

Predictability Loss in an Intermediate ENSO Model due to Initial Error and Atmospheric Noise*

ALICIA R. KARSPECK, ALEXEY KAPLAN, AND MARK A. CANE

Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York

(Manuscript received 17 March 2005, in final form 3 August 2005)

ABSTRACT

The seasonal and interannual predictability of ENSO variability in a version of the Zebiak–Cane coupled model is examined in a perturbation experiment. Instead of assuming that the model is “perfect,” it is assumed that a set of optimal initial conditions exists for the model. These states, obtained through a nonlinear minimization of the misfit between model trajectories and the observations, initiate model forecasts that correlate well with the observations. Realistic estimates of the observational error magnitudes and covariance structures of sea surface temperatures, zonal wind stress, and thermocline depth are used to generate ensembles of perturbations around these optimal initial states, and the error growth is examined. The error growth in response to subseasonal stochastic wind forcing is presented for comparison.

In general, from 1975 to 2002, the large-scale uncertainty in initial conditions leads to larger error growth than continuous stochastic forcing of the zonal wind stress fields. Forecast ensemble spread is shown to depend most on the calendar month at the end of the forecast rather than the initialization month, with the seasons of greatest spread corresponding to the seasons of greatest anomaly variance. It is also demonstrated that during years with negative (and rapidly decaying) Niño-3 SST anomalies (such as the time period following an El Niño event), there is a suppression of error growth. In years with large warm ENSO events, the ensemble spread is no larger than in moderately warm years. As a result, periods with high ENSO variance have greater potential prediction utility.

In the realistic range of observational error, the ensemble spread has more sensitivity to the initial error in the thermocline depth than to the sea surface temperature or wind stress errors. The thermocline depth uncertainty is the principal reason why initial condition uncertainties are more important than wind noise for ensemble spread.

1. Introduction

It is common practice in the field of numerical weather prediction (NWP) to equate confidence in a model prediction with its sensitivity to initialization errors. This is a useful supposition because the weather system is thought to be chaotic, and thus future states are highly sensitive to small differences in the initial state (Lorenz 1963). That weather forecasting systems are also dimensionally very large (a typical model state has a dimension of order 10^6 or higher) makes the ac-

curate specification of the initial condition that much more challenging. As a result, ensemble forecasting is becoming the operational standard in weather prediction. Forecasts are made from a distribution of initial conditions that span the observational uncertainty. The relative spread of the ensemble forecasts at a later time is then considered to be an indication of the confidence in a given event outcome.

This reasoning relies on the assumption that the uncertainty in a forecast outcome can be distilled into the uncertainty in the initial state. This is the equivalent of assuming that the forecasting models are without deficiencies (and that stochastic processes are not important). While there have been arguments that this is a poor assumption (e.g., Swanson et al. 2000; Orrell et al. 2001), the study of forecast error in NWP has recently focused more on chaos-related uncertainty (e.g., Palmer 2000) than on model error.

However, it is difficult to ignore model error in sea-

* Lamont-Doherty Earth Observatory Contribution Number 6884.

Corresponding author address: Alicia R. Karspeck, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307.

E-mail: aliciak@ucar.edu

sonal-to-interannual forecasting of the El Niño–Southern Oscillation (ENSO). For an illustration of this, one need only look to the range of operational climate forecasting models, which are an embodiment of the most insightful scientific understanding of the system. Intercomparison studies, such as Barnston et al. (1999) and Kirtman et al. (2001) cite large differences between the forecasts of tropical Pacific sea surface temperature made with different models.

There is little disagreement that ENSO arises as a result of coupled feedbacks between the ocean and atmosphere. The fundamental physical processes that govern the growth and decay of ENSO events are summarized in the delayed-oscillator theory, first put forth by Suarez and Schopf (1988) and Battisti and Hirst (1989). Other theories of ENSO behavior are discussed in Wang (2001).

What is still a matter of considerable debate is the intensity of the ocean–atmosphere feedbacks. Fedorov et al. (2003) showed that as the strength of the ocean–atmosphere coupling increases, the nature of the oscillation is altered. In the limit of low coupling, the system is strongly damped and stochastic “noise” is necessary to initiate an ENSO event. What we classify as noise are a host of phenomena that are separated from the ENSO dynamics (as we understand them) by either space or time scales. These smaller-scale dynamics are not resolved in most intermediate-complexity coupled models and must be treated as an external forcing. Typically, these higher-frequency signals are thought to be generated in the atmosphere, but there is some evidence that tropical instability waves in the ocean could also be considered a source of high-frequency system noise (Jochum and Murtugudde 2004). This “internal variability” of the atmosphere is defined as behavior that is uncoupled from the ocean. This issue is complicated by evidence suggesting that some forms of high-frequency variability are modulated by the larger ENSO patterns (Batstone and Hendon 2005). Atmospheric general circulation models (GCMs) should, in theory, be able to reproduce this noise, as they contain more comprehensive dynamics. Unfortunately, this fact alone does not necessarily lead to more realistic temporal ENSO variability (Deser et al. 2005).

If a strong coupling between the ocean and atmosphere is assumed, the ENSO system behaves in a chaotic fashion—oscillations are irregular and self-sustained through the nonlinearities in the model dynamics. As such, although the system may respond to external forcings, it is not essential for simulating the large-scale variability.

This issue is of practical importance for ENSO prediction strategies. If the system is a damped oscillator,

then the expectation is that the predictability of the system is limited by the timing, intensity, and spatial structure of the atmospheric noise that imposes itself on the system. ENSO predictability under this paradigm has been addressed in a number of studies (e.g., Penland and Sardeshmukh 1995; Kleeman and Moore 1999; Thompson and Battisti 2000; Flugel et al. 2004). The paradigm of a stochastically forced ENSO system was introduced in the literature a decade ago (e.g., Penland and Magorian 1993; Penland and Sardeshmukh 1995) and has persisted as an explanation for the initiation of ENSO events. In particular, this view has gained support since the 1997/98 El Niño event, which most models failed to forecast prior to the preceding summertime (Fedorov et al. 2003). One of the hallmarks of this event was a series of strong (and unpredicted) westerly wind bursts in the western and central Pacific, which were thought to initiate the eastern equatorial SST warming (McPhaden 1999). There have been a host of other studies indicating that an essential component of the stochastic forcing comes in the form of “subseasonal” wind variability (e.g., Kessler and Kleeman 2000; Moore and Kleeman 1999a,b; Vecchi and Harrison 2000). This variability is thought to be driven in part by the Madden–Julian oscillation (MJO; Madden and Julian 1994).

On the other extreme, if the system dynamics are chaotic (as is the case in weather prediction), then we expect the predictability to be limited by the amplification of errors in the initial conditions. This regime has been studied by Goswami and Shukla (1991), Chen et al. (1997), Xue et al. (1997a,b), Karspeck et al. (2004), and others.

In either case, most of the predictability studies in the literature have taken the form of “perfect model” studies. In experiments such as these, the “truth” is assumed to be a trajectory that a model is capable of producing exactly. Deviations from this true trajectory are forced by either perturbations to the model initial conditions or by the continuous application of stochastic forcing to the model. The metric of predictability is defined as some measure of the dispersion of these perturbed trajectories about the idealized trajectory. In the most idealistic of these, the “true” trajectory is generated from a long run of a model that is not constrained by the observed ENSO system (e.g., Kleeman and Moore 1999; Xue et al. 1997a). Other studies have aimed at estimating the actual (as opposed to the theoretical) model predictability using forecast models that are initialized using observations of the real system (Fan et al. 2000; Xue et al. 1997b; Chen et al. 1997). In these studies, the predictive skill of the forecast model forms part of the argument for the relevance of the study.

This study focuses on the problem of evaluating predictability loss in the context of an intermediate coupled ENSO forecasting model that was used operationally at the Lamont-Doherty Earth Observatory from 1999 to 2003 (LDEO4 is described in Chen et al. 2000). LDEO4 is a version of the Zebiak and Cane model (Zebiak and Cane 1987, hereafter ZC) that contains a statistical bias correction that works interactively with the model dynamics. The bias correction scheme serves to bring the model spatial variability in closer agreement with observations, making the assimilation of observations possible. The LDEO4 model resides in a parameter regime that is characterized by self-sustained variability. It is often cited as an example of a chaotic system that is highly sensitive to errors in initial conditions (e.g., Fedorov et al. 2003).

Here, we look at the predictability in the context of a semi-idealized experiment: instead of assuming that LDEO4 is a perfect model, we assume instead that we have access to a set of “best” initial conditions (henceforth referred to as best IC, following Cañizares et al. 2001). These states, arrived at through a low-dimensional inverse method (also called an “adjoint method”), begin model forecasts that correlate well with the observations. By perturbing these initial states in accordance with realistic observational error estimates, we can look at error growth along model trajectories that are tracking the observed system. In this way, the sensitivity of the model to errors is more likely to be consistent with that of a realistic and skillful prediction system. Additionally, we evaluate the effect that internal atmospheric variability has on predictability. In our experiments this missing variability is simulated as stochastic zonal wind forcing. We compare it to the predictability loss due to perturbations of initial conditions.

In section 2 we briefly present the adjoint method used to find the best IC for the LDEO4 model. Section 3 defines the ensemble prediction experiments that are carried out for perturbation of initial conditions by realistic error covariance estimates and for stochastic perturbations of the model trajectory with zonal wind noise. Section 4 describes the results of the predictability experiment by exploring the ensemble spread over a 1-yr forecast period as a function of season, ENSO cycle, and error parameters. Finally, section 5 discusses the results and the consequences for model predictions and predictability.

2. Finding a set of best initial conditions

In this section we describe a low-dimensional adjoint approach for finding LDEO4 trajectories that are a least squares best fit to observations of SST anomalies,

thermocline depth anomalies, and wind stress anomalies.

