PREDICTABILITY OF AN OCEAN/ATMOSPHERE MODEL USING ADJOINT MODEL ANALYSIS

M. Benno Blumenthal

Yan Xue

Mark Cane

Lamont-Doherty Geological Observatory of Columbia University
Palisades, NY, USA

This paper presents an analysis of a coupled-ocean atmosphere model used to study ENSO (El Niño-Southern Oscillation) Our interest here is in the predictability of ENSO. The analysis proceeds by constructing a linear model that optimally fits the behavior of the original non-linear coupled model. By construction, this approximate linear model has only a few degrees of freedom. Because the linear model is so much smaller than the original, it is possible to understand it in much finer detail, indirectly offering insight into the properties and behavior of the original model.

As it turns out, even linear models with only a few degrees of freedom can have rather elaborate and surprising short-term error behavior. Even a linear system that is not self-adjoint admits the possibility of error growth in a mode unrelated to the classic notion of a fastest growing linearly unstable mode.

Constructing a linear approximation to the coupled model gave us two major results: a cyclical Principal Oscillation Pattern (POP-cycle) that gives a clear ENSO cycle, and error growth that is due to the non-self-adjoint nature of the system. The POP-cycle method is a generalization of the established POP method that allows the model to have an explicit seasonal cycle. The POP-cycle isolates and clarifies features previously noted in the literature, such as the off-equatorial storage and westward propagation that appears 90° out of phase with ENSO events. The non-self-adjoint error growth explains both the previously noted seasonality of error growth in the coupled model and the previously noted 6 month initial growth rate.

1. INTRODUCTION

Given that a model system starts out with some initial error, what happens to that error as the system evolves according to the given model? In general the error evolution is a combination of two factors: the evolution of the initial error under the model dynamics, and the introduction of error at each time step due to the model being an approximate representation of the true dynamics.

As we construct better and better models, the error added at each time step due to model inadequacies is reduced. However, if it is the nature of the system to amplify the initial error structures, then even if we achieve a perfect representation of the underlying dynamics, we still have difficulty in making a prediction due to error growth in the model.

This paper reviews an analysis of a coupled-ocean atmosphere model used to study ENSO (El Niño-Southern Oscillation), that of Cane and Zebiak (CZ,1987). Parts of this analysis previously

appeared in Blumenthal (1991). Our interest is in the predictability of ENSO given what we might know at any given time: here we analyze the growth of error in the CZ model. The analysis proceeds by constructing a linear model that optimally fits the behavior of the original non-linear coupled model. By construction, this approximate linear model has only a few degrees of freedom. This procedure can be related to the linearizing commonly done in stability analysis (Blumenthal, 1991). Because the linear model is so much smaller than the original, it is possible to understand it in much finer detail, indirectly offering insight into the properties and behavior of the original model.

This technique of constructing a small linear model is the basis of the Principal Oscillation Pattern (POP) analysis of Hasselmann (1988). The POP technique has been used extensively to analyze both data and models (von Storch et al., 1988; von Storch et al., 1987; Xu and von Storch, 1990; Latif et al., 1990). Here we generalize some of that work by using a POP-cycle analysis which allows the model to have an explicit seasonal cycle.

As it turns out, even linear models with only a few degrees of freedom can have rather elaborate and surprising short-term error behavior. Farrell, in a long series of papers (1990,1989,1988, and earlier) demonstrates that perturbation growth in a system that is not self-adjoint can be dominated by a mode unrelated to the classic notion of a fastest growing linearly unstable mode. Lacarra and Talagrand (1988) and Farrell (1990) discuss the resulting lack of predictability in such a non-self-adjoint system. This sort of growth is found in simple linear models as well. The work here explores error growth in simple non-self-adjoint systems, showing that the error growth and the resulting predictability problems are inherent parts of the system.

In particular, we can investigate how error evolves in the model. While we call it error evolution, it is actually more general than that term suggests. Instead of the usual situation where one starts with the initial conditions (a point in state space) and studies how that point evolves in time, one starts with a small ball in state space and studies how that ball evolves with time. As it turns out, even in a system whose evolution consists of simple (but multi-dimensional) decay, the initial ball can be rapidly elongated, resulting in a direction in n-dimensional space that is relatively uncertain soon after integration begins. Note that this is a property of a purely deterministic model; there are no assumptions about statistical properties. As will be shown later, by making statistical assumptions, we can interpret the evolution of this n-ellipse in terms of uncertainty, thus relating uncertainty in the prediction to uncertainty in the initial conditions.

2. THE COUPLED MODEL

The model (henceforth CZ) analyzed here has been used for ENSO prediction (Cane et al., 1986) and the model and its performance are described in a number of places (Zebiak and Cane, 1987; Zebiak, 1989). In particular, the model runs analyzed here are 256 years runs similar to the 1024 year "Standard Parameter" run used to analyze the interannual variability of the model (Zebiak, 1989); the choice of parameters and discussion of the model's behavior is contained therein. By using the 256 years of model output as "ground truth", we analyze the system by constructing simple linear models that reproduce its essential behavior. Thus we intend to analyze more precisely the aspects of the model that lead to its predictability.

This coupled model is also used to make forecasts of El Niño events (Cane and Zebiak, 1987; Cane et al., 1986). In order to make forecasts, the model is initialized by forcing the ocean half

of the model with a filtered version of wind anomaly data (FSU) (Goldenberg and O'Brien, 1981. The filtering is designed to remove any long term trends in the data, cf. Cane et al., 1986). The atmospheric half of the model is then forced by the ocean, generating a series of state vectors for the coupled model, one for each month until the present. The coupled model can then be started from any of these state vectors, generating multiple predictions.

3. POP ANALYSIS: CONSTRUCTING EOF-REDUCED MODELS

The CZ model that we analyze here has seven dynamic fields and approximately $O(10^5)$ grid-points; [$\equiv$ degrees of freedom]. But this model is a numerical approximation to a continuous system and most of those dynamic variables are highly correlated. In fact, only a few dynamical patterns contain almost all of the energy.

Instead of working with the above system, with its many variables that have a fairly simple physical representation, we choose to work with an approximate system that contains only a few dynamical variables. Each of these variables has a complicated physical representation, but the factoring of the problem that this procedure allows makes it much easier to extract information from the system.

The procedure has three parts, and is identical to that done by the MPI group (Hasselmann, 1988; von Storch et al., 1988; von Storch et al., 1987) in computing Principal Oscillation Patterns (POPs). First the state space is reduced by computing a set of Empirical Orthogonal Functions (EOFs) and retaining only the patterns required to account for almost all of the variance. Then a linear model is computed which uses that reduced state space at any time step to predict the reduced state space at the following time step, i.e. a single-step auto-regressive process (Markov model). At this point the model reduction is complete. As an analysis tool, a third step is taken: the POPs are computed. These are the independent modes of the reduced (Markov) model, and can provide insight into the dynamical system.

The Markov model's eigenvectors are the POPs for the system. A purely real mode simply decays in time. A complex mode with eigenvalue λ_i and eigenvector $\mathbf{e}_i$ oscillates between two patterns (the real and imaginary parts) as it decays. Thus the POPs allow us to decompose the behavior of the model into a set of patterns that simply oscillate and decay. To the extent that one POP dominates, the behavior of the model becomes relatively simple to understand. They have been used in the past to investigate the 30-60 day oscillation in the equatorial troposphere (von Storch et al., 1987) and to construct a predictor for ENSO (Xu and von Storch, 1989; 1990). Here we make a different use of these eigensolutions.

3.1 EOF reduction

The reduction by EOFs is straightforward, since most of the systems we attempt to model have variance structures that can be fit with relatively few modes. We are here using the EOFs to reduce the size of the system, not as analysis tools in-and-of themselves.

The state space of the coupled model is given by an array $(u_{ij}, v_{ij}, h_{ij}, T_{ij}, \tau_{ij}^{(x)}, \tau_{ij}^{(y)}, P_{ij}) = \mathbf{w}$ where (u, v) is ocean surface velocity, h is thermocline displacement, T is sea surface temperature (SST), $(\tau^{(x)}, \tau^{(y)})$ is the wind stress, P is the surface pressure, and ij index latitude and

longitude. The evolution of the model is given by the set of time series $\mathbf{w}(t)$.

We calculate a reduced state space $\mathbf{d}(t)$ by using a singular value decomposition (Strang, 1988, appendix A) of $\mathbf{w}(\mathbf{x},t)$, i.e. a decomposition into a set of spatial patterns $\mathbf{S}(\mathbf{x})$ and their respective normalized time series $\mathbf{T}(t)$.

$$\mathbf{w} = \mathbf{T}^T \lambda \mathbf{S} \tag{1}$$

where $\mathbf{S}\mathbf{S}^T = \mathbf{1}$ and $\mathbf{T}\mathbf{T}^T = \mathbf{1}$ and λ is diagonal. We refer these structures as EOFs. This decomposition is done by doing an eigenanalysis of $\mathbf{w}\mathbf{w}^T = \mathbf{T}^T\lambda^2\mathbf{T}$ or $\mathbf{w}^T\mathbf{w} = \mathbf{S}^T\lambda^2\mathbf{S}$, for efficiency the smaller of the two systems (Strang, 1988).

We truncate the system by dropping the vectors that correspond to small λ_i, effectively setting them to zero. We are then left with a small set of time series $\mathbf{d} = \mathbf{T}\lambda'$ such that $\mathbf{w} \approx \mathbf{d}\mathbf{S}$. By construction this new set of time series $\mathbf{d}$ contains most of the variance of the original set of time series $\mathbf{w}$.

3.2 Linear Markov Model

We start with a first-order non-linear model $\mathbf{M}$ such that

$$\mathbf{d}_{t+1} = \mathbf{M}(\mathbf{d}_t) + \epsilon_t \tag{2}$$

This model is presumed to approach a statistical equilibrium, i.e. we can define a model spectrum $\mathbf{D}$

$$\mathbf{D} \equiv \langle \mathbf{d}\mathbf{d}^T \rangle + \langle \epsilon\epsilon^T \rangle \tag{3}$$

The model spectrum $\mathbf{D}$ differs from the true spectrum $\langle \mathbf{d}\mathbf{d}^T \rangle$ because of the error ϵ in the evolution equation. In our case we can think of that error being primarily due to the EOF truncation, i.e. to the extent that the part of the state space we discarded affects the part that we kept, there is some error in the evolution of the part that we kept.

We construct a linear approximation to this model in the hopes of better understanding its short-term behavior: we seek a linear model $\mathbf{A}$ satisfying

$$\mathbf{d}_{t+1} = \mathbf{A}\mathbf{d}_t + \mathbf{e}_t \tag{4}$$

where $\mathbf{e}_t$ is some error.

Given a linear model, we can also characterize the evolution of the error in our calculation of the current state $\mathbf{d}_t$ by computing the evolution of its covariance matrix $\mathbf{P}_t$,

$$\mathbf{P}_{t+1} = \mathbf{A}\mathbf{P}_t\mathbf{A}^T + \mathbf{Q}_t \tag{5}$$

Here we have characterized the error in the linear evolution equation by its covariance $\mathbf{Q}$, i.e.

$$\mathbf{Q}_t \equiv \langle \mathbf{e}_t\mathbf{e}_t^T \rangle = \langle (\mathbf{d}_{t+1} - \mathbf{A}\mathbf{d}_t)(\mathbf{d}_{t+1} - \mathbf{A}\mathbf{d}_t)^T \rangle \tag{6}$$

While we can hope the linear model can reproduce the short term behavior of the non-linear model, eventually it will cease tracking the original model, and we expect the error $\mathbf{P}$ to grow. With an optimal linear model, however, this growth is bounded, because there is no need to have an error greater than the model spectrum $\mathbf{D}$ (This is because we can construct a model

that always predicts zero: its error will be $\mathbf{D}$, and we expect our optimal model to do as well or better than this model).

