Tellus (1994), 46A, 512–528
Printed in Belgium – all rights reserved

TELLUS
ISSN 0280–6495

On the prediction of ENSO: a study with a low-order Markov model*

By YAN XUE[1], M. A. CANE, S. E. ZEBIAK and M. B. BLUMENTHAL, *Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY 10964, USA*

(Manuscript received 13 August 1993; in final form 22 February 1994)

ABSTRACT

A linear model best fit to the Zebiak and Cane (1987) ENSO forecast model (ZC) is used to study the model's prediction skill. Multivariate empirical orthogonal functions (MEOFs) obtained from the sea surface temperature anomaly, sea level and wind stress anomaly fields in a suite of 3-year forecast runs of ZC starting from the monthly initial conditions in the period January 1970 to December 1991, are used to construct a series of seasonally varying linear Markov models. It is found that the model with 18 MEOFs fits the original nonlinear model reasonably well and has comparable or better forecast skill. Assimilating the observed SST into the initial conditions further improves forecast skill at short lead times (<9 months). The transient initial error growth in the model's prediction is attributed to the non-self-adjoint property as in Farrell and Blumenthal. Initial error grows fastest starting from spring and slowest starting from late summer and is sensitive to the initial error structures. Two singular vectors (SVs) of the linear evolution operator have significant transient growth dominating the total error growth. Since the optimal perturbation (fastest SV) has mostly high MEOF components, the error growth tends to be larger when there are more high mode components in the initial error fields. This result suggests a way to filter the initial condition fields: the MEOFs higher than the 18th in the initial fields are mostly noise and removing them improves prediction skill. The forecasts starting from late summer have the best predictability because the fastest growth season (summer) is just avoided. The well known, very rapid decline in forecast skill in the boreal spring (the "spring barrier") is here attributed to the smallness of the signal to be forecast: the standard deviation of the NINO3 SST anomaly is smallest in spring.

1. Introduction

ENSO, the strongest interannual signal in the tropics, has a significant impact on global climate including killing drought in Australia, prolific rainfall on desert lands in Peru and catastrophic failures of the Indian monsoon. Prediction of ENSO is valuable both for scientific interests and economic benefits. Several prediction schemes are now used routinely for ENSO prediction. Among them, the non-linear anomaly model of Zebiak and Cane (1987) (referred to as ZC hereafter) has been recognized as among the most successful. ZC is an anomaly model having all climatological fields specified. This tremendous simplification avoids the difficulties in simulating the correct climatological field, which constitutes a serious problem in many models, including GCMs. However, studies with coupled GCMs are beginning to show good prediction skill (Latif et al., 1993). The statistical models, challenging the numerical models, are best represented by the Canonical Correlation Analysis (CCA) procedure (Barnett et al., 1988; Graham et al., 1987a, b; Barnston and Ropelewski, 1992). The ZC forecasts have been reported (Cane et al., 1986; Barnett et al., 1988; Cane, 1991) in terms of the NINO3

* Contribution Number 5209 of Lamont-Doherty Earth Observatory of Columbia University.

[1] Corresponding author.

index (the averaged SST anomaly in the region 5°N–5°S and 90°W–150°W) with demonstrated skill at 1 to 2 years lead time. The CCA scheme can predict ENSO events 3 seasons ahead.

Many efforts have been devoted to understanding ENSO forecast skill and to improving the predictions (Cane et al., 1986; Cane, 1991; Graham et al., 1992; Barnett et al., 1993; Latif et al., 1993; Webster and Yang, 1992). The surprising successes in ENSO prediction by both simple numerical and statistical models are believed attributable to the low frequency characteristics of ENSO. The skill of the simple anomaly model (ZC) suggests that ENSO system is well approximated as an internal oscillation about the climatological mean state in the tropical Pacific. Thus far, additional errors introduced by coupled GCMs hurt more than the added complexity helps.

The useful forecast skill in ZC in terms of the NINO3 index is 1 year to 2 years and is seasonally varying. An important feature is the rapid decline in forecast skill in the spring. Although the ZC forecast skill is better than most ENSO prediction models at long lead times, its short term skill is relatively poor because of the poor initialization procedure.

With increasing numbers of ENSO forecasts, the seasonality of predictability (ENSO is least predictable in spring and most predictable in winter) has become well established. Cane et al. (1986) suggested that initial noise could be rapidly amplified in the summer to lead to a poor forecast; but in the winter, it may be readily dispersed by wave motions to lead to a good forecast. They attributed the phenomenon to the seasonally varying stability of the coupled ocean-atmosphere system in the tropical Pacific: the system is most unstable in summer and stable in winter. Webster and Yang (1992) focus on the rapid decline in forecast skills in the boreal spring, the so-called "spring barrier", and view ENSO as an interactive system with the monsoon circulation. They point out that the summer monsoon circulation develops fastest from April to May, at the same time the Walker circulation is the weakest and most susceptible to external noise. They go on to argue that since the difference between strong and weak summer monsoons has a significant component in the trade wind system, it forms a source of noise and possibly modulates the coupled system through changes in the annual cycle. Thus the monsoon is thought to play an important role at this special season for the ENSO cycle.

Error growth is a key feature in all model forecasts and has been studied thoroughly in numerical weather prediction (Lacarra and Talagrand, 1988; Farrell, 1989; Molteni and Palmer, 1993). It is found that if a system is not self-adjoint there is a possibility of transient growth in a mode which is not a growing normal mode of classical stability analysis. These modes are the singular vectors of the evolution operator and their transient growth rate are given by the singular values (Molteni and Palmer, 1993). Since the growth of the fastest transient mode is often more rapid than the growth of the fastest growing normal mode, it becomes the greatest concern for weather forecasting models. Molteni and Palmer (1993) point out that the normal mode growth rate could not explain the observed value of about 2 days for the doubling-time of small errors in numerical weather predictions. In contrast, the transient mode can double in less than 12 h. Using the growing transient modes as the initial perturbations, ensemble prediction is being done on sophisticated numerical weather prediction models (Mureau et al., 1993).

Similar transient modes are found in a linear Markov model best fit to ZC (Blumenthal, 1991). It is suspected that the growing transient modes due to the non-self-adjoint property are the candidates causing the fast initial error growth in ZC. As the correlation between the NINO3 indices in the linear Markov model and ZC is as high as 0.8 even after 2 years of integration (Blumenthal et al., 1991), we use the linear Markov model as a tool to understand the predictability of ZC. The questions we will address are: how well does a low order linear model fit the high order nonlinear model (ZC)? What controls the initial error growth in ZC? Can we understand the ZC ENSO forecast skill better and improve its predictions? What factors cause the "spring barrier" in ENSO prediction?

Section 2, describes the construction of the Markov model. In Section 3, the divergence between the linear and the nonlinear model is briefly discussed. Section 4 shows the ENSO prediction skill of a series of linear Markov models with variable dimensions. Section 5 deals with the non-self-adjoint transient initial error growth and

Section 6 with predictability. An improved forecast with SST assimilation in the initial conditions is presented in Section 7. Section 8 contains a summary and conclusions.

2. Model

2.1. Model created data

Most theories of ENSO suggest that sea surface temperature (SST), sea level (h) and surface wind stress (τ) are the 3 key fields for sustaining the ENSO cycle. These three fields from the real forecast runs of ZC are taken to be a sufficient description of the model evolution. The oceanic initial condition in each forecast run is from the ocean model component, which is driven by the FSU wind stress anomaly (Goldenberg and O'Brien, 1981) continuously from January 1964 to the time when the coupled model forecast starts. The atmospheric initial conditions are obtained by running the atmospheric model with the SST anomalies taken from the ocean model simulation. Starting from the initial conditions for each month between January 1970 and December 1991, ZC was integrated for 3 years, yielding a data set with 22x12x37 monthly values.

2.2. Space reduction

For each of the 3 key fields in ZC, the number of grid points is $O(10^3)$. It is essential to reduce the state space first. Empirical orthogonal function (EOF) analysis is applied here to represent the model physical fields by the first few EOFs, which maximize the variance representation. For example, a physical field represented by vector $\mathbf{v}(t)$ is decomposed into EOFs $\boldsymbol{e}_j$ and principle components (PCs) $a_j(t)$ and filtered by truncating at the Jth EOF,

$$\mathbf{v}(t) = \sum_{j=1}^{J} a_j(t)\, \boldsymbol{e}_j. \tag{1}$$

In the present context, 20, 40 and 10 EOFs account for 97.4%, 97.6% and 96.3% of variance in the SST, h and τ fields, respectively. It is probable that some of the retained EOFs mostly represent noise as pointed out by Graham et al. (1992). We return to this issue in Section 4 where truncation is discussed in terms of ENSO prediction skill.