In a multivariate empirical orthogonal (MEOF) space that spans the coupled modes of variability of LDEO4, we seek to minimize the cost function:

$$J = \sum_{k=0}^m (\mathbf{x}_k - \mathbf{o}_k)^T (\mathbf{x}_k - \mathbf{o}_k), \quad (1)$$

subject to the dynamical constraint:

$$\mathbf{x}_{k+1} = \mathcal{N}\mathbf{x}_k. \quad (2)$$

Here \mathbf{x}_k and \mathbf{o}_k are vectors of model states and observational states (respectively) within the analysis window of length m months, and \mathcal{N} represents a 1-month integration of the LDEO4 model. Since the model dynamics are a strong constraint in the minimization, the entire problem can be reduced to finding \mathbf{x}_0 , the model state at the beginning of the analysis window.

To reduce the computational burden of the problem, both the model and observational states have been projected into the MEOF space of the model prior to the minimization. If \mathcal{N} were linear, the precise \mathbf{x}_0 that minimizes J could be easily found using matrix algebra. Cañizares et al. (2001) approximated the operation of \mathcal{N} in the reduced space by 12 seasonally dependent linear Markov models and found \mathbf{x}_0 for this case. Here we strive to represent the nonlinear nature of \mathcal{N} more precisely by the interactive computation of perturbation forecast models (PFMs; Lorenz 1965; Xue et al. 1997a,b). PFMs are numerical approximations to the tangent linear models. As their name implies, the PFMs evolve small perturbations along some background trajectory of the full, nonlinear model \mathcal{N} . By choosing the perturbations in the direction of the leading MEOFs of the model, relatively few perturbations can be used to (approximately) map the linearization of the full model into a matrix operator of manageable size. The sequential multiplication of k PFMs approximates a k -month integration of the tangent linear model.

The number of MEOFs to retain in the projection of observations and the calculation of PFMs is a practical consideration; the increased computational cost of retaining more modes must be balanced against the performance improvement that it could yield. Karspeck (2004) shows that in MEOF-based Markov model forecasts, retaining more than 30 MEOFs does not result in greater agreement between the low-order models and the full LDEO4. We also expect that the higher modes cannot be specified from observations and there is little to gain in including them (Cane et al. 1996). Henceforth, the dimension of the reduced state space is 30.

The minimization of J can be thought of as a series of incremental adjustments to the model state at the be-

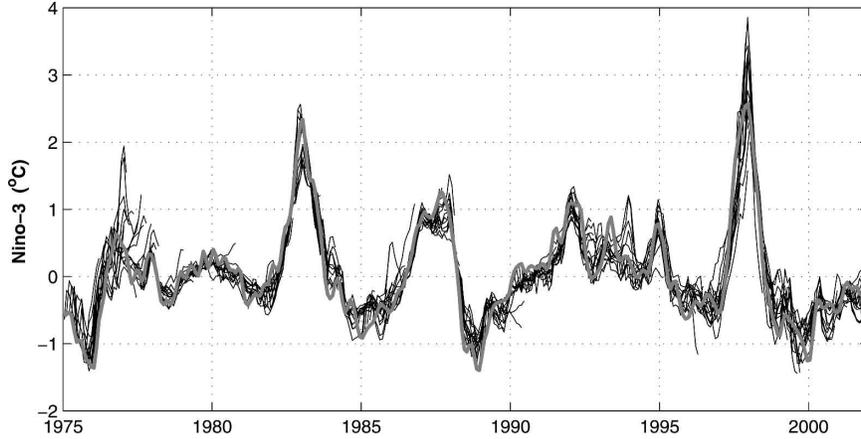


FIG. 1. Twelve-month LDEO4 forecasts initialized with best IC from 1975 to 2002. Initial states are found using the adjoint method. The thick gray line corresponds to the observed Niño-3 index.

ginning of the analysis window. We apply a gradient descent method, such that the adjustments (or analysis increments) are in a direction that tends to reduce the difference between the model forecast and the observations. A first guess can be obtained by using MEOF-based Markov models, trained on a long run of LDEO4, to constrain the minimization of J . Then the n th incremental correction in the minimization is given by

$$\delta \mathbf{x}_0^n = \left(\sum_{k=0}^m \mathbf{L}_k^T \mathbf{L}_k \right)^{-1} \sum_{k=0}^m (\mathbf{L}_k^T \delta_k^n); \quad \delta_k^n \equiv \mathbf{o}_k - \mathcal{N}_k \mathbf{x}_0^n, \quad (3)$$

where \mathbf{L}_k is the PFM approximation to the tangent linear from 0 to k and superscript T denotes matrix transposition. The δ_k are the innovations (i.e., the difference between the state that the model has forecast and the observations at that time). The new analysis vector at each iteration is

$$\mathbf{x}_0^{n+1} = \mathbf{x}_0^n + \delta \mathbf{x}_0^n. \quad (4)$$

At every iteration, the full model is run with the current estimate of the initial conditions that will lead to the best fit with the observations. New PFMs are constructed with respect to this trajectory and we use (3) to transform the observation/forecast mismatch into our next analysis increment. This is repeated until convergence.

The data that are used for the assimilation and the verification of model performance are from the following sources: 1) zonal and meridional components of wind stress from The Florida State University (Stricherz et al. 1997); 2) SST anomalies from the Climate Prediction Center of the National Centers for Environmental

Prediction (NCEP); and 3) analyzed ocean fields obtained from the assimilation of tide gauge data and satellite altimetry into the Cane–Patton model (Cane and Patton 1984) using a reduced space Kalman filter (Cane et al. 1996; Kaplan et al. 2004). Note that these are analysis products and not raw observations. The analyzed fields of the SST and wind stress anomalies were interpolated onto the ZC grid. The ocean fields were originally produced on the grid used in LDEO4. The observed fields were then projected on the reduced space spanned by the leading 30 MEOFs (see appendix for details). Henceforth the references to the “observations” are to the MEOF projection of the observations of SST, ocean states, and wind stress.

The minimization of (1) subject to the constraint (2) was done for every start month from 1975 to 2002, using a 12-month analysis window. The initial model states determined through this procedure will be referred to as “best IC” and designated by x^* . They are best in the sense that LDEO4 trajectories initialized with these states should agree well with the observations in a least squares sense. These states give good forecasts by using future states as a constraint, so this method does not represent a real-time forecasting procedure. The model trajectories do not perfectly shadow the observations because our model is not a perfect representation of the observed ENSO system and because the minimization is performed in a reduced state space on a nonlinear model. However, the skill the model exhibits when initialized with the best IC is near the upper bound of the model’s forecasting abilities.

Figure 1 shows 12-month LDEO4 trajectories of the Niño-3 index initialized with x^* at each month from 1975 to 2002. Superimposed in bold gray is the histori-

cal record of the Niño-3 index. The Niño-3 index is defined as the average SST anomaly in the region 5°S–5°N, 150°–90°W. This is computed from the monthly SST analysis by Kaplan et al. (1998, 2003). Over all calendar start months, these trajectories correlate with the observations with coefficients greater than 0.85 at all leads from 0 to 12 months. Note that all of the ENSO events are captured by these initial conditions. There are only two false forecasts, in 1976 and 1994. Generally, the incremental formulation of the low-dimension inversion employed here tends to give poor results when the system behaves in a highly nonlinear fashion. As a result, we expect that there will be instances when the technique is unsuccessful.

3. Experimental setup

a. Perturbation of best initial conditions

A 500-member ensemble of LDEO4 forecasts was initialized at each month from 1975 to 2002. For each member of the ensemble, the best ICs were perturbed with a random sample (zero mean, normally distributed) from the large-scale observed error covariance estimates of SST, wind stress, and thermocline depth (described in the appendix). Thus, at each start month the initial ensemble perturbation covariance is consistent with the error covariance estimates. Each ensemble member was integrated forward for 12 months. There is no temporal correlation between perturbations in adjacent start months, nor do we assume any seasonal variations in the error patterns.

b. Perturbation with continuous, stochastic wind forcing

Here we consider the effect that stochastic zonal wind forcing has on the spread of LDEO4 trajectories. It should be noted that the anomaly variance in LDEO4 is self-sustained through nonlinearities in the model dynamics, and as such the model does not require any external stochastic forcing.

An ensemble of LDEO4 forecasts (each 12 months in length) was initialized with the best IC at every month from 1975 to 2002. At every 10-day interval (the time step of the model), a perturbation was added to the zonal wind stress field. The perturbation was held constant for a total of 30 days (three time steps), and then a new temporally uncorrelated perturbation was applied. This was repeated for the duration of the 12-month forecasts. The perturbation fields were uncorrelated between model points, drawn from a zero-mean Gaussian distribution. Since the atmospheric component of the model has $5^\circ \times 2^\circ$ resolution, the noise is roughly equivalent to mesoscale variability with decor-

relation scales of 500 and 200 km in the zonal and meridional directions. The standard deviation of the perturbation has a maximum at the equator, falling off with latitude with an e -folding distance of 6° (recommended by Kessler and Kleeman 2000).

It is not altogether clear what the proper magnitude of the perturbation should be. Hourly wind data from the European Centre for Medium-Range Weather Forecasts (ECMWF) show that the gridpoint scale distribution of zonal wind anomalies in the central equatorial Pacific has a standard deviation of $\sim 2 \text{ m s}^{-1}$ (G. Vecchi 2004, personal communication). In their studies of the effect of intraseasonal variability on ENSO, Zebiak (1989) uses wind stress perturbations that map to wind speeds of $\sim 1.5 \text{ m s}^{-1}$ and Kessler and Kleeman (2000) estimate wind perturbations of $\sim 3.5 \text{ m s}^{-1}$, for a coherent temporal and spatial pattern. As pointed out by Moore and Kleeman (1999a, hereafter MK99), much of the coherent intraseasonal variability in the observations can be generated by the model dynamics within a few weeks. (The intermediate coupled model used in their study is dynamically similar in most respects to the ZC model.) In essence, low-magnitude stochastic perturbations to the model can grow and develop into the coherent structures that we associate with intraseasonal wind patterns. Because of this, only a fraction of the observed intraseasonal wind variability can be treated as internal atmospheric variability. MK99 estimate that the standard deviation of uncoupled atmospheric noise that projects onto the MJO-type structure is only $\sim 0.4 \text{ m s}^{-1}$. As we will discuss in section 4a, this is consistent with a grid-scale forcing amplitude in our model of $\sim 4 \text{ m s}^{-1}$. While this is our best estimate of the appropriate standard deviation, we explore a generous range from 2 to 10 m s^{-1} .

c. Measuring predictability

A number of metrics have been used in the literature to quantify predictability. Most of them are related in some way to the size of the ensemble spread relative to the variance of the system, or the “signal-to-noise” ratio. There are also a host of predictability measures that can be useful for looking at the utility of individual forecasts (e.g., Kleeman 2002; Schneider and Griffies 1999; DelSole 2004). In this study we use two measures to examine predictability. The forecast ensemble spread is used as a simple indicator of our confidence in a given forecast and the “potential prediction utility” (PPU; Kleeman and Moore 1999) is used as a measure of the usefulness of the forecast. The PPU can be shown to be analogous to a correlation coefficient in the sense that it takes into consideration the magnitude of the signal.

