

Reliability of ENSO Dynamical Predictions

YOU MIN TANG AND RICHARD KLEEMAN

Courant Institute of Mathematical Sciences, New York University, New York, New York

ANDREW M. MOORE

Program in Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado

(Manuscript received 6 June 2004, in final form 28 October 2004)

ABSTRACT

In this study, ensemble predictions were constructed using two realistic ENSO prediction models and stochastic optimal. By applying a recently developed theoretical framework, the authors have explored several important issues relating to ENSO predictability including reliability measures of ENSO dynamical predictions and the dominant precursors that control reliability. It was found that prediction utility (R), defined by relative entropy, is a useful measure for the reliability of ENSO dynamical predictions, such that the larger the value of R , the more reliable the prediction. The prediction utility R consists of two components, a dispersion component (DC) associated with the ensemble spread and a signal component (SC) determined by the predictive mean signals. Results show that the prediction utility R is dominated by SC.

Using a linear stochastic dynamical system, SC was examined further and found to be intrinsically related to the leading eigenmode amplitude of the initial conditions. This finding was validated by actual model prediction results and is also consistent with other recent work. The relationship between R and SC has particular practical significance for ENSO predictability studies, since it provides an inexpensive and robust method for exploring forecast uncertainties without the need for costly ensemble runs.

1. Introduction

Generally, there are two kinds of climate predictability studies associated with different sources of prediction errors. The first addresses how uncertainties in an initial state of the climate system affect the prediction of a later state, whereas the second addresses how the growth of the parameterization errors, including the uncertainties of boundary conditions, evolves in a dynamical system. ENSO prediction is an initial value problem, and the future evolution of the system depends critically on the initial state from which it started, so initial-condition errors have a large impact on model skill and the growth of forecast errors. The first kind of predictability has attracted a lot of attention (e.g., Moore and Kleeman 1998; Kleeman and Moore 1999; Chen et al. 1997; Xue et al. 1997). This is particularly interesting from a practical point of view since certain types of ocean states are known to be more predictable than others.

Predicting the first kind of forecast uncertainty is equivalent to solving the Liouville equation for the probability density function (pdf) of the climate state (Epstein 1969; Palmer 1999). However, it is impractical to solve such an equation due to the huge dimensionality of the climate system (e.g., 10^6 variables for a typical climate model) and because the initial pdf is generally unknown. A practical solution is to approximately approach the pdf using a finite number of ensembles by a specific technique (Kleeman and Majda 2005). Especially, for a Gaussian or approximate Gaussian process, ensemble prediction could generate a good approximation for the pdf.

An important issue in the study of predictability is to seek a measure of forecast uncertainty. Typically, there are two kinds of measures widely used in the study of ENSO predictability. One is ensemble spread, and Moore and Kleeman (1998) have shown that, when spread is small, skill is invariably good whereas, when it is large, skill is much more variable. A similar relationship has also been noted in ensemble numerical weather prediction (Buizza and Palmer 1998) and in other ENSO models (e.g., Xue et al. 1997). On the other hand, Kirtman and Shukla (1998) have shown that ensemble spread is not a good indicator of skill, and both Buizza and Palmer (1998) and Moore and

Corresponding author address: Dr. Youmin Tang, Environmental Science and Engineering, 3333 University Way, Prince George, University of Northern British Columbia, V2N 4Z9, Canada.
E-mail: ytang@CIMS.nyu.edu

Kleeman (1998) note that such relationships can be norm dependent. An alternate criterion that has been used for determining forecast skill is the leading eigenmode amplitude (signal size) of the forecast initial conditions (Kleeman and Moore 1999, hereafter KM99). KM99 showed that periods in which the long period, slowly decaying normal modes of the dynamical system are present with large amplitude should be intrinsically more predictable because such modes are able to resist dissipation by the more chaotic or stochastic components of the system.

Using information theory, Kleeman (2002) recently proposed a general framework to explain why the two reliability measures discussed above are central to predictability studies. With simple conceptual models, he found that the prediction utility, defined by the relative entropy of the prediction and climatological pdfs, is a good measure of the reliability of predictions. When the pdfs are Gaussian, the utility consists of two components, one is the dispersion component associated with the ensemble spread and the other is the signal component related to the leading eigenmode amplitudes present in the initial conditions.