The variance distributions among the retained EOFs for the initial and overall (initial conditions plus coupled model forecast) SST, h and τ fields are presented in Fig. 1. It is seen that the first EOF accounts for 81% and 75% of variance in the overall SST and τ fields, respectively, while in the overall sea level field the first and second EOFs account for 58% and 15% of variance. These variance distributions follow from the fact that the SST and τ fields are mostly stationary and the h field is propagating because of oceanic wave dynamics. Cane (1991) pointed out that the model initial conditions differ from the coupled model forecasts because the model winds differ from the FSU winds. The difference is clearly visible in the variance distributions among the retained EOFs (Fig. 1), with the major difference being in the sea level field. The 1st, 5th and 6th EOFs account for 14%, 14% and 10% of the variance in the initial sea level field, which is in contrast to the first two dominant EOFs in the overall sea level fields. In order to build a Markov model suitable for ENSO prediction studies, both the initial states and forecast states must be well represented in the reduced EOF space. As about 90% of the variance in the initial SST, sea level and wind stress fields are retained by 20, 40 and 10 EOFs, respectively, this reduced EOF space is sufficient for the current purpose.

In order to reduce the space further and to have a consistent multivariate basis, a second EOF analysis is conducted. A vector of dimension 70 is constructed from the PCs of SST, h and τ. Each of the 3 fields is given equal weight:

$$\boldsymbol{b} = \left[\frac{a_1^1}{\sigma_1} \cdots \frac{a_{J_1}^1}{\sigma_1}, \frac{a_1^2}{\sigma_2} \cdots \frac{a_{J_2}^2}{\sigma_2}, \frac{a_1^3}{\sigma_3} \cdots \frac{a_{J_3}^3}{\sigma_3}\right]^T. \tag{2}$$

Here, a_j^1, a_j^2 and a_j^3 are the PCs for SST, h and τ with dimensions $J_1 = 20$, $J_2 = 40$ and $J_3 = 10$; σ_1^2, σ_2^2 and σ_3^2 are the total variance described by each set of PCs.

$$\sigma_v^2 = \sum_{j=1}^{J_v} \sum_{t} [a_j^v(t)]^2, \qquad v = 1, 2, 3. \tag{3}$$

Then $\boldsymbol{b}$ is decomposed into a set of space components $\boldsymbol{f}_j$ and time components $d_j(t)$ and filtered

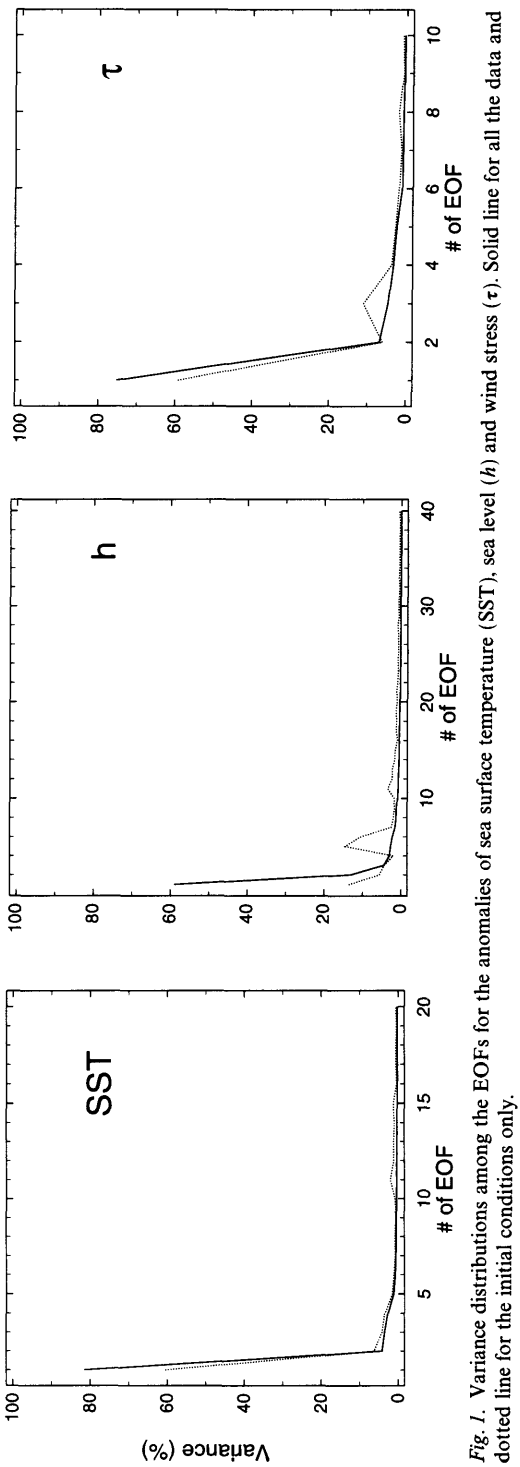

Fig. 1. Variance distributions among the EOFs for the anomalies of sea surface temperature (SST), sea level (h) and wind stress (τ). Solid line for all the data and dotted line for the initial conditions only.

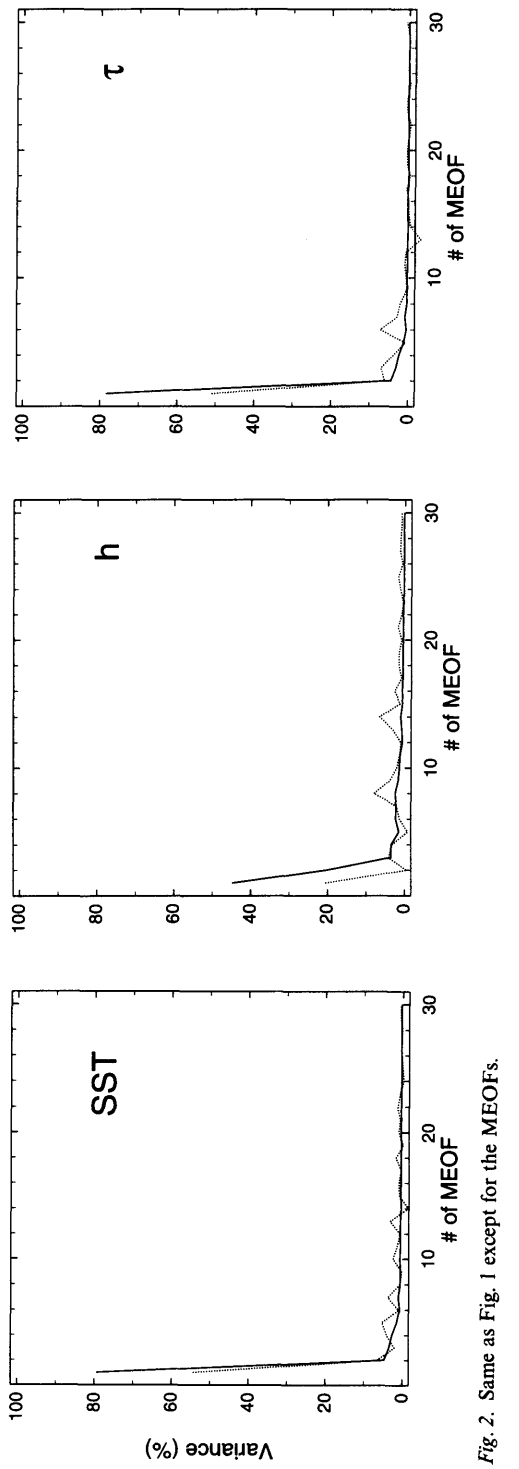

Fig. 2. Same as Fig. 1 except for the MEOFs.

by truncating at the Kth multivariate EOF (MEOF),

$$\boldsymbol{b}(t) = \sum_{j=1}^{K} \boldsymbol{f}_j d_j(t). \tag{4}$$

30 MEOFs account for 96%, 94% and 94% of variance in the total SST, h and τ fields, respectively. The time components $d_j(t)$, $j = 1, 2, ..., 30$, represent all 3 physical fields in the phase space spanned by the retained MEOFs. As in the EOF space, it is useful to see the different variance distributions of the initial and overall fields in the MEOF space. Fig. 2 shows that the first, the eighth and fourteenth MEOFs account for 21%, 8% and 8% of variance in the initial sea level field, while the first and second MEOF account for 46% and 20% of variance in the overall sea level field. The multivariate EOF analysis is essential to the space reduction and helpful in combining the model SST, h and τ fields together physically. The retained 30 MEOFs are sufficient for ENSO prediction studies, as shown in Section 4.