If we evaluate (5) for large time we can then write ($\mathbf{P}_t = \mathbf{P}_{t+1} = \mathbf{D}$)

$$\mathbf{D} = \mathbf{A}\mathbf{D}\mathbf{A}^T + \mathbf{Q} \tag{7}$$

Using an expanded version of (6),

$$\mathbf{D} = \mathbf{A}\mathbf{D}\mathbf{A}^T + \mathbf{D} + \mathbf{A}\mathbf{D}\mathbf{A}^T - \mathbf{C}\mathbf{A}^T - \mathbf{A}\mathbf{C}^T \tag{8}$$

where $\mathbf{C} \equiv \langle \mathbf{d}_{t+1}\mathbf{d}_t^T \rangle$ is the one-step covariance matrix. This equation determines the linear model $\mathbf{A}$,

$$\mathbf{A} = \mathbf{C}\mathbf{D}^{-1} \tag{9}$$

We can now calculate the single step error $\mathbf{Q}$ by either using (6) or by

$$\mathbf{Q} = \mathbf{D} - \mathbf{C}\mathbf{D}^{-1}\mathbf{C}^T \tag{10}$$

(9) and (10) are an interesting pair of equations because they determine both the model $\mathbf{A}$ *and* the single step error $\mathbf{Q}$ from properties of the signal, i.e. the model spectrum $\mathbf{D}$ and the single step covariance $\mathbf{C}$. This shows the signal covariance, the error covariance and the linear model are quite tightly intertwined.

3.3 Seasonal Linear Model

While it is possible to follow the above procedure and calculate a linear model $\mathbf{A}$ that is independent of season (Blumenthal, 1991), the seasonal cycle is a rather important part of this system, and we do better if we take explicit account of it. So instead of calculating a single transition matrix from the single-step covariance averaged over all time, we calculate a set of transition matrices, one per month, from the single-step covariance matrices averaged over all years. Thus

$$\mathbf{A}^{(i)} = \mathbf{C}^{(i)}[\mathbf{D}^{(i)}]^{-1} \tag{11}$$

where $\mathbf{A}^{(1)}$ gives the transition from January to February, $\mathbf{C}^{(1)}$ gives the single step covariance between January and February, and $\mathbf{D}^{(1)}$ gives the covariance for January, etc.

4. COMPARISON OF THE LINEAR AND CZ MODELS

While the last section shows that one can construct a linear model based on the non-linear CZ model, there is no guarantee that such a linear approximation will be a sufficiently close approximation.

Normally, one would do the single eigenanalysis described earlier, which groups all the dynamical variables together into a single set of EOFs. In the analysis presented here, eigenvectors were computed separately for seasonal (every three months) values of $(\tau^{(x)}, \tau^{(y)})$, h, and T, and 25 eigenvectors were kept for each field (the two components of wind stress were done together). While a single eigenanalysis that encompassed all the dynamical variables would represent the variance more efficiently (i.e. more variance for a given number of modes), this system is both sufficiently small and contains a significant fraction of the variance.

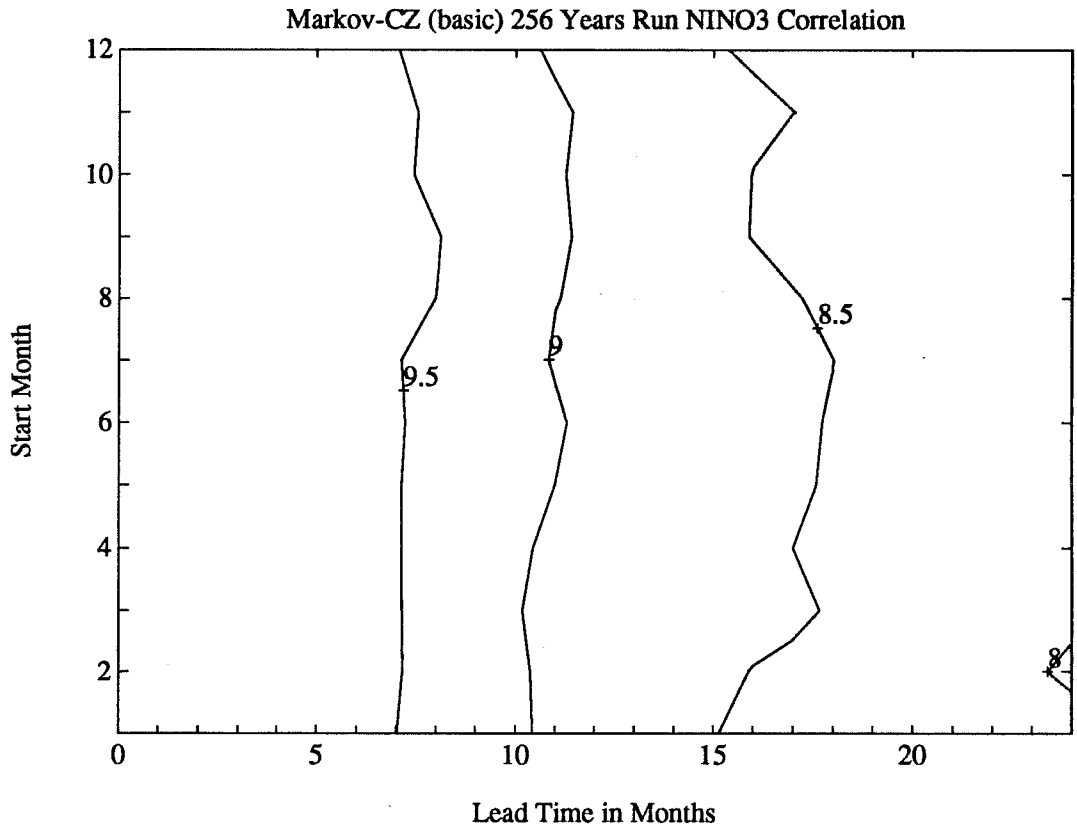

Figure 1: NINO3 correlation between the linear model and the CZ model as a function of starting calendar month (1 = Jan) and lead time. Initial conditions are taken from the 256 year CZ model run that the linear model is calculated from.

Most of the variance is in the first few modes. In particular, the first five modes account for 93% $(\tau^{(x)}, \tau^{(y)})$, 86% (h), and 94% (T) of the total variance respectively. This leaves us with a system with 15 degrees of freedom that represents most of the variance in the original set of time series. This was the initial EOF reduction in Blumenthal (1991).

We have now calculated monthly time series by rerunning the model and projecting onto the above set of EOFs at monthly intervals. This procedure captured 99.2% in $(\tau^{(x)}, \tau^{(y)})$, 96.6% in h, and 98% in T using the above set of 75 EOFs. Using only the first five EOFs in each variable gives 93% $(\tau^{(x)}, \tau^{(y)})$, 85% (h), and 93% (T) of the monthly variable, results that are almost identical to the seasonal results of the original run.

In what follows, we show correlations and time series based on the NINO3 index, i.e. the average sea surface temperature (SST) between 5°S and 5°N, 90°W and 150°W. NINO3 is used because it is a fairly standard index for characterizing El Niños: being able to predict NINO3 indicates one is able to predict ENSO events.

Figure 1 shows the correlation between the 75 mode linear model and the CZ model. Initial conditions were taken from the 256 year run that the model was fit to. The correlations are very high: even after 24 months the correlations are above 0.8. These correlations are also readily apparent in the time series (figure 2) which show a very high correspondence between the linear

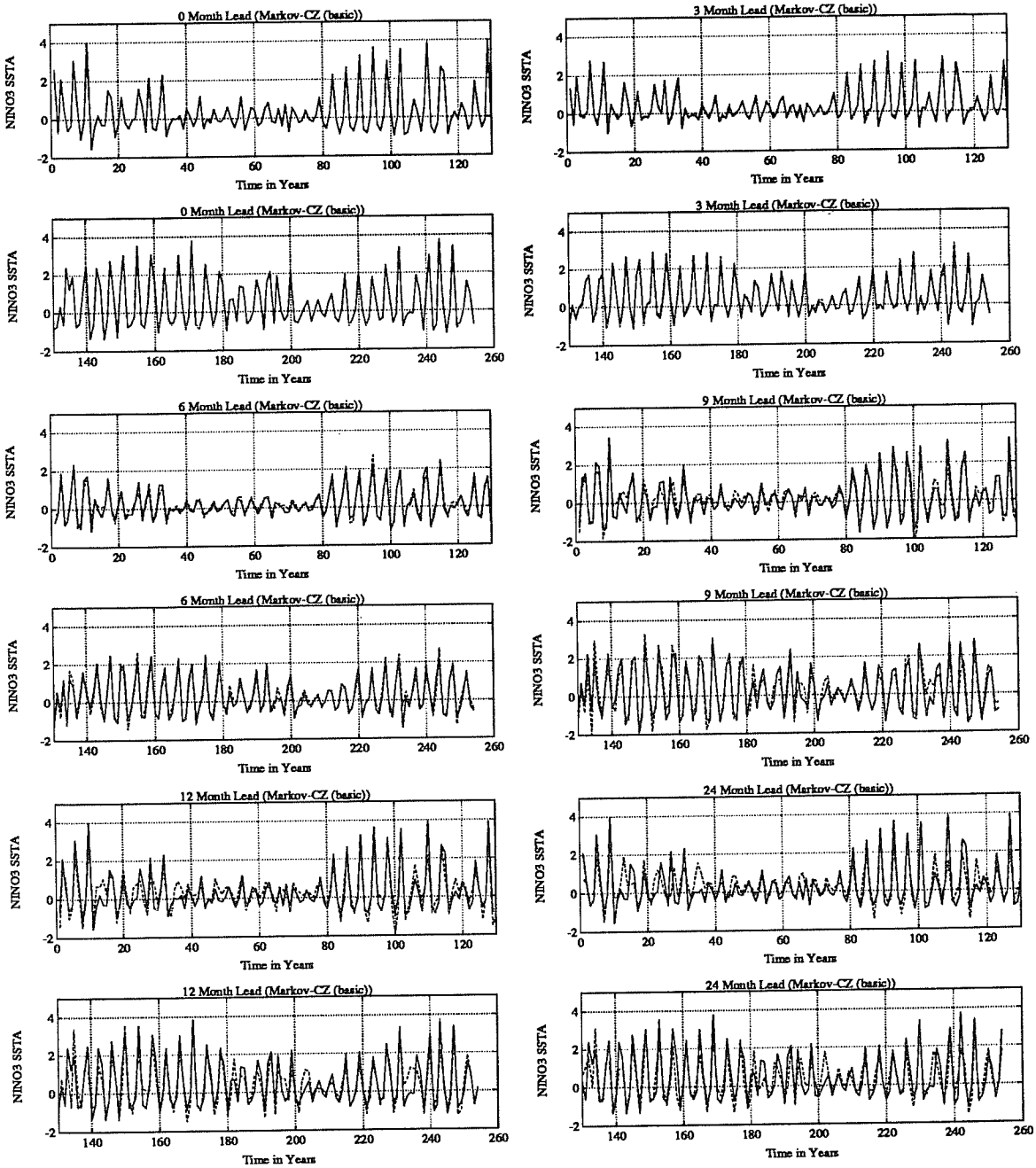

Figure 2: NINO3 time series for a series of simulations of the CZ model by the linear model at 0, 3, 6, 9, 12, and 24 month leads. The solid line is the CZ model; the dashed line is the linear model. Initial conditions are taken from the 256 year CZ model run that the linear model is calculated from.