Of course, to measure ensemble spread, a choice must be made as to which variables we think best define the state of the system. This choice ultimately depends on the purpose of the experiment. Since we are motivated by the practical problem of predicting the ENSO system, we choose the Niño-3 index, as it is commonly used to verify operational forecast systems. It is worth noting, however, that there are other measures that may be as useful for categorizing the state of the system. The zonal average of the equatorial upper-ocean heat content, for example, is highly correlated with the Niño-3 index at a 7–9-month lead time (McPhaden 2003).

The straightforward measure of forecast ensemble spread is defined as

$$\sigma_{\epsilon}^2(t, k) \equiv \frac{1}{N} \sum_{i=1}^N [f_i(t, k) - \bar{f}(t, k)]^2, \quad (5)$$

where $f_i(t, k)$ are members of the ensemble forecasts at start times t and forecast lengths k , and $\bar{f}(t, k)$ is the ensemble mean over the N ensemble members. The forecasts made from the best IC do not differ significantly from \bar{f} . In theory, LDEO4 will correctly simulate the variance of the ENSO system regardless of the forecast lead time. However, there are small differences between the forecast variance at leads up to 12 months. These are less than 10% of the system variance. Nonetheless, we postprocess the ensemble members by scaling their variances to match the observations. This processing did not alter the results presented in this study.

Kleeman (2002) and Kleeman and Moore (1999) make note of the fact that an ensemble prediction with a large spread may still have considerable utility if one is forecasting a relatively large signal. The potential prediction utility, which takes this into consideration, can be expressed as

$$\text{PPU}(t, k) = \frac{1}{1 + \sigma_{\epsilon}^2(t, k)/\bar{f}(t, k)^2}. \quad (6)$$

Like the commonly used correlation coefficient, the PPU also varies from zero to one, with a value of one being a perfect forecast. The PPU naturally tends toward unity as the ensemble spread tends to zero. Using this measure, it is easy to see that the condition for a useful prediction is not simply low ensemble spread. Predictions can be equally useful when the forecasted signal is very large.

An issue to keep in mind as the results are presented is that the predictability can be viewed from two different perspectives: as a function of the time at which the forecast was initialized (i.e., “initialization time”), or as a function of the time at the end of the forecast (i.e.,

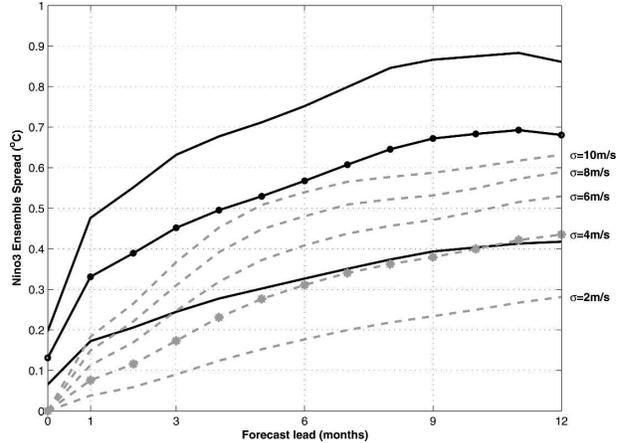


FIG. 2. The ensemble spread in LDEO4 forecasts as a function of lead time (averaged from 1975 to 2002). The dark solid line marked with filled circles corresponds to the spread generated with our standard estimate of the magnitude of the initialization error. The other two dark solid lines show the spread when the initialization error is reduced or increased by 50%. Dashed gray lines correspond to ensembles generated by initializing the model with the best IC and then perturbing the model trajectory with stochastic zonal wind noise. The std devs of the wind perturbations are indicated on the left-hand side of the plot. Our most realistic estimate for wind forcing is marked with stars.

“verification time”). Categorizing the results in terms of the initialization state can seem like a natural choice, and if a forecast length is specified a priori, then this will appear to be a reasonable categorization scheme. However, this practice would be misleading. In the predictability experiment done here it masks the primary influence on the ensemble spread: the state of the system at the end of the forecast (see section 4b). Therefore, the results here are presented as a function of the verification time.

4. Results: LDEO4 predictability from 1975 to 2002

a. Uncertainty in initial conditions versus continuous stochastic forcing

The dot-marked solid black line in Fig. 2 shows the spread of the forecast ensemble as a function of forecast lead time averaged over all start months using our best estimate of the magnitude of the initial error. The other two solid black lines correspond to the spread when the initial error is increased and reduced by 50%.

For comparison, the dashed lines show the resulting ensemble spreads when stochastic zonal wind stress is applied continuously to the model forecast initialized from the best IC. On the low end of our estimate of the magnitude of the forcing ($\sigma = 2 \text{ m s}^{-1}$) it is clear that

the stochastic wind forcing is not very effective in generating forecast spread in this model. Even with the highest value shown, $\sigma = 10 \text{ m s}^{-1}$, which is five times the usual estimate for σ , the spread is smaller than with the best estimate of the initialization uncertainty. However, we consider only a spatial structure for which each $5^\circ \times 2^\circ$ grid box is uncorrelated with all other boxes. It is far from clear what the appropriate spatial structure should be for the wind noise.

What matters for the size of the spread is the amplitude of the leading stochastic optimal \mathbf{S} that the wind noise of variance σ^2 excites (MK99). MK99 show that this pattern is reminiscent of MJO episodes. Suppose the typical scale of \mathbf{S} (i.e., the characteristic distance over which it changes sign) is s grid points and the correlation scale of the wind noise is w grid points. Then the stochastic optimal sees roughly s/w independent samples within a characteristic scale. The average of these samples, to which the stochastic optimal effectively responds, has the variance of $\sigma^2 w/s$. This is a factor of w higher than what occurs in the case of uncorrelated noise (when $w = 1$). Therefore, the existence of wind noise correlations over w grid points will raise the ensemble spread by the factor \sqrt{w} . The largest amplitude variance results when $w \approx s$, which increases the ensemble spread by approximately \sqrt{s} . This one-dimensional argument easily carries to higher dimensions. Specifically, if a stochastic optimal has decorrelation scales s_x and s_y in zonal and meridional directions, respectively, the maximum ensemble spread due to the correlated noise can be a factor $\sqrt{s_x s_y}$ larger than for uncorrelated noise. The details of the response depend on the shape of the correlation pattern, but the size of the spread will not be sensitive to it.

If one takes the leading stochastic optimal in MK99 as representative, then the characteristic scale of \mathbf{S} is $\sim 50^\circ \times 20^\circ$; that is, $s_x = s_y = 10$, with $\sqrt{s_x s_y} = 10$. Thus, a 10-fold increase in the standard deviation of the forcing will be equivalent to forcing at the scale of the stochastic optimal. MK99 estimate that the standard deviation of this pattern over the tropical Pacific (that does not result from coupled model feedbacks) is $\sim 0.4 \text{ m s}^{-1}$. Thus, our imposition of a value of $\sigma = 4 \text{ m s}^{-1}$ is consistent with the amplitude of MJO-type forcing studied in MK99, and the higher amplitudes we test easily envelope the possible range of variability. While different spatial structures for the wind noise would yield somewhat different answers, it is unlikely that they could be outside the range shown in Fig. 2.

At any rate, unless the errors are vastly different from our best estimates, it is clear that initialization error is the primary contributor to error growth. However, a reduction in the initialization error of 50% puts

us squarely in the ballpark of the error we might expect from stochastic forcing. These results suggest that while stochastic processes may not be large contributors to predictability loss at this point, they may increase in importance as initialization procedures are improved.

These results are consistent with those of Zebiak (1989), who studied the impact of perturbing the wind stress fields in the original ZC model with intraseasonal variability. He also concluded that intraseasonal variability had only a marginal impact on forecast spread relative to differences in initial state. Based on these results, we restrict our remaining analysis to the loss of predictability due to initialization uncertainty. We will return to stochastic wind error in the discussion.

b. Seasonal cycle of predictability

Viewing the historical spread at lead times from 1 to 12 months, we see that there is a strong seasonal cycle at all leads (Fig. 3). Note that the spreads are not plotted at their initialization time, but at their verification time. To emphasize this point, Fig. 4 shows contours of the ensemble spread as a function of lead time (x axis) and of initialization time (y axis). In general, the verification calendar months of June have the lowest forecast ensemble spread and the calendar months of December (indicated by the dashed line) have the greatest. While there is an overall trend toward greater spread at longer leads, it is far from monotonic, being dominated by a seasonal expansion and contraction of ensemble spread. The idea that the ensemble spread is better understood as a function of the calendar month at the end of the forecast (regardless of the lead time) than as a function of the initialization month departs from what is typically presented in the literature (e.g., Fan et al. 2000; Xue et al. 1997a,b).