2.3. *Markov model*

The Markov model is computed as in Blumenthal (1991). Blumenthal noted that the seasonal model fits the original ZC model much better than the nonseasonal model. The importance of including seasonality in ENSO modeling has been much discussed in the literature (e.g., Cane et al., 1986; Barnett et al., 1993). Thus we construct 12 Markov models, representing the monthly transition from each calendar month. The suite of 3-year runs starting from each of the monthly initial conditions for the period January 1970 to December 1991 are concatenated to form a long series. Then the data set is grouped into 12 subsets, one for each calendar month. So each subset has 814 data points. The data in subset i and $i+1$ are used to calculate the monthly transition matrix from calendar month i to the next month $i+1$. Denoting the data in subset i by $\boldsymbol{d}_i$, the formula is:

$$\boldsymbol{d}_{i+1} = \boldsymbol{A}^{(i)}\boldsymbol{d}_i + \boldsymbol{e}_i, \tag{5}$$

where $\boldsymbol{A}^{(i)}$ is the transition matrix and $\boldsymbol{e}_i$ is the residue. Multiplying by the transpose of vector $\boldsymbol{d}_i$ on both sides of (5), then averaging on all samples gives:

$$\langle \boldsymbol{d}_{i+1}\boldsymbol{d}_i^{\mathrm{T}} \rangle = \boldsymbol{A}^{(i)}\langle \boldsymbol{d}_i\boldsymbol{d}_i^{\mathrm{T}} \rangle + \langle \boldsymbol{e}_i\boldsymbol{d}_i^{\mathrm{T}} \rangle, \tag{6}$$

where $\langle \cdots \rangle$ means the average over all samples in subset i (except those at the end of each 3 year run). For the best fit model $\boldsymbol{A}^{(i)}$, $\boldsymbol{e}_i$ does not correlate with $\boldsymbol{d}_i$, so

$$\boldsymbol{A}^{(i)} = \langle \boldsymbol{d}_{i+1}\boldsymbol{d}_i^{\mathrm{T}} \rangle \langle \boldsymbol{d}_i\boldsymbol{d}_i^{\mathrm{T}} \rangle^{-1} = \boldsymbol{C}_i\boldsymbol{D}_i^{-1}; \tag{7}$$

here $\boldsymbol{C}_i$ is the lag one covariance matrix, while $\boldsymbol{D}_i$ the autocovariance matrix.

We are not interested in the eigenvectors of any single monthly transition; rather we want an analysis that encompasses all the behaviors of the model. One way is to do the eigenanalysis of the yearly transition matrix $\boldsymbol{Y}$ (cf., Blumenthal, 1991),

$$\begin{aligned} \boldsymbol{Y}^{(1)} &= \boldsymbol{A}^{(12)} \cdots \boldsymbol{A}^{(2)}\boldsymbol{A}^{(1)}, \\ \boldsymbol{Y}^{(2)} &= \boldsymbol{A}^{(1)}\boldsymbol{A}^{(12)} \cdots \boldsymbol{A}^{(2)}, \end{aligned} \tag{8}$$

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

$$\boldsymbol{Y}^{(i)}\boldsymbol{e}_j^{(i)} = \lambda_j \boldsymbol{e}_j^{(i)}, \tag{9}$$

then

$$\boldsymbol{Y}^{(i+1)}(\boldsymbol{A}^{(i)}\boldsymbol{e}_j^{(i)}) = \boldsymbol{A}^{(i)}\boldsymbol{Y}^{(i)}\boldsymbol{e}_j^{(i)} = \lambda_j(\boldsymbol{A}^{(i)}\boldsymbol{e}_j^{(i)}), \tag{10}$$

so that for all i, we get a single set of eigenvalues λ_j, where the eigenvectors for different months are related by

$$\boldsymbol{A}^{(i)}\boldsymbol{e}_j^{(i)} = \alpha_j^{(i)}\boldsymbol{e}_j^{(i+1)}. \tag{11}$$

$\alpha_j^{(i)}$ is such that $\boldsymbol{e}_j$ is properly normalized, i.e., $\|\boldsymbol{e}_j\| = 1$ and its phase is adjusted so that the real and imaginary parts are orthogonal and the norm of the real part is larger than that of the imaginary part. Then this implies

$$\lambda_j = \alpha_j^{(12)} \cdots \alpha_j^{(1)}; \tag{12}$$

here, $\|\lambda_j\| < 1$ because the yearly transition matrix is a best fit to the data. The eigenvector $\boldsymbol{e}_j$ evolves from month to month by (11) and $\|\alpha_j^{(i)}\|$ gives the amplitude change for the mode from month i to month $i+1$. Although the eigenvector $\boldsymbol{e}_j$ decays

after a year, the amplitudes of some of the e_j do grow in summer and fall; cf., Fig. 4 of Blumenthal (1991). The consequences of these seasonal variation of the stability will be discussed later.

The eigenvectors e_j are often called the principle oscillation patterns (POPs) (Hasselman, 1988). Since the matrices are not symmetric, the eigenvalues and eigenvectors may be complex and the eigenvectors are not necessarily orthogonal. So the adjoint of any POP, which is orthogonal to all other POPs, is not parallel to the POP itself (non-self-adjoint). The fast transient error growth possible in this non-self-adjoint system will be discussed in Section 5.

Xu and von Storch (1989) used a single POP model to predict ENSO events. There are several possible criteria to select the POP, often called the ENSO POP. It might be chosen by its period and by having a long decay time (see Table 1). It might be chosen by variance explained. Note, however, that since the POPs are not orthogonal, their combined variance will be larger than 100%. A useful measure due to Xu and von Storch (1989) is defined as $\mu = 1 - \|\boldsymbol{d} - \tilde{\boldsymbol{d}}\|^2 / \|\boldsymbol{d}\|^2$, where $\boldsymbol{d}$ is the original field and $\tilde{\boldsymbol{d}}$ is the field represented by a single POP and the norm includes the average over all time. A value of 1 means the field is fully represented by the POP and a value $\leqslant 0$ means it is poorly represented. Table 1 shows that only the first two POP pairs project strongly on the data, with the first POP pair having the largest value of μ. The real component of the POP with period of 3.1 years simulates an ENSO mature phase while the small imaginary component represents a transition phase (not shown; see Blumenthal (1991)). The real component of this POP's time series correlates with the NINO3 index at 0.69, so it is the ENSO POP. However, it is not guaranteed that a single ENSO POP can explain all dynamic characteristics. We find that the POP pair with period of 6.8 years is also very important in predictions. Penland and Magorian (1993) also noticed that some POP time series in their Markov model are highly correlated and probably represent related physical processes. They concluded that the SST field can not be described by a single pair of POPs. We found that each individual eigenvector (POP) of the yearly transition matrix is very sensitive to the number of MEOFs retained and each individual POP does not necessarily represent a physical pattern, but all of them together are able to describe the ENSO cycle reasonably well. So the whole set of POPs is used to do prediction. Still, we have to decide how many MEOFs are retained when the Markov model is constructed.

Table 1. *The eigenvalues of the yearly transition matrix in MK18 (Markov model with 18 retained MEOFs) expressed by period and decay time (1st 2 columns), measure μ by a single POP (3rd column), the correlation between the real component of each POP's time series (4th column) and the NINO3 index and the correlation between the imaginary component of each POP's time series (5th column) and the NINO3 index*

	Period (years)	Decay (years)	μ (%)	Correlation (real-NINO3)	Correlation (image-NINO3)
1	3.1	2.7	32	0.7	−0.4
2	6.8	4.0	19	0.6	0.3
3	4.8	1.5	−1	−0.2	−0.1
4	2.1	1.2	−2	0.2	−0.2
5	3.7	0.6	−5	−0.1	−0.1
6	∞	1.0	−6	−0.2	0.1
7	6.1	0.9	−8	0.0	−0.1
8	∞	2.0	−12	0.0	−0.0
9	23.4	1.7	−14	−0.0	−0.1
10	2.9	0.8	−17	0.3	0.0

3. Divergence between linear Markov models and ZC

A series of linear Markov models are constructed by retaining different numbers K of MEOFs. They are initialized with the same monthly initial conditions in the period January 1970 to December 1991 as in the ZC runs but truncated with the corresponding K MEOFs. Each model is integrated for 3 years, and the correlations between each linear model NINO3 and the ZC NINO3 are calculated. The correlations look quite similar for the linear models with $K = 5$ to $K = 30$. As an example, the correlation of the linear Markov model with 18 MEOFs (MK18 hereafter) is presented in Fig. 3. MK18 can not follow ZC forever because of the decaying of the Markov model and the absence of nonlinearities. However for lead times up to about one year, which is about the limit of the useful predictability in ZC, MK18 is a good approximation to ZC.

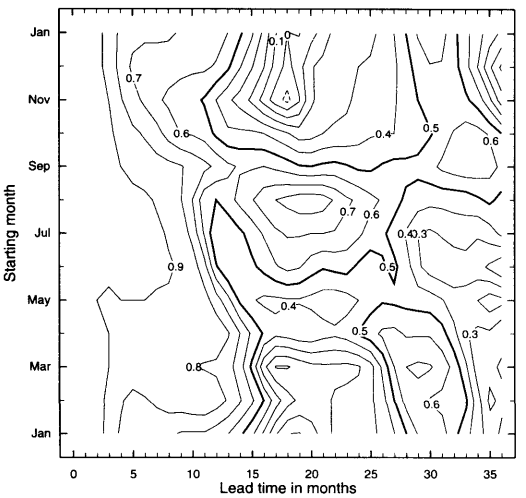

Fig. 3. MK18-ZC NINO3 correlation as a function of start months and lead months. Each correlation is on a sample of 22 (all years from 1970 to 1991).