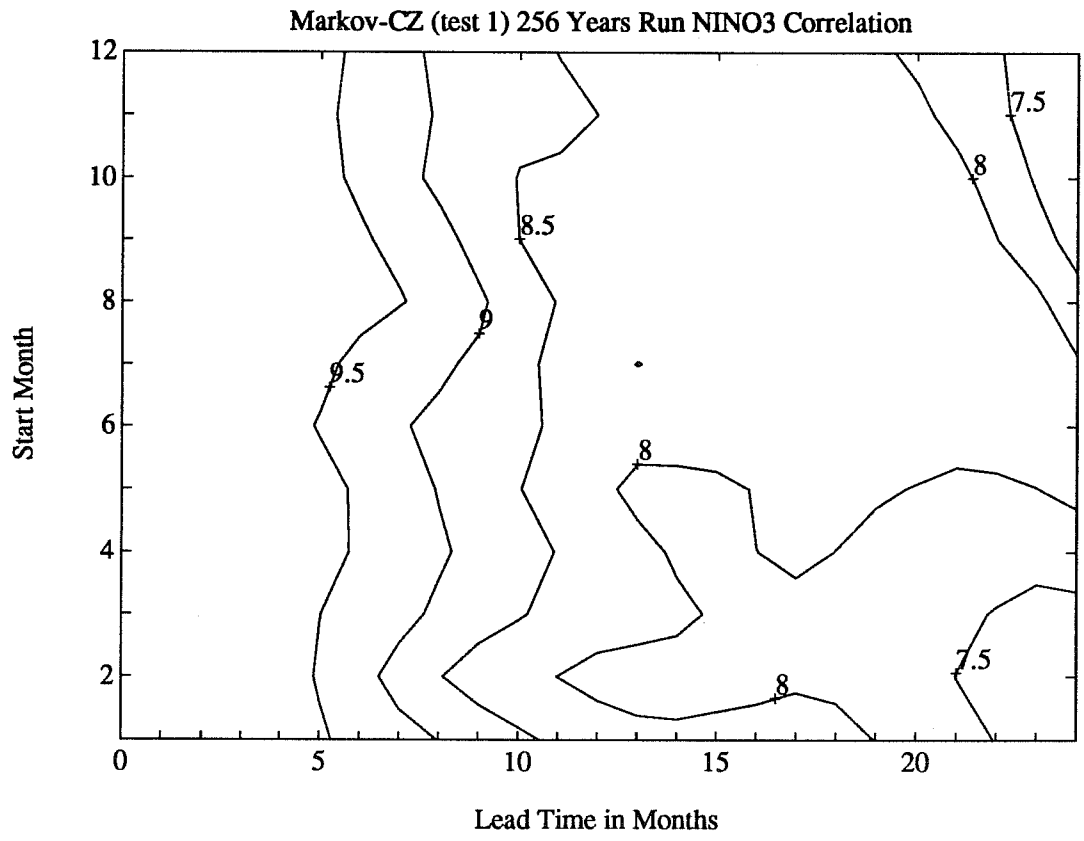

Figure 3: NINO3 correlation between the linear model and the CZ model as a function of starting month and lead time. Initial conditions are taken from a 256 year CZ model run independent of the run that the linear model is calculated from.

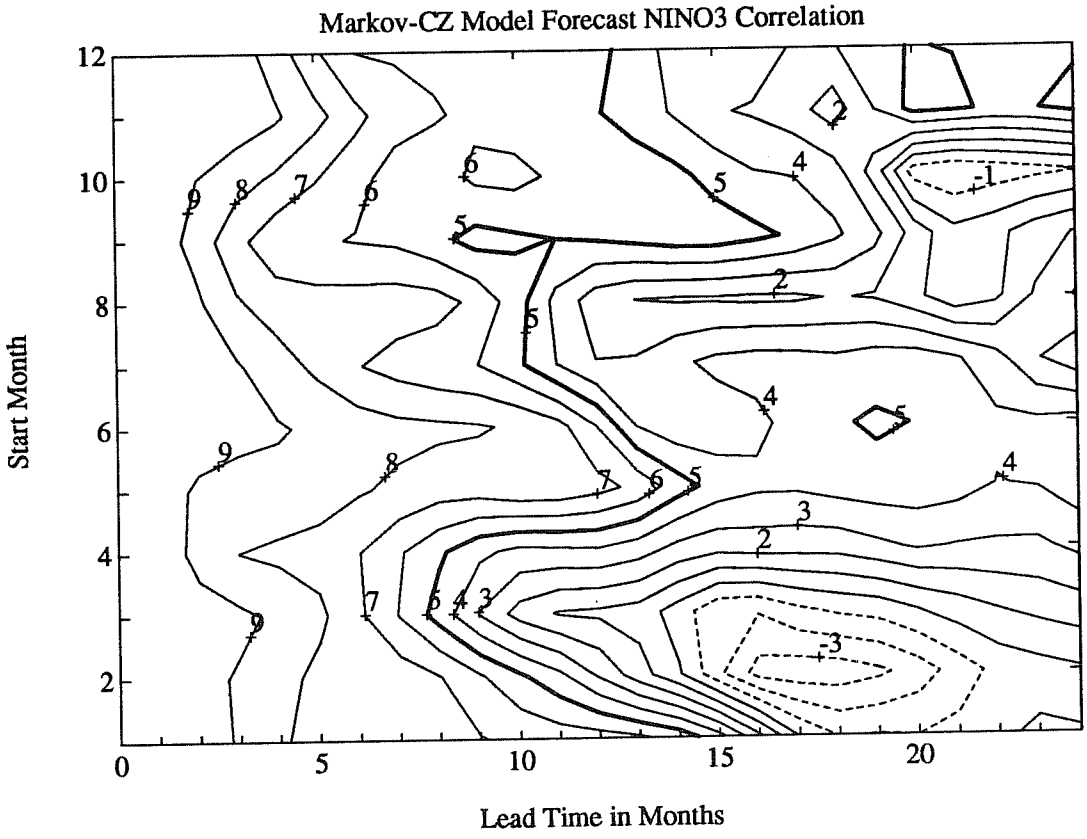

Figure 4: NINO3 correlation between the linear model and the CZ model as a function of starting month and lead time. Initial conditions are taken from the ocean model forced by a filtered FSU wind data set.

and non-linear models.

These correlations correspond precisely to the error growth expressed in (5), i.e. the initial error is zero, and the error added at each step is the average difference between the CZ model and the linear model. Presumably this is a slight underestimate of the error, since we are checking against the same data that we fit.

Figure 3 shows the correlation between these two models when the initial conditions are taken from an independent CZ model run. Again, the correlations are very high. These correlations indicate that the linear model does in fact reproduce the behavior of the CZ model to at least two years, beyond the range that we believe the CZ model has reliable skill at forecasting the real ocean/atmosphere system.

While the above comparison shows that the linear model tracks the coupled model well under relatively noise-free conditions, under more noisy conditions the correlation decreases. Figure 4 again shows the correlation between the 75 mode linear model and the CZ model, but this time initial conditions are taken from the ocean model forced by a filtered FSU wind data set. With these initial conditions, the linear model does not track the CZ model as well. Here the correlation is about .7 after 5 months of integration; the correlation varies from .3 to .7 after ten months. The initialization procedure tends to generate a model state vector that has adjusted

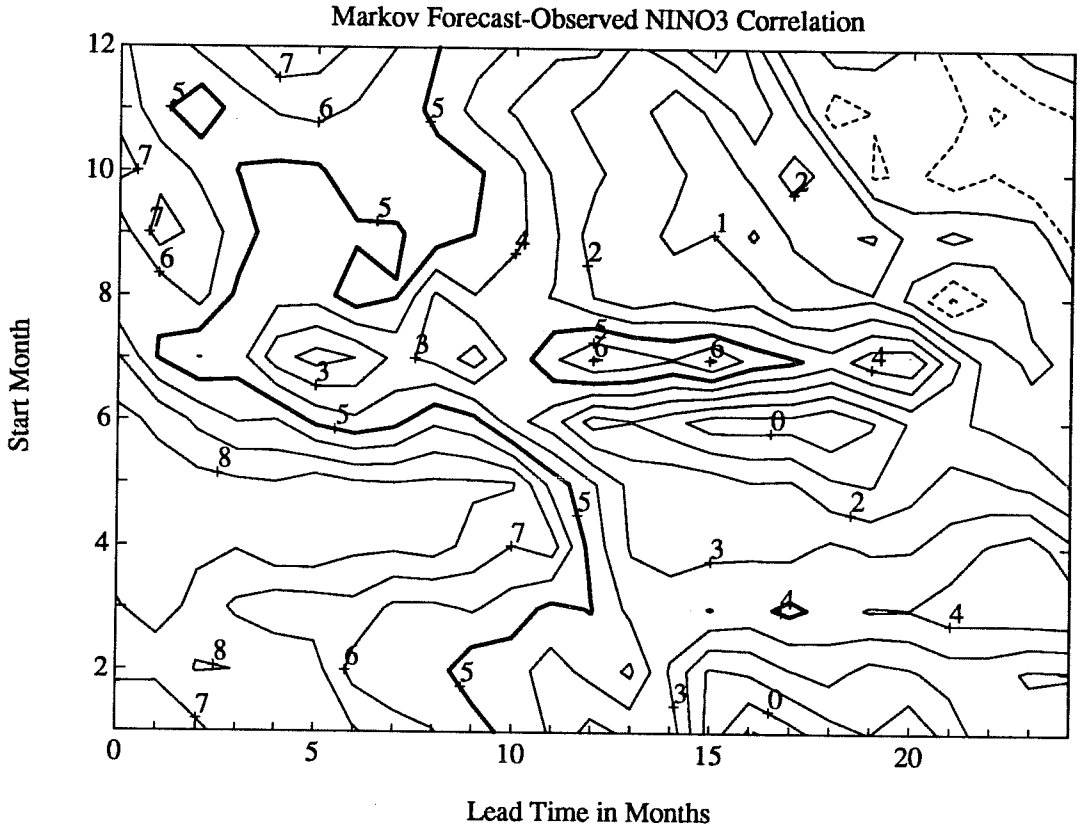

Figure 5: NINO3 correlation between the linear model and data as a function of starting month and lead time. Initial conditions are taken from the ocean model forced by a filtered FSU wind data set.

dynamically (unlike sticking SST data directly into the model), but this is only a tendency: the last time step was forced by a fairly noisy wind field. So the state vector has energy in it that is not entirely appropriate for the model, though the CZ model does a pretty good job of filtering it out. The differences between the CZ model and the linear model suggest, however, that more noise could be filtered out of the initialization procedure.

Given that the linear has trouble reproducing the CZ model given noisy initial conditions, it is not surprising that it has trouble predicting the actual data. It does, however, have some skill. Figure 5 shows the correlation between the linear model and the observed NINO3. The model does particularly well with forecasts that start between April and May, with a correlation of better than .8 after a nearly ten-month prediction. This compares with the the CZ model performance shown in figure 6: the linear model is fairly close in the spring, particularly worse in July. Note also that in July when the linear model does quite badly, the initialization makes the model start out quite badly, which is not really giving the linear model a chance to do a prediction. The CZ model manages to avoid suffering from the bad start, retaining correlation above .5 out to fairly long lead times.

The linear model predictions are presented in figure 7. Results for six different lead times are given, ranging from 0 to 24 months. Each forecast is actually the mean of forecasts from six consecutive months, adjusted to have the same mean and standard deviation as the observed.