The Niño-3 anomalies from the analysis of observations by Kaplan et al. (1998, 2003) for the 1856–2003 period have the same distinct dependence on season (Fig. 5). The seasonal dependence of both the model ensemble spread and the standard deviation of Niño-3 anomalies have a similar “U” shape, with minimum variance in the spring and early summer, and the greatest variance in winter. On average, both model perturbations and observed anomalies of Niño-3 decay in January through March have nearly no growth through the springtime and grow rapidly through the fall and early winter.

c. Predictability and ENSO phase

The strong relationship between the phase of the ENSO cycle and the ensemble spread is evident in Fig. 3. A sharp suppression of error growth coincides with

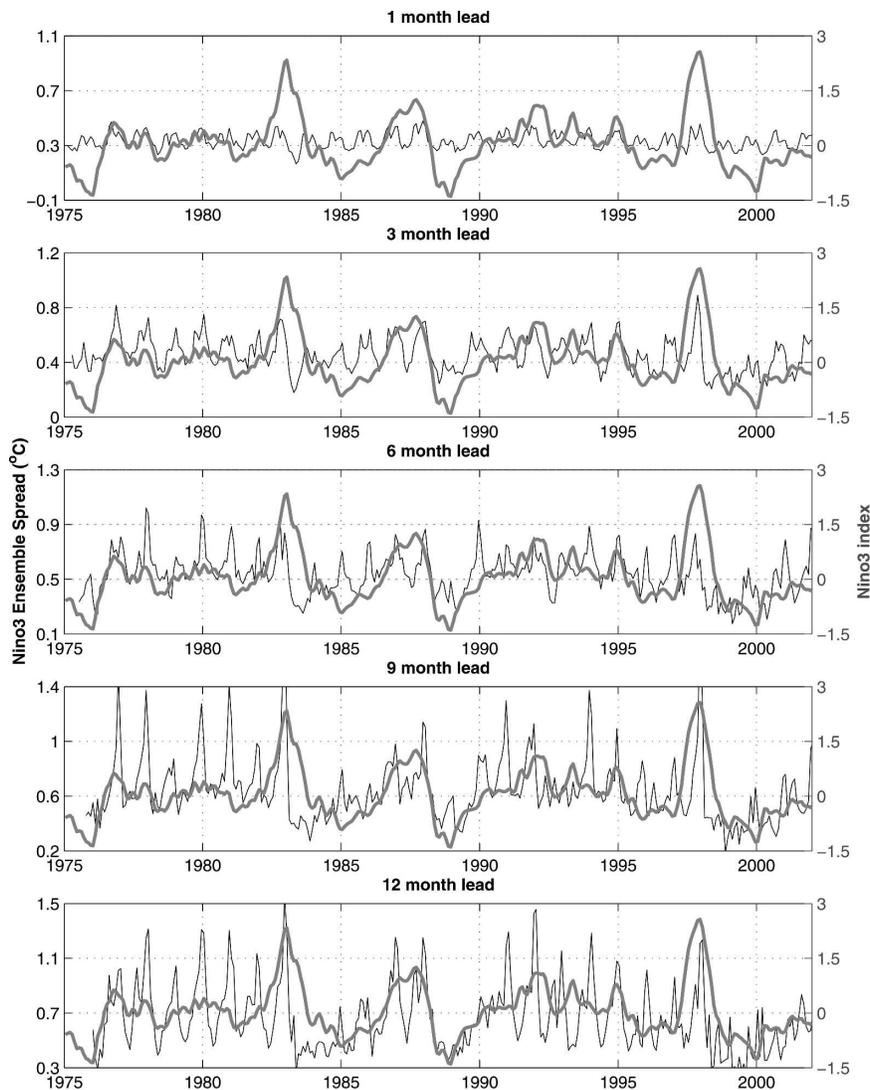


FIG. 3. Ensemble spread in LDEO4 forecasts of Niño-3 from 1975 to 2000 for 1-, 3-, 6-, 9-, and 12-month lead times (in black). Time axis shows the month of verification. The Niño-3 index is plotted in bold gray for reference.

the decay of strong El Niño events and continues through the following cool ENSO phase (La Niña). We see this pattern very clearly in the 1982/83 event, the protracted 1986–88 event, and the 1997/98 event. This is consistent with the perturbation growth suppression during the cool ENSO phase found by Xue et al. (1997a).

The coupled feedbacks between positive SST anomalies and increased surface wind convergence can also be responsible for faster perturbation growth when the SST anomaly is positive. Warmer SSTs induce convergent surface wind fields that, in turn, increase the atmospheric heating, leading to further convergence. This positive feedback enhances the growth of perturba-

tions. This feedback is less active when SST anomalies are negative (Tziperman et al. 1997).

The scatterplot in Fig. 6 illustrates the relationship between the ensemble spread and the sign of the Niño-3 index. Plotted in the figure are the 1-yr running means of the ensemble spread centered at the verification time versus the 1-yr running mean of the observed Niño-3 index. The annual averaging is applied in order to highlight the interannual variability. The spread is low when Niño-3 is negative and higher when Niño-3 is positive. The points with squares correspond to the 1997–98 event (from January 1997 to December 1998), and the points with circles correspond to the 1982–83 event (from January 1982 to December 1983). The lin-

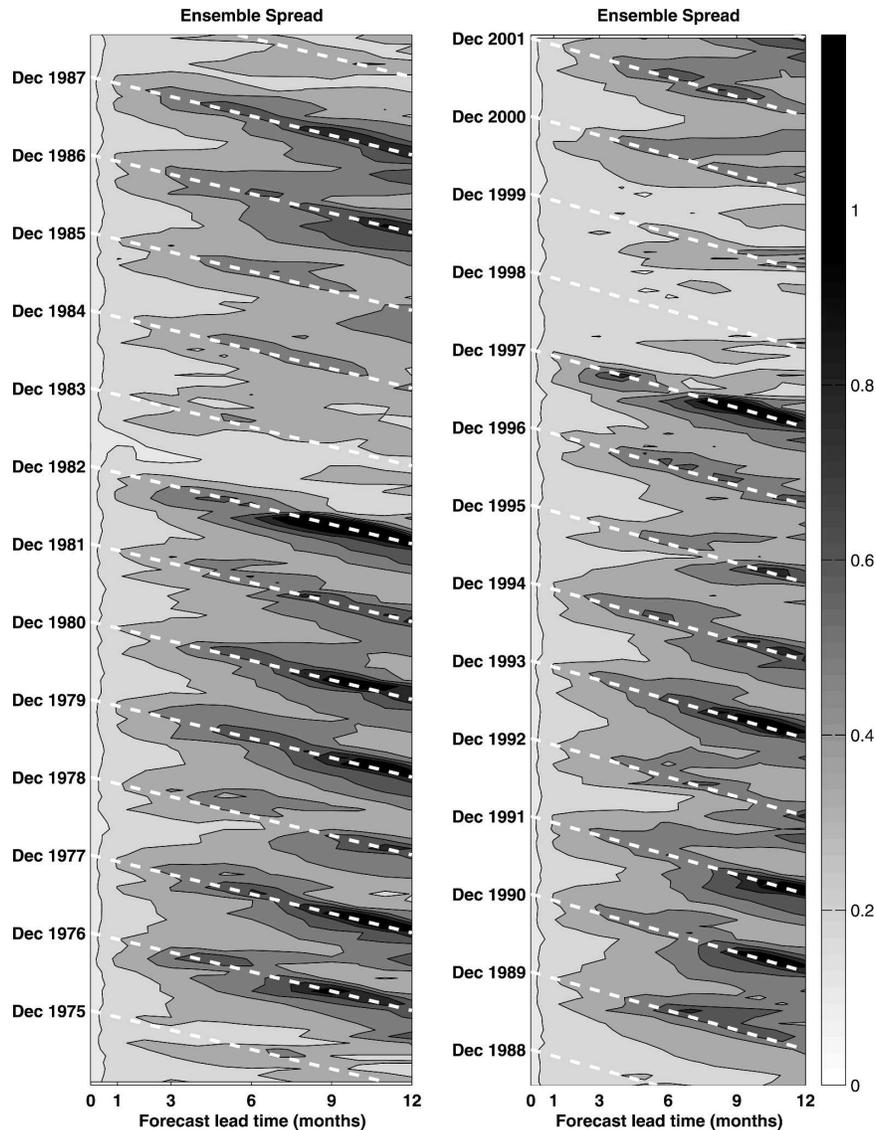


FIG. 4. Contours of the ensemble spread in Niño-3 index as a function of the initialization month and lead time. The dashed white line corresponded to the verification months of December. Units are $^{\circ}\text{C}$.

ear relationship between SST anomalies and the ensemble spread does not hold for larger events. It appears that the spread in Niño-3 saturates at an annual average of $\sim 1^{\circ}\text{C}$, even as Niño-3 continues to grow. This is likely due to the subsurface temperature parameterization in the model. As in the real world, the SST cannot increase indefinitely. When the thermocline is very deep, the time tendency of the surface temperature due to subsurface upwelling is greatly reduced. When we exclude these large events, the annually averaged spread has higher than 0.75 correlation with Niño-3 at all leads.

The use of the PPU [Eq. (6)] as a predictability mea-

sure can help illuminate the relationship between the system state and the ensemble spread that we discuss above. The contour plot in Fig. 7 shows the PPU from 1975 to 2002. PPU values near one are indicative of useful predictions, while values near zero indicate regions where the ensemble predictions may be wildly misleading. Dashed lines have been plotted at the verification month of December. A visual inspection of this plot reveals that once we have normalized the spread by the squared anomaly value (as is done in the PPU measure) there is no longer a robust seasonal cycle. We see, instead, multiyear swaths of high and low predictability.

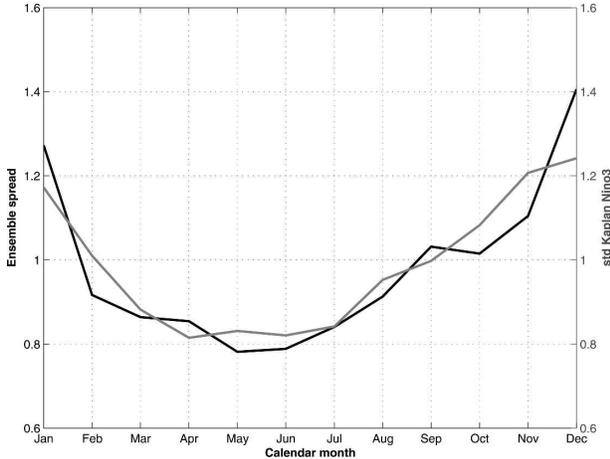


FIG. 5. Gray line: observed standard deviation of Niño-3 anomalies for each calendar month (normalized by the total standard deviation in observed Niño-3). Black line: Ensemble spread of LDEO4 forecasts for each calendar month averaged over all leads and all initialization times and normalized by the total ensemble spread for all months at all leads.