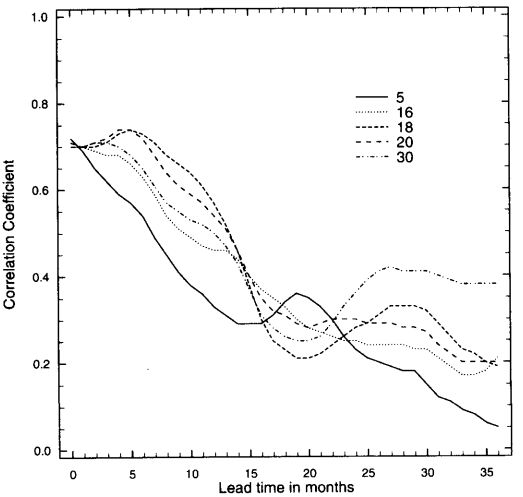

Fig. 4. Prediction-observation NINO3 correlations versus lead months, based on 228 individual forecasts (all starting months from January 1970 to December 1988). The line labels represent the dimensions of the different Markov models.

4. Forecast results: the effects of MEOF truncation

When the Markov models are tested against observations the verification data is not the data used to construct them, so the forecast skills obtained should be reliable. Fig. 4 shows the monthly averaged prediction-observation correlation for the models with 5, 16, 18, 20 and 30 retained MEOFs. The observed NINO3 anomaly in the period January 1970 to December 1991 from the Climate Analysis Center SST product (Reynolds, 1988) is used as verification. It is seen that the forecast skills of models with different dimensions are different and that of the model with 18 MEOFs (MK18) is the best. With either fewer or more MEOFs, the model's forecast skill gets worse. For typical statistical models skill always increases as the number of model parameters increase. That this does not happen here is because no observed fields are used directly in constructing the transition matrixes in the Markov models. Only the initial conditions (taken from the ocean model driven by the observed winds) are related with the observations indirectly. So there is little artificial skill in the Markov models. In order to confirm this point, another Markov model with 18 MEOFs (MK18') is constructed using only half of the data (all the 3 year runs from each of the monthly initial conditions from January 1970 to December 1980) and the observed NINO3 from January 1981 to December 1991 is used to test the model. When ZC, MK18 and MK18' are verified against the same observations, we find that their forecast skills are similar (not shown).

A way to understand the difference in the predictions of the models with different dimensions is to run them with identical initial conditions. Here the models with 18, 25 and 30 dimensions (MK18, MK25 and MK30) are initialized by the truncated initial conditions: $d_j = d_j(0)$, $j = 1, 2, ..., 18$ and $d_j = 0$, $j = 19, ..., 30$. It is seen that the forecast skills are quite similar (Fig. 5). Since the forecast by MK30 has been improved after its initial conditions are truncated at the 18th MEOF, the MEOFs higher than the 18th in the initial conditions do more harm than good. It is seen in Fig. 2 that the variance in the MEOFs higher than the 18th are all very small in both the initial and overall fields. It appears that the high MEOFs (>18) in the initial conditions are largely noise. No matter what the dimension of the models, the forecast skill has been largely determined by the initial conditions. When the high MEOFs (>18) are

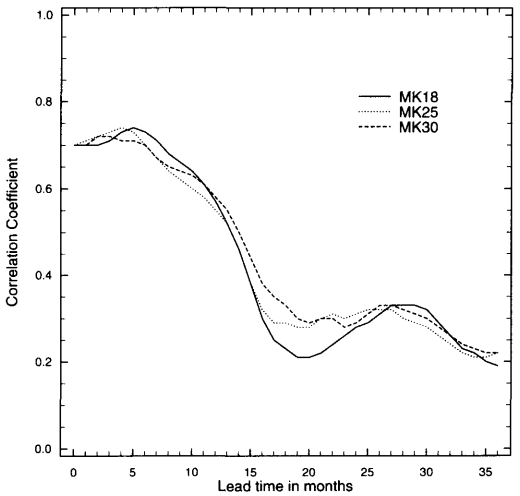

Fig. 5. Same as Fig. 4 except the solid line is for MK18, the dotted line for MK25 and the dashed line for MK30 (MKnn are Markov models with nn dimensions). The same initial conditions truncated at the 18th MEOF are used for all models.

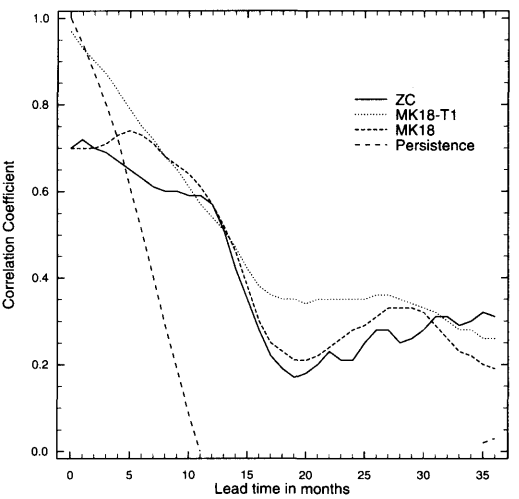

Fig. 6. Same as Fig. 4 except the solid line for ZC, short dashed line for MK18, long dashed line for persistence forecast and dotted line for MK18 with SST assimilation (MK18-T1, see text for details).

kept in the initial conditions, the main effect is to cause fast error growth.

Fig. 6 shows that MK18 has a forecast skill even better than ZC (though the difference can not be said to be statistically significant). Both ZC and MK18 overtake the persistence forecast at lead times beyond 5 months. Figs. 7a, b show that the monthly forecast skills by ZC and MK18 are quiet similar in overall structure. Both the plateau at lead time 13 months and the sharp drops of forecast skill in April in ZC are very well preserved in MK18. The smoother structure and higher correlation at the long lead times are due to the fact that MK18 is less noisy than ZC.

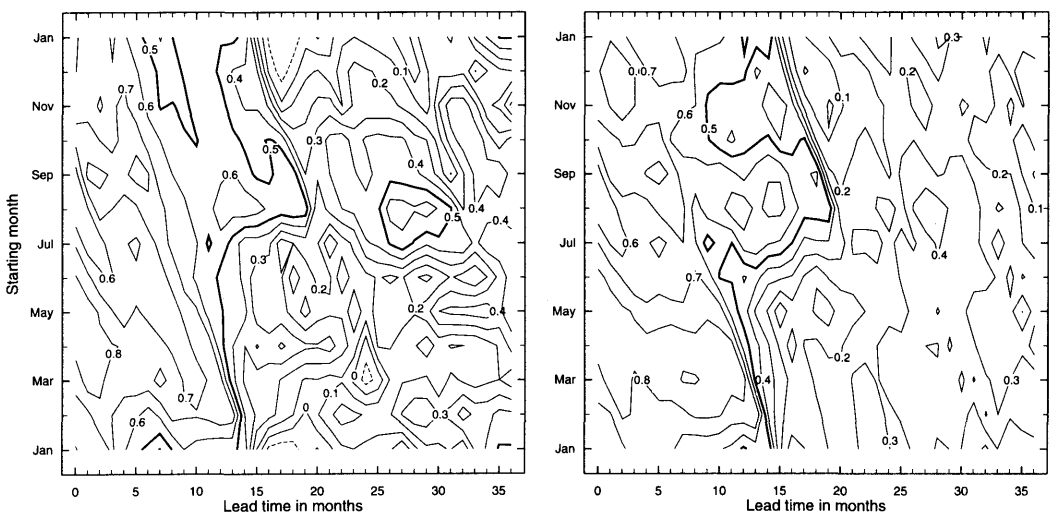

Fig. 7. Prediction-observation NINO3 correlation as functions of start months and lead months, by (a) ZC and (b) MK18. Each correlation is on a sample of 19 (all years from 1970 to 1988).