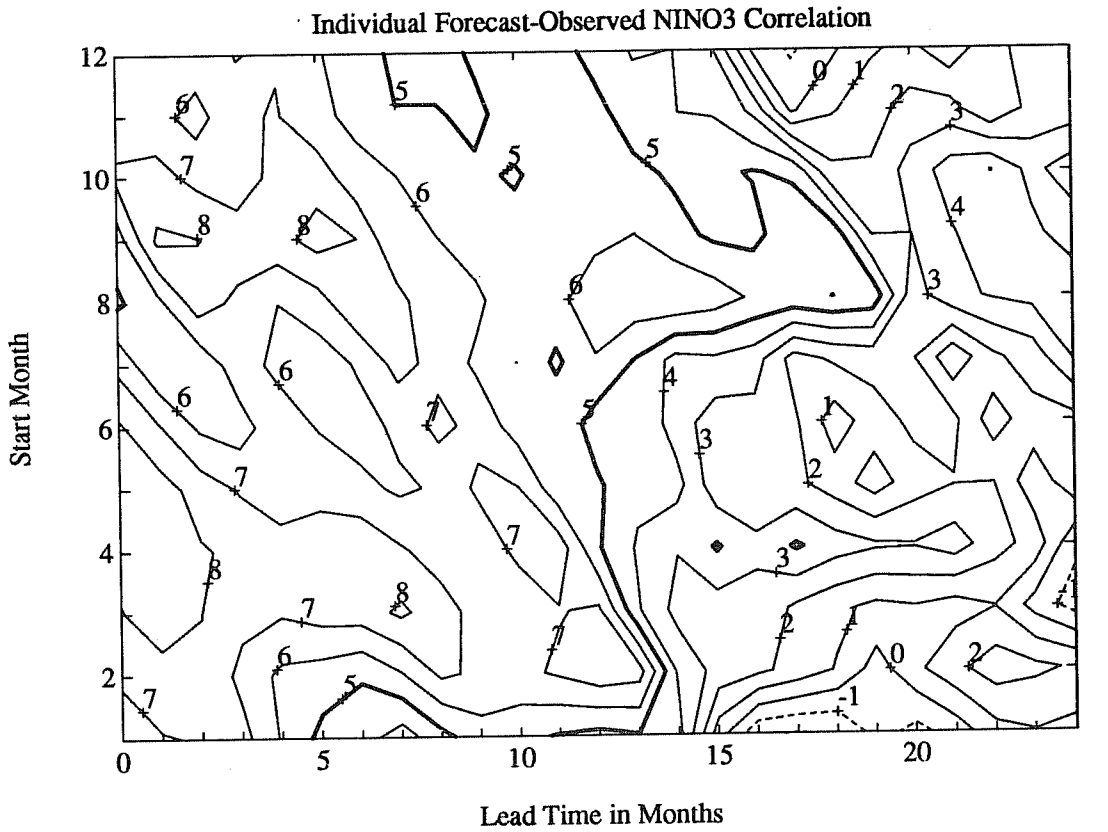

Figure 6: NINO3 correlation between the CZ model and data as a function of start month and lead time. Initial conditions are taken from the ocean model forced by a filtered FSU wind data set.

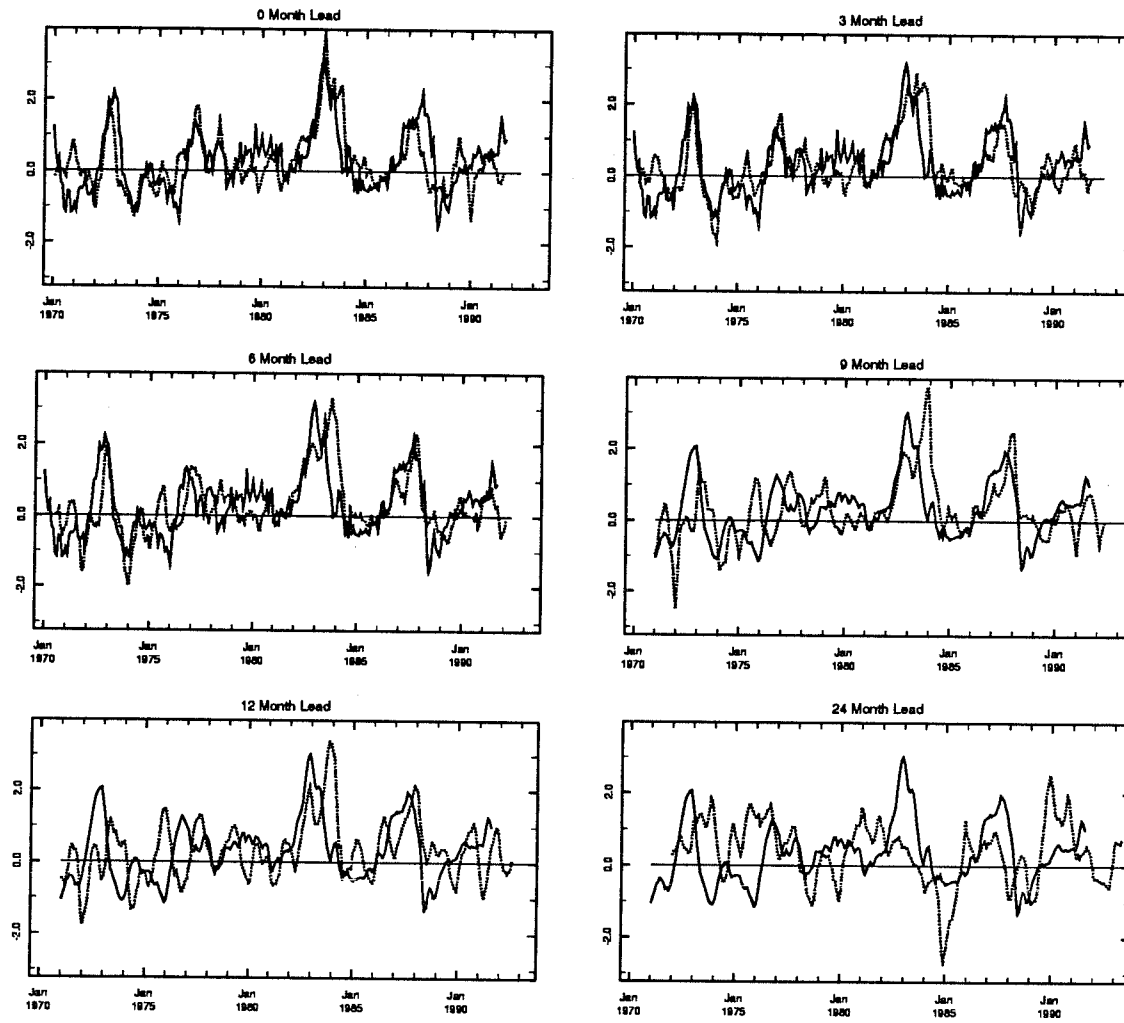

Figure 7: NINO3 time series for a series of predictions at 0, 3, 6, 9, 12, and 24 month leads. 0, 3, and 6 month predictions are averaged over one month; 9, 12, and 24 month predictions are averaged over three months. The solid line is the observed NINO3 signal; the dashed line is the rescaled linear model. Initial conditions are taken from the ocean model forced by a filtered FSU wind data set.

Note that even at 0 month lead the forecast can differ from the observed: this is because only wind data is used in the initialization, SST data is not used. For each lead time, forecast values are indicated by a dashed line and observed values by a solid line (for 9,12 and 15 month lead times, observed values are 3 month averages; otherwise they are monthly averages). The time series indicate some skill at the lower lead times; by 24 months the results are pretty ragged.

5. ERROR GROWTH

The results of the last section can be understood by considering the error growth in the system. The system has the property that error growth can be very rapid, even though the part of the system that best models the signal has relatively slow error growth. This can be seen by considering two different error problems: forced error growth with zero initial error, and the evolution of an initial error ball.

To more easily compare error, we are going to use two measures (See Blumenthal (1991) for a more extensive discussion of error measures). First we use what we call SST *residue* (χ^2): this is the total error variance in the 25 EOFs of SST.

$$\chi^2(t) = \sum_{1}^{25} p_{ii}(t) \tag{12}$$

where $\mathbf{P}(t)$ is the error covariance at time t, and the first 25 EOFS correspond to SST.

The second error measure will be an estimate of the correlation between the true NINO3 signal and the model NINO3 signal. This will allow comparisons with the correlation plots we discussed earlier. The details of the procedure will be given in the next section.

5.1 Forced error growth

In section 3.2 we show that, in calculating a linear approximation to a non-linear model, we determine both a linear model $\mathbf{A}$ and a single step error $\mathbf{Q}$. Furthermore, $\mathbf{A}$ and $\mathbf{Q}$ are such that if we run the error evolution equation (5) to equilibrium, the equilibrium solution $\mathbf{D}$ is the signal covariance, i.e. the full spectrum of the signal.

Figure 8 shows the rather slow growth of error in this case, where the error is initially zero and evolves using (5). The figure is composed of twelve curves, which represent the forced error growth starting from different calendar months. In this case the signal spectral level is reached only after many years of integration. Note that here we have both a model $\mathbf{A}^{(i)}$ and an error $\mathbf{Q}^{(i)}$ with explicit seasonal dependence.

Figure 9 expands the time scale to show the early firced error evolution. This slow error growth corresponds to the very high correlation we saw between the linear and CZ models in figure 3. It shows that with zero initial error, error growth in the model is quite slow, and the linear model matches the CZ model rather well.