The late 1970s (through 1981) are characterized by low prediction utility. This is attributable to the fact that there is significant wintertime error growth throughout this period while the system remains in a near-neutral state. By construction, PPU gives less credit for reasonable forecasts of normal conditions than for anomalous conditions. As with correlation scores, it implicitly measures against a strawman forecast of climatology. As we saw in Fig. 3 and Fig. 6, the error growth is modulated by the slight positive–negative swings in the Niño-3 index, but these changes are a relatively small percentage of the total wintertime ensemble spread. In practical terms, this period would be one in which there would be a danger of making false forecasts nearly every winter. The 5-yr period from 1990 to 1995 has similar characteristics.

Contrast this to the period 1982 through 1990. This period is characterized by relatively high prediction utility. Even though there is significant wintertime error growth during the warm events of 1982–83 and 1986–88, the utility measure downscales it by the high signal. In the years following these events, the wintertime error growth is largely suppressed. The same relationship holds beginning in 1997, through the growth and decay of the 1997–98 El Niño, following La Niña, and into the beginning of 2001.

d. Error growth and the relative uncertainty in different model fields

We would like to investigate the relative roles of initialization error in SST, thermocline depth, and zonal

wind stress on error growth in LDEO4. We use the reduced-space, empirically calculated linear approximations to LDEO4 (PFMs; described in section 2) in lieu of the full model. The PFMs were constructed along the model trajectories initialized at every month with the best IC. Since each PFM was constructed over a 1-month interval, an m -month forecast is performed by a sequential application of m PFMs.

In the previous experiments, we calculated the ensemble spread by integrating a 500-member ensemble of initial conditions forward using LDEO4. To avoid the computational burden of doing full ensemble integrations, we approximate the spread after m months by projecting the initial error covariance ($\hat{\mathbf{R}}$) forward in time using the PFMs:

$$\hat{\sigma} = \sqrt{\hat{\mathbf{n}}^T \mathbf{L}_m \hat{\mathbf{R}} \mathbf{L}_m^T \hat{\mathbf{n}}}. \quad (7)$$

In this notation \mathbf{L}_m is the sequential multiplication of m PFMs. Multiplication by \mathbf{L}_m simulates an m -month integration of LDEO4 from an initial condition that is in the vicinity of the set of best IC. The vector $\hat{\mathbf{n}}$ is a 30-element column vector that contains the Niño-3 index for each MEOF mode. The bookend multiplication by $\hat{\mathbf{n}}$ and its transpose projects the spread of the MEOF covariance into the spread of Niño-3.

In general, the PFM approximations do a fair job of estimating the seasonal cycle of error growth in LDEO4, and the suppression of growth during time periods of extended negative Niño-3 anomalies (such as after an El Niño event). However, they do not contain the full model nonlinearities that act to limit growth. As a result, their wintertime error growth tends to overestimate the growth in LDEO4. Because of this, we do not wish to make a direct comparison of the temporal average error growth calculated by the PFM and the error growth in LDEO4. Instead, we use the median error growth over all months from 1975 to 2002 as it tends to be in close agreement with the median spread of LDEO4 and is not affected by the unrealistically high error growth in the winter season.

The total error covariance (in MEOF space) can be expressed as the summed contribution of the individual error fields

$$\hat{\mathbf{R}} = a\hat{\mathbf{R}}_w + c\hat{\mathbf{R}}_{sst} + b\hat{\mathbf{R}}_h, \quad (8)$$

where the scalars a , b , and c can be used to vary the magnitude of the error covariances, while keeping the patterns constant. When they are all equal to one, the error covariance is our most realistic estimate of current error levels (explored in the previous sections). We vary a , b , and c from 0 to 2 in increments of 0.25.

Figure 8 shows the level contours of the PFM-

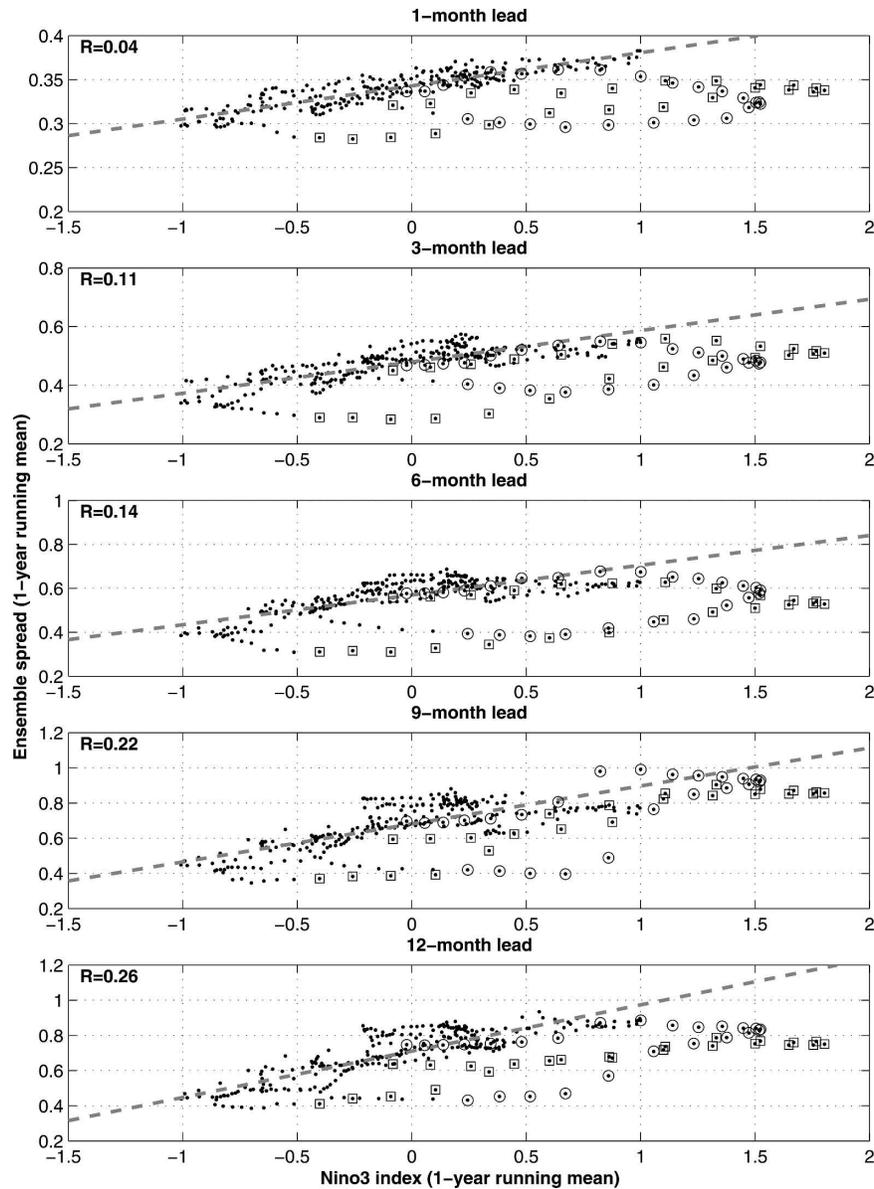


FIG. 6. Scatterplots of the 1-yr running mean of the ensemble spread versus the 1-yr running mean value of the Niño-3 index. The points are centered at their verification time, not their initialization time. The points with squares coincide with the 1997–98 event (from January 1997 to December 1998), and the points with circles coincide with the 1982–83 event (from January 1982 to December 1983). Linear regression coefficients are given in upper-left-hand corner. Units are $^{\circ}\text{C}$.

generated ensemble spread in Niño-3 (the median value over the 27-yr record) at a 6-month lead. The levels range from 0.1° (blue) to 1.0°C (red) in intervals of 0.05°C . Altering the magnitude of the errors in thermocline depth has the greatest impact on the 6-month ensemble spread. On the other hand, the spread is relatively insensitive to changes in the SST initial error covariance. The spread is only sensitive to variations in

wind stress error when the thermocline depth error is small. In Fig. 9 this can be seen in a simplified format. In Fig. 9 (top), each dark line corresponds to holding two of the error fields constant (with their scaling parameter equal to one), while the remaining one is allowed to vary. For comparison, the gray line represents the spread if the magnitudes of the fields were scaled simultaneously.