In this paper, we will use realistic ENSO models to further explore and examine the notion of the utility and to attempt to answer two central questions: 1) What are appropriate measures of the reliability of ENSO dynamical predictions and 2) what are the dominant precursors that control variations in reliability? This paper is structured as follows: Section 2 briefly describes the models and initialization scheme used. Section 3 introduces the ideas central to prediction utility and proposes a practical algorithm for its computation. Section 4 describes the strategy and methodology used in the ensemble experiments, while a reduced space in which the utility is evaluated is discussed in section 5. Utility analyses for two coupled models are presented in sections 6 and 7. A summary and discussion are given in section 8.

2. The coupled models and initialization scheme

a. The coupled models

Two hybrid coupled models (HCMs), an ocean general circulation model (OGCM) coupled to a statistical atmosphere (hereafter HCM1) and the same ocean model coupled to a dynamical atmospheric model of intermediate complexity (hereafter HCM2), were used for this study. The different atmospheric components in the two HCMs allow us to examine a theoretical framework to measure reliability of ENSO predictions in more general terms and to confirm the robustness of results across model formulations.

The ocean model used is based on the Océan Parallélisé (OPA) version 8.1 (Madec et al. 1998), a primitive equation OGCM. The model uses an Arakawa C grid and was configured for the tropical Pacific Ocean

between 30°N–30°S, 120°E–75°W. The horizontal resolution in the zonal direction is 1°, while the resolution in the meridional direction is 0.5° within 5° of the equator, smoothly increasing to 2.0° at 30°N and 30°S. There were 25 vertical levels with 17 concentrated in the top 250 m of the ocean. The time step of integration was 1.5 h and all boundaries were closed, with no slip conditions. A turbulent closure hypothesis was used to parameterize subgrid-scale physical processes where small-scale horizontal and vertical transports are evaluated in terms of diffusion coefficients and derivatives of the large-scale flow as described by Blanke and Delcluse (1993). The detailed formulation and configuration of the ocean model and its performance in simulating the tropical Pacific can be found in Vialard et al. (2002).

The statistical atmospheric model is a linear model identical to that of Barnett et al. (1993) and Tang et al. (2004), which predicts the contemporaneous surface wind stress anomalies from sea surface temperature anomalies (SSTA). The seasonal variations of the responses of wind stress to SST were also included so that for each month there is essentially a different atmospheric model. The model was trained using National Centers for Environmental Prediction (NCEP) atmospheric reanalysis wind products and Reynolds–Smith SST observations (Smith et al. 1996) from 1951 to 1980. Therefore, the ensemble experiments performed for the period 1981–98 in the next sections are completely independent of the construction of the atmospheric model. This strategy eliminates any artificial skill when evaluating the hindcast skills.

The dynamical atmospheric model consists of a Gill-type steady-state model, which has been used for routine ENSO prediction and for the study of climate predictability, developed by Kleeman (1991, referred to as the Kleeman model hereafter). The model computes global anomalies relative to the observed seasonal cycle of surface wind and mean atmospheric wind at 850 mb. When the Kleeman model was coupled to the OGCM, the OGCM provides SST anomalies to the atmospheric model. The atmosphere is heated by Newtonian cooling/relaxation to the SST anomaly and by latent heating due to deep penetrative convection via a simple moist static energy-dependent convection scheme. In both coupled models, the OGCM was forced by the sum of the associated wind anomalies computed by the atmospheric model and the observed monthly mean climatological winds.

b. The initialization scheme

A very important task in ENSO prediction is to determine the oceanic initial conditions. It has been found that initialization with subsurface in situ temperature observations can significantly improve ENSO prediction skill (e.g., Ji et al. 2000; Segschneider et al. 2001; Tang and Hsieh 2003). However, due to relatively spatially sparse and temporally sporadic subsurface in situ