5. Initial error growth in MK18

Since MK18 simulates ZC well up to one year lead time (Fig. 3), presumably the initial error growth in MK18 would represent that in ZC. Blumenthal (1991) points out that the error evolution is a combination of 2 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 to the true dynamics. It is clear that both ZC and MK18 are only approximate representations of the ENSO system. However, the results given below indicate that the initial error growth in ZC due to the very poor initial conditions is the dominant source of error in short lead time forecasts. Only the initial error growth is studied here. Neglecting the added error at each time, the error vector in MK18 evolves as

$$\boldsymbol{r}_{i+1} = \boldsymbol{A}_i \boldsymbol{r}_i. \tag{13}$$

Hence the error covariance $\boldsymbol{R}_i = \langle \boldsymbol{r}_i \boldsymbol{r}_i^{\mathrm{T}} \rangle$ evolution is described by the formula

$$\boldsymbol{R}_{i+1} = \boldsymbol{A}_i \boldsymbol{R}_i \boldsymbol{A}_i^{\mathrm{T}}. \tag{14}$$

The error norm, defined by $\|\boldsymbol{r}\|^2 = \boldsymbol{r}^{\mathrm{T}}\boldsymbol{r}$, equals the trace of the error covariance. The residue ε^2 (cf., Blumenthal, 1991) is used to describe the initial error growth.

$$\varepsilon^2 = \frac{\|\boldsymbol{r}_\tau\|^2}{\|\boldsymbol{r}_0\|^2} = \frac{\mathrm{trace}(\boldsymbol{R}_\tau)}{\mathrm{trace}(\mathbf{R}_0)}. \tag{15}$$

Since we have little knowledge of the initial error structures, an uniform error ball in the MEOF space is assumed first. Physically, it means that the initial error is uniformly distributed among all MEOFs and uncorrelated. The residue growth shown in Fig. 8a has a similar seasonal dependence to that in Fig. 12 of Blumenthal (1991). For February starts, the error grows slowly to May, then grows rapidly in summer and fall and reaches a plateau in November; for May starts, the error grows very rapidly all the way to November and then grows slowly to a maximum in February. The equivalent e-folding times are 10 and 6 months respectively. The integration starting from August, on the other hand, shows a much slower error growth, with a peak value that does not occur until the following November. The equivalent e-folding time is 18 months. The error growth for November starts is similar to February starts with an e-folding time of 10 months. The e-folding times

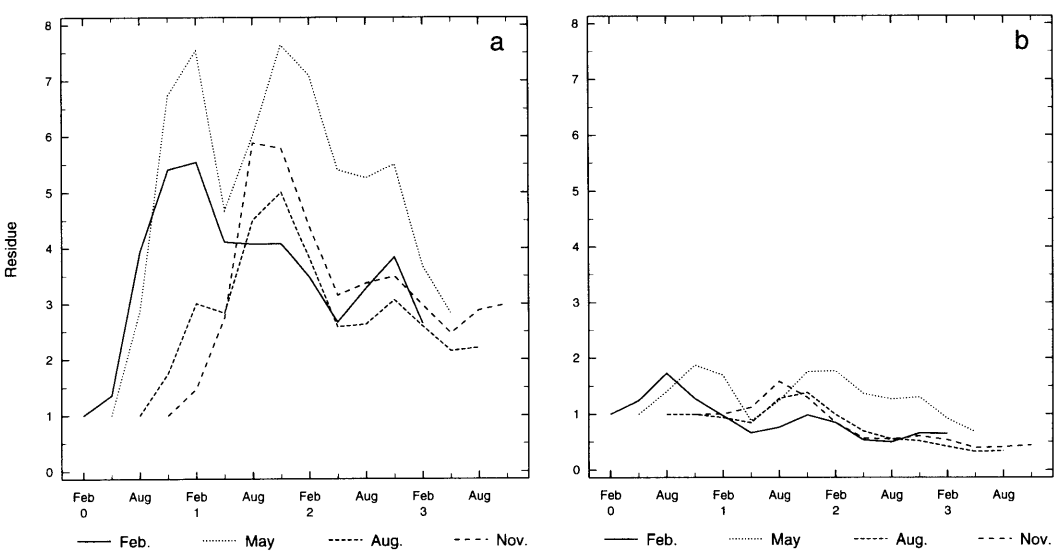

Fig. 8. Residue (initial error growth in MK18) versus verification months. Curves are noted by the starting months. (a) Uniform initial error ball; (b) initial error covariance proportional to signal covariance.

found by Blumenthal (1991) are 6 months for February and May starts and 12 months for August and November starts. In Blumenthal (1991) the Markov model is constructed from a long climatology run of ZC keeping only 5 EOFs for each of the SST, sea level and wind stress anomaly fields. As discussed above, the rate of error growth is sensitive to the number of MEOFs retained. However, the different initial error growth starting from different seasons is better resolved by our monthly model.

The error growth is sensitive to the initial error structures. When the initial error covariance is proportional to the signal covariance (diagonal, with the most variance distributed in the first MEOF), the error growth is much smaller (Fig. 8b). However, the seasonal variations have more or less the same pattern as before.

The error covariance evolution (14) also describes how the error ellipse evolves. Let the eigenvalues of $\boldsymbol{R}_\tau$ be λ_j, $j=1, ..., 18$. Then $\sqrt{\lambda_j}$ are the lengths of the axes of the 18-dimensional error ellipse. Fig. 9 shows that the evolution of the lengths of the longest five axes of the error ellipse starting from May with an uniform error ball. The error along one axis expands very rapidly and dominates the initial error growth. The longest axis of the error ellipse after nine months of integration is 11, which corresponds to a residue

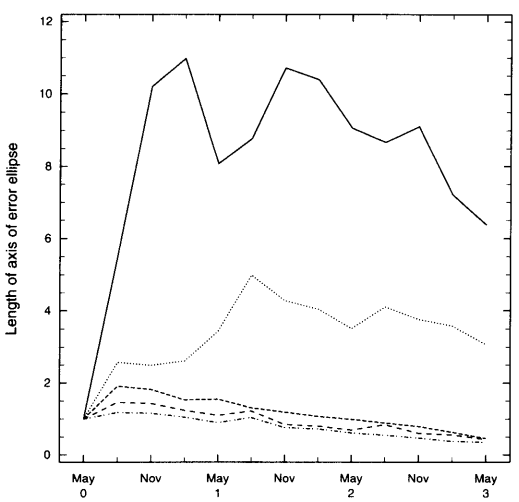

Fig. 9. Evolution of length of the longest five axes of the error ellipse in MK18 starting from May with an uniform error ball. Note that only 2 axes expand significantly.

$11^2/18 = 6.7$, almost the total residue value in Fig. 8a. It is interesting that only the errors along the two axes have significant growth. When $\boldsymbol{R}_0$ is the identity matrix, $\boldsymbol{R}_\tau = \boldsymbol{A}_\tau \boldsymbol{A}_\tau^{\mathrm{T}}$, where $\boldsymbol{A}_\tau(t)$ is the evolution operator at lead time τ starting from month t. So for an uniform initial error, $\sqrt{\lambda_j}$ are the singular values of $\boldsymbol{A}_\tau$. The largest singular value gives the optimal growth rate and the corresponding singular vector (SV) is the optimal perturbation (cf., Molteni and Palmer, 1993). The two growing singular vectors in Fig. 9 make a subspace of optimal perturbations, i.e., they maximize the error at lead time τ. The SVs of the 6 months evolution operators starting from other calendar month are calculated. We found that the structure of the fastest growing SV does not change much with start seasons and has mostly high MEOF (>2) components. Similarly only two singular vectors grow rapidly. The fastest growing SV of the 6 month evolution operator staring from May evolves into a mature phase of ENSO in November (Fig. 10). The starting state has been scaled to give a SST anomaly 3–4°C in the eastern Pacific in November. A very small perturbation in May can quickly develop into a mature phase of ENSO in 6 months. When the initial perturbation has this structure, energy can be most efficiently drawn from the basic state to support a quick development of ENSO. For a November start, the fastest growing SV of the 6-month evolution operator and its final state in May are similar to those in Fig. 10. It seems that for any start season the fastest growing SV evolves into a mature phase of ENSO in about 6 months.

Sensitivity to the initial error structure can be more clearly shown in the singular vector space. Decompose an initial error vector $\boldsymbol{r}_0$ into a set of singular vectors (SVs) $\boldsymbol{S}$, of $\boldsymbol{A}_\tau$. If $\boldsymbol{r}_0 = \boldsymbol{S}\boldsymbol{a}$, then $\langle \boldsymbol{a}\boldsymbol{a}^{\mathrm{T}} \rangle = \boldsymbol{S}^{\mathrm{T}} \boldsymbol{R}_0 \boldsymbol{S}$. The error at time τ is $\boldsymbol{A}_\tau \boldsymbol{S}\boldsymbol{a}$ and the residue is:

$$\varepsilon^2 = \frac{\|\boldsymbol{r}_\tau\|^2}{\|\boldsymbol{r}_0\|^2} = \frac{\boldsymbol{a}^{\mathrm{T}} \boldsymbol{S}^{\mathrm{T}} \boldsymbol{A}_\tau^{\mathrm{T}} \boldsymbol{A}_\tau \boldsymbol{S}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{T}} \boldsymbol{S}^{\mathrm{T}} \boldsymbol{S}\boldsymbol{a}}$$

$$= \frac{\boldsymbol{a}^{\mathrm{T}} \Lambda \boldsymbol{a}}{\boldsymbol{a}^{\mathrm{T}} \boldsymbol{a}}$$

$$= \frac{\lambda_1 a_1^2 + \lambda_2 a_2^2 + \cdots + \lambda_{18} a_{18}^2}{a_1^2 + a_2^2 + \cdots + a_{18}^2}. \qquad (16)$$

