

ENSO Prediction with Markov Models: The Impact of Sea Level

YAN XUE,* ANTS LEETMAA, AND MING JI

NOAA/NWS/National Centers for Environmental Prediction, Washington, D.C.

(Manuscript received 10 November 1998, in final form 8 April 1999)

ABSTRACT

A series of seasonally varying linear Markov models are constructed in a reduced multivariate empirical orthogonal function (MEOF) space of observed sea surface temperature, surface wind stress, and sea level analysis. The Markov models are trained in the 1980–95 period and are verified in the 1964–79 period. It is found that the Markov models that include seasonality fit to the data better in the training period and have a substantially higher skill in the independent period than the models without seasonality. The authors conclude that seasonality is an important component of ENSO and should be included in Markov models. This conclusion is consistent with that of statistical models that take seasonality into account using different methods.

The impact of each variable on the prediction skill of Markov models is investigated by varying the weightings among the three variables in the MEOF space. For the training period the Markov models that include sea level information fit the data better than the models without sea level information. For the independent 1964–79 period, the Markov models that include sea level information have a much higher skill than the Markov models without sea level information. The authors conclude that sea level contains the most essential information for ENSO since it contains the filtered response of the ocean to noisy wind forcing.

The prediction skill of the Markov model with three MEOFs is competitive for both the training and independent periods. This Markov model successfully predicted the 1997/98 El Niño and the 1998/99 La Niña.

1. Introduction

The El Niño–Southern Oscillation (ENSO) phenomenon has been extensively studied in the past two decades. Significant progress in both understanding ENSO and achieving useful prediction skill for the sea surface temperature (SST) anomaly in the tropical Pacific has been achieved during the Tropical Oceans Global Atmosphere decade (1985–94; see review papers by McCreary and Anderson 1991; Latif et al. 1994). Both statistical and dynamical models are now used to make experimental ENSO forecasts in real time (see the “Experimental Long Lead Forecast Bulletin” issued quarterly by the Center for Ocean–Land–Atmosphere Studies). The forecast SSTs have been used as boundary forcings for atmospheric general circulation models in producing seasonal climate outlooks (Barnett et al. 1994; Ji et al. 1994). Accurate prediction of the SST in the tropical Pacific several seasons in advance is becoming increasingly important because significant so-

cio-economic benefits can be achieved from improved prediction of large-scale temperature and precipitation on seasonal timescales.

The prediction skill of ENSO is often measured by temporal correlation of large-scale indices, such as area-averaged SST anomalies. Barnston et al. (1994) concluded that the 6-month lead prediction skill of ENSO during 1982–93 was modest (0.6 correlation) and that the skill of statistical and dynamical models was comparable. Recently, Barnston et al. (1998) compared the performance of statistical and dynamical models during the 1997/98 El Niño episode and the 1998 La Niña onset. They concluded that although many models forecasted some degree of warming prior to the onset of the El Niño in boreal spring 1997, none predicted its strength until the event was already quite strong in late spring. Neither the dynamical nor the statistical models, as groups, performed significantly better than the other during this episode. This study suggests that both statistical and dynamical models have significant errors and more work is needed to improve the models. Since statistical models are much easier to construct than dynamical models are, they will continue to serve as useful references with which more sophisticated dynamical models can be compared.

Statistical models are often based on observed surface winds, sea surface temperature, and near-global sea level pressure fields (Graham et al. 1987a,b; Barnston and

* Additional affiliation: RDC Research Scientist, RDC, Greenbelt, Maryland.

Corresponding author address: Dr. Yan Xue, Room 807 WWB/ NCEP/NOAA, 5200 Auth Road, Camp Springs, MD 20746.
E-mail: yxue@ncep.noaa.gov

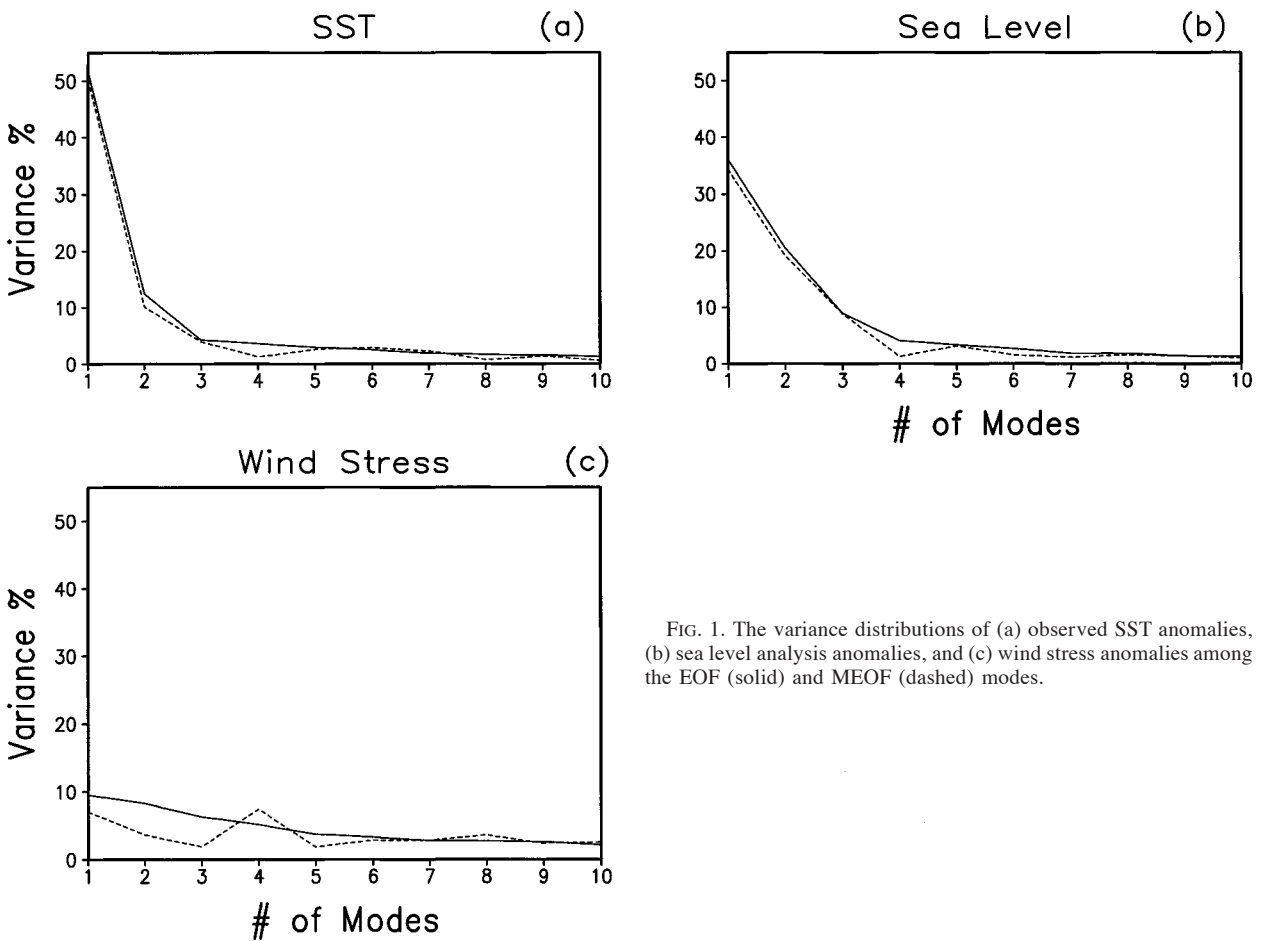


FIG. 1. The variance distributions of (a) observed SST anomalies, (b) sea level analysis anomalies, and (c) wind stress anomalies among the EOF (solid) and MEOF (dashed) modes.

Ropelewski 1992; Penland and Magorian 1993). Since oceanic heat content is an important component of ENSO (Wyrtki 1985; Zebiak and Cane 1987; Zebiak 1989), subsurface ocean data are also valuable for statistical models. However, until recently there was no ocean analysis in the tropical Pacific that covered a sufficiently long period for this purpose. The ocean data assimilation system at the National Centers for Environmental Prediction (NCEP) began to produce an ocean analysis in 1992 (Ji et al. 1995). Recently this analysis system has been improved significantly (Behringer et al. 1998) and a retrospective analysis that covered about 18 yr has been produced.

We choose to use the sea level field from the NCEP ocean analysis in addition to observed SST and surface winds to construct Markov models. Sea level is chosen because the tide gauge network [the Integrated Global Ocean Services System sea level program in the Pacific available online at <http://uhslc.soest.hawaii.edu>] provides an independent dataset for validation of the sea level analysis. The accuracy of the sea level analysis at NCEP is about 3–4 cm in the equatorial belt (Behringer et al. 1998). These three variables were also used by Blumenthal (1991) and Xue et al. (1994) to construct

Markov models that were best fit to the outputs from the Zebiak–Cane (ZC) model (Zebiak and Cane 1987). Xue et al. (1994) found that the Markov model constructed with these three variables fits the ZC model quite well up to a year, and the prediction skill of the Markov model was comparable to that by the ZC model.

The Markov model in Xue et al. (1994) was constructed in a multivariate EOF (MEOF) space in which SST, surface winds, and sea level were equally weighted. An interesting question is how the prediction skill of Markov models varies with the weightings among the three variables. When one variable is heavily weighted and other variables are much less weighted, the impact of that variable on the prediction skill of Markov models can be investigated. Smith et al. (1995) showed that using the subsurface temperature data from an earlier version of the NCEP ocean analysis, in addition to using sea level pressure and SST data, improved the ENSO forecast skill of the canonical correlation analysis model (Barnston and Ropelewski 1992). Johnson et al. (2000) used observed SST anomalies and oceanic heat content anomalies of the upper ocean to construct Markov models. They concluded that the skill of the Markov models was improved by including oceanic heat content, but

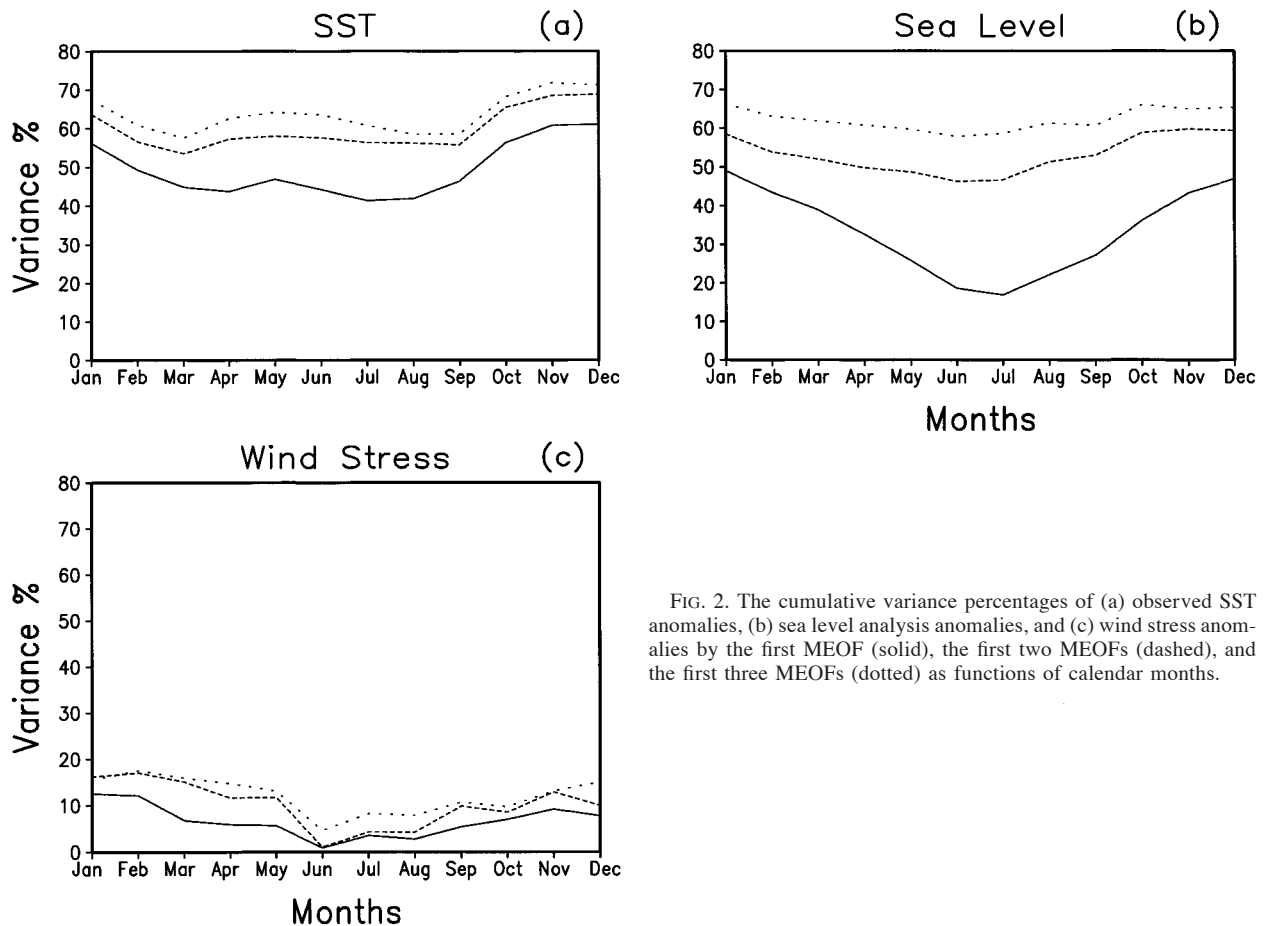


FIG. 2. The cumulative variance percentages of (a) observed SST anomalies, (b) sea level analysis anomalies, and (c) wind stress anomalies by the first MEOF (solid), the first two MEOFs (dashed), and the first three MEOFs (dotted) as functions of calendar months.

the improvement was not significant with high certainty. They suggested that the usefulness of including subsurface temperature data in Markov models for ENSO prediction was underestimated due to the limitation of their data. In this paper we will use the latest sea level analysis data to study the impact of sea level on the prediction skill of Markov models. We will attempt to demonstrate that the prediction skill of Markov models can be improved significantly by including sea level information.

Since El Niño is phase locked with the annual cycle (Philander 1983; Zebiak and Cane 1987), the statistics of the anomalies fields associated with ENSO have a strong seasonal dependency. Because of this seasonal nonstationarity, Hasselmann and Barnett (1981) suggested that the normal statistical analysis techniques based on time-invariant models are inappropriate for ENSO prediction. In the problem of predicting El Niño off South America they developed phase-averaged models and concluded that El Niño is predicted with considerably more confidence and accuracy using phase-averaged models than with time-invariant models. In a similar approach to that of Hasselmann and Barnett (1981), Blumenthal (1991) developed seasonally vary-

ing Markov models to best fit the outputs from the Zebiak–Cane model (Zebiak and Cane 1987). More recently, Kim (2000) developed a unique prediction scheme based on cyclostationary EOFs (Kim and North 1997). Since cyclostationary EOFs are specifically designed for describing periodic statistics, they are presumably better basis functions than standard EOFs based on which statistical models can be constructed. However, cyclostationary EOFs are much more expensive to compute than standard EOFs are. We choose to use the method of Blumenthal (1991). We will argue that MEOFs are useful basis functions based on which seasonally varying Markov models are constructed to account for the seasonal nonstationarity of ENSO.