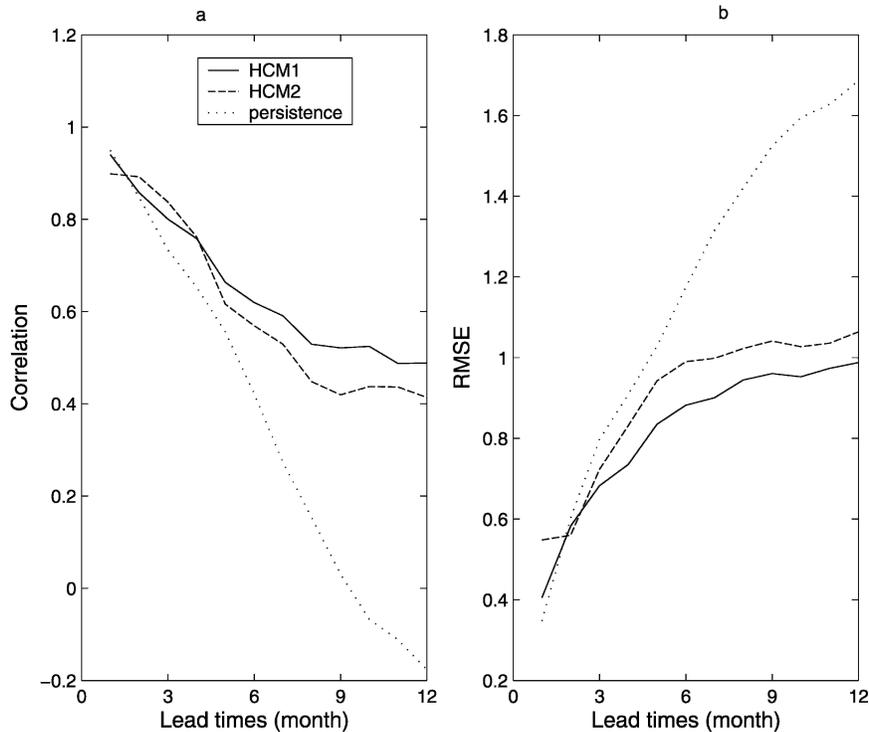


FIG. 1. (a) Prediction skill based on anomaly correlation and (b) rmse of Niño-3 SSTA for HCM1 and HCM2, and persistence. The prediction skill is evaluated for the single realization without the noise forcing.

observations, a great deal of effort is required to process and assimilate the data in models. For simplicity, we use the NCEP ocean analysis dataset to initialize our prediction models. Compared with the sparse and sporadic real observations, existing analysis products are easier and more convenient to use since they are regular gridded datasets. The possibility of initializing ENSO prediction models by assimilating NCEP subsurface temperature analyses via a 3DVAR algorithm has been explored in details by Tang et al. (2003, 2004). The results show that the NCEP analysis product can effectively improve the prediction of Niño-3 (5°N – 5°S , 150° – 90°W) SSTA at all lead times up to 12 months (in particular for lead times over 4 to 6 months). The oceanic analysis from the assimilation with existing NCEP analysis products can be as good as those generated by directly assimilating subsurface in situ temperature observations.

Using an assimilation scheme identical to that of Tang et al. (2003), we obtained oceanic analyses for the period 1981–98. Figure 1 shows the hindcast skills for HCM1 and HCM2 initialized with the oceanic analyses for period from 1981 to 1998. As can be seen, both HCM1 and HCM2 have a reasonable prediction skill for Niño-3 SSTA, especially for prediction lead times within 6 months, the time scale of interest in this study. This allows us to apply both models for further study on ENSO predictability.

3. Prediction utility and ensemble prediction

a. Relative entropy and prediction utility

In a dynamical system there is a certain amount of a priori information available before an ensemble (or statistical) prediction is made. This consists of historical information about the system such as climatological mean, standard deviation, and so on. A prediction further augments this prior information, and the additional information provided by the prediction constitutes what we should refer to as the utility of the prediction. Information theory (Cover and Thomas 1991) allows us to precisely quantify this additional information using a function known as relative entropy that, crudely speaking, measures the distance between the prediction pdf and the climatological pdf. If a discrete set of states is being predicted, the relative entropy, R , is given by

$$R = \sum_i p_i \ln \frac{p_i}{q_i}, \quad (1)$$

where q_i is the climatological distribution and p_i is that for the prediction.

In the case where the pdfs are Gaussian, which is approximately the case in many practical cases (including ENSO prediction), the relative entropy may be cal-

culated exactly in terms of the predictive and climatological variances and the difference between their means. The resulting analytical expression for the relative entropy (or utility as we have interpreted it) is given by (Kleeman 2002):

$$R = \frac{1}{2} \left[\ln \left(\frac{\det(\boldsymbol{\sigma}_q^2)}{\det(\boldsymbol{\sigma}_p^2)} \right) + \text{tr}(\boldsymbol{\sigma}_p^2(\boldsymbol{\sigma}_q^2)^{-1}) + (\boldsymbol{\mu}^p - \boldsymbol{\mu}^q)^T (\boldsymbol{\sigma}_q^2)^{-1} (\boldsymbol{\mu}^p - \boldsymbol{\mu}^q) - n \right]. \quad (2)$$