When $\boldsymbol{R}_0$ is the identity, $\langle \boldsymbol{a}\boldsymbol{a}^{\mathrm{T}} \rangle$ is the identity

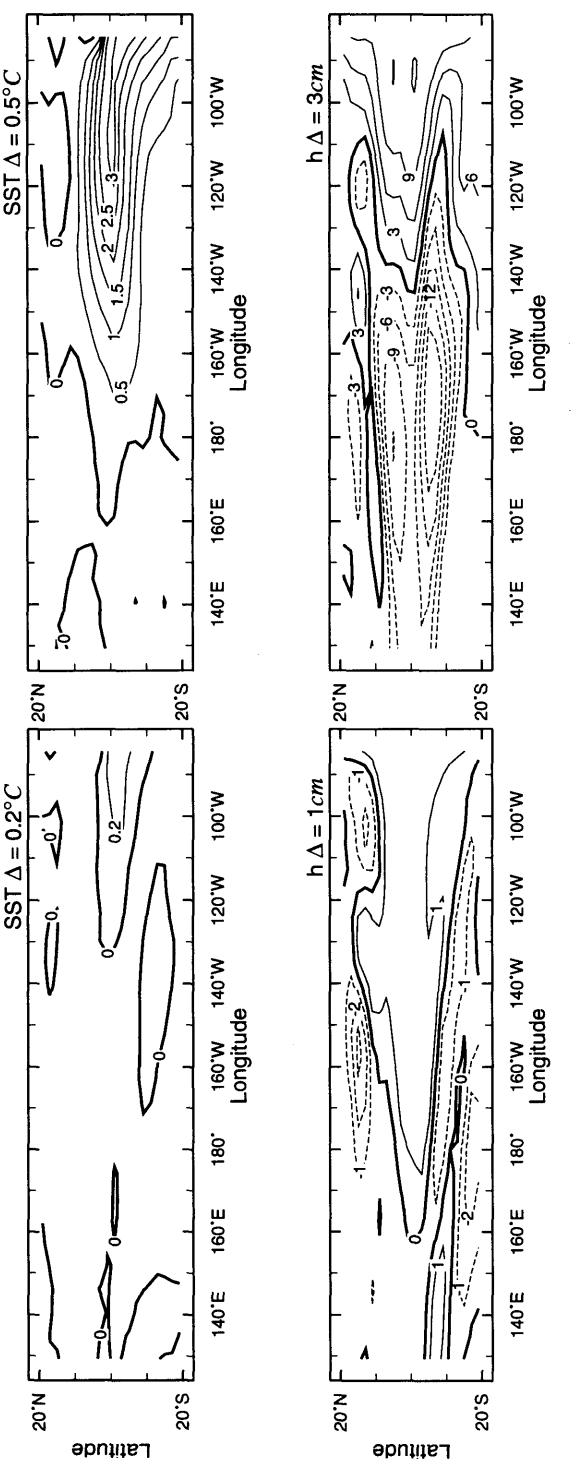

Fig. 10. Structure of the fastest growing singular vector of the 6 months evolution operator starting from May and structure after 6 months evolution. From top to bottom are SST, sea level and wind stress fields.

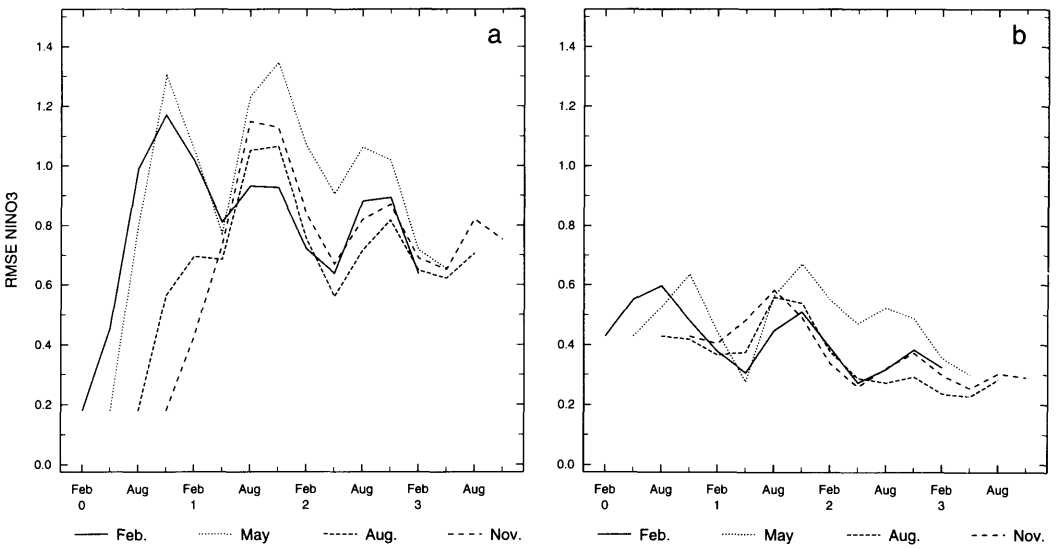

Fig. 11. Growth of root mean square error (RMSE) NINO3 starting from (a) an uniform initial error ball and (b) an initial error covariance in proportional to the signal covariance versus verification months. Curves are noted by the starting months and the initial error variance is 10% of the total signal variance.

too. The residue is reduced to $\varepsilon^2 = \sum_k \lambda_k/18$. Initially the error is equally distributed on all singular vectors, but soon after it is dominated by the fastest growing SV (Fig. 9). When $\boldsymbol{R}_0 = \alpha \boldsymbol{D}$, where $\boldsymbol{D}$ is the signal covariance and α is constant, the diagonal elements of $\langle \boldsymbol{a}\boldsymbol{a}^{\mathrm{T}} \rangle$ will not be equal. If the SVs of the 6 months evolution operator starting from May are ordered by their corresponding singular values, all the diagonal elements of $\langle \boldsymbol{a}\boldsymbol{a}^{\mathrm{T}} \rangle$ are very small except the 16th (a_{16}). Since only the first two SVs grow, the error growth (Fig. 8b) is much smaller than that for the uniform initial error (Fig. 8a).

Fig. 11 shows the growth of root mean square error (RMSE) of NINO3, the most widely used prediction index. Fig. 11a starts from an uniform initial error ball and Fig. 11b from an initial error covariance proportional to the signal covariance. In both cases the total initial error variance is chosen to be 10% of the total signal variance, but the different distributions of error covariance result in different initial RMSEs of NINO3. In Fig. 11a, the initial RMSE of NINO3 for May starts is only 0.18 but grows rapidly to the maximum value 1.3 in November; while in Fig. 11b, the initial RMSE of NINO3 for May starts is 0.42 but grows slowly to the maximum value 0.63 in November. The reason is that the fastest growing SV does not include the first two MEOFs, but these two contribute the most to NINO3.

Clearly, the growth of the RMSE of NINO3 has no single relationship with its starting value but depends on the initial error structure. If the error growth rates are measured by e-folding time as in Goswami and Shukla (1991), they are 3.6, 2, 5 and 4 months for February, May, August and November starts from an uniform initial error. They are much faster than those measured by residue, which measures the sum of error growth in all fields. It was found that the residue is dominated by the error growth in the SST and wind stress fields and the error growth in the sea level field is much slower. So those two indices give consistent results. The e-folding time of error growth estimated by Goswami and Shukla is 6.7 months. Since the e-folding time should be longer when the initial field is perturbed with non-white noise, the e-folding time in Goswami and Shukla is in reasonable agreement with our results.

6. Seasonal dependency of predictability

The ZC prediction skill (Fig. 7a) shows a sharp decline in the boreal spring (hereafter just "spring"). This is true of MK18 (Fig. 7b), in

Tellus A/46/4-13

common with other prediction schemes (cf., Latif et al., 1993; Barnston and Ropelewski, 1992). It has been noticed that nature has a similar characteristic, as seen from the rapid drop in the auto correlation of observed NINO3 in spring (Fig. 12). This feature implies that the ENSO signal before the spring season does not correlate well with that after the spring season. It seems that both the models and reality are experiencing a similar process, which changes the ENSO cycle every year in spring.

The mean square error (MSE) between two variables f_1 and f_2 is $\mathrm{MSE} = \langle (f_1 - f_2)^2 \rangle = 2\sigma^2 - 2\langle f_1 \cdot f_2 \rangle$, where σ, the standard deviation, is assumed the same for both. Then the correlation of f_1 and f_2 is

$$\rho = \frac{\langle f_1 f_2 \rangle}{\sigma^2} = 1 - \frac{\mathrm{MSE}}{2\sigma^2}. \tag{17}$$

Since the observed variance of NINO3 is seasonally dependent, smallest in spring and largest in winter, it contributes to the seasonality of predictability as measured by ρ. If the MSE between the model and observed NINO3 is roughly the same for each month, the correlation will be low in spring and high in winter because of the seasonally varying NINO3 variance.