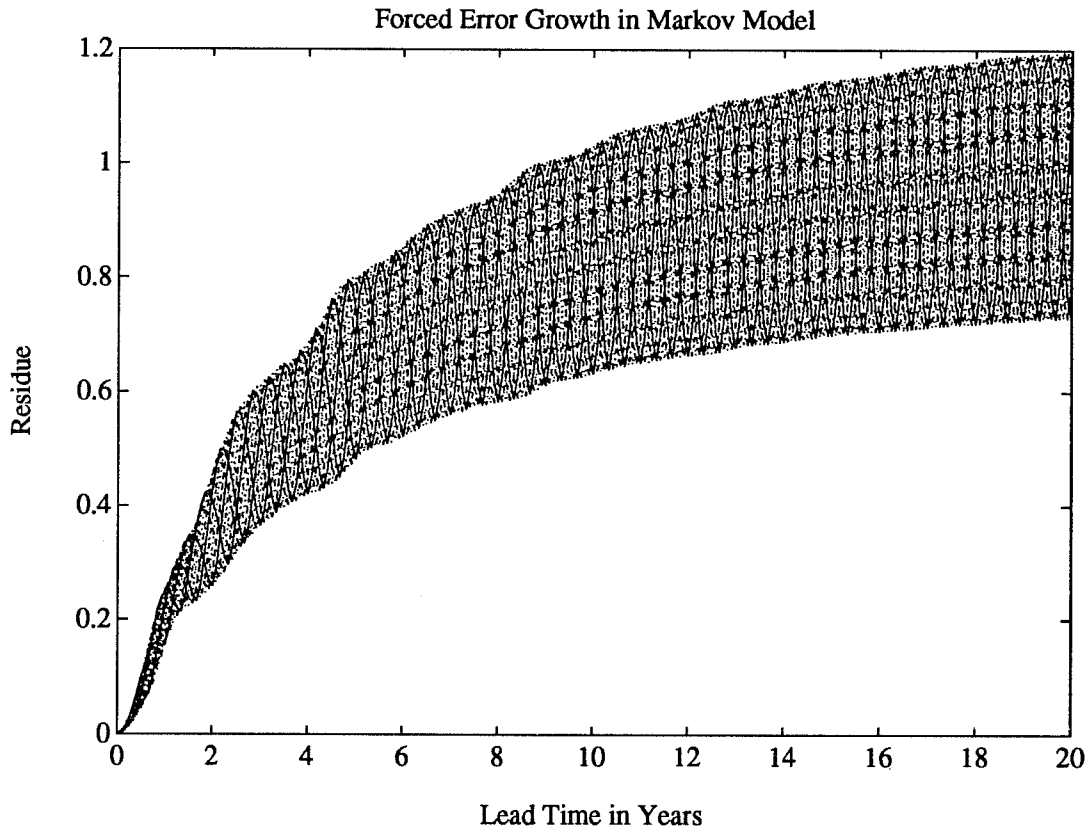

Figure 8: Time evolution of the linear model error residue

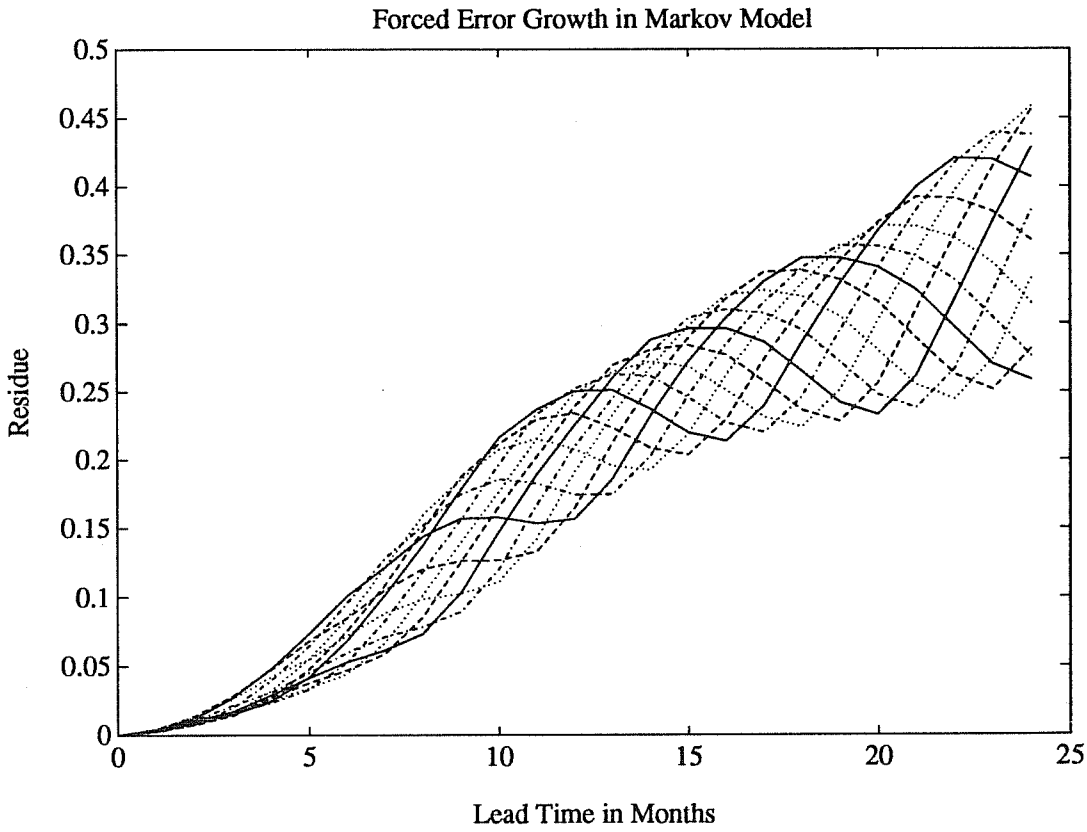

Figure 9: Early time evolution of the linear model error residue

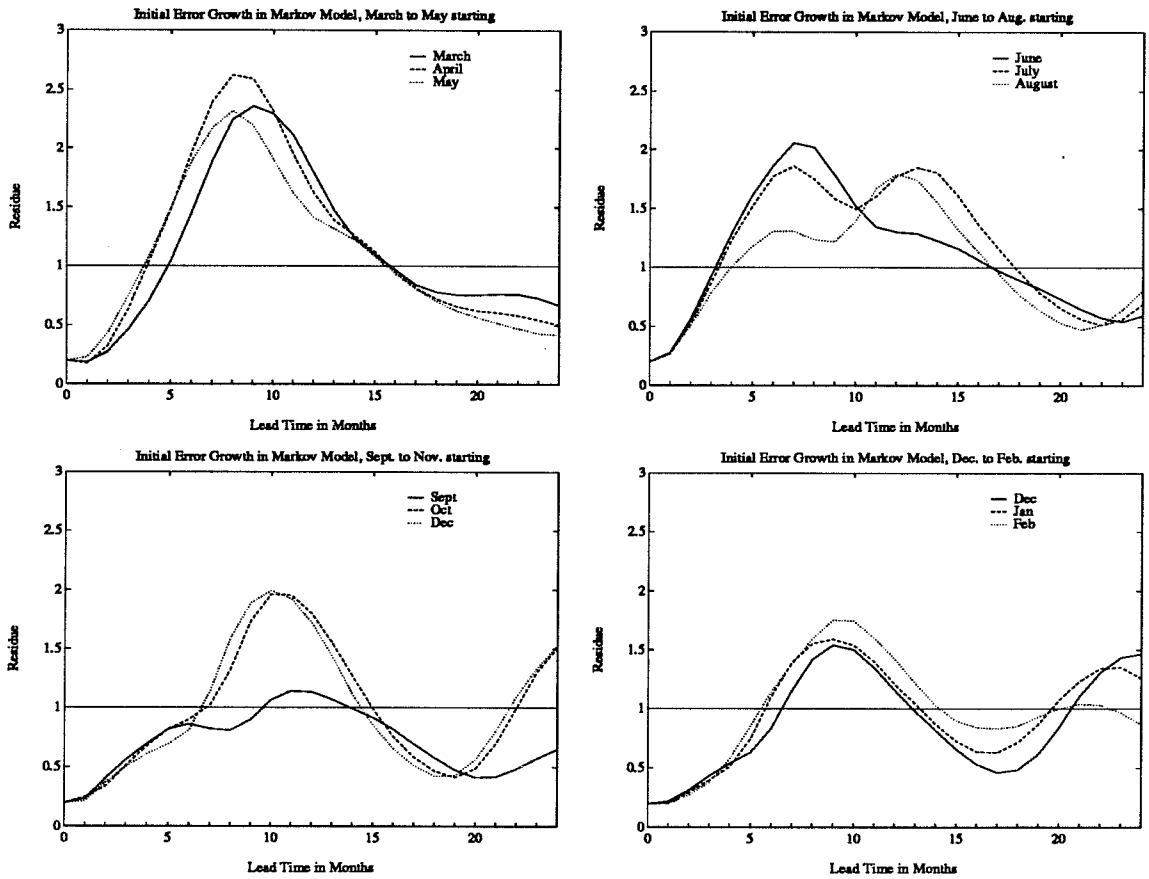

Figure 10: Initial error growth: the time evolution of an error ball initially uniform in SST. Each line is labeled with its starting month.

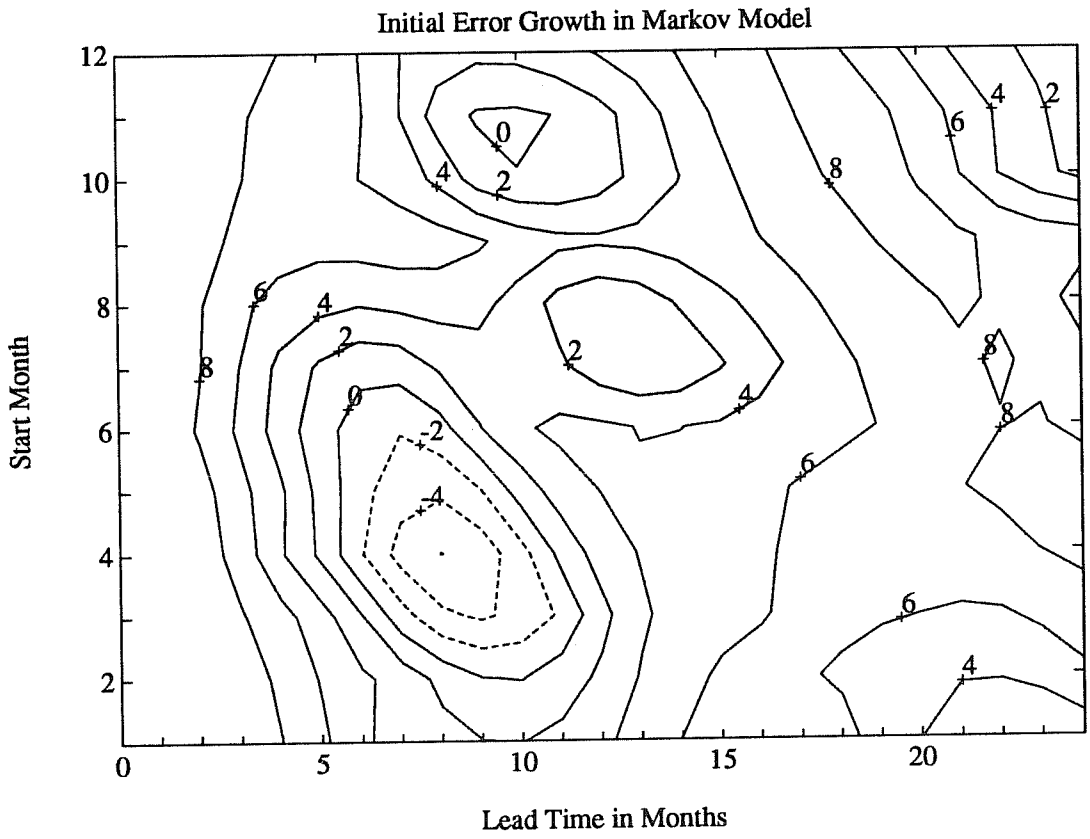

Figure 11: Correlation between the linear model and the true state, as estimated by using an initial error ball with error uniform in SST, sea level, and wind stress

5.2 Initial error growth

It seems clear from the slow forced error growth seen above, that the relatively poor results when compared to data must be due to other causes. If we now consider the unforced error growth of a uniform error ball, (i.e. error evolution under (5) with $\mathbf{Q}$ taken to be zero), we can gain some understanding of the growth of initial error in this system.

We choose an initial value of the error covariance $\mathbf{P}$ to be diagonal and white in SST, sea level, and wind stress, i.e. error is uniformly distributed over the 75 EOFs. This initial error is scaled to be 20% of the signal variance in each field.

Figure 10 shows the evolution of this initial error ball as measured by residue. The residue grows by a factor of as much as ten (for April starts) and as little as four (for September starts), all within a period of six months. This shows that the linear system can in fact have rapid error growth, unlike what we saw in the forced case. This is consistent with our earlier results where the linear model was used to make predictions of the measured NINO3.