Here, the $\boldsymbol{\sigma}_q$ and $\boldsymbol{\sigma}_p$ are the climatological and predictive covariance matrices while $\boldsymbol{\mu}^q$ and $\boldsymbol{\mu}^p$ are the climatological and predictive mean vectors of the system; n is the number of degrees of freedom. From (2), we can deduce that R is composed of two components: (i) a reduction in climatological uncertainty by the prediction [the first two terms on the rhs of (2)] and (ii) a difference in the predictive and climatological means [the last two terms on the rhs of (2)]. These components can be interpreted respectively as the dispersion and signal components of the utility of a prediction (Kleeman 2002). A larger R indicates that more useful information is being supplied by the prediction, which could be interpreted as making it more reliable.

b. The stochastic optimal perturbations and ensemble prediction

To compute the utility R it is necessary to estimate the first and second momentums of the climatology and prediction pdfs. The estimates of the mean vector ($\boldsymbol{\mu}^q$) and variance matrix ($\boldsymbol{\sigma}_q$) of climatology are straightforward using a relatively long run of the coupled model. To estimate $\boldsymbol{\mu}^p$ and $\boldsymbol{\sigma}_p$ for a specific lead time, we can use ensemble prediction, which allows us to increase the prediction samples for the specific lead time when only a limited observation set is available. There are at present two typical methods used to perturb the initial conditions for constructing ensemble forecasts, breeding vectors and singular vectors (Toth and Kalnay 1993; Molteni and Palmer 1993). A recent comparison between the two methods has revealed that bred-vector ensembles provide an average error distribution more similar to a Monte Carlo distribution, while the singular-vector ensemble provides a more reliable estimate of the upper bound on error growth (Trevisan et al. 2001). Studies of ENSO predictability have so far mainly considered singular vectors.

While singular vectors can measure the fast error growth associate with uncertainties in the initial conditions, these fail to consider the influence of stochastic processes on predictions. Such stochastic processes are not described by our hybrid coupled models in which a steady-state atmosphere is used. However the impact of stochastic processes such as the Madden–Julian oscillation and westerly wind bursts on ENSO may be signifi-

cant (Zavala-Garay et al. 2003). Stochastic processes therefore cannot be ignored when forecasting the real coupled system.

Farrell et al. (1993) and Kleeman and Moore (1997) introduced the idea of stochastic optimals (SOs) to represent the effect of stochastic processes on prediction errors. For white noise in time, the stochastic optimals are defined by the eigenvectors of the operator S

$$S = \int_0^T \mathbf{A}^*(t, 0) \mathbf{U} \mathbf{A}(t, 0) dt. \quad (3)$$

Here T is the forecast interval of interest and is assumed to be 12 months in this study, $\mathbf{A}(t, 0)$ is the forward tangent propagator of the linearized dynamical model that advances the state vector of the system from time 0 to time t , $\mathbf{A}^*(t, 0)$ is the adjoint of $\mathbf{A}(t, 0)$, and the matrix \mathbf{U} defines the norm of interest. In this study, we use a seminorm defined as the square of the Niño-3 SSTAs.

In the present study, the SOs are taken to represent uncertainties associated with stochastic events in the coupled ocean–atmosphere system that can be amplified by the dynamical model during the forecast interval T , which in turn leads to forecast error growth. A detailed description of the stochastic optimals used in the present study can be found in Moore et al. (2005, manuscript submitted to *J. Climate*, hereafter M05).

It is worth noting that the traditional singular vectors are defined by the eigenvectors of the operator $\mathbf{A}^*(t, 0) \mathbf{U} \mathbf{A}(t, 0)$. The methodology described in Kleeman and Moore (1997) and M05 thus enables us to compute the SOs for all times up to T at the same computational cost as the singular vectors.

After computing the SOs, a series of new model states, denoted \mathbf{X} , were obtained by perturbing the model with a stochastic forcing composed of the leading SOs, that is,

$$\frac{d\mathbf{X}_i}{dt} = \mathbf{N}(\mathbf{X}_i) + \alpha \mathbf{Q} \mathbf{T}_i(t). \quad (4)$$

Here $\mathbf{X}_i(t)$ are the model state vectors at time t ; \mathbf{N} is the coupled model nonlinear operator, \mathbf{Q} is the matrix of the leading SOs with unit variance; $\mathbf{T}_i(t)$ is a random red noise with the unit variance, and α is a dimensional factor. Since the variances of \mathbf{Q} and \mathbf{T} are set to unity, α actually represents the variance of stochastic forcing, i denotes a different random red noise as i changes from 1, 2, . . . to M , where M represents ensemble size.