The paper is organized into five sections. Section 2 describes the space reduction by MEOFs. Section 3 describes the construction of Markov models and the importance of including seasonality in Markov models. In section 4 we discuss how the prediction skill of Markov models varies with the weightings among SST, wind stress, and sea level. The prediction skill of the Markov models is estimated using a cross-validation scheme and an independent period. The Markov model's predictions

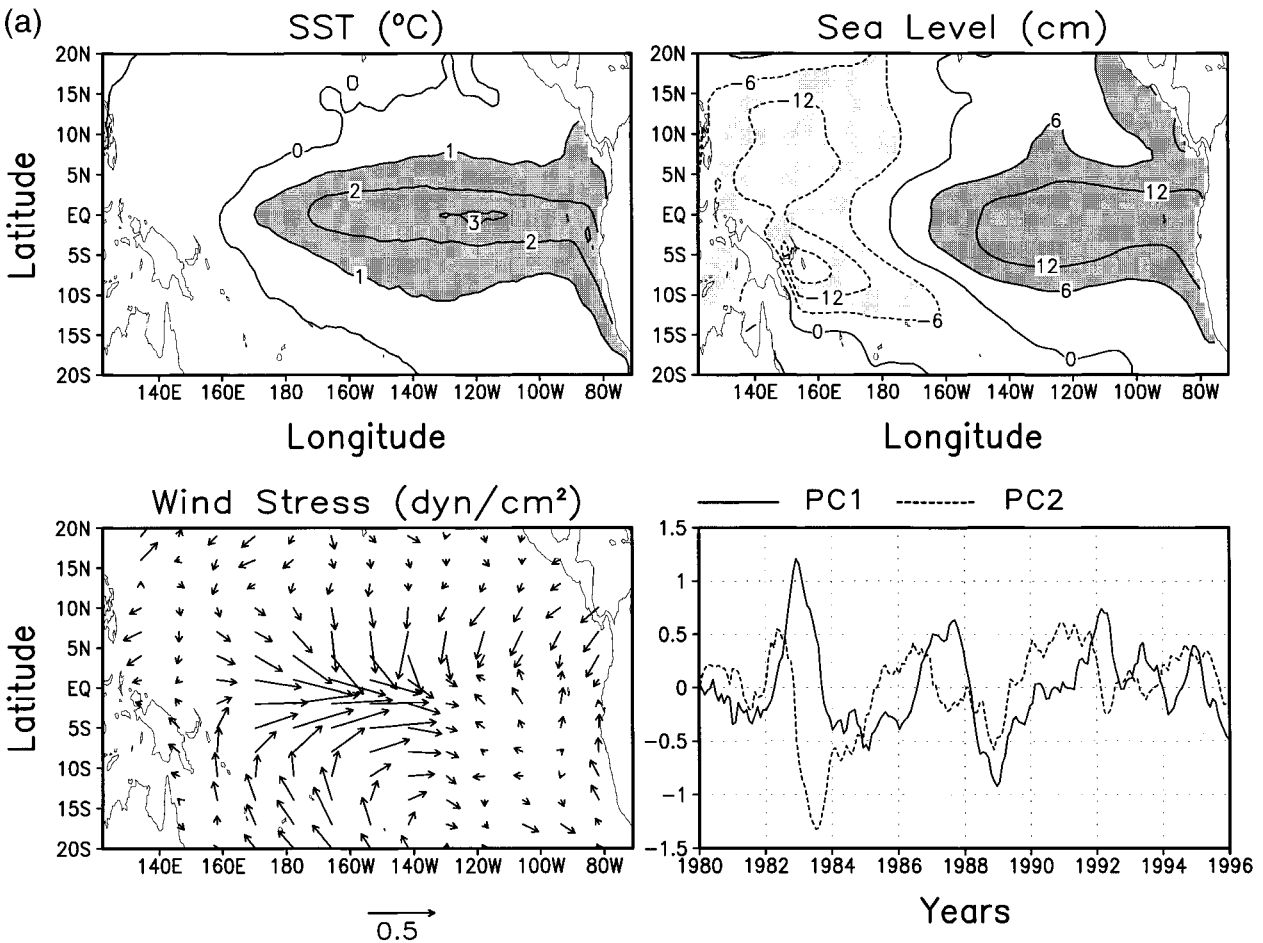


FIG. 3. The (a) first, (b) second, and (c) third MEOF patterns and PCs. Note that the PC of the second MEOF (dash) is shown in (a) as well.

for the 1997/98 warm event and the 1998/99 cold event are also discussed. Section 5 summarizes the paper.

2. Space reduction

a. Data

The SST data consist of two parts: from January 1964 to December 1981 it is the reconstruction of historical SST by Smith et al. (1996); from January 1982 to January 1999 it is the SST analysis by Reynolds and Smith (1994). The wind stress data are The Florida State University (FSU) pseudo-wind stress product from January 1964 to January 1999 (Goldenberg and O'Brien 1981). Two sea level datasets are used. One is the sea level dataset from the ocean analysis at NCEP, which covers the period from January 1980 to January 1999 (Behringer et al. 1998). The second sea level dataset is obtained from a ocean model simulation that uses the Geophysical Fluid Dynamics Laboratory MOM1 model forced by the FSU wind stress analysis. This sea level dataset covers the period from January 1964 to December 1995. The model configuration for this simulation

is similar to that of Ji and Smith (1995) except that the total FSU wind stress (converted from the FSU pseudo-wind stress with a drag coefficient of 1.3×10^{-3}) are used instead of a combination of the Hellerman and Rosenstein (1983) climatology and the FSU wind stress anomalies. All the datasets are monthly values and have been interpolated onto a common grid $1^\circ \text{ lat} \times 1.5^\circ \text{ long}$ with approximately 4600 grid points covering the tropical Pacific region 20°S – 20°N .

The Markov models are trained in the 1980–95 period using the anomalous fields of observed SST, wind stress, and sea level analysis obtained by removing the annual cycles for 1980–95. The prediction skill of the Markov models is estimated using the independent 1964–79 period during which the anomalous fields of observed SST, wind stress, and model simulated sea level data are used as initial conditions. These anomalous fields are obtained by removing the annual cycles for 1980–95.

b. Space reduction

As in Xue et al. (1994), two-step EOF analyses are used to calculate MEOF modes based on which Markov

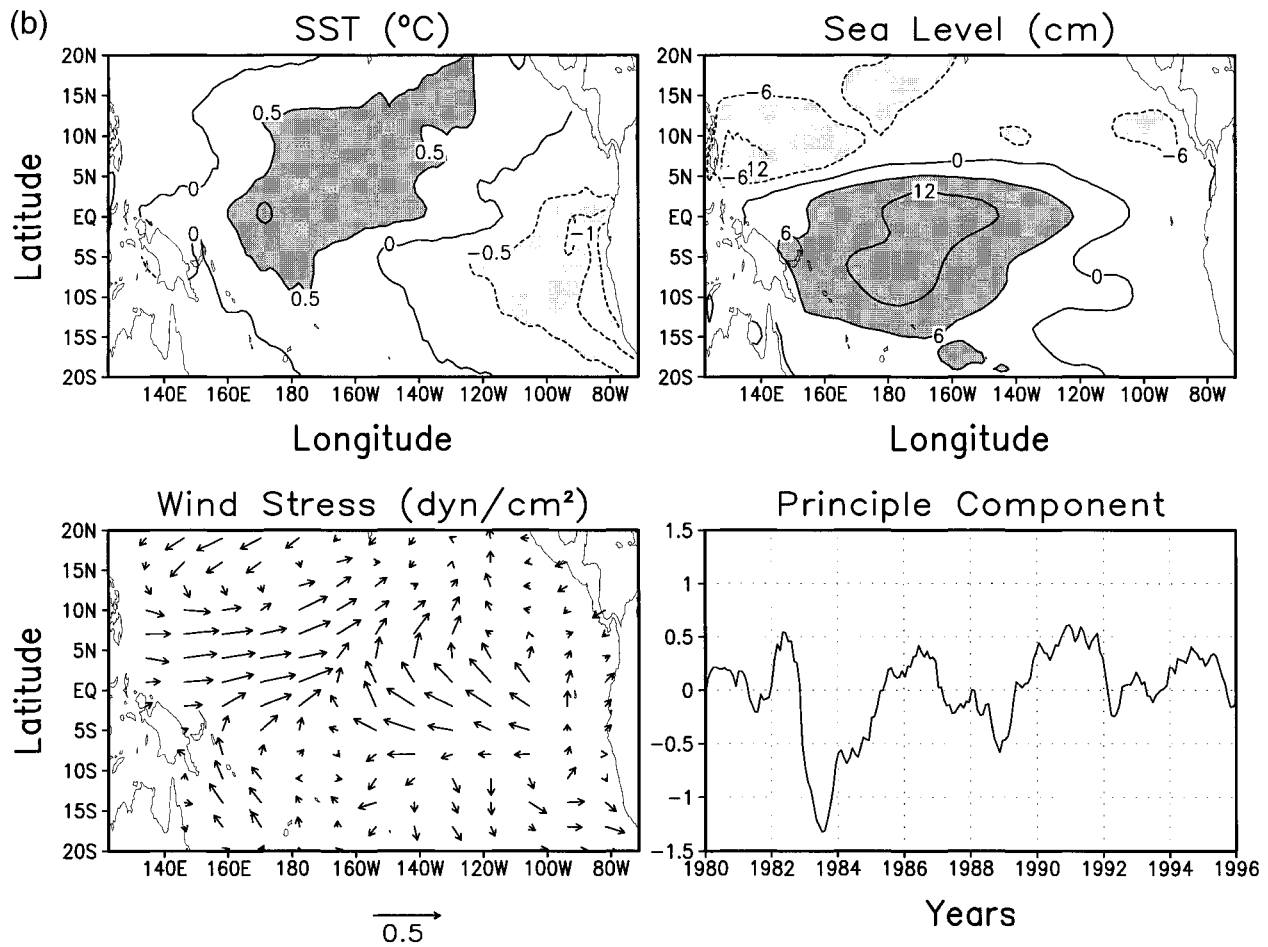


FIG. 3. (Continued)

models are constructed. In the first step, the SST, wind stress, and sea level anomaly fields are subject to EOF analysis separately. Each field is represented 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 \mathbf{e} and principal components (PCs) $a_j(t)$ and filtered by truncating at the J th EOF:

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

In the present context, 91, 154, and 85 EOFs account for 99% of the variance in the observed SST, wind stress, and sea level anomaly fields, respectively. The variance distributions of the EOFs of SST, wind stress, and sea level are shown in Fig. 1. It is seen that the first two EOFs of SST account for 52% and 13% of the total variance and each of the higher-order EOFs account for less than 5% (Fig. 1a); the first three EOFs of sea level account for 36%, 21%, and 9% of the total variance, and each of the higher-order EOFs account for less than 5% (Fig. 1b). It is interesting that the variance percentage of the second EOF of SST is much smaller than that of the first EOF, while the variance percentages of

the first two EOFs of sea level are comparable. The variance distribution of wind stress is relatively flat among the EOFs (Fig. 1c), indicating that the signal and noise in the wind stress field are not well separated.

In order to reduce the space further and to have a consistent multivariate basis, a second EOF analysis is conducted. A vector \mathbf{b} is constructed from the PCs of SST, wind stress, and sea level:

$$\mathbf{b} = \left(\epsilon_1 \frac{a_1^1}{\sigma_1}, \dots, \epsilon_1 \frac{a_{J_1}^1}{\sigma_1}, \epsilon_2 \frac{a_1^2}{\sigma_2}, \dots, \epsilon_2 \frac{a_{J_2}^2}{\sigma_2}, \epsilon_3 \frac{a_1^3}{\sigma_3}, \dots, \epsilon_3 \frac{a_{J_3}^3}{\sigma_3} \right). \quad (2)$$

Here, a_j^1 , a_j^2 , and a_j^3 are the PCs of SST, wind stress, and sea level with dimensions $J_1 = 91$, $J_2 = 154$, and $J_3 = 85$; ϵ_1 , ϵ_2 , ϵ_3 are the weights assigned to SST, wind stress, and sea level, respectively; σ_1^2 , σ_2^2 , and σ_3^2 are the total variance described by each set of PCs:

$$\sigma_\mu^2 = \sum_{j=1}^{J_\mu} \sum_t [a_j^\mu(t)]^2, \quad \mu = 1, 2, 3. \quad (3)$$

The vector \mathbf{b} is then decomposed into EOFs \mathbf{f}_k and PCs $d_k(t)$ and filtered by truncating at the K th EOF,

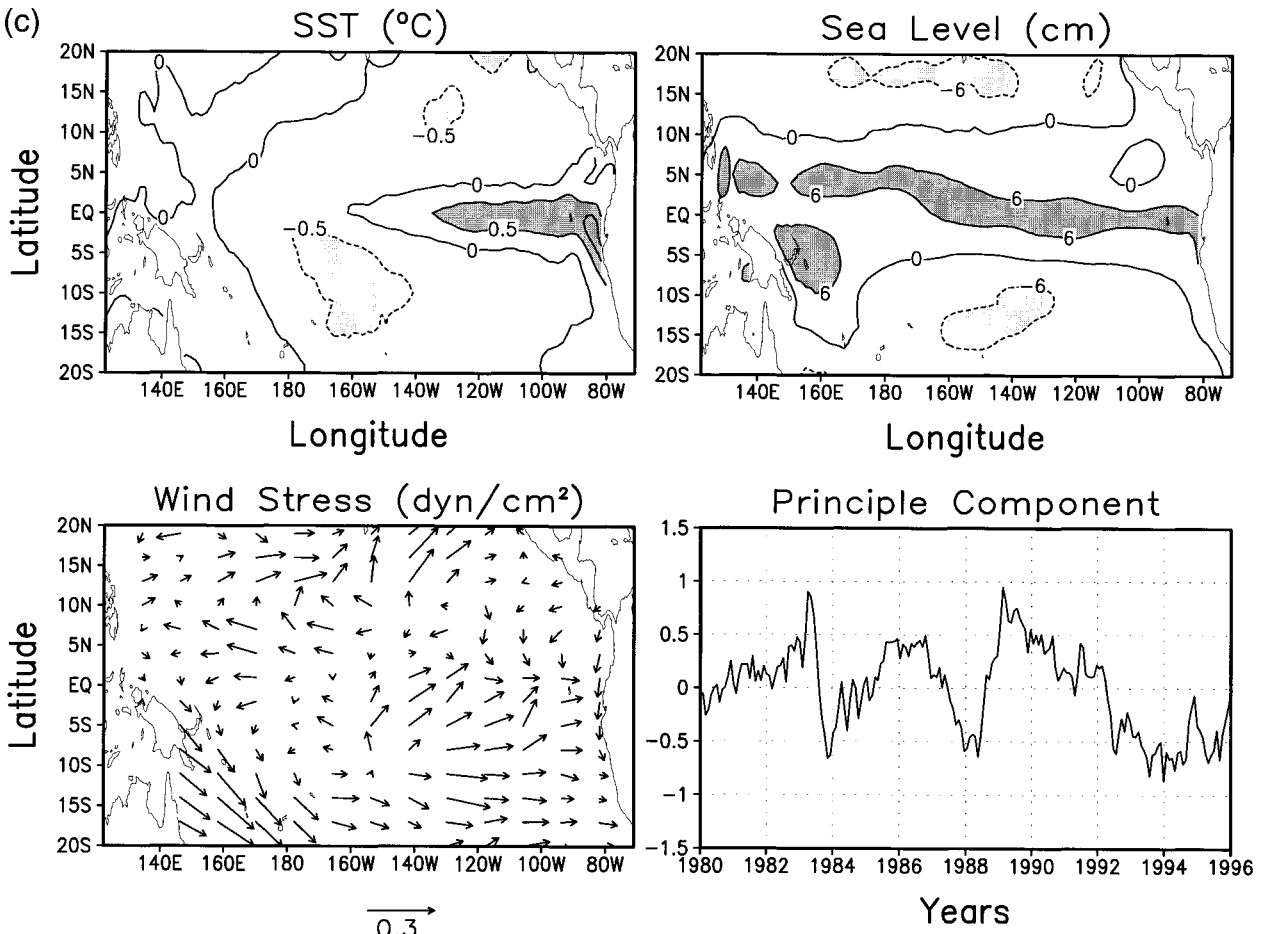


FIG. 3. (Continued)

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

The PCs $d_j(t)$ represent the three physical fields in the phase space spanned by the MEOFs and will be used to construct Markov models in next section. The corresponding spatial pattern \mathbf{g}_j , containing three physical fields, can be derived from the EOF patterns of each variable \mathbf{e}_j , the MEOF functions \mathbf{f}_j , and the weights ϵ_1 , ϵ_2 , ϵ_3 . So the combined fields of SST, wind stress, and sea level are decomposed into spatial patterns \mathbf{g}_j and PCs $d_j(t)$.