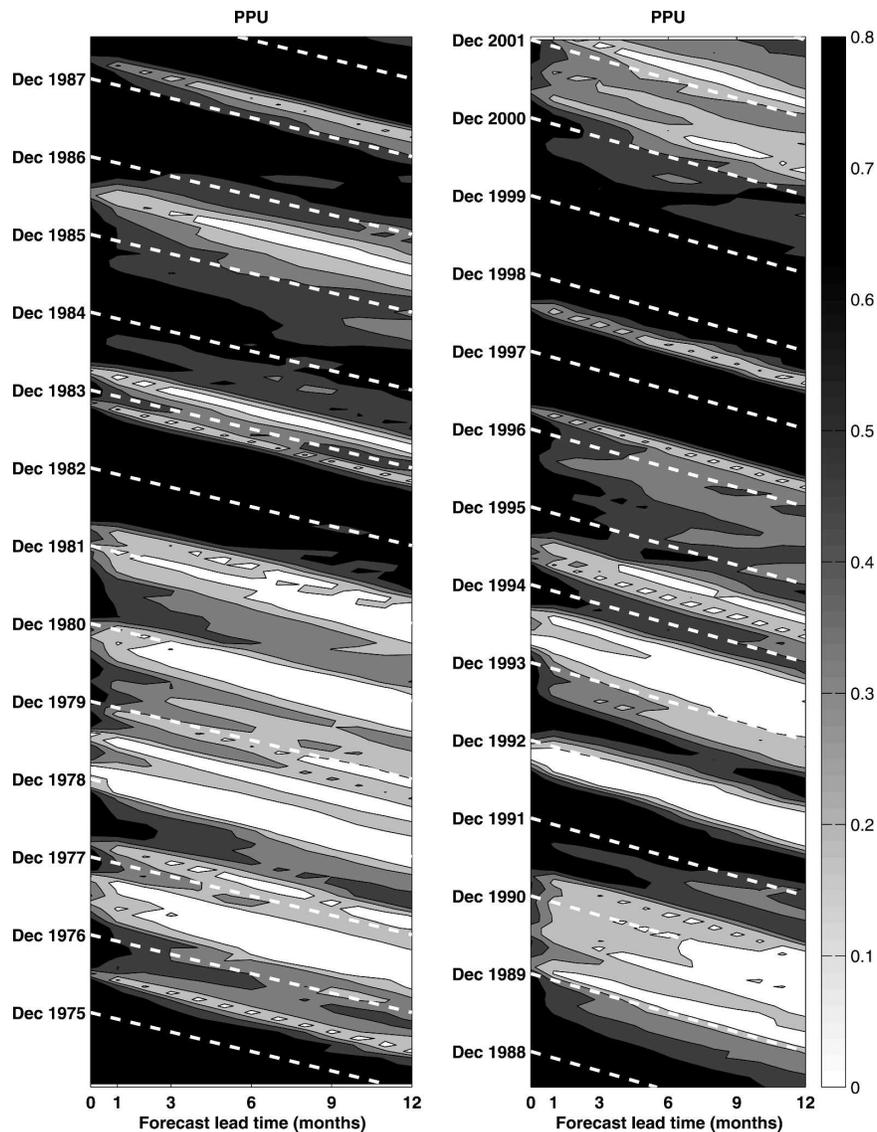


FIG. 7. Contours of the potential prediction utility in Niño-3 index as a function of the initialization month and lead time. The dashed white line corresponded to the verification months of December.

In Fig. 9 (bottom) the dark lines correspond to holding two of the error fields to zero, as the magnitude of the third is varied. Again, the solid gray line shows the spread if the magnitudes of the fields were scaled simultaneously. The remaining gray dashed/dot-marked line shows the spread if only the SST and wind stress error fields were varied. We can see from this plot that if there was no error in the thermocline depth field, the predictability loss associated with initial error would be comparable to that associated with stochastic wind forcing. These general results hold for all lead times from 1 to 12 months.

5. Discussion and conclusions

The purpose of this study was to consider the predictability of LDEO4 out to 1-yr lead times. In particular, we explored predictability loss due to two error sources: large-scale uncertainty in the initial conditions and model sensitivity to subseasonal stochastic zonal wind forcing.

In general, from 1975 to 2002 our best estimate of the large-scale uncertainty in the initial conditions leads to larger ensemble spreads than continuously applied stochastic wind forcing (at all leads from 1 to 12 months).

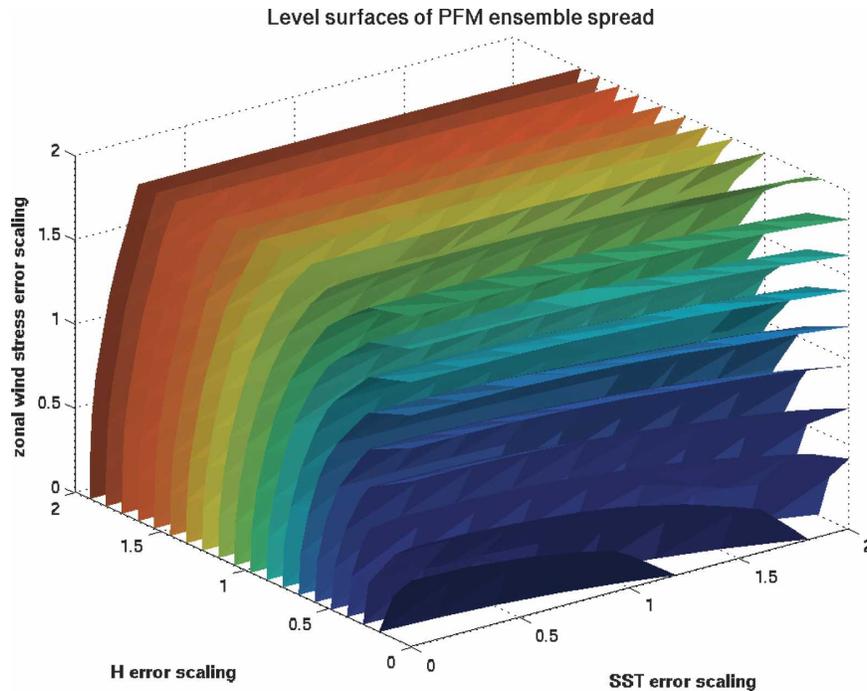


FIG. 8. Level contours of the 6-month spread in Niño-3 as calculated with the PFM approximations to LDEO4. Contours range from 0.1° (blue) to 1°C (red) in intervals of 0.05°C . The scales represent the magnitude of each error field relative to our standard estimate.

However, the effects of internal atmospheric noise are not negligible. We would expect that with improvements in observing systems and data assimilation methods, the “noise” effects will become of comparable importance to initial condition errors.

The forecast ensemble spread has a very strong seasonal dependence, with greatest spread in the winter and the least spread in the spring and early summer. This seasonal dependence closely matches the seasonal cycle of observed Niño-3 anomalies independently of forecast lead time or calendar month of initialization. These results support earlier findings by Thompson and Battisti (2000).

Our results suggest that reducing the uncertainty in our representation of the initial thermocline depth (or sea level height) would be the most beneficial for reducing the uncertainty of LDEO4 forecasts. We additionally conclude that were it not for the error in the thermocline depth, the stochastic wind forcing associated with atmospheric internal variability would dominate the spread.

This does not mean, however, that there is no utility in improving the observing systems of SST and wind stress. On the contrary, because the tropical ocean/atmosphere system is so tightly coupled, SST and wind stress, if known well enough over a period of a few months to a year, can effectively be used to infer the

thermocline depth in the tropical Pacific. This helps us better understand the results of Chen et al. (2004), who find that skillful ENSO forecasts are possible even if only SST data are assimilated into the initial conditions. Their success results from using a coupled nudging scheme, where the thermocline depth fields have time to adjust to the SST information.

Averaged over the seasonal cycle, there is an approximately linear relationship between the Niño-3 index and the ensemble spread. More precisely, during time periods when Niño-3 is negative the error growth is reduced and when Niño-3 is positive the error growth is increased. However this relationship breaks down during the strong El Niño years, when the spread is not significantly larger than in moderately warm years. The most dramatic feature of the state dependence of spread is the suppression of wintertime growth in the years following an El Niño event.

These two features of the error behavior have important implications for interannual-scale prediction utility in the sense that it leads us to conclude that both the warm and cold phases of ENSO are potentially very predictable, but for different reasons. While El Niño events have a large signal, their wintertime growth is typical of states that are only moderately warm. Hence, prediction utility is high. Cool events, on the other hand, have suppressed wintertime growth, also result-

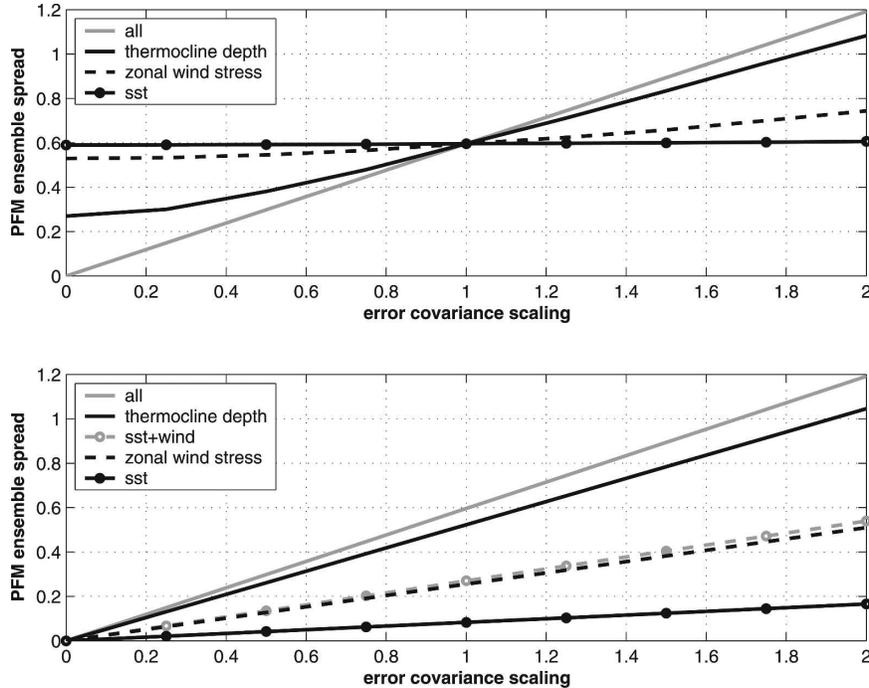


FIG. 9. The 6-month spread in Niño-3 as calculated with the PFM approximations to LDEO4 (units are $^{\circ}\text{C}$). (top) The magnitude of the initialization error covariance for each of the fields is multiplied by a scaling factor from 0 to 2 (in increments of 0.25) while keeping the remaining two error covariances fixed at their standard magnitude (corresponding to a scaling factor of 1). For reference, the gray line represents the spread if the magnitudes of the fields were scaled simultaneously. (bottom) Same as top, except that the unscaled error fields are held to zero. Gray lines show simultaneous scaling of all the error fields and just the SST and zonal wind stress.

ing in high prediction utility. It is the neutral phase of ENSO that is least predictable, as wintertime error growth leaves us vulnerable to false predictions.

In terms of the seasonal time scale and ENSO phase, it is important to note that the error growth is not well described as a function of either the initialization state or the initialization season. Instead, it depends more on the state of the system (or the calendar month) at the end of the forecast. The literature, however, has tended to focus on categorizing error growth in terms of the initial state. This is clearly a pragmatic approach, as it suggests that there could be advance knowledge of the predictability of the system. (Of course, if we are only interested in the seasonal cycle of error growth, this is still possible.) In the paradigm developed here, however, anticipating the ensemble spread on interannual time scales requires that we have a forecasting system that can accurately follow the system trajectory. Thus, knowing the predictability a priori is no less (or more) difficult than predicting the state of the system.