By a relatively simple construct, we can actually turn these error growth calculations into predictions of NINO3 correlation patterns such as are plotted in figure 5. First define the NINO3 time series as $\overline{T(t)}$, where the overbar denotes an average over the NINO3 region, 5°S to 5°N,

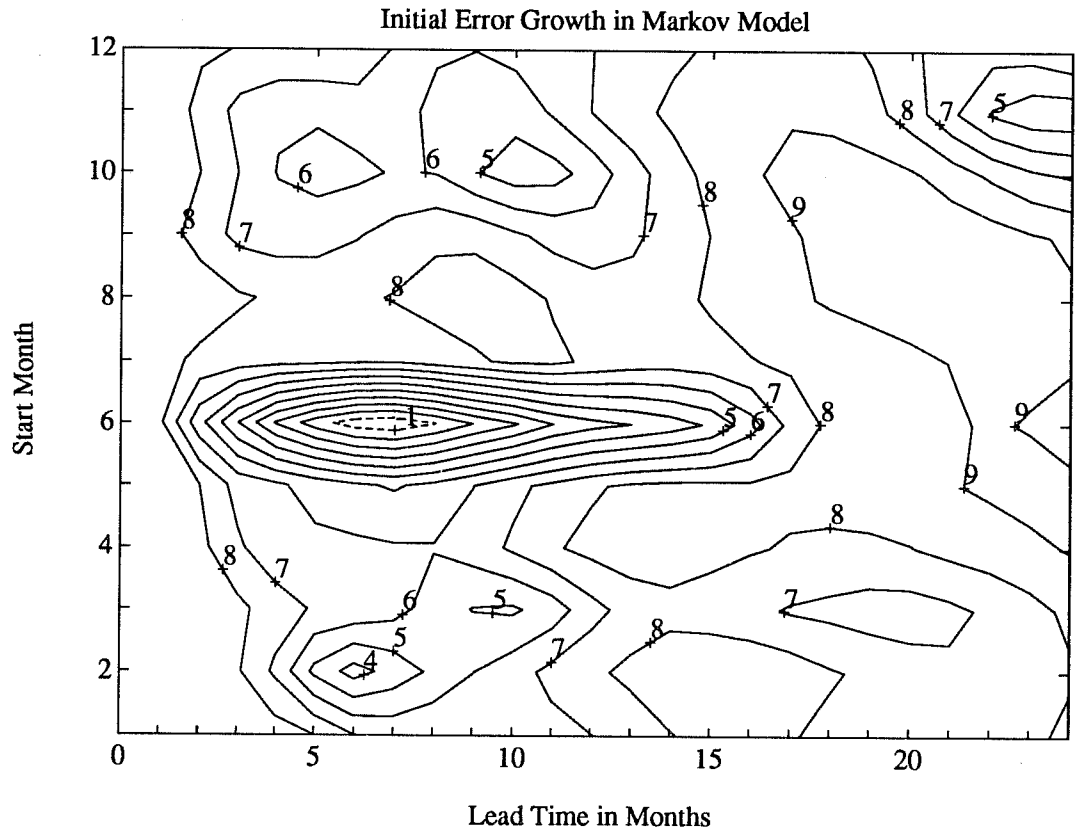

Figure 12: Correlation between the linear model and the true state, as estimated by using an initial error ball with error distributed according to the signal (i.e. the eigenvalues).

90°W to 150°W. We can calculate the mean square difference between two realizations of $\overline{T(t)}$ in terms of EOFs, namely

$$\langle(\overline{T}-\overline{T}')^2\rangle = \sum_i\sum_j\langle\epsilon_i\epsilon_j\rangle\overline{s_i s_j} \tag{13}$$

$$= \sum_i\sum_j P_{ij}\overline{s_i s_j} \tag{14}$$

where S are the spatial EOFs of SST, and ϵ_i is the error in the i^{th} eigenvector's amplitude. On the hand, by definition we have

$$\langle(\overline{T}-\overline{T}')^2\rangle = \langle\overline{T}^2\rangle + \langle(\overline{T}')^2\rangle - 2\langle\overline{T}\overline{T}'\rangle \tag{15}$$

If we assume that T and T' are close enough so that they have the same variance, then we can rewrite this as

$$\langle(\overline{T}-\overline{T}')^2\rangle = 2\langle\overline{T}^2\rangle(1-\gamma) \tag{16}$$

where γ is the correlation. Such a correlation is plotted in figure 11, which shows a rapid drop in correlation due to this error growth. Unfortunately it is difficult to compare to figure 5, because the initial error is in fact much more variable than the initial error ball we have used in this analysis. As an example, consider figure 12, which shows the correlation computed from an error ball whose initial error is distributed according to the signal covariance, i.e. proportional to each EOF's eigenvalue. Here the error growth rates are reduced, and correlation drops more slowly. This suggests that a properly designed initialization filter may have a significant impact on the error growth rates in the model.

6. ANALYSIS

While the last section successfully characterized the error growth in the linear system, there was no explanation of how such error growth can come about. In this section we see that rapid error growth is possible, even likely, in non-self-adjoint systems, even when the signal is primarily decaying.

6.1 POPs

The eigenvectors $\mathbf{e}_j$ of the linear model A are the POPs (Hasselmann, 1988) of the system. This means that

$$\mathbf{A}\mathbf{e}_j = \lambda_j\mathbf{e}_j, \tag{17}$$

where λ_j is the j^{th} eigenvalue. Because A is not symmetric, the eigenvalues and eigenvectors are possibly complex; because A is real, the complex eigenvectors and values come in conjugate pairs. We thus use the term POP to either refer to a single real eigenvector with a real eigenvalue (a vector that will simply decay with time) or to refer to the pair of real components that form a complex eigenvector with complex eigenvalue (a pair of vectors that oscillate and decay with time). For the most part, the pair of complex conjugate POPs will be presented together: the plot vs. time will show both the real and imaginary components, and the spatial patterns will be given for both the real and imaginary parts.

6.2 POP-cycle method

It is possible to calculate a linear model that takes explicit account of the seasonality of the system. In the case of monthly data we have a set of twelve transition matrices $\{\mathbf{A}^{(i)}\}$, where $\mathbf{A}^{(1)}$ is the transition from Jan to Feb, $\mathbf{A}^{(2)}$ is the transition from Feb to Mar, etc. In the case of seasonal data we have a set of four transition matrices $\{\mathbf{A}^{(i)}\}$, where $\mathbf{A}^{(1)}$ is the transition from Feb to May, $\mathbf{A}^{(2)}$ is the transition from May to Aug, etc.

When the model has an explicit seasonal dependence, then we need to modify how we calculate POPs. We are not interested particularly in the eigenvectors of any single seasonal transition; rather we want an analysis that encompasses all seasons. One possibility is to do the eigenanalysis of the yearly transition matrices $\mathbf{Y}^{(i)}$,

$$\begin{aligned} \mathbf{Y}^{(1)} &= \mathbf{A}^{(n)} \ldots \mathbf{A}^{(1)} \\ \mathbf{Y}^{(2)} &= \mathbf{A}^{(1)} \mathbf{A}^{(n)} \ldots \mathbf{A}^{(2)} \end{aligned} \tag{18}$$

etc, i.e. $\mathbf{Y}^{(i)}$ is the transition matrix from month i in the current year to month i in the next year. It turns out that the eigenvalues of the $\mathbf{Y}^{(i)}$ do not depend on month i, since if

$$\mathbf{Y}^{(i)} \mathbf{e}_j^{(i)} = \lambda_j \mathbf{e}_j^{(i)}, \tag{19}$$

then

$$\mathbf{Y}^{(i+1)} (\mathbf{A}^{(i)} \mathbf{e}_j^{(i)}) = \mathbf{A}^{(i)} \mathbf{Y}^{(i)} \mathbf{e}_j^{(i)} = \lambda_j (\mathbf{A}^{(i)} \mathbf{e}_j^{(i)}). \tag{20}$$

So that for all i we get a single set of eigenvalues λ_j, where the eigenvectors do depend on month. A complex eigenvalue corresponds to a decay and phase–shift. This means we have found a set of patterns such that evolution from one month to the same month a year later is given by a phase–shift and decay, and that same phase–shift and decay gives the evolution for the other seasons as well. Thus instead of the single POP pattern for a particular eigenvalue that we found in calculating non-seasonal POPs, we have a POP-cycle of patterns that vary within each year, yet all patterns have the same year–to–year behavior. The eigenvectors for different months are related by

$$\alpha_j^{(i)} \mathbf{e}_j^{(i+1)} = \mathbf{A}^{(i)} \mathbf{e}_j^{(i)} \tag{21}$$

where $\alpha_j^{(i)}$ is such that $\mathbf{e}_j^{(i+1)}$ is properly normalized, i.e. $||e_j^{(i)}|| = 1$ and phase is adjusted so that the real and imaginary parts are orthogonal. Note that this implies

$$\lambda_j = \alpha_j^{(n)} \ldots \alpha_j^{(1)} \tag{22}$$

Just as λ gives the yearly phase–shift and decay for the modes, as long as the patterns do not change much, the $\alpha^{(i)}$ give the phase–shift and decay from season to season. Conceptually, it is much easier to fully understand the time evolution of the complex scalar $\alpha^{(i)}$ than the complex vector field $\mathbf{e}^{(i)}$. The α give the growth/decay and phase evolution from season to season, with the proviso that, if the patterns $\mathbf{e}^{(i)}$ change dramatically, the meaning of a phase shift is less clear.

6.3 POP-cycle analysis of the ENSO model

As it turns out, for the ENSO POP-cycle in the CZ model most of the seasonal variation is in the $\alpha_j^{(n)}$—the changes in the $\mathbf{e}_j^{(i)}$ are relatively minor, being primarily a seasonal intensification

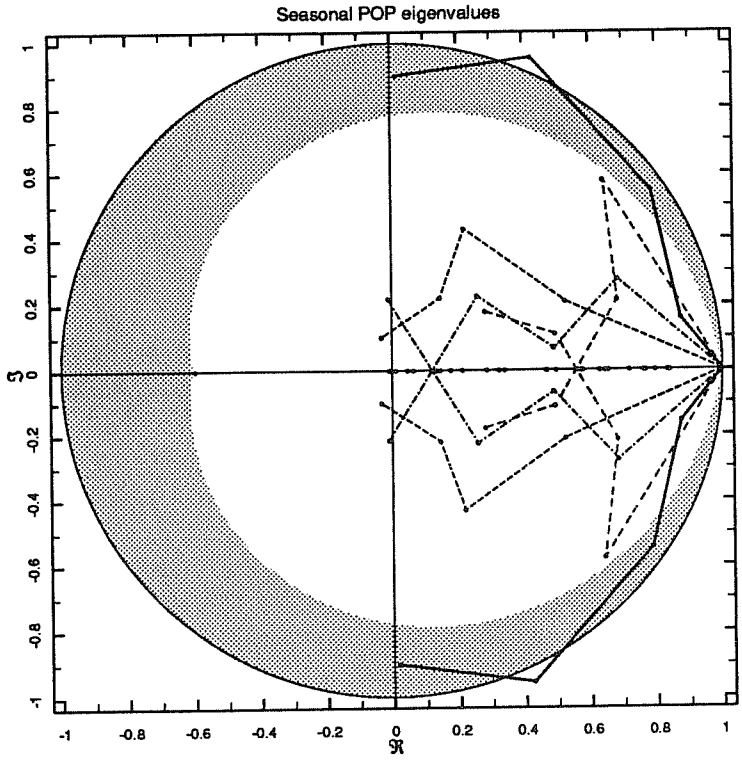

Figure 13: Eigenvalues for the POPs of the seasonal model: the initial point (1,0) is February, the evolution is given for May, August, November, and the following February. Final February points in the shaded area have e-folding (decay) times longer than their period.

of SST, with the eastern extreme in the real part being more tightly confined and stronger in November. Aside from the ability of the linear combination of two orthogonal patterns (such as sine and cosine) to represent a propagating structure, there is no obvious reason why the patterns in $\mathbf{e}^{(i)}$ should be so seasonally independent. Because the patterns change so little, most of the seasonality in the model can be thought of as changes in phase and decay rather than changes in structure.