We now examine the variance distribution among the MEOFs where SST, wind stress, and sea level are equally weighted, for example, $\epsilon_1 = 1$, $\epsilon_2 = 1$, $\epsilon_3 = 1$. Figure 1 shows that the variance distributions of SST and sea level among the MEOFs are very similar to those among the EOFs, while the variance distribution of wind stress among the EOFs and MEOFs are somewhat different. This result indicates that the dominant EOFs of SST vary coherently with those of sea level, but the dominant EOFs of wind stress do not. This is understood since SST and sea level contain mainly low-frequency vari-

abilities associated with ENSO, while wind stress contains a significant amount of high-frequency variabilities not associated with ENSO. Figure 1 also shows that the variance distributions of SST and sea level become flat when the order of MEOFs is greater than 3. This suggests that the appropriate number of MEOFs for the construction of Markov models is 3.

The MEOF modes calculated for the whole time series do not necessarily maximize the variance representation for the time series at a specific calendar month. Since the PC time series at a specific calendar month are not orthogonal to each other, the variances among the MEOFs are not separable. So we used cumulative variance percentages, defined as $\gamma = 1 - \|Z - Z'\|^2 / \|Z\|^2$, to describe how well each physical field is represented by the first few MEOFs. Here Z represents the original physical field, Z' is the field represented by the first few MEOFs, and $\|\cdot\|^2$ is the sum of the variance. Figure 2 shows that the cumulative variance percentages of SST, wind stress, and sea level all vary with seasons, and the seasonal variations are most prominent in sea level. For sea level, the first MEOF accounts for 49% of the variance in winter, but only 17% in summer, for

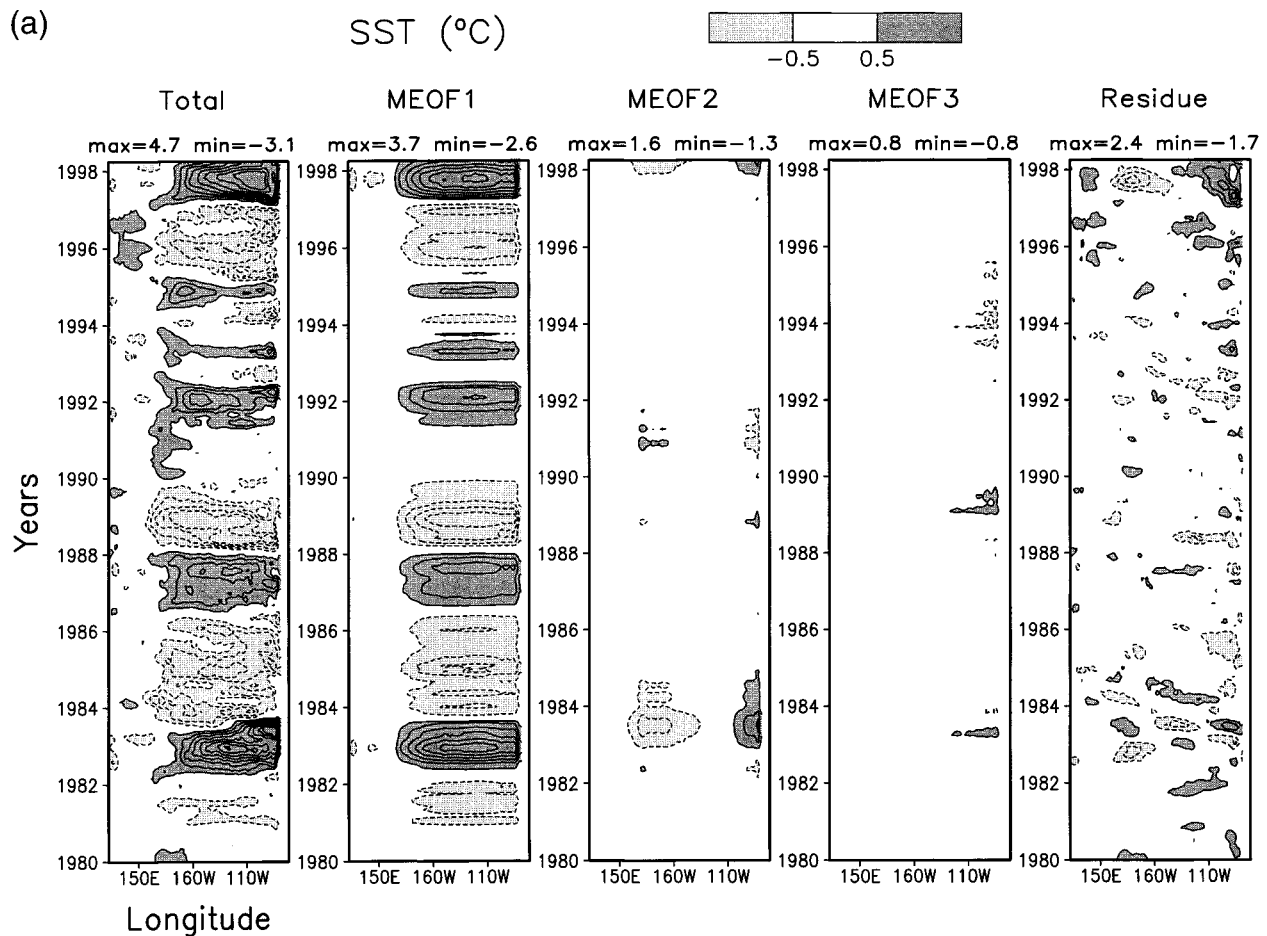


FIG. 4. Hovmoeller plots of (a) SST anomalies, (b) zonal wind stress anomalies, and (c) sea level anomalies averaged in the equatorial belt 2°S – 2°N . For each variable the total field; the first, second, and third MEOF; and the residual field are shown in the first, second, third, fourth, and fifth columns, respectively. In (a) the contour interval is 0.5; in (b) the contour interval is 0.1 except it is 0.05 for the third MEOF; in (c) the contour interval is 3.

example, a change of 65% (Fig. 2b). However, when the second MEOF is included, the seasonal variation is much less. This suggests that the second MEOF contributes to the variance more in summer than in winter, which is opposite to what the first MEOF does. Including the third MEOF increases the cumulative variance percentage about 10% evenly for all the seasons. The three MEOFs together account for about 63% of the variance of sea level relatively evenly through the year.

For SST the cumulative variance percentage of the first MEOF varies about 33% in a year, which is much less than that of sea level (Fig. 2a). The three MEOFs together explain about 60% of the variance of SST relatively evenly through the year. In contrast, the three MEOFs do not represent wind stress well, especially in early summer (Fig. 2c). If we assume that the unrepresented variability of wind stress is not critical for ENSO development, the three MEOFs span a useful reduced space in which Markov models can be built.

c. Spatial and time variability of MEOFs

Figure 3 shows the spatial patterns and the normalized PCs of the first three MEOFs discussed above. The first MEOF pattern resembles a mature phase of ENSO where the SST in the eastern and central Pacific is abnormally warm (Fig. 3a). Associated with the positive SST anomalies are convergent winds on the equator where the wind stress anomaly maximum is located to the west of the SST anomaly maximum. In addition, the sea level is anomalously high in the eastern Pacific and low in the western Pacific. The PC shows maxima for all the warm and cold events.

The second MEOF describes an onset phase of ENSO because the PC of the second MEOF leads that of the first MEOF by 9–12 months (Figs. 3a,b). Here the sea level has a maximum on the equator east of the date line. Associated with the sea level maximum are positive SST anomalies near the date line. It is noted that the negative SST anomalies along the west coast of South

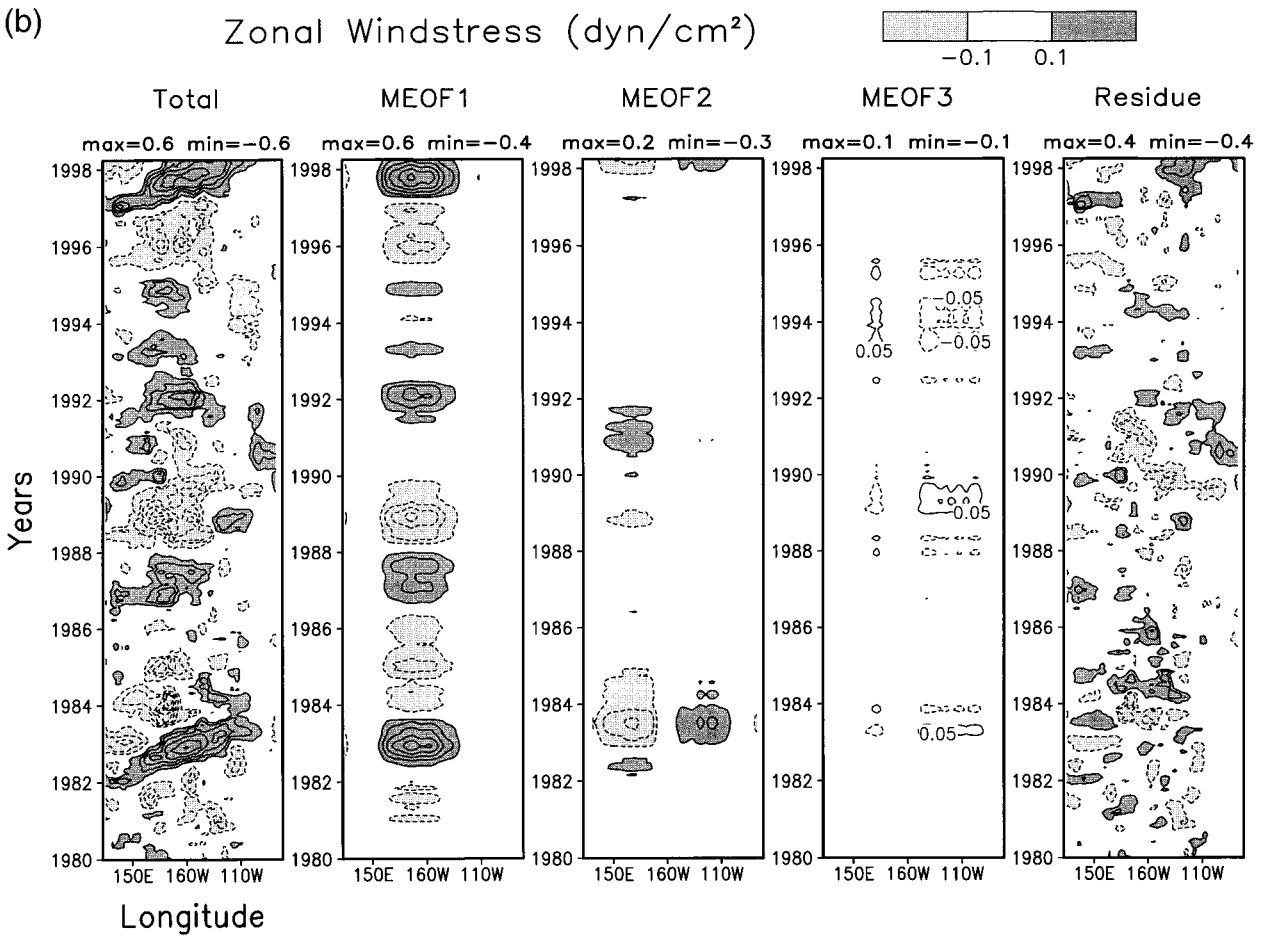


FIG. 4. (Continued)

America are mainly due to the 1982/83 event. The wind stress anomalies on the equator are convergent toward the positive SST anomalies; off the equator an anomalous cyclonic flow in the northwest Pacific is associated with the SST anomalies in the subtropics. These cyclonic winds contribute to the sea level minimum in the northwest Pacific.

The third MEOF pattern is characterized by a narrow belt of positive SST anomalies on the equator in the eastern Pacific surrounded by negative SST anomalies off the equator (Fig. 3c). In this case the wind stress is generally weak on the equator, while the sea level is anomalously high in the whole equatorial belt. The PC of the third MEOF indicates that this MEOF accounts for the changes of structures from event to event.

The contribution of each MEOF to the variabilities of SST, wind stress, and sea level on the equator are shown in Fig. 4. For comparison, the total and residual fields (defined as the differences between the total fields and the fields spanned by the three MEOFs) are also shown. For better displaying the signals, a 3-month running mean has been applied to the total and residual fields of SST and sea level, and a 5-month running mean

has been applied to the total and residual fields of wind stress. It is seen in Fig. 4a that the first MEOF accounts for most of the variability of SST on the equator and the second MEOF is usually weak except during the 1982/83 and 1997/98 events. The third MEOF has some weak variability in the far eastern Pacific. The amplitude of the residual field is typically 0.5°C except it is much larger during the 1982/83 and 1997/98 events.

The variations of the first two MEOFs in zonal wind stress are important while the variations of the third MEOF is negligible (Fig. 4b). The amplitude of the residual wind stress is as large as 0.4 dyn cm⁻². Since this is the 5-month running mean of the residual field, the amplitude of the actual residual field can be even larger.