At present, however, the degradation of useful prediction may not necessarily result from stochastic forc-

ing or from initial condition error, but from systematic inadequacies in either the model dynamics or the initialization system (biases). In this study, we have circumvented this issue through the artificial construction of “best” initial conditions. Systematic deficiencies in the model dynamics constitute an especially difficult problem, as the resolution of the bias component of model error requires that we actually have a deeper understanding of the dynamics of the system. At the very least it requires an observational record that is long enough to allow for a statistical diagnosis of the bias (Chen et al. 2000, 2004). While there may be predictability loss that is inherent in the system (i.e., stemming from small-scale errors or external influences), there is clearly a large component of predictability loss that can be reduced through better specification of initial conditions. But the fruits of these advances can only be realized if we have a model that is capable of tracking the ENSO system trajectory. As such, the findings of this work serve to underscore the need for better models, assimilation procedures, and observing systems.

Acknowledgments. The first author was supported by the NASA Earth Systems Science Graduate Fellowship and the NOAA Global Change Postdoctoral Fellowship. The other authors were funded by NOAA-CICAR NA03OAR432017. Helpful discussions with Peter Gent, Joe Tribbia, and Martin Visbeck are gratefully acknowledged, as are comments from Gerrit Burgers and an anonymous reviewer.

APPENDIX

Reduced Space Representation of Observations and Their Error

a. Projection of observations into the model state space

The model state space (\mathbf{x}) is only partially observed (via SST, wind stress, and thermocline depth). For projecting these observations into the model state we define sampling matrices, \mathbf{H}_w , \mathbf{H}_{sst} , and \mathbf{H}_h , which map the model variables to the observed fields (\mathbf{o}_w , \mathbf{o}_{sst} , and \mathbf{o}_h):

$$\mathbf{o}_w = \mathbf{H}_w \mathbf{x}; \quad \mathbf{o}_{sst} = \mathbf{H}_{sst} \mathbf{x}; \quad \mathbf{o}_h = \mathbf{H}_h \mathbf{x}. \quad (\text{A1})$$

The subscripts w , sst , and h refer to SST, wind stress, and thermocline depth anomaly fields. Each sampling matrix is a submatrix of the model state identity matrix with only rows corresponding to the variables that are observed. The reduced space representation of the model state can be expanded via $\mathbf{x} = \mathbf{E}\hat{\mathbf{x}}$, where \mathbf{E} is a matrix whose 30 columns are the leading multivariate empirical orthogonal functions (MEOFs) and $\hat{\mathbf{x}}$ is the 30-dimensional vector of their amplitudes. In computing the MEOF amplitudes, the model variables are nondimensionalized such that each field has approximately equal contribution to the total variance (see Cañizares et al. 2001 for details).

To project the observations onto the reduced space, we minimize the squared difference between the observations (\mathbf{o}) and the expansion from their reduced space representation ($\hat{\mathbf{o}}$):

$$\begin{aligned} & (\mathbf{H}_w \mathbf{E} \hat{\mathbf{o}} - \mathbf{o}_w)^T (\mathbf{H}_w \mathbf{E} \hat{\mathbf{o}} - \mathbf{o}_w) + (\mathbf{H}_{sst} \mathbf{E} \hat{\mathbf{o}} - \mathbf{o}_{sst})^T \\ & \times (\mathbf{H}_{sst} \mathbf{E} \hat{\mathbf{o}} - \mathbf{o}_{sst}) + (\mathbf{H}_h \mathbf{E} \hat{\mathbf{o}} - \mathbf{o}_h)^T (\mathbf{H}_h \mathbf{E} \hat{\mathbf{o}} - \mathbf{o}_h), \end{aligned} \quad (\text{A2})$$

leading to

$$\begin{aligned} \hat{\mathbf{o}} &= [\mathbf{E}^T (\mathbf{H}_w^T \mathbf{H}_w + \mathbf{H}_{sst}^T \mathbf{H}_{sst} + \mathbf{H}_h^T \mathbf{H}_h) \mathbf{E}]^{-1} \mathbf{E}^T \\ & \times (\mathbf{H}_w^T \mathbf{o}_w + \mathbf{H}_{sst}^T \mathbf{o}_{sst} + \mathbf{H}_h^T \mathbf{o}_h). \end{aligned} \quad (\text{A3})$$

b. Error covariance estimates

By reducing the dimension of the error covariance in SST, thermocline depth, and zonal wind stress, we can

decrease the number of realizations needed to adequately sample it. We do this by projecting them into the model MEOF space, as in

$$\hat{\mathbf{R}} = \hat{\mathbf{C}}^{-1} \mathbf{E}^T (\hat{\mathbf{R}}_w + \hat{\mathbf{R}}_{sst} + \hat{\mathbf{R}}_h) \mathbf{E} \hat{\mathbf{C}}^{-1}, \quad (\text{A4})$$

where

$$\hat{\mathbf{C}} = \mathbf{E}^T (\mathbf{H}_w^T \mathbf{H}_w + \mathbf{H}_{sst}^T \mathbf{H}_{sst} + \mathbf{H}_h^T \mathbf{H}_h) \mathbf{E} \quad (\text{A5})$$

and

$$\hat{\mathbf{R}}_w = \mathbf{H}_w^T \mathbf{R}_w \mathbf{H}_w; \quad \hat{\mathbf{R}}_{sst} = \mathbf{H}_{sst}^T \mathbf{R}_{sst} \mathbf{H}_{sst}; \quad \hat{\mathbf{R}}_h = \mathbf{H}_h^T \mathbf{R}_h \mathbf{H}_h. \quad (\text{A6})$$

Here, \mathbf{R}_w , \mathbf{R}_{sst} , and \mathbf{R}_h are the full space error covariances of the observations. The relations in (A4)–(A6) follow easily from (A3) if you consider that the full space errors in each field (\mathbf{e}_h , \mathbf{e}_{sst} , and \mathbf{e}_w) can be projected into the reduced space in exactly the same way in which observations are projected. Then $\hat{\mathbf{R}}$ is defined as the covariance of the reduced space error (i.e., $\hat{\mathbf{R}} \equiv \langle \hat{\mathbf{e}} \hat{\mathbf{e}}^T \rangle$), and it is assumed that there is no cross covariance between the three error fields.

In general, data analysis products (such as those used for initialization of ENSO forecasts) have a theoretical error associated with them that stems from the error in the raw observations (this is both instrumental and sampling error), from covariance patterns that are not resolved by the analysis scheme, and from the large-scale statistical or dynamical constraints applied in the analysis. Each of the error covariance estimates used in this study are derived from analysis schemes that are performed in a reduced EOF space. As such, the “unresolved” component derives from the loss of covariance information due to EOF truncation. However, the large-scale component of the analysis error that projects onto the space of the leading EOFs is our concern in this study. A description of the large-scale error covariance estimates for zonal surface wind stress, SST, and ocean upper-layer depth are presented below. Note that the surface wind velocity and SST anomaly error covariances that we use for the perturbation experiment were not derived from the same analyzed products that were used as the data source in our adjoint scheme. However, the raw data used in the analyzed products come from the same sources, and we assume that the large-scale spatial patterns in the errors are still relevant.

1) ZONAL WIND STRESS

Error estimates for analyzed wind fields were taken from a reduced space optimal interpolation (OI) analysis described in Evans and Kaplan (2004). Although the

analysis error is time dependent, we take the average error from January 1856 to September 2001. The magnitude of the pattern is adjusted such that the error magnitude in the central equatorial Pacific is of comparable magnitude to the small-scale variability reported in daSilva et al. (1994), that is, $\sim 3 \text{ m s}^{-1}$. We calculate zonal wind stress errors from the zonal surface wind errors (u') using the bulk formula

$$\tau'_x = \rho_a C_d W u', \quad (\text{A7})$$

where ρ_a is the density of air and C_d is a homogeneous bulk drag coefficient. The product $\rho_a C_d$ is estimated to be $0.02 \times 10^4 \text{ dyn s}^2 \text{ m}^{-4}$ (Zebiak and Cane 1987), and the absolute wind speed (W) was taken as the annual mean from the climatology in the *Atlas of Surface Marine Data* (daSilva et al. 1994).

The standard deviation of the zonal wind stress error is fairly uniform through most of the central tropical Pacific, with an amplitude of $\sim 0.14 \text{ dyn cm}^{-2}$. The errors tend to be a bit smaller over the far western and eastern parts of the equatorial basin. The largest errors occur about 12° north of the equator in the western part of the basin, where the mean winds are very strong. These maximum errors have amplitudes of 0.24 dyn cm^{-2} .

2) SEA SURFACE TEMPERATURE

Kaplan et al. (1998) describe a reduced space optimal smoothing analysis method for global SST anomaly fields. The data used in the analysis are from the U.K. Met Office historical sea surface temperature dataset (MOHSST5) of the Global Ocean Surface Temperature Atlas (GOSTA). Their optimal smoothing technique uses a first-order autoregressive model as a weak constraint to smooth the interpolation in time. We take the average error covariance from January 1956 to December 1991. The errors tend to be greatest along the equator, with a standard deviation of $\sim 0.2^\circ\text{C}$. Regions in the equatorial eastern and central Pacific have the largest errors, with standard deviations of 0.25°C .

3) UPPER-LAYER THICKNESS

Estimates of the error in the depth of the upper layer (and the corresponding geostrophic surface current) are based on the reduced space Kalman filter scheme described in Cane et al. (1996) and Kaplan et al. (2004). The data assimilated into the analysis are the gridded TOPEX/Poseidon sea level height altimetry product of Cheney et al. (1994). The forecast model used in the Kalman filter scheme is a version of the Cane and Patton (1984) algorithm for solving the shallow water equations on an equatorial beta plane. The model is

forced with wind stress fields from the Florida State University analysis (Stricherz et al. 1997). The errors in the wind stress fields are assumed to be the sole source of error in the shallow water model. Analysis error estimates for the thermocline depth anomaly are taken as the temporal average from October 1992 to November 2003. These estimates are comparable to those reported by a variety of sources. See Kaplan et al. (2004) for a comparison of different altimetry assimilation products.