Figure 13 gives the eigenvalues for the original seasonal model, i.e. each step is 3 months—one season. Each set of connected points gives the seasonal (single year) evolution for one POP-cycle: the initial point (1,0) corresponds to February, the next point is in May, the last is February of the next year. Note that the final February value is the yearly eigenvalue: it gives the evolution from one February to the next February. The other points give the (May-August-November) values from the initial February value, so that in general they are maps between different structures. Thus the correlation constraint that forces all modes to decay (i.e. the eigenvalues have to be less than one) only applies to the yearly transition, and we do indeed see values outside the unit circle (the November value is slightly greater than one).

The figure shows that only one POP-cycle has a clean seasonal evolution: it has a 4 year period and a 9.5 year decay time. This POP-cycle's structures show relatively little seasonal dependence, thus the behavior of the eigenvalues is a good first order approximation to its behavior. The eigenvalues show growth from May to November, and decay from November to May. The decay is greater than the growth so that on a yearly basis the POP-cycle does in fact

decay.

There is only one other POP-cycle whose period is even close to being shorter than its decay rate: it has period 2 years and decay time of 1.9 years. Because it is purely real, there is no way to consider its seasonal phase evolution (unless two of its structures happen to be simple linear combinations of the other two seasonal structures). In any case, this paper will focus on the 4 year POP-cycle, the ENSO POP.

Figure 14 shows the time series for the POP-cycle that has period 4 years. The timeseries shows a steady evolution of phase for the most part, the evolution going as $(\Im) \rightarrow (\Re) \rightarrow (-\Im) \rightarrow (-\Re)$. This corresponds to clockwise phase evolution on the phase diagram (figure 13). To show the correspondence with ENSO, plotted along with the time series of this POP is the model's NINO3 index. Comparison of the POP component time series with NINO3 shows that the real component is highly correlated with NINO3. Thus the POP plotted is the POP that corresponds to ENSO.

Figure 15 shows the ENSO POP components for November over the three ranges given in the time series plots in figure 14. It shows that the phase tends to be $0 \pm n90$, i.e. one tends to see the real part or the imaginary part rather than a mixture of the two. This means the spatial structures of figure 16 tend to be what is actually observed in any given November.

Figure 16 gives the November spatial structure of the ENSO POP for this explicitly seasonal model. 90° out of phase with NINO3 ($\Re$) there are essentially flat patterns in SST and wind (a slight cooling in the east), while sea level shows a depression in the west and elevation in the eastern off-equatorial regions (poleward of 15°N and 15°S).

These two patterns in sea level suggest a scheme where one year preceding an ENSO event, winds and SST are essentially normal, but sea level shows some elevation in the west and some deficit in the off-equatorial east. This is followed by the height of an ENSO event, when sea level is high throughout the east, SST is warm in the equatorial east, and winds show some convergence above that warm anomaly. One year later, eastern equatorial sea level is back to normal while eastern off-equatorial sea level remains relatively high between 140°W and 120°W. This suggests that the off-equatorial warm anomaly is propagating westward relatively slowly compared to the equatorial response. Because this is a linear model, the rest of the cycle follows with the negative (though slightly weaker since the mode is decaying) version of the above. Thus there are relatively cool temperatures in the east and a matching eastern sea level depression. Note that the westward–propagating off–equatorial anomaly merges nicely with the elevation anomaly in the western Pacific that appears two years after an El Niño event. This scenario, with a slowly propagating off-equatorial elevation anomaly appearing to separate ENSO events, corresponds with earlier observations of how the ENSO cycle behaves (Cane et al., 1986).

Because POPs oscillate by transferring energy from one structure to another, it is very important that enough dynamical variables are analyzed to retain the energy at all times. In this analysis, wind and SST have little expression in the structures out-of-phase with El Niño, while sea level has two distinct structures, one in-phase with El Niño and one not. Thus if one were to do a POP analysis using only winds or only SST, there would be no energy in the out-of-phase component and the POP analysis would have difficulty recognizing an oscillation. Sea level, on the other hand, would have no such problem. Latif et al.(1990) do a POP analysis of sea level anomaly data. They find a POP with period 40 months and decay time of 16 months; the period, at least, is quite similar to what we have found (and the more rapid decay is expected

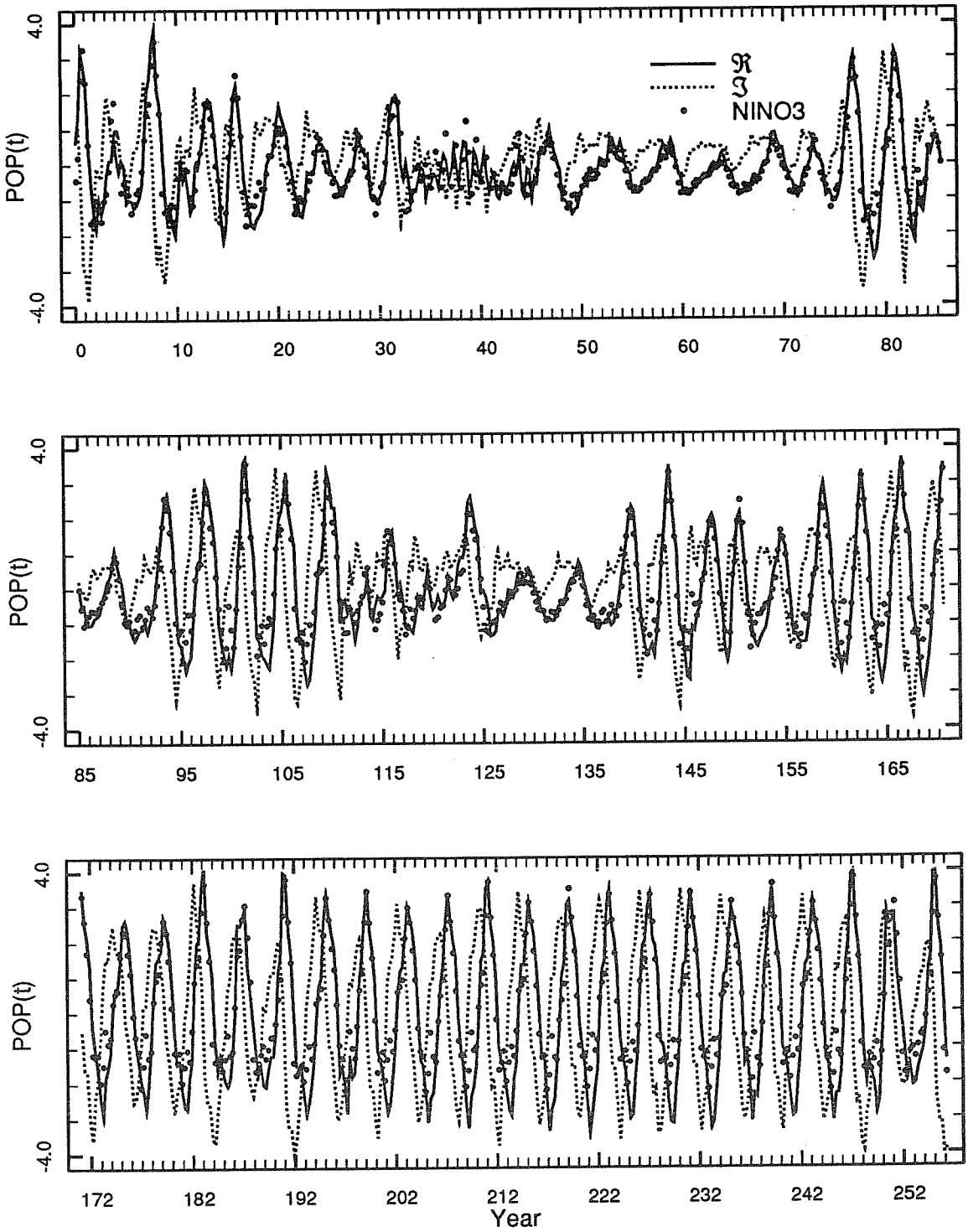

Figure 14: Time series of ENSO POP for the seasonal model: the real part shows a strong correlation with the NINO3 index.

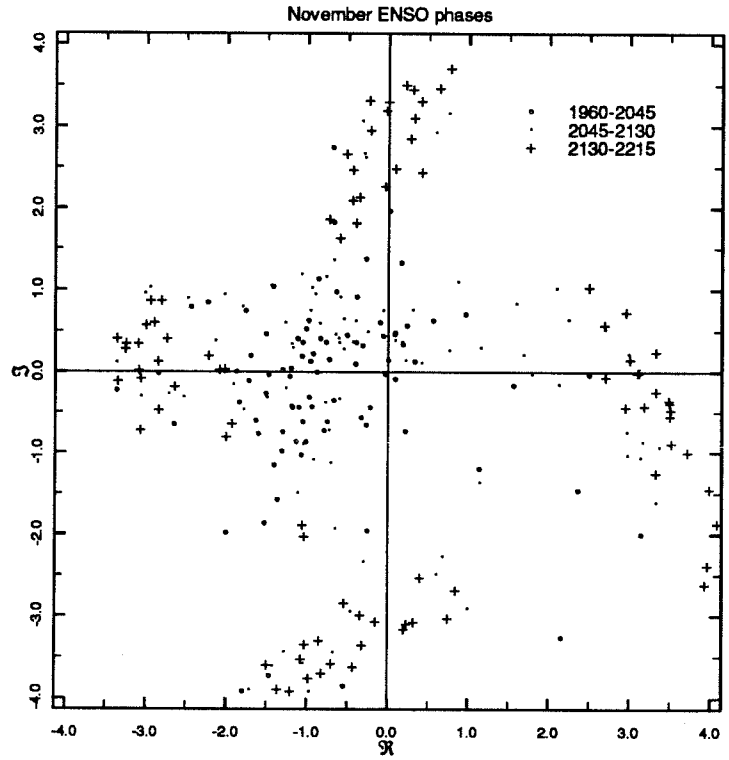

Figure 15: November ENSO POP components for the explicitly seasonal model: phase tends to be near $\pm n90°$.

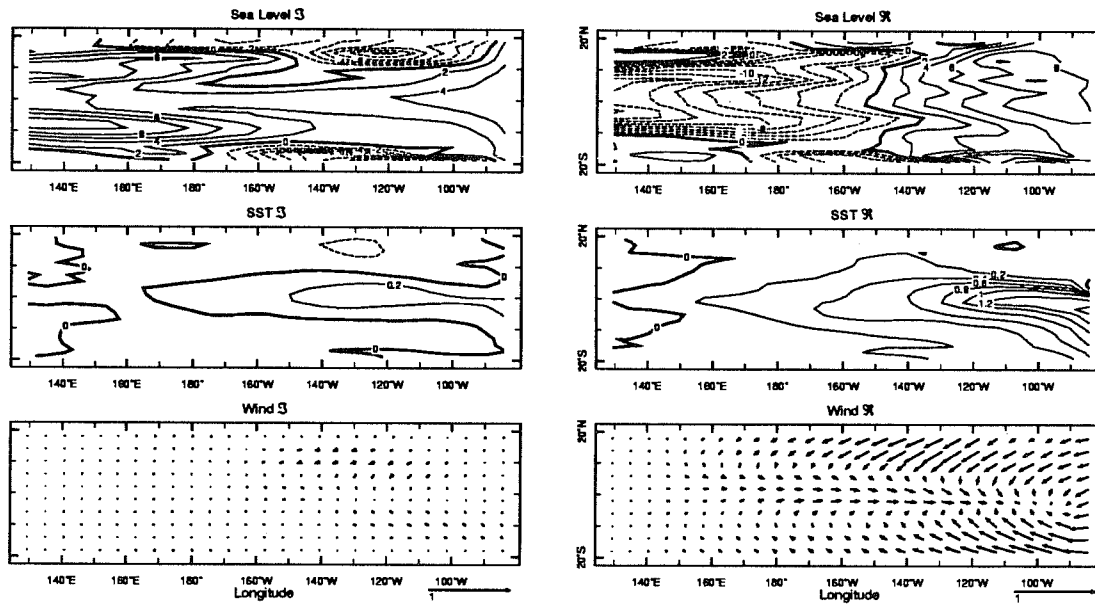

Figure 16: November spatial structure of ENSO POP for seasonal model. SST and winds show little signal out-of-phase with El Niño; sea level has structures that manifest westward propagation between El Niño events.