Blanke et al. (1997) derived a residual wind stress similar to that discussed above and suggested that such residual wind stress anomalies can have a significant impact on the predictability of ENSO. Similar results are also obtained by Eckert and Latif (1997) and Moore and Kleeman (1999). When a Markov model is constructed in the MEOF space spanned by the three MEOFs, the potential impact of the residual wind stress

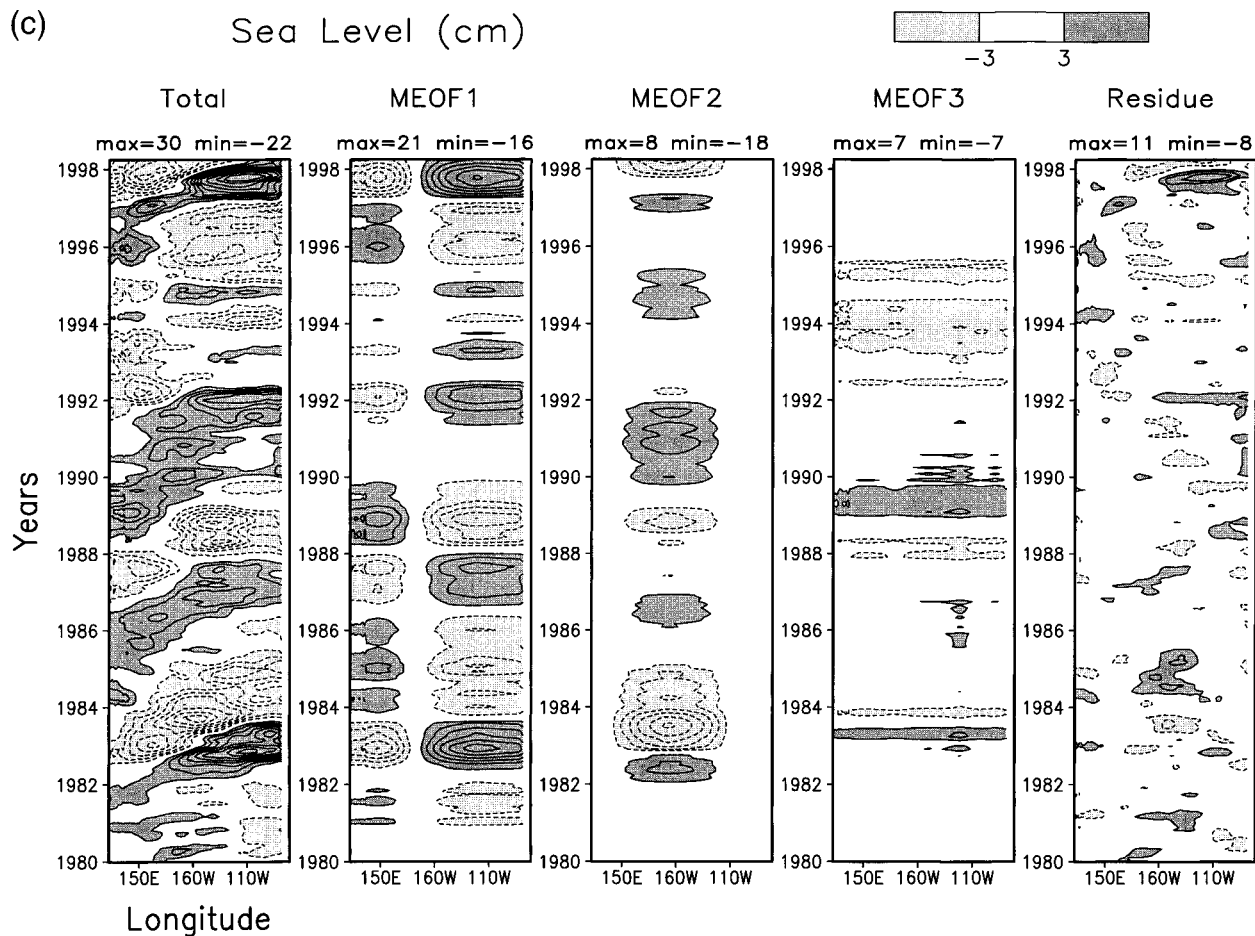


FIG. 4. (Continued)

on predictability of ENSO is ignored. This is also true in most of dynamical coupled models in which atmospheric components do not simulate such high-frequency winds well.

Figure 4c shows that all the three MEOFs contribute to the equatorial sea level variabilities significantly. The first MEOF describes an east–west oscillation of sea level, and the second MEOF describes a buildup of sea level in the central equatorial Pacific. The third MEOF describes an uniform sea level change across the equatorial belt. The amplitude of the residual sea level is typically 3 cm, except it is much larger during the 1997/98 event.

3. Markov models and seasonality

a. Markov models

Two kinds of Markov models are constructed, referred as nonseasonal and seasonal Markov models. The deviations of the Markov models are provided in appendix. The nonseasonal Markov model contains one monthly transition matrix, which is a first-order linear regression based on the monthly PC time series. The

nonseasonal Markov model is very similar to the linear inverse model by Penland and Magorian (1993). Xue et al. (1994) constructed seasonal Markov models that contain 12 monthly transition matrixes for month to month evolution. Since the sample size for training each of the 12 monthly transition matrixes is one-twelfth of that for training the monthly transition matrix for the nonseasonal Markov model, the statistical significance of the seasonal Markov model is lower than that of the nonseasonal Markov model. In order to increase the statistical significance, we sacrificed some of the annual cycle resolution by developing phase-averaged Markov models (see the appendix for details). This approach is similar to that of Hasselmann and Barnett (1981), who used three harmonics to approximate the annual cycle.

b. Seasonality

The importance of seasonality in the ENSO system has been discussed extensively using simple and intermediate couple models (Philander 1983; Zebiak and Cane 1987; Battisti 1988). It is also considered important in statistical modeling for ENSO (Hasselmann and

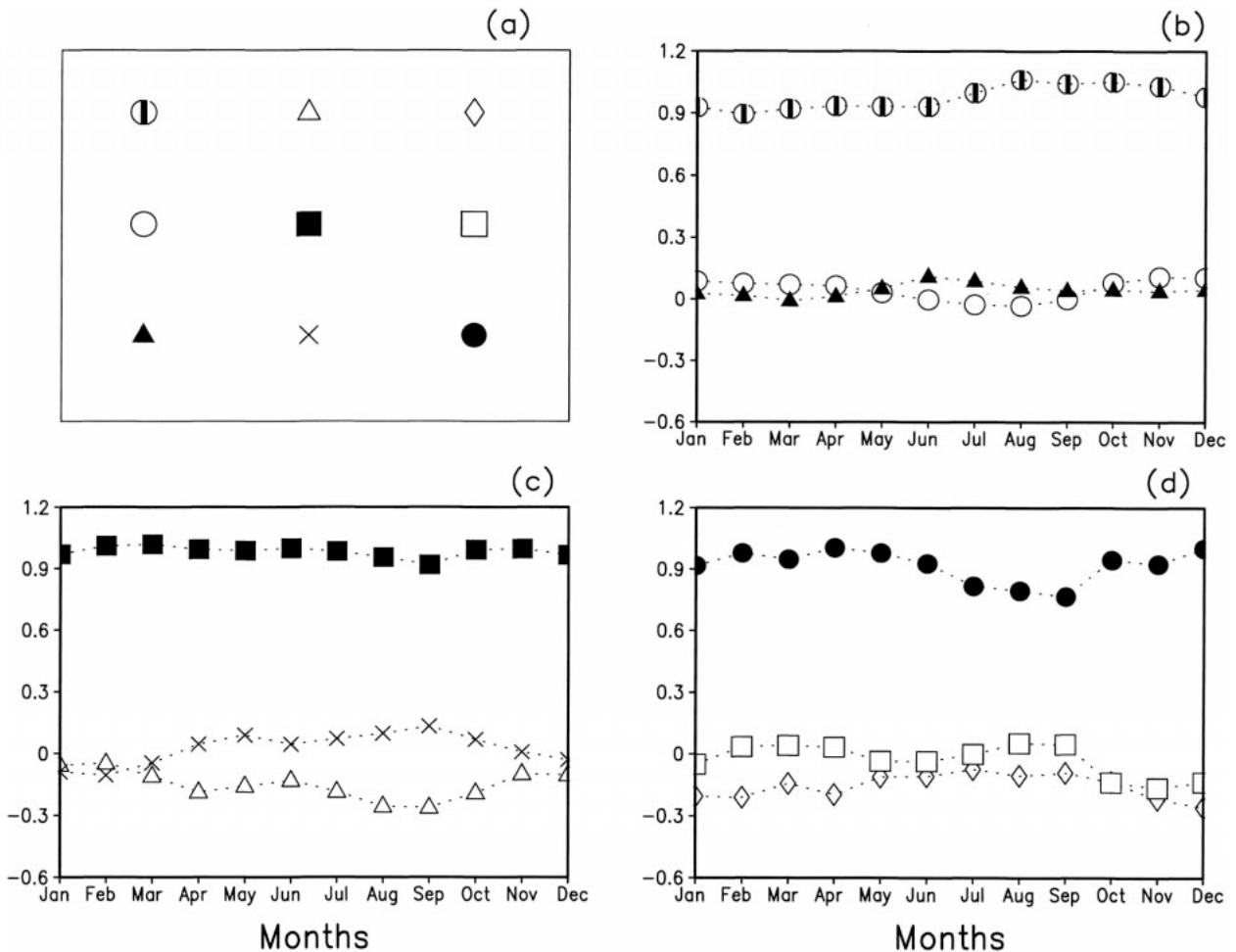


FIG. 5. Seasonality in the seasonal Markov model with 3 retained MEOFs. (a) A schematic plot of the elements of the monthly transition matrices. The elements in the (b) first, (c) second, and (d) third columns of the monthly transition matrices are shown as functions of calendar months.

Barnett 1981; Barnston and Ropelewski 1992). However, there is a debate on whether seasonality should be included in Markov models. The linear inverse model by Penland and Magorian (1993) did not include seasonality. Blumenthal (1991) and Xue et al. (1994), however, included seasonality in their Markov models, which are best fits to the ZC model. Recently, Johnson (2000) revisited some of the analyses in Penland and Magorian (1993), and concluded that seasonality is an important component of the deterministic dynamics of ENSO. In a subsequent study, Johnson et al. (2000) found that including seasonality in Markov models did not significantly improve prediction skill because there is a trade-off between the improvement of prediction by including seasonality and the reduction in significance of the model due to the reduction of sample size.

The seasonal variations of the model parameters of the seasonal Markov models are shown in Fig. 5. It is seen that the monthly transition matrices of the seasonal Markov models are largely diagonal. This suggests that

persistence is dominant in monthly timescales. The seasonal variations of the diagonal elements are small, while the seasonal variations of the off-diagonal elements are substantial. In the 6-month transition matrices, the off-diagonal elements are as large as the diagonal elements, indicating energy exchanges between the MEOF modes (Fig. 6). Most of the matrix elements have substantial seasonal variations. These results indicate that seasonality is generally small in monthly timescales but is large in 6-month timescales. So including seasonality in Markov models should improve prediction skill more at long lead times than at short lead times. Next we will compare the skill of the nonseasonal and seasonal Markov models.

c. Skill comparison

A series of nonseasonal and seasonal Markov models are constructed with different number of retained MEOFs. The skill of the Markov models is measured

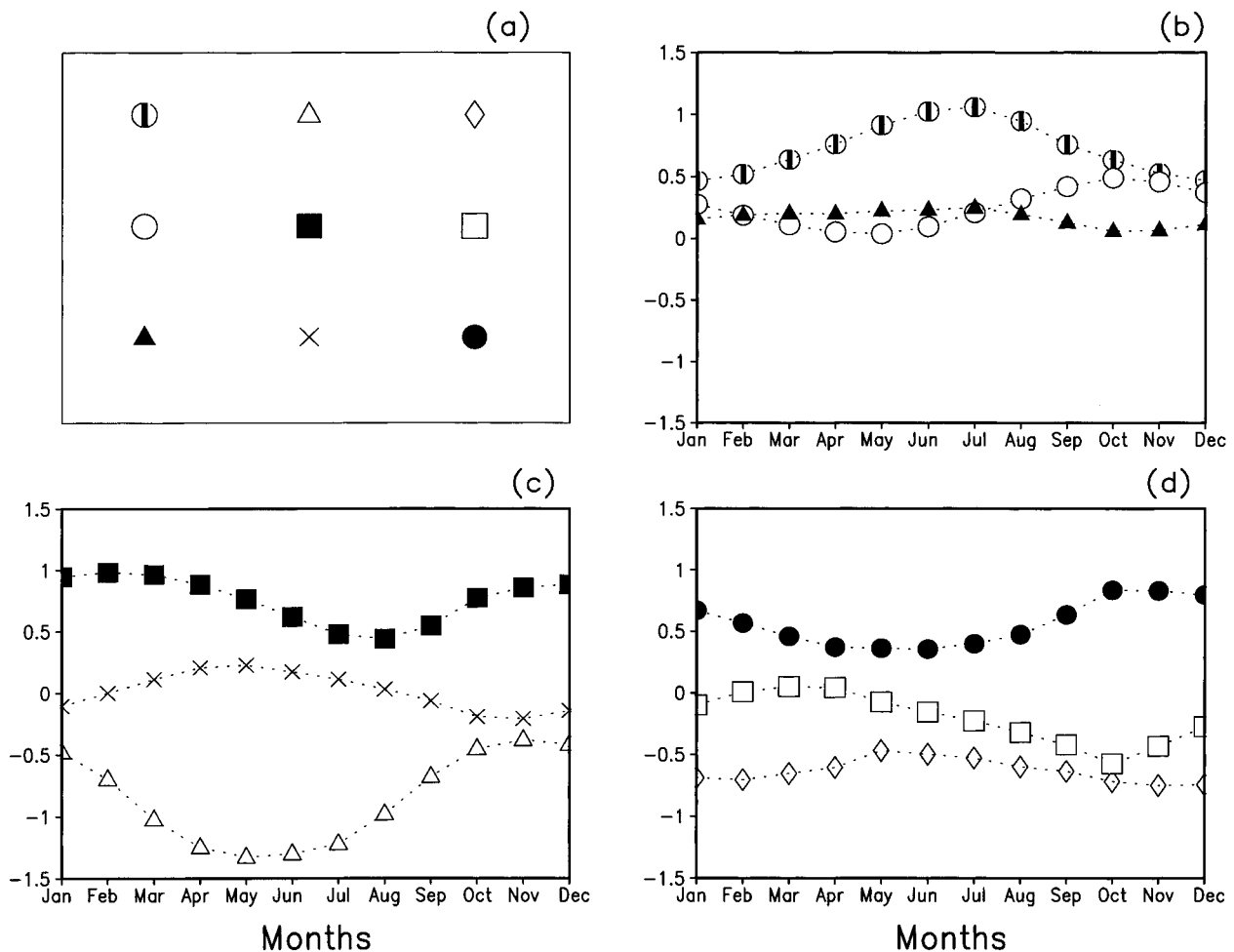


FIG. 6. Same as Fig. 5 except for the 6-month transition matrixes.

by model–observation correlation of the averaged SST anomaly in the Niño-3.4 region (5°S–5°N, 170°–120°W). Figure 7 shows a comparison between the skills of the nonseasonal and seasonal Markov models for 1980–95. Since the data used for verification are also used for constructing the models, this skill is referred as noncross-validated skill later. When three MEOFs are retained, the noncross-validated skill of the seasonal Markov model is better than that of the nonseasonal Markov model (Fig. 7a). As expected, the improvement of skill at long lead times is larger than that at short lead times. However, this improvement of skill might not be statistically significant since the number of the model parameters in the seasonal Markov models are 12 times that in the nonseasonal Markov model. With 10 retained MEOFs, the number of the model parameters in the nonseasonal Markov model is about the same as that in the seasonal Markov model with three retained MEOFs. Figure 7a shows that the skill of the nonseasonal Markov model with 10 retained MEOFs is still lower than that of the seasonal Markov model with three retained MEOFs. The above results suggest that sea-

sonal Markov models generally fit the data better than nonseasonal Markov models.