However, the MSE between the ZC forecast and observed NINO3 is seasonally varying. For all start months it grows from April and reaches a maximum in December and then decays dramatically in late winter, going to a minimum in March. Thus the seasonal cycle of MSE is almost in phase with the seasonal cycle of $2\sigma^2$. Fig. 13 illustrates for a January start (in Fig. 13: the MSE was calculated after the raw ZC output was rescaled to have the same variance as the observations in order to comply with the assumptions of (17)). When the MSE and $2\sigma^2$ are comparable, there is no forecast skill. Fig. 13 shows this condition is approached in April, although the MSE is small. We propose that the rapid drop of correlation in spring is a consequence of the low variance at that time and thus is characteristic of the ENSO cycle. As long as model error exists, a rapid decline of *correlation* in spring is to be expected, even when, as with ZC, the MSE is lowest then.

Knowledge of the patterns of initial error growth is helpful in understanding the seasonal predictability of ZC. For late summer starts the initial error actually decays, so predictions are able to survive through the next spring barrier (Fig. 7a) and gain some skill in the following winter when the NINO3 variance is again high. For spring starts, the initial error grows rapidly through the

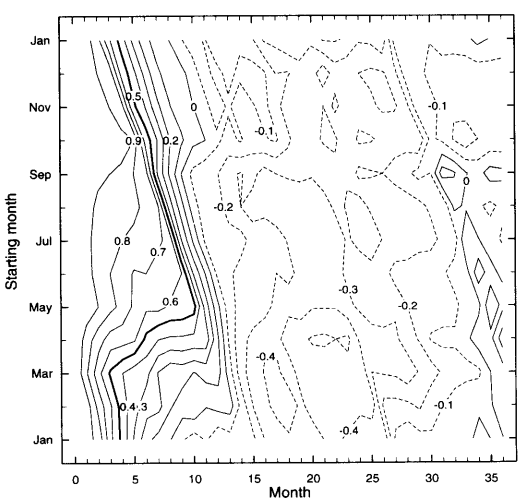

Fig. 12. Autocorrelation of the monthly observed NINO3 from January 1970 to December 1991 as functions of starting months and lead months. Notice the rapid correlation drops around April.

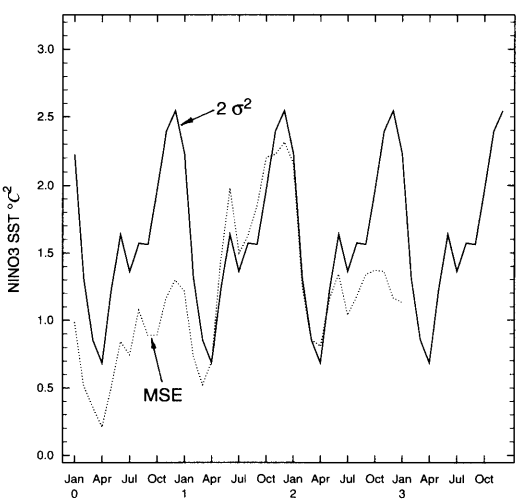

Fig. 13. Mean square error (MSE) (dotted) between the scaled ZC NINO3 and observed NINO3 for January start and the observed NINO3 variance multiplied by 2 (solid) versus verification months.

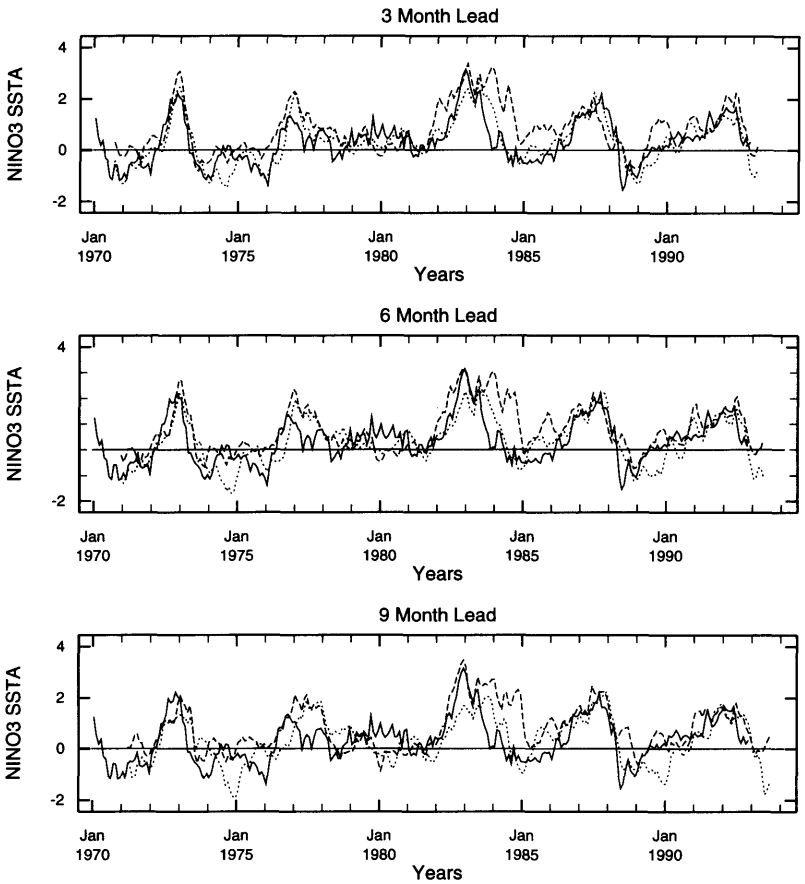

Fig. 14. The observed (solid), ZC (dashed) and MK18-T1 (dotted) composite NINO3 by the forecasts initiated in 6 consecutive months in the period January 1970 to November 1992 at lead months 3, 6 and 9 months. MK18-T1 referred to the 18 MEOFs model with SST assimilation in the initial conditions.

summer and the NINO3 variance increases all the way to December. If only the initial error growth mattered, the predictions would degrade rapidly. However the increase of the NINO3 variance keeps the prediction skill useful until the next spring barrier (Fig. 7a).

7. Assimilation of SST into initial conditions

One way to increase the predictability is to reduce the initial error. It has been mentioned that the initial conditions have significant errors in the low MEOFs, evident in the low correlation between the model initial NINO3 and observed NINO3. These systematic errors can be partially corrected by assimilating the observed SST into the initial conditions. The observed sea surface temperature anomalies in the Pacific basin in the period January 1970 to December 1991 from the Climate Analysis Center (Reynolds, 1988) are decomposed by the EOFs $\boldsymbol{e}_j$ from before as in (1). The first EOF accounts for 41% of the total variance; the next 19 EOFs account for 27% of the variance. Even if only one EOF is kept, the represented NINO3 correlates with the total observed NINO3 at 0.97. Vector $\boldsymbol{b}$ is constructed by replacing the model initial SST PCs $a_j^1(0)$ by the observed SST PCs $a_j^{(\mathrm{obs})}(0)$ and leaving the sea level and wind components $a_j^2(0)$ and $a_j^3(0)$ as before in (2). Then the new initial state vector

$\boldsymbol{d}^{(\mathrm{obs})}$ in the MEOF space is calculated by decomposing $\boldsymbol{b}$ as in (4). By trying several truncations on $a_j^{(\mathrm{obs})}(0)$, we found that when only the first EOF is retained in $a_j^{(\mathrm{obs})}(0)$ and the rest are set to zero (referred to as MK18-T1 hereafter) the forecast had the best skill. The monthly averaged prediction-observation correlation versus lead time is shown in Fig. 6, together with ZC, MK18 and persistence forecasts. It is seen that MK18-T1 is superior to persistence after about 2 months lead and superior to ZC up to 10 months lead. The skill is a little lower than ZC for lead times 11–13 months and higher again for lead times longer than 13 months. The seasonal dependence of forecast skill is similar to that in Figs. 7a, b, except that there is much higher correlation at short lead times in MK18-T1 than that in ZC and MK18.

Since MK18-T1 has better forecast skill than ZC at short lead times, it is useful to show the retrospective forecasts for the period January 1970 to November 1992. To make the comparison with the ZC forecasts easier, the composites by the forecasts initiated in 6 consecutive months (see Cane et al. (1986) for details) in MK18-T1 are calculated and shown with the observed and ZC NINO3 in Fig. 14. The 3- and 6-month lead forecasts are better than the ZC's but the forecasts with lead times beyond 9 months are not obviously superior to ZC. The recent 91–92 El Niño is well predicted for lead time up to 12 months.