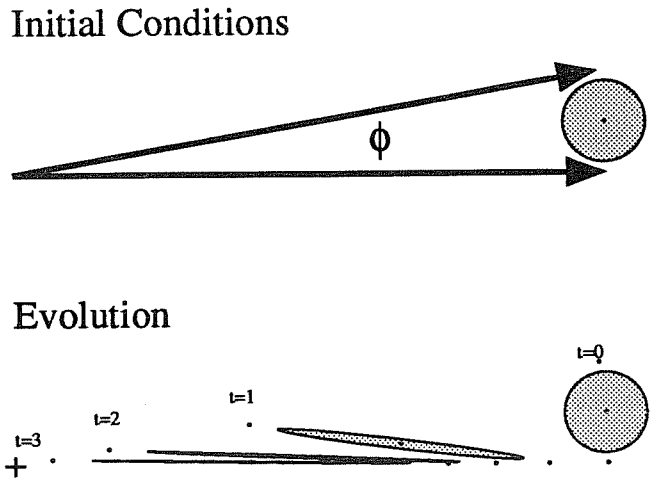

Figure 17: Error evolution in a 2D model: the error (shaded circle) has one dimension that elongates with time, despite the fact that the two dynamically independent dimensions of the system (the arrows) are steadily shrinking.

in a dataset that is both shorter and contains more noise). The in-phase component of the POP structure is quite similar, showing the shift of water between the east and west. The out-of-phase component, on the other hand, is peaked on the equator, not at higher latitudes. This difference could in part be due to the fact that Latif et al.(1990) do not have any off-equatorial data in the eastern Pacific, so that they would not be able to resolve the sort of peaks seen in the model. The correspondences found in the period and in-phase structures suggests that the ocean and the model do have this POP structure in common.

As a stability check, we redid the analysis using 10 EOFs for each of the three dynamical variables, giving us 30 EOFs in all. We will call this the 3×10 case. In this case (τ, h, T) have 97%, 93%, and 97% of the variance, respectively. The POP analysis results were very much the same as the 3×5 case—these two POPs remained the only POPs with a reasonably slow decay rate: one had period 4.01 years with a 17.5 year decay time (the ENSO POP), and the other had a 2 year period with a 8.8 year decay time. The time series for the ENSO POP was quite similar to the 3×5 mode case, with a strong correlation between the NINO3 index and the real part of the ENSO POP.

In the next section, we investigate the initial error growth (and thus the predictability) of the CZ model by analyzing the corresponding properties of the linear models. While the complete linear model is used to make the predictions (rather than using one or two of the linear model's POPs to make the predictions), the dominant POPs (particularly the ENSO POP) are helpful in understanding the behavior of the complete linear model. In particular, the ENSO POP will make it possible to imagine an analogous 2D linear system that has the same mechanism for error growth as the complete linear systems.

6.4 Error growth in a 2D non-self-adjoint system

By analyzing a simple 2D system, we can understand how error growth comes about in this ENSO model. Error growth in a simple 2D non-self-adjoint system is discussed extensively in Blumenthal (1991). Here we summarize briefly. This problem of initial error is illustrated in figure 17. The upper part of the figure shows the initial condition; the lower part the time evolution. The two-dimensional system is represented by two vectors that shrink at different rates ($\lambda = .9$/step, $\gamma = .4$/step, respectively) as the system evolves. Since the two vectors evolve independently, they are in fact the POPs of the system. The two vectors are 10° apart, i.e. they are not orthogonal. Thus we have dynamically independent vectors (POPs) which are spatially somewhat similar.

Our measured value of the initial condition is represented by a dot; we will call that the nominal initial condition. The initial uncertainty is represented by a shaded circle about that dot. This shaded circular area can be thought of as a collection of initial conditions, all of which have some probability of being the true initial condition given our uncertain measurement of the nominal initial condition. Calculating the evolution of this collection of points is exactly equivalent to calculating the evolution of covariance according to (5), but is more intuitive: instead of the evolution of a single state, we are going to calculate the evolution of a collection of states, states which start close to our nominal initial condition.

The two dots not contained in the shaded region give the evolution of unit eigenvectors, i.e. they decay with the one at 10° to the x-axis decaying much faster (.9/.4).

At the initial time, even though the error circle encompasses a small area, it includes points that are nearly purely one POP and points that are nearly the purely the other POP. As the system evolves, the error circle deforms to the point where one ellipse axis is much longer than the initial radius of the circle. Thus error has grown in that particular direction. Because one of the POPS has decayed rapidly while the other has not, it makes a big difference in the *future* state how much energy was put into each vector *initially.* Consequently the future state is relatively uncertain. In the long run, of course, both POPs have decayed completely, but by then we have lost interest in the problem.

The 2D model analysis shows that error growth comes about when two conditions on the model are satisfied:

1. two POPs have similar spatial structure, and
2. the two POPs have different time behavior.

Both conditions are necessary for error growth, and if the initial error has significant projection on both POPs, sufficient as well. The 2D analysis also shows that the elongation of the ellipse points approximately along the slowly decaying component, especially after the fast component has been given time to decay. Figure 13 shows that the ENSO POP-cycle is much slower to decay than the remaining POP-cycles. It is also true that the remaining POP-cycles taken as a whole are spatially similar to the ENSO POP-cycle (see Blumenthal (1991) for a more detailed analysis). Thus both criteria are satisfied, and one would expect error growth due to this non-self-adjoint error growth mechanism.

7. SUMMARY

Constructing a linear approximation to the CZ model gave us two major results: a POP-cycle that gives a clear ENSO cycle, and error growth that is due to the non-self-adjoint nature of the system, unlike the 'normal mode' (Lyapanov exponent) instability more frequently analyzed in such systems. The POP-cycle isolates and clarifies features previously noted in the literature, such as the slowly propagating off-equatorial elevation anomaly appearing to separate ENSO events (Cane et al., 1986). The apparent six month growth rate found also corresponds to what was found in earlier work (Goswami, B.N. and J. Shukla, 1991).

The high correlations found between the linear model and the CZ model in the comparisons with zero initial error (figures 1 and 3) show that the linear model does accurately reproduce the CZ model's short-term behavior under such conditions. Thus the linear model can be used to help us understand the CZ model's predictability.

The decreased correlation found when the two models were started with states that contain noise suggests that a substantial fraction of the prediction error comes from initialization. The fact that a redder initial noise spectrum led to slower error growth suggests that an initialization procedure that filters higher EOF noise should improve predictability.

In the future, we would like to do calculations with better estimates of the initial error. Clearly there is a wide range of behaviors possible with different initial error covariances, and estimates with errors closer to what is observed would be helpful. Once we are successful in accurately modeling the error in our current initialization procedure, we will be able to determine what kind of data will improve the initial conditions, thus improving our predictions. We would also like to improve the EOF-reduced model by incorporating the major non-linearities. While it is not certain that this will improve things significantly, it is quite possible.

Our most striking result is that we find slow error growth in the forced error problem starting with zero initial error, yet we found relatively rapid error growth with the two initial error evolutions we tried. This splits the state space of the model into two manifolds. One manifold contains the set of states the model tends to stay in when run by itself: there it apparently has relatively slow error growth. Yet there is a second manifold: the states that we found with our initialization procedure have rapid error growth. This leaves two possibilities:

- There could be a defect in the model, a defect such that the model does not go be itself into the states necessary to do predictions. This would be saying whatever skill the model has in prediction is distinct from its behavior when run by itself for long periods of time.
- There could be a defect in the initialization, a defect such that the state found by the initialization is not in the slowly evolving manifold of the model where it ought to be. Given the defects in the data, it is virtually certain that the initial state is not the correct one. Whether we can correct it by moving to the nearest point in the slowly varying manifold will only be clear after more work.

Acknowledgements. This work was funded by TOGA grant NA16RC-0432-01 and ONR grant N00014-90-J-1595.

References

Blumenthal, M. B., 1991: Predictability of a coupled ocean-atmosphere model. *J. Climate*, 4(8), 766–784.

Cane, M. A. and Zebiak, S., 1987: Deterministic prediction of El Niño events. In Cattle, H., editor, *Atmospheric and Oceanic Variability*, pp. 153–182. Royal Meteorological Society/American Meteorological Society.

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

Farrell, B., 1988: Optimal excitation of neutral Rossby waves. *J. Atm. Sci.*, **45**, 163–172.

Farrell, B., 1989: Optimal excitation of baroclinic waves. *J. Atm. Sci.*, **46**, 1193–1206.

Farrell, B., 1990: Small error dynamics and the predictability of atmospheric flows. *J. Atm. Sci.*, **47**(20), 2409–2416.

Goldenberg, S. and O'Brien, J., 1981: Time and space variability of tropical Pacific wind stress. *Mon. Wea. Rev.*, **109**, 1190–1205.

Goswami, B.N. and J. Shukla, 1991: Predictability of a Coupled Ocean-Atmosphere Model. *J. Climate*, **4**, 3–22.

Hasselmann, K., 1988: PIPs and POPs: the reduction of complex dynamical systems using principal interaction and Oscillation patterns. *J. Geophys. Res.*, **93**, 11015–11021.

Lacarra, J.-F. and Talagrand, O., 1988: Short-range evolution of small perturbations in a barotropic model. *Tellus*, **40A**, 81–95.

Latif, M., Flügel, M., and Xu, J.-S., 1990: An investigation of short range climate predictability in the tropical pacific. Technical Report 52, Max-Planck-Institut für Meteorologie.

Strang, G., 1988: *Linear algebra and its applications.* Harcourt, Brace, and Jovanovitch.

von Storch, H., Bruns, T., Fischer-Bruns, I., and Hasselmann, K., 1987: Principal Oscillation Pattern analysis of the 30-60 day oscillation in a GCM equatorial troposphere. Technical Report 7, Max-Planck-Institut für Meteorologie.

von Storch, H., Bruns, T., Fischer-Bruns, I., and Hasselmann, K., 1988: Principal Oscillation Pattern analysis of the 30-60 day oscillation in a GCM equatorial troposphere. *J. Geophys. Res.*, **93**.

Xu, J.-S. and von Storch, H., 1989: Principal Oscillation Pattern–Prediction of the state of ENSO. Technical Report 35, Max-Planck-Institut für Meteorologie.

Xu, J.-S. and von Storch, H., 1990: Predicting the state of the Southern Oscillation using principal oscillation pattern analysis. *J. Climate*, **3**, 1316–1329.

Zebiak, S. E., 1989: On the 30-60 day oscillation and the prediction of El Niño. *J. Climate*, **2**, 1381–1387.

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