The hindcast skill of the Markov models for the independent 1964–79 period is expected to be lower than the noncross-validated skill for 1980–95 (Fig. 7b). Besides, the hindcast skill is hurt due to the inaccuracy of the sea level simulation that is used as initial conditions for Markov models in this period. Behringer et al. (1998) showed that the inaccuracy of the sea level simulation is 4–5 cm in the equatorial belt, which is 1–2 cm larger than that of the sea level analysis. In addition, interdecadal changes of predictability of ENSO also contribute to the lower skill in the pre-1980 period than in the post-1980 period (Balmaseda et al. 1995).

The hindcast skill of the seasonal Markov model with three retained MEOFs is significantly better than that of the nonseasonal Markov model with three retained MEOFs (Fig. 7b). Including more MEOFs does not help the hindcast skill of the nonseasonal Markov models much (Fig. 7b).

The above results suggest that seasonality is an important component of ENSO and should be included in

Markov models. From now on only seasonal Markov models are discussed.

4. Impact of sea level on prediction skill

a. Noncross-validated skill

The Markov models discussed above are built in an MEOF space in which SST, wind stress, and sea level are equally weighted. An interesting question is how the prediction skill of Markov models changes when the weightings among the three variables are changed. In particular, we study the impact of sea level on the prediction skill of Markov models.

Four sets of Markov models are built in which the weightings among SST, wind stress, and sea level are different. MK(SST) stands for a model built in an MEOF space in which wind stress and sea level are weighted substantially less than SST, specifically $\epsilon_1 = 1$, $\epsilon_2 = 10^{-20}$, $\epsilon_3 = 10^{-20}$. So the MEOFs are essentially determined by SST. MK(SST, TAU) stands for a model built in an MEOF space in which the weights of SST, wind stress, and sea level are $\epsilon_1 = 1$, $\epsilon_2 = 1$, $\epsilon_3 = 10^{-20}$, respectively. MK(SST, TAU, SL) stands for a model built in an MEOF space where $\epsilon_1 = 1$, $\epsilon_2 = 1$, $\epsilon_3 = 1$. MK(SL) stands for a model built in an MEOF space where $\epsilon_1 = 10^{-20}$, $\epsilon_2 = 10^{-20}$, $\epsilon_3 = 1$.

For each of the model sets a series of Markov models are constructed with different number of retained MEOFs. The noncross-validated skill of the Markov models with 2, 3, 5, and 7 retained MEOFs for 1980–95 is shown in Fig. 8. Although this skill is expected to increase with the number of retained MEOFs, it is not always true because the Markov models are best fits for monthly evolutions only. For example, the skill of the model MK(SST) with 3 retained MEOFs is higher than that with 2 retained MEOFs (Fig. 8a). However, when 5 MEOFs are retained the skill of MK(SST) is slightly lower than that with 3 retained MEOFs. When 7 MEOFs are retained the skill of the model is improved further. The overall skill of MK(SST, TAU) is comparable to that of MK(SST) (Fig. 8b). This is understood since wind stress anomalies are well correlated with SST anomalies and they do not add much new information. The skill of the model MK(SST, TAU, SL) with 3 retained MEOFs is substantially higher than that with 2 retained MEOFs (Fig. 8c). When more than 3 MEOFs are retained the skill changes little. The overall skill of MK(SL) is comparable to that of MK(SST, TAU, SL) (Fig. 8d).

Figure 8 shows that the Markov models that include sea level information generally fit the data better than the Markov models without sea level information. For example, with 2 retained MEOFs the skill of the model MK(SL) is as high as that of the model MK(SST) with 7 retained MEOFs (cf. Figs. 8a and 8d). This can be explained by referring to the results in section 2. There the first EOF of SST accounts for about 52% of the

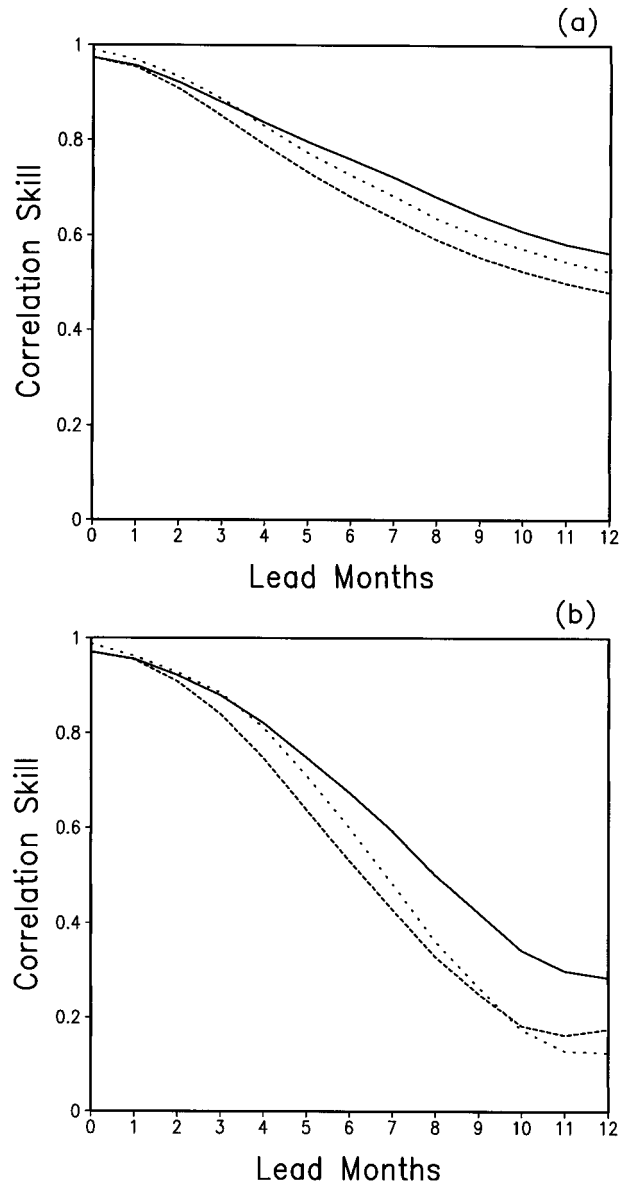


FIG. 7. Model–observation correlation of Niño-3.4 SST anomalies for (a) 1980–95 and (b) 1964–79 as functions of lead months for the seasonal (solid) and nonseasonal (dashed) Markov models with 3 retained MEOFs, and the nonseasonal Markov models with 10 retained MEOFs (dotted).

total variance and the higher-order EOFs (> 1 st) contain much less variance (Fig. 1a). So the signal to noise ratio for the higher-order EOFs of SST is low. However, the variance of sea level is more evenly distributed among the EOFs (Fig. 1b). The signal to noise ratio for the first few EOFs of sea level is relatively high. With the same number of retained MEOFs the model based on sea level contains less noise than the model based on SST. The fact that the overall skill of MK(SST, TAU, SL) is comparable to that of MK(SL) suggests that when SST, wind stress, and sea level are equally weighted the skill of the Markov models is mostly controlled by sea level.

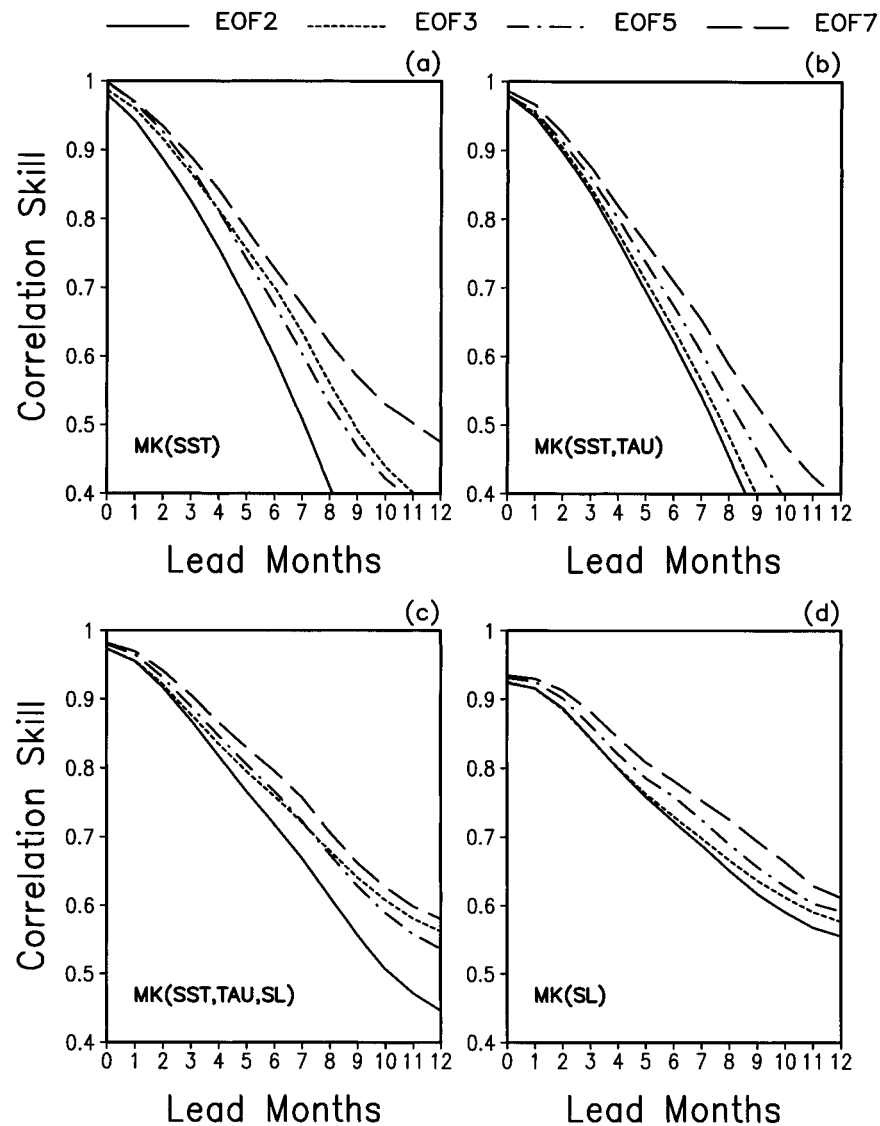


FIG. 8. Model-observation Niño-3.4 correlation for 1980-95 (not cross validated) as functions of lead months for (a) MK(SST), (b) MK(SST, TAU), (c) MK(SST, TAU, SL), and (d) MK(SL). The different curves are for models with 2 (solid), 3 (short dashed), 5 (dot-dashed), and 7 (long dashed) retained MEOFs.

The skill of the Markov models has a seasonal dependency. It is seen in Fig. 9a that the skill is highest when models start from the boreal spring, and it is lowest when models start from the boreal winter. The skill appears drop substantially when forecasts pass through the spring season. This sharp drop of skill in spring is common in various prediction schemes (Latif et al. 1993; Barnston and Ropelewski 1992) and is often referred as “spring barrier.” A similar characteristic is seen in the autocorrelation of observed Niño-3.4 anomalies (Fig. 9b). Xue et al. (1994) proposed that the rapid drop of skill in spring is due to the small variance at that time and thus is characteristic of the ENSO cycle.

It is noticed that the decline of the skill in spring is not as sharp as that in the autocorrelation, and it is often followed by a substantial recovery. Balmaseda et al. (1995) suggested that the correlation skill decline and recovery are due to seasonal changes in variance, because low variance tends to decrease correlations. In addition, skill recovery is assisted by a physical mechanism, that sea level and SST information are not lost simultaneously and the coupled system can keep part of its memory. A comparison between Figs. 9a and 9b shows that the skill is improved most by knowing sea level information for spring starts when SST persistence is the weakest.

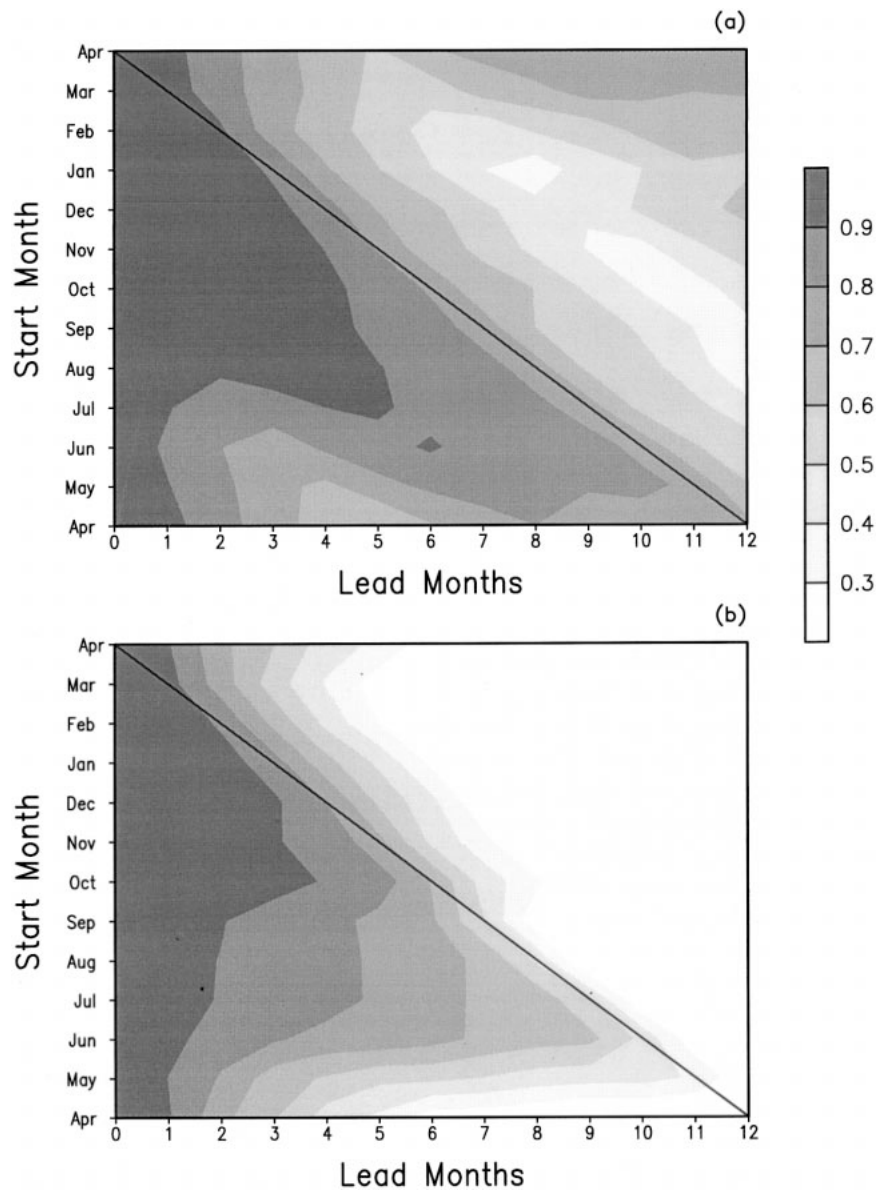


FIG. 9. Model–observation Niño-3.4 correlation for 1980–95 (not cross validated) as functions of start months and lead months by (a) MK(SST, TAU, SL) with three retained MEOFs. (b) Autocorrelation of the monthly observed Niño-3.4 for 1980–95 as functions of starting months and lead months. Notice there is a rapid decline of correlation around Apr in (a) and (b).

b. Hindcast skill

The hindcast skill of the Markov models is estimated using the independent 1964–79 period. Figure 10 shows that the hindcast skill of MK(SST) does not vary much with the number of retained MEOFs and is slightly higher than the skill of the persistence forecasts. The hindcast skill of MK(SST, TAU) is not sensitive to the number of retained MEOFs either and is also slightly better than the skill of the persistence forecasts. The hindcast skill of MK(SST, TAU, SL) with 3 retained MEOFs is substantially higher than that with 2 retained MEOFs and keeping more MEOFs (>3) does not help the skill

much. The hindcast skill of MK(SL) is similar to that of MK(SST, TAU, SL) and both models have a significantly higher skill than that of the persistence forecasts.