An obvious question to ask is whether the 18 MEOFs model is still the best one when the observed SST is assimilated. As we have done before (Fig. 4), the Markov models with 5, 16, 18, 20 and 30 MEOFs initialized by the initial conditions with SST assimilation were compared (not shown). We found that the model with 18 MEOFs is still the best one and the sequence in terms of forecast skill is the same as in Fig. 4.

The SST assimilation in the initial conditions only helps forecasts at short lead times ($\leqslant 6$ months), while the forecasts at long lead times remain the same. We propose that the forecast skill at long lead times is largely determined by the initial set up of the oceanic heat content. So improvement in the initial oceanic heat content is expected to help forecasts at longer lead times. By improving the ocean model, using better winds to force it, or incorporating the ocean heat content information directly, a better forecast skill than MK18-T1 is potentially possible.

8. Summary and conclusions

Based on the assumption that the ENSO is a low order, mainly linear and low frequency system, a seasonally varying first order auto regressive model (Markov model) is constructed from an ensemble of forecast data generated by ZC. The data is made up of a suite of 3-year coupled model forecasts starting from each of the monthly initial conditions for the period January 1970 to December 1991. These initial conditions are generated by forcing the ocean component alone with the observed wind stress (FSU) beginning with January 1964 and then forcing the atmosphere component alone with the hindcast SSTs. The SST anomaly, sea level anomaly (h) and surface wind stress anomaly (τ) fields which characterize the model state space are combined into a reduced state space by finding multivariate EOFs (MEOFs). Eventually 30 MEOFs are kept.

Various Markov models utilizing between 5 and 30 MEOFs all simulate ZC quite well for one year. When the observed NINO3 is the reference, the model with 18 MEOFs (MK18) has the best forecast skill, comparable with that of ZC. It is demonstrated that the Markov model has little artificial skill; this is to be expected, since the model is fit to ZC not observational data. The truncation of the initial conditions is responsible for the increased forecast skill of the 18 MEOF model compared to those with more MEOFs, which suggests that the MEOFs higher than the 18th in the initial conditions are mostly noise. When the observed SST is assimilated into the initial conditions, the forecast skill is improved at short lead times ($\leqslant 6$ months).

Analysis of MK18 shows that its eigenmodes (POPs) are non-self-adjoint, which could lead to a fast transient initial error growth as in Farrell (1989) and Blumenthal (1991). It is found that the initial error growth is fastest starting from spring and slowest starting from late summer and is sensitive to the initial error structures. Two singular vectors (SVs) of the linear evolution operator have significant transient growth dominating the total error growth. Since the fastest growing SV has mostly high MEOF (>2) components, the error growth tends to be larger when there are more high components in the initial error fields. Removing them improves prediction skill. Since the singular vectors control the error growth, in future work we

plan to use knowledge of them to filter the initial conditions. We were not able to use the POPs to filter the initial conditions because we did not find a good way to differentiate them for truncation. The fastest growing SV has components on all POPs and its fast growth lies in the collective action of all the POPs. Blumenthal (1991) has vividly demonstrated it in a 2-dimension system. The relationship between the singular vectors and POPs in MK18 is more difficult to demonstrate and understand. More analyses are necessary.

The seasonal variation of the initial error growth due to the non-self-adjoint property, fastest starting from spring and slowest starting from late summer, is consistent with the seasonally varying stability of the coupled ocean-atmosphere system as established by linear stability analysis (Battisti and Hirst, 1989). It is found that the coupled system is the most unstable in summer and least unstable in winter. However the singular vectors in a non-self-adjoint system are different from the normal modes in linear stability analysis. The normal modes have fixed spatial structures and fixed growth rate while the structures and the growth rate of singular vectors change with evolution time. The optimal growth rate of singular vectors can be much faster than the fastest growing normal mode. For an example, the optimal growth of amplitude of the normal modes (POPs) is 1.2 from May to November in MK18, while the optimal growth of singular vectors is 10.2. This thansient growth is the greatest concern in short term ENSO prediction.

The seasonality of predictability is characteristic of ENSO. The observed standard deviation of NINO3 is variable with season, smallest in (northern) spring and biggest in winter. Since the NINO3 variance is the smallest in spring, the correlation is especially sensitive to the change of mean square error (MSE) in this season. This sensitivity causes the correlation lines to squeeze together, indicating a fast decline in forecast skills. This phenomenon is expected to be common with all ENSO forecast models, since it depends on a property of nature, the low variance in spring. Webster and Yang (1992) attribute this special feature in spring to the influence of the summer monsoon circulation. Their argument has 2 factors: one is that an increase in external noise causes the spring barrier and another one is that the strong monsoon circulation could modulate the weak Walker circulation at spring to make the system especially sensitive to external noise. The 1st could be checked form data, but it is not straightforward to do so. The 2nd is partly addressed by us since the climatological monsoon is included in the ZC background state. We see that spring is not an especially favorable time; summer is. We can not now address just how the anomalous monsoon may modulate the background state. Our point is that, even assuming the system noise is unaffected by the monsoon and noise levels are the same for spring as for other seasons, the spring barrier would nonetheless occur due to the smallness of NINO3 variance.

The seasonal variations analyzed here, including the spring barrier, are characteristic of ENSO and so effect all model predictions. Predictability will vary from model to model because of different non-self-adjointness and different initializations. The rapid error growth through the summer season imposes an upper limit on predictability, but more sophisticated models and better initialization procedures may well achieve better forecast skills than those obtained to date.

9. Acknowledgments

We would like to thank Gerd Bürger who read our draft carefully and had interesting discussions with the authors. This work was supported by grant no. NOOO14-90J-1595 from the Office of Naval Research and grant no. NA16-RC-0432 from NOAA.

REFERENCES

Barnett, T., Graham, N., Cane, M. A., Zebiak, S. E., Dolan, S., O'Brien, J. J. and Legler, D. 1988. One the prediction of the El Niño of 1986–1987. *Science* **241**, 192–196.

Barnett, T. P., Latif, M., Graham, N., Flügel, M., Pazon, S. and White, W. 1993. ENSO and ENSO related predictability, part I, Prediction of equatorial Pacific sea surface temperature with a hybrid coupled ocean-atmosphere model. *J. Climate* **6**, 1545–1566.

Barnston, A. G. and Ropelewski, C. F. 1992. Prediction

of ENSO episodes using Canonical Correlation Analysis. *J. Climate* **5**, 1316–1345.

Battisti, D. S. and Hirst, A. C. 1989. Interannual variability in a tropical atmosphere/ocean model: Influence of the basic state, ocean geometry and nonlinearity. *J. Atmos. Sci.* **46**, 1687–1712.

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

Blumenthal, M. B., Xue, Y. and Cane, M. A. 1991. Predictability of an ocean/atmosphere model using adjoint model analysis. *Proceedings of a workshop held at ECMWF on New developments in predictability*, (November), pp. 277–304.

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

Cane, M. A. 1991. *Forecasting El Niño with a geophysical model. Teleconnections linking worldwide climate anomalies*. M. Glantz, R. W. Katz, and N. Nicholls (eds.). Cambridge University Press.

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

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

Goswami, B. N. and Shukla, J. 1991. Predictability of a coupled ocean-atmosphere model. *J. Climate* **4**, 3–22.

Graham, N. E., Michaelsen, J. and Barnett, T. P. 1987a. An investigation of the El Niño-southern Oscillation cycle with statistical models (1). Predictor field characteristics. *J. Geophys. Res.* **92**, C13, 14251–14270.

Graham, N. E., Michaelsen, J. and Barnett, T. P. 1987b. An investigation of the El Niño-southern Oscillation cycle with statistical models (2). Model results. *J. Geophys. Res.* **92**, C13, 14271–14289.

Graham, N. E., Barnett, T. P. and Latif, M. 1992. Considerations of the predictability of ENSO with a low order coupled model. *TOGA Notes*, 11–15 April.

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., Barnett, T. P., Cane, M. A., Flügel, M., Graham, N. E., Von Storch, H., Xu, J.-S. and Zebiak, S. E. 1994. A review of ENSO prediction studies. *Climate Dynamics* **9**, 167–179.

Mureau, R., Molteni, F. and Palmer, T. N. 1993. Ensemble prediction using dynamically-conditioned perturbations. *Q. J. R. Meteor. Soc.* **119**, 299–324.

Molteni, F. and Palmer, T. N. 1993. Predictability and finite-time instability of the northern winter circulation. *Q. J. R. Meteor. Soc.* **119**, 269–298.

Penland, C. and Magorian, T. 1993. Prediction of NINO3 sea surface temperatures using linear inverse modeling. *J. Climate* **6**, 1067–1076.

Reynolds, R. W. 1988. A real-time global sea surface temperature analysis. *J. Climate* **1**, 75–86.

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.

Webster, P. J. and Yang, S. 1992. Monsoon and ENSO: selectively interactive systems. *Q. J. R. Meteor. Soc.* **118**, 877–926.

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