños in 1962–1976. *ibid.*, **113**, 1970–1996

Lau, K-M. W. and Chan, P. 1985 Aspects of the 40–50-day oscillation during the northern winter as inferred from outgoing longwave radiation. *ibid.*, **113**, 1889–1909

McCreary, J. P. 1983 A model of tropical ocean–atmosphere interaction. *ibid.*, **111**, 370–387

1985 Modeling equatorial ocean circulation. *Ann. Rev. Fluid Mech.*, **17**, 359–409

McWilliams, J. C. and Gent, P. R. 1978 A coupled air–sea model for the tropical Pacific. *J. Atmos. Sci.*, **35**, 962–989

Matsuno, T. 1966 Quasi-geostrophic motions in the equatorial area. *J. Meteorol. Soc. Jap.*, **44**, 25–43

Pazan, S. E. and White, W. 1986 Off-equatorial influence upon Pacific equatorial dynamic height variability during the 1982–83 ENSO event. *J. Geophys. Res.*, **91**, 8437–8449

Philander, S. G. H. 1979 Variability of the tropical oceans. *Dyn. Atmos. Ocean.*, **3**, 191–208

1985 El Niño and La Nina. *J. Atmos. Sci.*, **42**, 2652–2662

1987 'General circulation models of the oceans'. Pp. 103–114 in *Atmospheric and oceanic variability*. Royal Meteorological Society, Bracknell

Philander, S. G. H., Yamagata, T. and Pacanowski, R. C. 1984 Unstable air–sea interactions in the tropics. *ibid.*, **41**, 604–613

Phillips, N. A. 1970 Models for weather prediction. *Ann. Rev. Fluid Mech.*, **2**, 251–292

Ramage, C. S. 1977 Sea surface temperature and local weather. *Mon. Weather Rev.*, **105**, 540–544

1986 El Niño. *Scientific American*, **254**, 55–61

Ramage, C. S. and Hori, A. M. 1981 Meteorological aspects of El Niño. *Mon. Weather Rev.*, **109**, 1827–1835

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

Rasmusson, E. M. and Wallace, J. M. 1983 Meteorological aspects of the El Niño/Southern Oscillation. *Science*, **222**, 1195–1202

Richardson, L. F. 1922 *Weather Prediction by Numerical Process*. Cambridge University Press, London; reprint, Dover Publication, New York, 1965

Shukla, J. and Wallace, J. M. 1983 Numerical simulation of the atmospheric response to equatorial Pacific sea surface temperature anomalies. *J. Atmos. Sci.*, **40**, 1613–1630

Vallis, G. K. 1986 El Niño: A chaotic dynamical system? *Science*, **232,** 243–245

Weare, B. C. 1983 Interannual variation in net heating at the surface of the tropical Pacific Ocean. *J. Phys. Oceanogr.*, **13,** 873–885

1986 A comparison of shallow water model results for three estimates of a composite El Niño forcing. *J. Atmos. Sci.*, **43,** 162–170

White, W. B., Pazan, S. E. and Inoue, M. 1987 Hindcast/forecast of ENSO events based upon the redistribution of observed and model heat content in the western tropical Pacific, 1964–1986. *J. Phys. Oceanogr.*, **17**, 264–280

Wright, P. B. 1985 The Southern Oscillation: An ocean–atmosphere feedback system? *Bull. Am. Meteorol. Soc.*, **66,** 398–412

Wyrtki, K. 1985 Water displacements in the Pacific and the genesis of El Niño cycles. *J. Geophys. Res.*, **90,** 11 710–11 725

Zebiak, S. E. 1982 A simple atmospheric model of relevance to El Niño. *J. Atmos. Sci.*, **39,** 2017–2027

1984 'Tropical atmosphere–ocean interaction and the El Niño/Southern Oscillation phenomenon'. Ph.D. thesis, Massachusetts Institute of Technology

1986 Atmospheric convergence feedback in a simple model of El Niño. *Mon. Weather Rev.*, **114,** 1263–1271

Zebiak, S. E. and Cane, M. A. 1987a A model ENSO. *ibid.*, **115,** (in press)

1987b A simulation of sea surface temperature anomalies during El Niño. *J. Phys. Oceanogr.* (to appear)