Generally speaking, the models that include sea level information have a higher skill than the models without sea level information. For example, at 6-month lead the model MK(SST, TAU, SL) with 5 retained MEOFs has a skill of 0.7; the model MK(SL) with 5 retained MEOFs has a skill of 0.8; while the skills of the models MK(SST) and MK(SST, TAU) are only 0.5.

We wonder whether the poor skill of MK(SST) was due to the use of monthly SST anomalies, which are

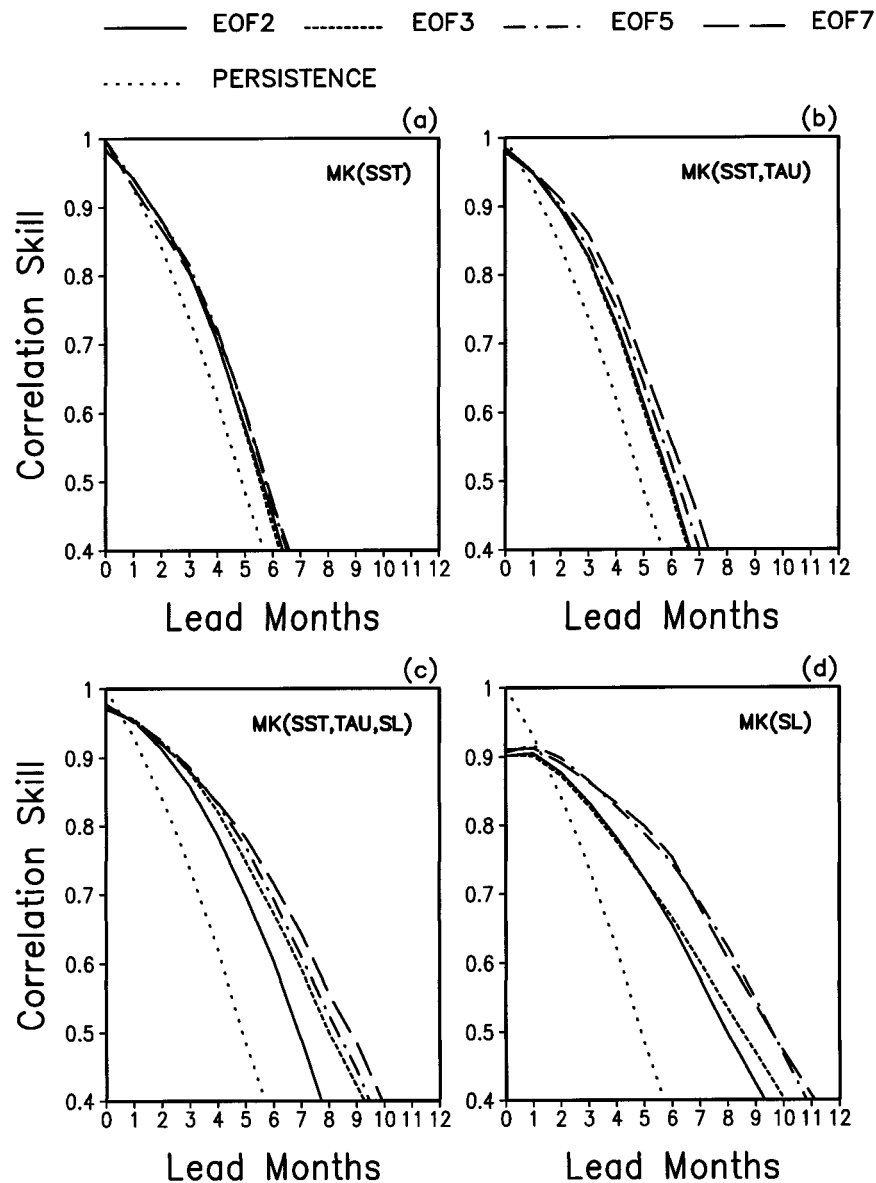


FIG. 10. Same as Fig. 8 except for the independent 1964–79 period. The skill of the persistence forecasts is also shown (dot).

often considered too noisy for Markov models (Penland and Magorian 1993; Johnson et al. 2000). We reconstructed Markov models using the 3-month running mean of SST anomalies and found that the new Markov models fit the data slightly better for the training period, but the skill for the independent 1964–79 period is about the same as before.

The hindcast skill of the Markov models has a strong seasonal dependency, similar to that of the autocorrelation of observed Niño-3.4 anomaly (Fig. 11). As for the 1980–95 period, the hindcast skill is improved most by knowing sea level information for spring starts when SST persistence is the weakest. A comparison between Figs. 9 and 11 suggests that the spring barrier

for 1964–79 is stronger than that for 1980–95. This result is consistent with that of Balmaseda et al. (1995), who studied the seasonal dependency of prediction skill of an intermediated coupled model in the pre- and post-1980 periods. They suggested that this difference depends substantially on the degree of phase locking of El Niño to the annual cycle, as well as on stability conditions associated with the background seasonal cycle.

c. Cross-validated skill

The hindcast skill of the Markov models for 1980–95 can be estimated using a cross-validation scheme

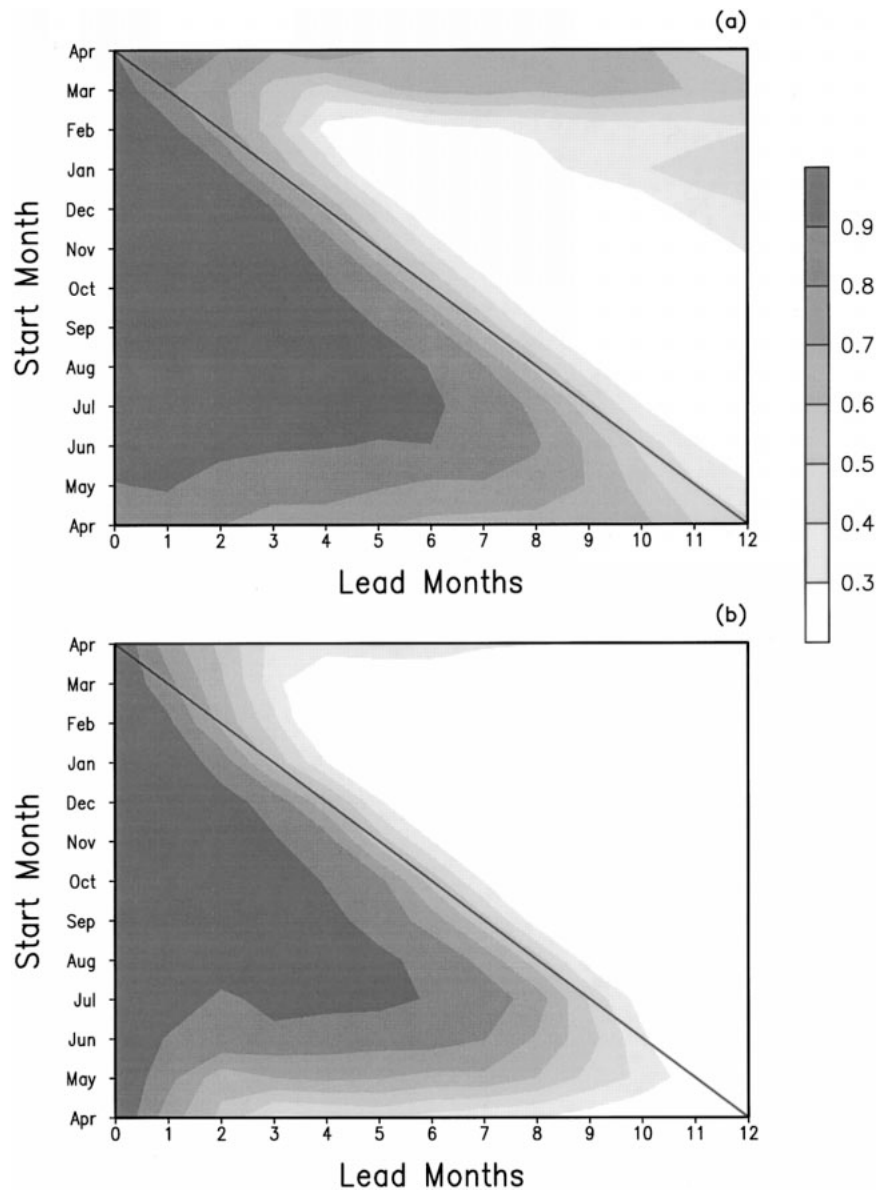


FIG. 11. Same as Fig. 9 except it is the hindcast skill for 1964–79.

(Barnston and Ropelewski 1992). In a cross-validation scenario, one year of data is removed, and a Markov model is trained upon the remaining years (15 yr) and verified at the removed year. The 1-yr window is moved forward month by month until the end of the time series is reached. So there are totally $192 = 16 \times 12$ multiple analyses with different years removed. Only MK(SST, TAU, SL) and MK(SL) are cross-validated because they have a modest skill for 1964–79. Figure 12 shows that the skill of MK(SST, TAU, SL) with 3 retained MEOFs is slightly higher than that of the models with 2, 5, and 7 retained MEOFs at long lead times. Although the cross-validated skills are low-

er than the noncross-validated skills at all the four MEOF truncations (cf. Figs. 12a and 8c), they all outperform the persistence forecasts at lead times longer than 4 months. The skill of MK(SL) with 2 retained MEOFs is the highest and keeping more MEOFs decreases the skill. But all the models with different MEOF truncations outperform the persistence forecasts at lead times longer than 5 months.

The correlation skills in Figs. 8c, 10c, and 12a suggest that the appropriate number of retained MEOFs for MK(SST, TAU, SL) is 3. The correlation skills in Figs. 8d, 10d, and 12b suggest that no more than 3 MEOFs are needed for MK(SL). Since the skill of MK(SST,

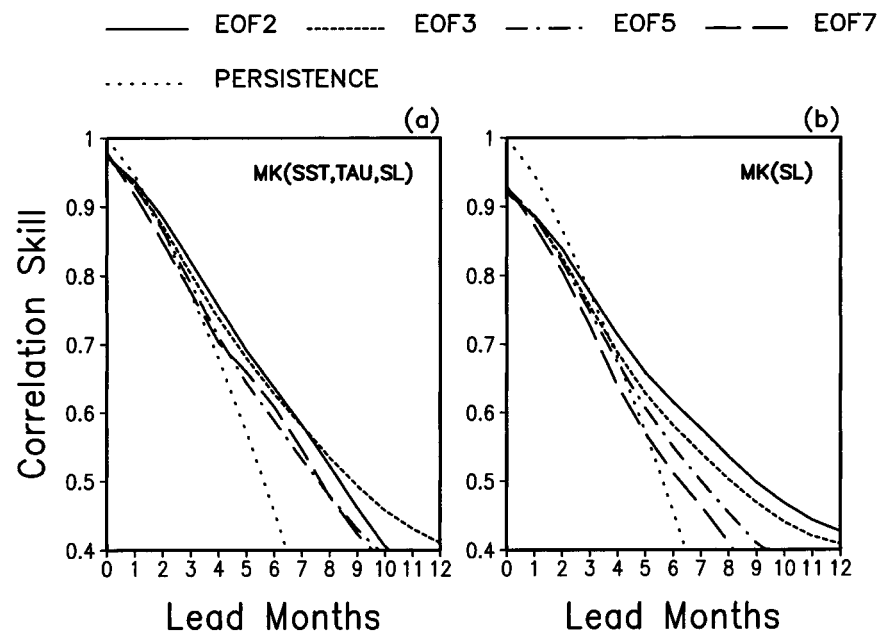


FIG. 12. Model-observation Niño-3.4 correlation for 1980-95 (cross validated) as functions of lead months for (a) MK(SST, TAU, SL) and (b) MK(SL). The different curves are for models with 2 (solid), 3 (short dashed), 5 (dot-dashed), and 7 (long dashed) retained MEOFs.

TAU, SL) and MK(SL) is modest for both the training and independent periods, we propose that ENSO can be approximately described as a low-order linear system with a dimension of 3.

The cross-validated skill of the Markov models is hampered by the small sample size. This can be seen from the comparison between the Niño-3.4 SST anomalies forecast by MK(SST, TAU, SL) with 3 retained MEOFs with and without cross validation (Fig. 13). It is seen that the Niño-3.4 SST anomalies forecasts are very similar in both cases, except during the 1982/83 event. There the fast growth and decay phases were seriously underestimated. Despite our short record, the model has a useful skill for up to a 6-month lead in forecasting moderate events such as those of 1984/85, 1986/87, 1988/89, and 1991/92. The lead time with usable skill for forecasting the short warm episode in 1994 and the 1995/96 weak cold event is only about 3 months.

d. "Real time" forecast skill

The model MK(SST, TAU, SL) with 3 retained MEOFs is used to predict the 1997/98 warm and the 1998/99 cold events. Figure 14 shows the individual 12-month forecasts of Niño-3.4 SST anomalies. The thin lines in each panel of the figure depict forecasts starting from three consecutive months for each year. For example, the top left panel depicts all the forecasts initiated in December-February (DJF) of each year. The phase transition from a mild cold to warm anomaly in the early

1997 is delayed by the model by 1-2 months (Fig. 14a). The fast growth in the spring and early summer of 1997 is seriously underestimated by the model (Fig. 14b). The predictions initiated from the late 1996 and early 1997 underestimated the peak phase of the event, but the predictions initiated later improve with shorter lead times. This is also the case for most of ENSO prediction models (Barnston et al. 1999). The decay phase of the warm event was predicted quite well except the transition to the cold phase was delayed by two months. The amplitude of the 1998/99 La Niña event was forecast well. The overall performance of the Markov model is competitive among the best performers of various dynamical and statistical models documented by Barnston et al. (1999).