Errors in the upper-layer thickness are on the order of 10 m ($\sim 4 \text{ cm}$ in sea level) throughout the tropical basin. The maximum errors occur just off the equator along the western boundary, with a standard deviation of $\sim 15 \text{ m}$. The position of these maximums results from the accumulation of wind errors transmitted by Rossby waves traveling westward across the basin.

REFERENCES

- Barnston, A., M. Glantz, and Y. He, 1999: Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997–98 El Niño episode and the 1998 La Niña onset. *Bull. Amer. Meteor. Soc.*, **80**, 217–243.
- Batstone, C., and H. Hendon, 2005: Stochastic variability associated with ENSO and the role of the MJO. *J. Climate*, **18**, 1773–1789.
- Battisti, D., and A. Hirst, 1989: Interannual variability in a tropical atmosphere–ocean model: Influence of the basic state, ocean geometry and nonlinearity. *J. Atmos. Sci.*, **46**, 1687–1712.
- Cane, M., and R. Patton, 1984: A numerical model for low-frequency equatorial dynamics. *J. Phys. Oceanogr.*, **14**, 1853–1863.
- , A. Kaplan, R. Miller, B. Tang, E. Hackert, and A. Busalacchi, 1996: Mapping tropical Pacific sea level: Data assimilation via a reduced state space Kalman filter. *J. Geophys. Res.*, **101**, 22 599–22 617.
- Cañizares, R., A. Kaplan, M. Cane, and D. Chen, 2001: Use of data assimilation via linear low-order models for the initialization of El Niño–Southern Oscillation predictions. *J. Geophys. Res.*, **106**, 30 947–30 959.
- Chen, D., S. Zebiak, M. Cane, and A. Busalacchi, 1997: Initialization and predictability of a coupled ENSO forecast model. *Mon. Wea. Rev.*, **125**, 773–788.
- , M. Cane, S. Zebiak, R. Canizares, and A. Kaplan, 2000: Bias correction of an ocean–atmosphere coupled model. *Geophys. Res. Lett.*, **27**, 2585–2588.
- , —, A. Kaplan, S. Zebiak, and D. Huang, 2004: Predictability of El Niño over the past 148 years. *Nature*, **428**, 733–736.
- Cheney, R., L. Miller, R. Agreen, N. Doyle, and J. Lillibridge, 1994: TOPEX/Poseidon: The 2-cm solution. *J. Geophys. Res.*, **99**, 24 555–24 564.
- daSilva, A., A. Young, and S. Levitus, 1994: *Algorithms and Procedures*. Vol. 1, *Atlas of Surface Marine Data*, NOAA Atlas NESDIS 6, 83 pp.
- DelSole, T., 2004: Predictability and information theory. Part I: Measures of predictability. *J. Atmos. Sci.*, **61**, 2425–2440.
- Deser, C., A. Capotondi, R. Saravanan, and A. Phillips, 2005:

- Tropical Pacific and Atlantic climate variability in CCSM3. *J. Climate*, **19**, 2451–2481.
- Evans, M., and A. Kaplan, 2004: The Pacific sector Hadley and Walker Circulations in historical marine wind analyses. *The Hadley Circulation: Present, Past and Future*, H. F. Diaz and R. S. Bradley, Eds., Kluwer Academic, 239–258.
- Fan, Y., M. Allen, D. L. T. Anderson, and M. A. Balmaseda, 2000: How predictability depends on the nature of uncertainty in initial conditions in a coupled model of ENSO. *J. Climate*, **13**, 3298–3313.
- Fedorov, A., S. Harper, S. Philander, B. Winter, and A. Wittenberg, 2003: How predictable is El Niño? *Bull. Amer. Meteor. Soc.*, **84**, 911–919.
- Flugel, M., P. Chang, and C. Penland, 2004: The role of stochastic forcing in modulating ENSO predictability. *J. Climate*, **17**, 3125–3140.
- Goswami, B., and J. Shukla, 1991: Predictability of a coupled ocean–atmosphere model. *J. Climate*, **4**, 3–22.
- Jochum, M., and R. Murtugudde, 2004: Internal variability of the tropical Pacific Ocean. *Geophys. Res. Lett.*, **31**, L14309, doi:10.1029/2004GL020488.
- Kaplan, A., M. Cane, Y. Kushnir, A. Clement, M. Blumenthal, and B. Rajagopalan, 1998: Analysis of global sea surface temperature 1856–1991. *J. Geophys. Res.*, **103**, 18 567–18 589.
- , Y. Kushnir, and M. Cane, 2003: Reduced space approach to the optimal analysis interpolation of historical marine observations: Accomplishments, difficulties, and prospects. *Advances in the Applications of Marine Climatology: The Dynamic Part of the WMO Guide to the Applications of Marine Climatology*. Tech. Rep. WMO/TD-1081, World Meteorological Organization, Geneva, Switzerland, 99–216.
- , M. Cane, D. Chen, D. Witter, and R. Cheney, 2004: Small-scale variability and model error in tropical Pacific sea level. *J. Geophys. Res.*, **109**, CO2001, doi:10.1029/2002JC001743.
- Karspeck, A., 2004: Predictability of ENSO on interannual and decadal timescales. Ph.D. dissertation, Lamont-Doherty Earth Observatory, Columbia University, 181 pp. [Available from ProQuest Digital Dissertations AAT3138346 at <http://wwwlib.umi.com/dxweb/>]
- , R. Seager, and M. Cane, 2004: Predictability of tropical Pacific decadal variability in an intermediate model. *J. Climate*, **17**, 2842–2850.
- Kessler, W., and R. Kleeman, 2000: Rectification of the Madden–Julian oscillation in the ENSO cycle. *J. Climate*, **13**, 3560–3575.
- Kirtman, B., J. Shukla, M. Balmaseda, N. Graham, C. Penland, Y. Xue, and S. Zebiak, 2001: Current status of ENSO forecast skill. Tech. Rep. 56, CLIVAR Working Group on Seasonal to Interannual Prediction, 26 pp.
- Kleeman, R., 2002: Measuring dynamical prediction utility using relative entropy. *J. Atmos. Sci.*, **59**, 2057–2072.
- , and A. Moore, 1999: A new method for determining the reliability of dynamical ENSO predictions. *Mon. Wea. Rev.*, **127**, 694–705.
- Lorenz, E., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141.
- , 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **3**, 321–333.
- Madden, R., and P. Julian, 1994: Observations of the 40–50-day tropical oscillation—A review. *Mon. Wea. Rev.*, **122**, 814–837.
- McPhaden, M., 1999: Climate oscillations—Genesis and evolution of the 1997–98 El Niño. *Science*, **283**, 950–954.
- , 2003: Tropical Pacific Ocean heat content variations and ENSO persistence barriers. *Geophys. Res. Lett.*, **30**, 1480, doi:10.1029/2003GL016872.
- Moore, A., and R. Kleeman, 1999a: Stochastic forcing of ENSO by the intraseasonal oscillation. *J. Climate*, **12**, 1199–1220.
- , and —, 1999b: The nonnormal nature of El Niño and intraseasonal variability. *J. Climate*, **12**, 2965–2982.
- Orrell, D., L. Smith, J. Barkmeijer, and T. Palmer, 2001: Model error in weather forecasting. *Nonlinear Processes Geophys.*, **8**, 357–371.
- Palmer, T., 2000: Predicting uncertainty in forecasts of weather and climate. *Rep. Prog. Phys.*, **63**, 71–116.
- Penland, C., and T. Magorian, 1993: Prediction of Niño 3 sea surface temperatures using linear inverse modeling. *J. Climate*, **6**, 1067–1076.
- , and P. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, **8**, 1999–2024.
- Schneider, T., and S. Griffies, 1999: A conceptual framework for predictability studies. *J. Climate*, **12**, 3133–3155.
- Stricherz, J., D. Legler, and J. O'Brien, 1997: TOGA pseudostress atlas, 1985–1994: II, tropical Pacific Ocean. COAPS Rep. 97-2, Center for Ocean–Atmosphere Prediction Studies, Tallahassee, FL, 155 pp.
- Suarez, M., and P. Schopf, 1988: A delayed oscillator for ENSO. *J. Atmos. Sci.*, **45**, 3283–3287.
- Swanson, K., T. N. Palmer, and R. Vautard, 2000: Observational error structures and the value of advanced assimilation techniques. *J. Atmos. Sci.*, **57**, 1327–1340.
- Thompson, C., and D. Battisti, 2000: A linear stochastic dynamical model of ENSO. Part I: Model development. *J. Climate*, **13**, 2818–2832.
- Tziperman, E., S. Zebiak, and M. Cane, 1997: Mechanisms of seasonal–ENSO interaction. *J. Atmos. Sci.*, **54**, 61–71.
- Vecchi, G., and D. Harrison, 2000: Tropical Pacific sea surface temperature anomalies, El Niño, and equatorial westerly wind events. *J. Climate*, **13**, 1814–1830.
- Wang, C., 2001: On the ENSO mechanisms. *Adv. Atmos. Sci.*, **18**, 674–691.
- Xue, Y., M. Cane, and S. Zebiak, 1997a: Predictability of a coupled model of ENSO using singular vector analysis. Part I: Optimal growth in seasonal background and ENSO cycles. *Mon. Wea. Rev.*, **125**, 2053–2056.
- , —, and —, 1997b: Predictability of a coupled model of ENSO using singular vector analysis. Part II: Optimal growth and forecast skill. *Mon. Wea. Rev.*, **125**, 2057–2073.
- Zebiak, S., 1989: On the 30–60 day oscillation and the prediction of El Niño. *J. Climate*, **2**, 1381–1387.
- , and M. Cane, 1987: A model El Niño–Southern Oscillation. *Mon. Wea. Rev.*, **115**, 2262–2278.