The spatial patterns of the observed and forecast SST initiated from March 1997 are shown in Fig. 15. The warming in the eastern Pacific is seriously underestimated by the model but the warming in the central Pacific is simulated well. The spatial patterns of the forecast SST by the Markov model initiated from March 1998 are shown in Fig. 16. A negative anomaly starts on the equator in the central Pacific in summer 1998 and amplifies quickly in fall, and at the same time the positive anomaly in the southeastern Pacific retreats quickly. This evolution of the 1998/99 cold event is well simulated by the Markov model.

The model MK(SL) was also used to forecast the 1997/98 El Niño and the 1998/99 La Niña events. The forecast NINO3.4 anomalies are actually indistinguish-

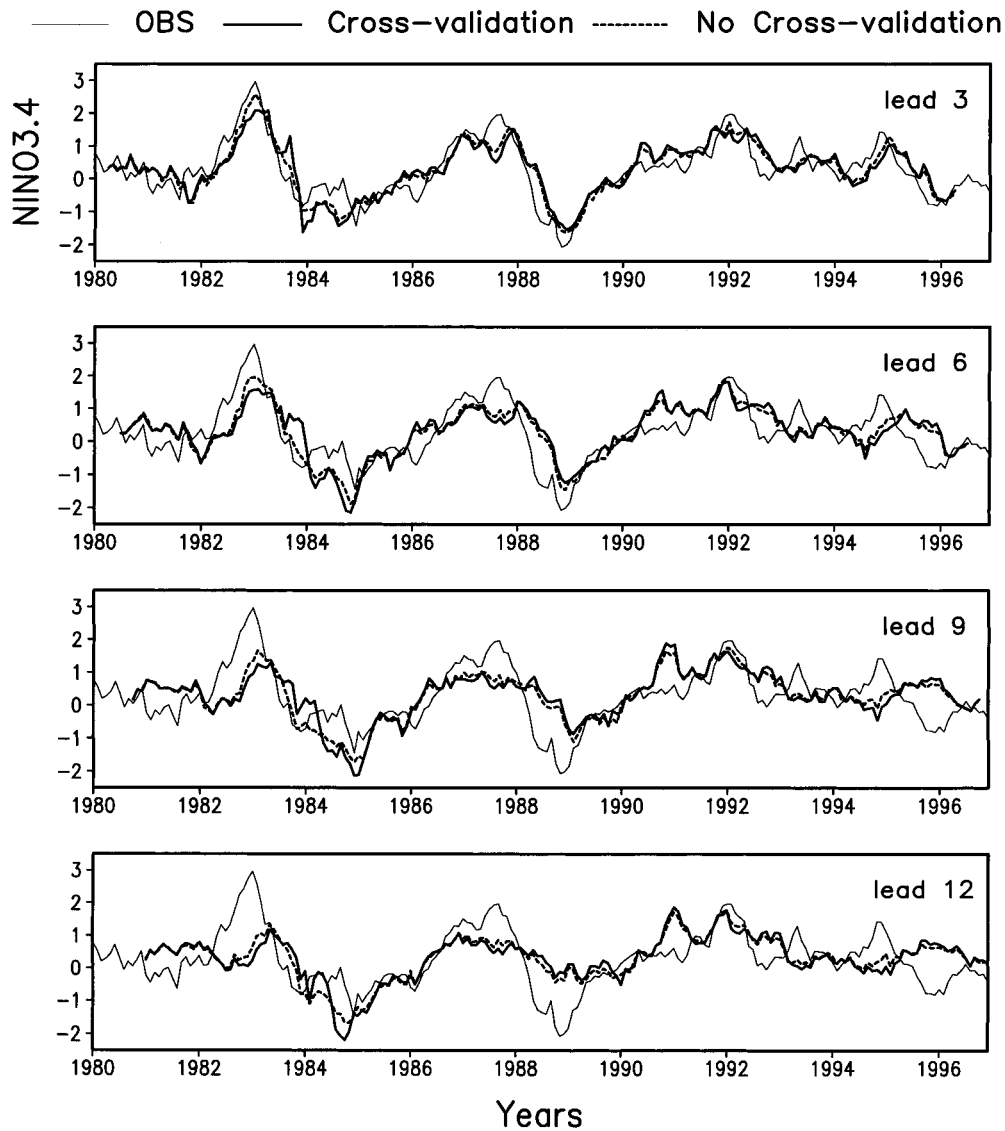


FIG. 13. A comparison of the forecast Niño-3.4 SST anomalies by MK(SST, TAU, SL) with three retained MEOFs with (solid) and without (dashed) cross validation at 3-, 6-, 9-, and 12-month leads. The observed Niño-3.4 SST anomalies are shown by the light lines.

able to those by the model MK(SST, TAU, SL) (not shown). This suggests that the forecast skill of the Markov models is determined by sea level, which contains the critical information for the development of these two major events.

5. Summary and discussions

Markov models are constructed in a multivariate EOF space of observed SST, wind stress, and sea level analysis in 1980–95. Two types of Markov models, which include and exclude seasonality, are discussed. It is found that seasonality is small in monthly timescales but large in 6-month timescales. So including

seasonality in Markov models improves prediction skill more at long lead times than at short lead times. This result is consistent with the conclusions of Hasselmann and Barnett (1981) and Johnson (2000) that seasonality is an important component of the ENSO system and should be included in statistical modeling for ENSO.

Four sets of Markov models are constructed in MEOF spaces where the weightings among SST, wind stress, and sea level are varied. It is found that the models MK(SST, TAU, SL) and MK(SL), which include sea level information, generally fit the data better than the models MK(SST) and MK(SST, TAU), which do not include sea level information. This is because the signal

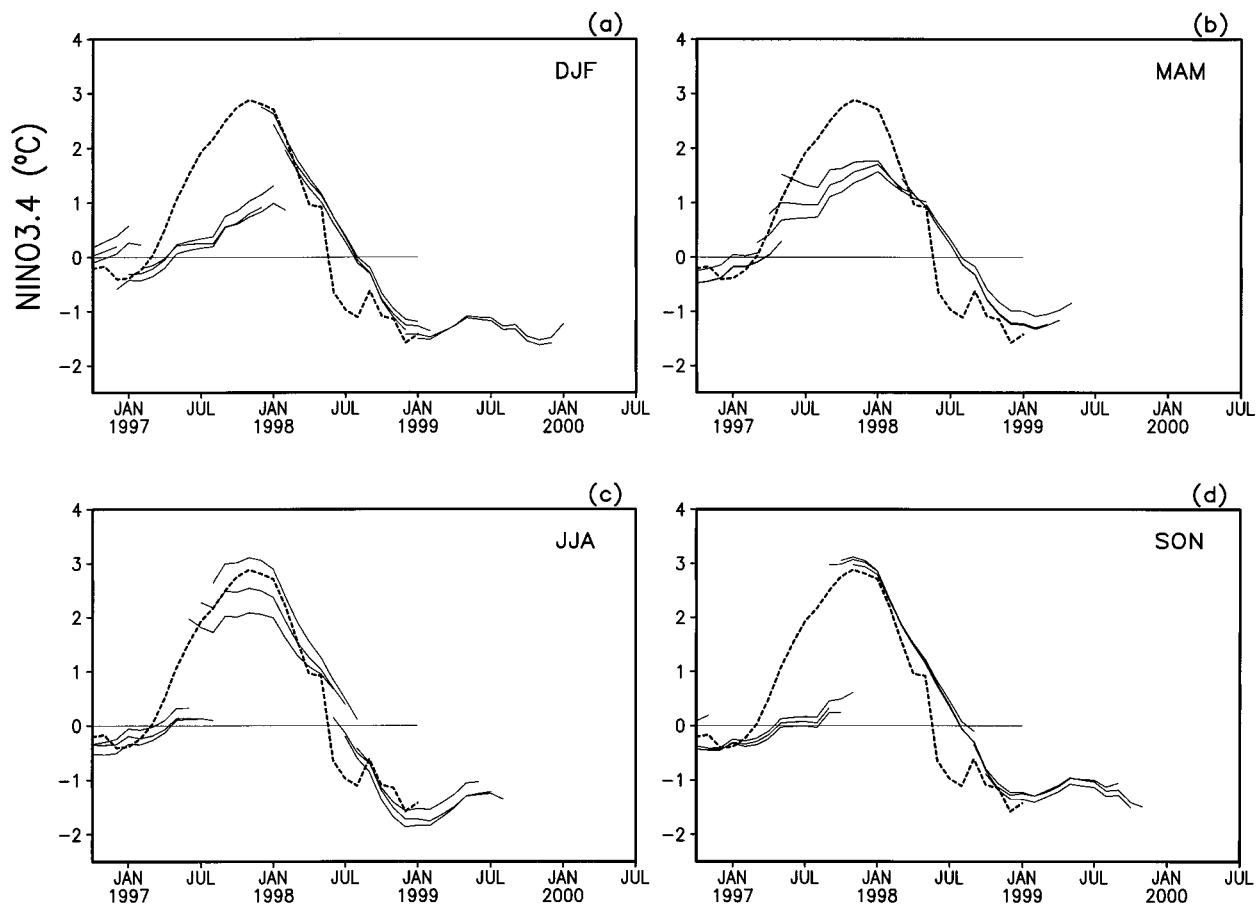


FIG. 14. Evolution of Niño-3.4 SST anomalies ($^{\circ}\text{C}$) from all forecasts initiated monthly using MK(SST, TAU, SL) with three retained MEOFs from Oct 1996 to Jan 1999. Shown in each panel are the forecasts grouped by three consecutive starting months: (a) DJF, (b) MAM, (c) JJA, and (d) SON. The observed Niño-3.4 SST anomalies are shown by the heavy-dashed lines.

to noise ratio for the first few EOFs of sea level is higher than that for SST and wind stress. With the same number of retained MEOFs, the models that include sea level information contain less noise than the models without sea level information.

The impact of sea level on the prediction skill of Markov models can be seen from two aspects. For the training period the models that include sea level information fit the data better than the models without sea level information; for the independent 1964–79 period the models that include sea level information have a much higher hindcast skill than the models without sea level information. Those results suggest that sea level carries the most critical information for ENSO. This is consistent with the fact that prediction skill of numerical models has been significantly improved by assimilating subsurface temperature data into ocean initial conditions, assuming sea level is equivalent to subsurface temperature (Kleeman et al. 1995; Ji and Smith 1995; Rosati et al. 1996).

Many studies suggest that oceanic heat content, which is equivalent to sea level, in the equatorial Pa-

cific is critical for the timing and strength of ENSO events (Wyrtki 1985; Zebiak and Cane 1987; Zebiak 1989). The oceanic heat content information can be obtained by either forcing ocean models with observed wind stress or directly assimilating subsurface temperature data into ocean models. Since Markov models can only use information in the last time step, they cannot simulate the integrated response of the ocean to wind forcings. So they rely on sea level data for oceanic heat content information. This is similar to the situation where ocean data assimilation systems are used to provide the most accurate oceanic heat content information.

The model MK(SST, TAU, SL) with 3 retained MEOFs successfully predicted the 1997/98 El Niño and the 1998/99 La Niña, and its performance is competitive among the best performers of various dynamic and statistical models documented by Barnston et al. (1999). Like many other models the Markov model predicted a warming for 1997/98 a year in advance, but the fast warming in spring and summer was significantly underestimated. An outstanding question is whether the

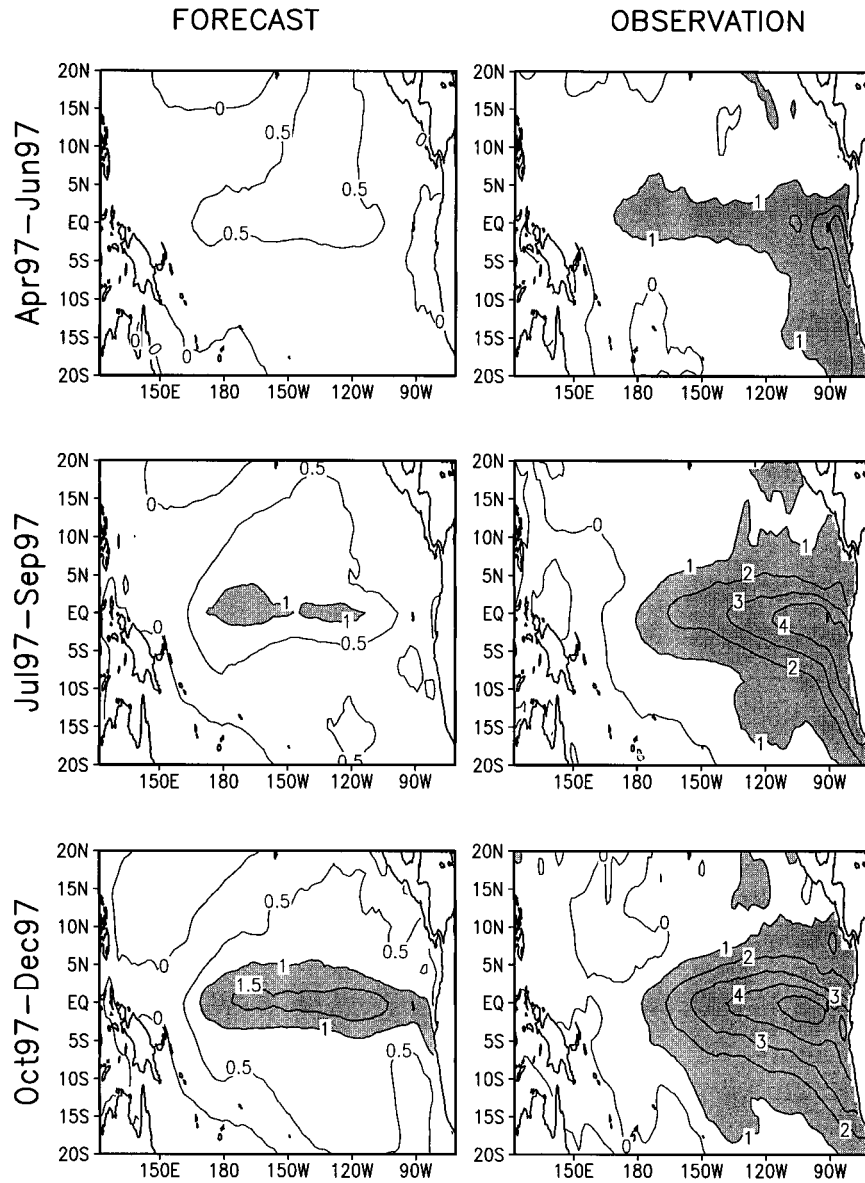


FIG. 15. A comparison of the seasonal mean SST anomalies forecast (left column) by MK(SST, TAU, SL) with three retained MEOFs initiated from Mar 1997 and the observation (right column).

amplitude of the event was predictable starting from late 1996 and early 1997. Moore and Kleeman (1999) suggest that the two strong Madden-Julian oscillations initiated in January and March of 1997 caused optimal perturbation growth that significantly increased the intensity of this event. If this is indeed the cause, then ENSO prediction models will need to simulate the influences of atmospheric high-frequency variability on ENSO development.

Acknowledgments. The authors are grateful to Dr. H. van den Dool for helping improve the construction of Markov models, Dr. A. G. Barnston for providing useful comments on the paper, and R. W. Reynolds for helping with the English. This research was supported in part by an appointment to the National Centers for Environmental Prediction sponsored by NCEP and administered by the University Corporation for Atmospheric Research.

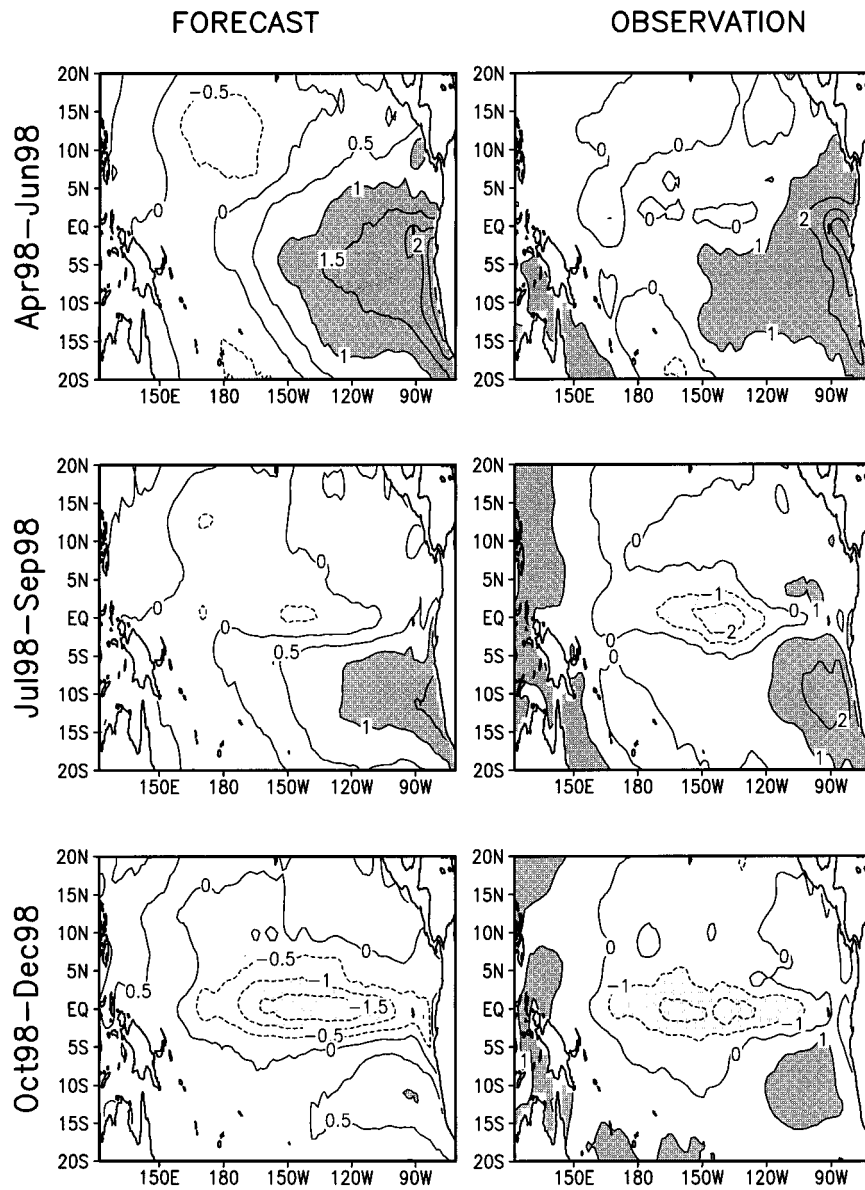


FIG. 16. A comparison of the seasonal mean SST anomalies forecast (left column) by MK(SST, TAU, SL) with three retained MEOFs initiated from Mar 1998 and the observation (right column).

APPENDIX

Construction of Markov Models

The formula for nonseasonal Markov models is

$$\mathbf{b}_{i+1} = \mathbf{A}\mathbf{b}_i + \mathbf{e}_i, \tag{A1}$$

where \mathbf{b}_i is the PC of the MEOFs at the i th month, \mathbf{A} is the monthly transition matrix and \mathbf{e}_i is the residue. Multiplying by the transpose of vector \mathbf{b}_i on both sides of (A1), then averaging on all samples gives

$$\langle \mathbf{b}_{i+1}\mathbf{b}_i^T \rangle = \mathbf{A}\langle \mathbf{b}_i\mathbf{b}_i^T \rangle + \langle \mathbf{e}_i\mathbf{b}_i^T \rangle, \tag{A2}$$

where $\langle \dots \rangle$ means the average over all samples. For the best fit model \mathbf{A} , \mathbf{e}_i does not correlate with \mathbf{b}_i , so

$$\mathbf{A} = \langle \mathbf{b}_{i+1}\mathbf{b}_i^T \rangle \langle \mathbf{b}_i\mathbf{b}_i^T \rangle^{-1} = \mathbf{C}_i\mathbf{D}_i^{-1}; \tag{A3}$$

here \mathbf{C}_i is the lag-1 covariance matrix, while \mathbf{D}_i the autocovariance matrix. Equation (A3) is often described as a Yule–Walker equation in the literature (Brockwell and Davis 1991).

For the construction of seasonal Markov models, the monthly PCs \mathbf{b}_i are grouped into 12 subsets, one for each calendar month. In Xue et al. (1994), the data in subset m and $m + 1$ were used to calculate the monthly transition matrix $\mathbf{A}^{(m)}$ from month m to $m + 1$, and in total 12 monthly transition matrixes were constructed. Since the data sample for training each of the 12 monthly transition matrixes is one-twelfth of that for training

the monthly transition matrix of the nonseasonal Markov model, the statistical significance of seasonal Markov models is lower than that of nonseasonal Markov models. One way to increase the statistical significance of the model is to sacrifice the annual cycle resolution (Hasselmann and Barnett 1981). We tripled the sample size by including the data samples one month before and after the month m in subset m . So $\mathbf{A}^{(m)}$ is a phase-averaged monthly transition matrix for the three months centered at m .

Denoting the data in subset m by \mathbf{b}_m , the formula is

$$\mathbf{b}_{m+1} = \mathbf{A}^{(m)}\mathbf{b}_m + \mathbf{e}_m, \quad (\text{A4})$$

where $\mathbf{A}^{(m)}$ is the monthly transition matrix and \mathbf{e}_m is the residue. Multiplying by the transpose of vector \mathbf{b}_m on both sides of (A4), then averaging on all the samples in subset m (16×3 data points) gives

$$\langle \mathbf{b}_{m+1}\mathbf{b}_m^T \rangle = \mathbf{A}^{(m)}\langle \mathbf{b}_m\mathbf{b}_m^T \rangle + \langle \mathbf{e}_m\mathbf{b}_m^T \rangle. \quad (\text{A5})$$

For the best fit model $\mathbf{A}^{(m)}$, \mathbf{e}_m does not correlate with \mathbf{b}_m , so

$$\mathbf{A}^{(m)} = \langle \mathbf{b}_{m+1}\mathbf{b}_m^T \rangle \langle \mathbf{b}_m\mathbf{b}_m^T \rangle^{-1} = \mathbf{C}_m\mathbf{D}_m^{-1}; \quad (\text{A6})$$

here \mathbf{C}_m is the lag-1 covariance matrix, while \mathbf{D}_m is the autocovariance matrix for subset m .

When a prediction is made for leads longer than one month, the monthly transition matrices are used sequentially through the forecast lead time. For example, a prediction starts from vector \mathbf{b}_m and lasts for six months, the final vector \mathbf{b}_{m+6} is

$$\mathbf{b}_{m+6} = \mathbf{A}^{(m+6)} \cdots \mathbf{A}^{(m+1)}\mathbf{A}^{(m)}\mathbf{b}_m. \quad (\text{A7})$$

REFERENCES

- Balmaseda, M. A., M. K. Davey, and D. L. T. Anderson, 1995: Decadal and seasonal dependence of ENSO prediction skill. *J. Climate*, **8**, 2705–2715.
- Barnett, T. P., and Coauthors, 1994: Forecasting global ENSO-related climate anomalies. *Tellus*, **46A**, 381–397.
- Barnston, A. G., and C. F. Ropelewski, 1992: Prediction of ENSO episodes using canonical correlation analysis. *J. Climate*, **5**, 1316–1345.
- , and Coauthors, 1994: Long-lead seasonal forecasts—Where do we stand? *Bull. Amer. Meteor. Soc.*, **75**, 2097–2114.
- , M. H. Glantz, and Y. He, 1999: Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997–98 El Niño episode and the 1998 La Niña onset. *Bull. Amer. Meteor. Soc.*, **80**, 217–243.
- Battisti, D. S., 1988: The dynamics and thermodynamics of a warming event in a coupled tropical atmosphere–ocean model. *J. Atmos. Sci.*, **45**, 2889–2919.
- Behringer, D. W., M. Ji, and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part I: The ocean data assimilation system. *Mon. Wea. Rev.*, **126**, 1013–1021.
- Blanke, R., J. D. Neelin, and D. Gutzler, 1997: Estimating the effect of stochastic wind stress forcing on ENSO irregularity. *J. Climate*, **10**, 1473–1486.
- Blumenthal, M. B., 1991: Predictability of a coupled ocean–atmosphere model. *J. Climate*, **4**, 766–784.
- Brockwell, P. J., and R. A. Davis, 1991: *Time Series: Theory and Methods*. 2d ed. Springer-Verlag, 577 pp.
- Eckert, C., and M. Latif, 1997: Predictability of a stochastically forced hybrid coupled model of El Niño. *J. Climate*, **10**, 1488–1504.
- Goldenberg, S. B., and J. J. O'Brien, 1981: Time and space variability of tropical Pacific wind stress. *Mon. Wea. Rev.*, **109**, 1190–1207.
- Graham, N. E., J. Michaelsen, and T. P. Barnett, 1987a: An investigation of the El Niño–Southern Oscillation cycle with statistical models. 1. Predictor field characteristics. *J. Geophys. Res.*, **92**, 14 251–14 270.
- , —, and —, 1987b: An investigation of the El Niño–Southern Oscillation cycle with statistical models. 2. Model results. *J. Geophys. Res.*, **92**, 14 271–14 289.
- Hasselmann, K., and T. P. Barnett, 1981: Techniques of linear prediction for systems with periodic statistics. *J. Atmos. Sci.*, **38**, 2275–2283.
- Hellerman, S., and M. Rosenstein, 1983: Normal monthly wind stress over the world ocean with error estimates. *J. Phys. Oceanogr.*, **13**, 1093–1104.
- Ji, M., and T. M. Smith, 1995: Ocean model response to temperature data assimilation and varying surface stress: Intercomparisons and implications for climate forecast. *Mon. Wea. Rev.*, **123**, 1811–1821.
- , A. Kumar, and A. Leetmaa, 1994: A multiseason climate forecast system at the National Meteorological Center. *Bull. Amer. Meteor. Soc.*, **75**, 569–577.
- , A. Leetmaa, and J. Derber, 1995: An ocean analysis system for seasonal to interannual climate studies. *Mon. Wea. Rev.*, **123**, 460–481.
- Johnson, S. D., 2000: Seasonality in an empirically derived Markov model of tropical Pacific sea surface temperature anomalies. *J. Climate*, in press.
- , D. S. Battisti, and E. S. Sarachik, 2000: Empirically derived Markov models and prediction of tropical Pacific sea surface temperature anomalies. *J. Climate*, **13**, 3–17.
- Kim, K.-Y., 2000: Statistical prediction of cyclostationary processes. *J. Climate*, in press.
- , and G. R. North, 1997: EOFs of harmonizable cyclostationary processes. *J. Atmos. Sci.*, **54**, 2416–2427.
- Kleeman, R., A. M. Moore, and N. R. Smith, 1995: Assimilation of subsurface thermal data into a simple ocean model for the initialization of an intermediate tropical coupled ocean–atmosphere forecast model. *Mon. Wea. Rev.*, **123**, 3103–3113.
- Latif, M., T. P. Barnett, M. A. Cane, M. Flugel, N. E. Graham, H. von Storch, J.-S. Xu, and S. E. Zebiak, 1994: A review of ENSO prediction studies. *Climate Dyn.*, **9**, 167–179.
- McCreary, J. P., Jr., and D. L. T. Anderson, 1991: An overview of coupled ocean–atmosphere models of El Niño and the Southern Oscillation. *J. Geophys. Res.*, **96**, 3125–3150.
- Moore, A. M., and R. Kleeman, 1999: Stochastic forcing of ENSO by the intraseasonal oscillation. *J. Climate*, **12**, 1199–1220.
- Penland, C., and T. Magorian, 1993: Prediction of Niño-3 sea surface temperatures using linear inverse modeling. *J. Climate*, **6**, 1067–1076.
- Philander, S. G., 1983: El Niño Southern Oscillation phenomena. *Nature*, **302**, 295–301.
- Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929–948.
- Rosati, A., K. Miyakoda, and R. Gudgel, 1996: The impact of ocean initial conditions on ENSO forecasting with a coupled model. *Mon. Wea. Rev.*, **125**, 754–772.
- Smith, T. M., A. G. Barnston, M. Ji, and M. Chelliah, 1995: The impact of Pacific ocean subsurface data on operational prediction of tropical Pacific SST at the NCEP. *Wea. Forecasting*, **10**, 708–714.
- , R. W. Reynolds, R. E. Livezey, and D. C. Stokes, 1996: Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *J. Climate*, **9**, 1403–1420.

- Wyrski, K., 1985: Water displacements in the Pacific and the genesis of El Niño cycles. *J. Geophys. Res.*, **90**, 7129–7132.
- Xue, Y., M. A. Cane, S. E. Zebiak, and M. B. Blumenthal, 1994: On the prediction of ENSO: A study with a low-order Markov model. *Tellus*, **46A**, 512–528.
- Zebiak, S. E., 1989: Oceanic heat content variability and El Niño cycles. *J. Phys. Oceanogr.*, **19**, 475–486.
- , and M. A. Cane, 1987: A model El Niño–Southern Oscillation. *Mon. Wea. Rev.*, **115**, 2262–2278.