There are several important issues in the construction of the stochastic forcing $f_i = \alpha \mathbf{Q} \mathbf{T}_i(t)$. One is the choice of the forcing fields that will be perturbed and the second is the variance assigned to f_i . For the former, we consider heat flux and wind stress, as both dominate the coupling behaviors of coupled models and play a crucial role in controlling forecast error growth (Moore and Kleeman 1998, 1999). For the second issue, there are several ways to extract the stochastic components

from observations as discussed in Kleeman and Moore (1997). Here, the variance of f_i is estimated from high-pass filtered estimates of heat flux and wind stress from NCEP–National Center for Atmospheric Research (NCAR) reanalysis. This is because the noise is relatively much stronger in high-frequency components than in low-frequency components although it could occur on all time and space scales. In fact, we find that the noise variance of f_i is only little sensitive to our predictability results that will be presented in following discussion (see section 7).

Figure 2 shows the variance of f defined in this way for wind stress and heat flux in 1997. As shown in the figure, their variances vary spatially with large amplitudes in subtropical region and small amplitudes in the region along the equator, indicative of strong stochastic forcing in the subtropics associated with synoptic variability. Over the equatorial Pacific Ocean, the stochastic forcing is relatively small since large amplitude interannual variability dominates in this region.

Since our interest here is ENSO, we only focus on the equatorial region. As shown in Fig. 2, the variance of f_i in this region is relatively small. Based on Fig. 2, we choose the variance of f_i to be $0.02 \text{ N}^2 \text{ m}^{-4}$ for wind stress and 30 W m^{-2} for heat flux. These estimates are also consistent with those of Zavala-Garay et al. (2003). It should be noted that the variance in Fig. 2 changes little in time.

Another issue relative to (4) is the number of SOs that should be used. Two factors will be explored here for determining the truncation, that is, the variance explained by each SO and the sensitivity of experiments to the number of SOs. For HCM1, the first two SO modes account for over 90% of the variability that would result in the coupled model (M05). In addition, M05 show that the first SO of each model is present in estimates of f_i based on NCEP reanalysis data with a significant amplitude. Sensitivity experiments indicated that after the first two SOs, additional SOs have a rather small influence on forecast error growth. Thus, the first five SOs were used to obtain f_i for HCM1. However for HCM2, the first ten SOs are required for the same purpose.

The procedure for ensemble generation for the period 1981 to 1998 can be summarized as follows:

- (i) NCEP subsurface analysis temperatures were assimilated in the OGCM using a 3DVar algorithm to generate initial conditions for each ensemble member.
- (ii) Hindcasts of 12 months duration were performed using HCM1 and HCM2, and the forecast trajectories were saved for each start date (i.e., 1 January, 1 April, 1 July, and 1 October of each year).
- (iii) The SOs of each hindcast trajectory in (ii) were computed using the tangent linear and adjoint operator of HCM1 and HCM2 according to (3).
- (iv) Each ensemble member was generated by rerun-

ning the hindcasts with a stochastic forcing term added to the wind stress and heat flux forcing with a different red-noise time series $\mathbf{T}_i(t)$ in (4). In this study, the ensemble size M was chosen to be 31.

4. Reduced space

Using the ensemble hindcasts, it is straightforward to calculate the prediction utility R for each start date using (2). However, the large state dimension of each coupled model somewhat complicates this calculation because of the computational cost of computing the covariance matrix σ_q^2 in (2) and its determinant. Therefore, for practical reasons, a reduced state space will be used that represents the ENSO characteristics of each model.

Obtaining a suitable reduced space from the original model space is an interesting question in its own right (Kleeman et al. 2003; M05). The basic procedure is to project the original space onto a set of specific spatial patterns that represent the most significant characteristics of interest in the original space. The simplest method is a regional average (e.g., Niño-3 index) to reduce the original space to one 1D index. Another method widely used is EOF analysis by which the original space can be reduced to a few leading principal components. In this study, we will use principal oscillation patterns (POPs; Hasselmann 1988) to obtain a reduced space for (2). In contrast to EOFs, which describe the stationary patterns that account for differing fractions of the variance, POPs describe the oscillatory behavior of the field since they actually represent eigenmodes of a filtered linear stochastic process (Xu and von Storch 1990). As such, the POPs better characterize the behaviors of ENSO and its physics than do traditional EOFs.

As the propagator matrix (\mathbf{L}) of a linear system (e.g., the tangent linear operator of the original model) used to derive POPs is usually asymmetric, the POPs do not form a set of orthogonal patterns, so the POPs coefficients characterizing the reduced space are not given as the dot product of the patterns with the original field. This complicates the calculation of obtaining the reduced space. One effective method is to calculate adjoint POPs (APOPs), which are the eigenmodes of the adjoint matrix of \mathbf{L} . By definition, the APOPs and conventional POPs form a biorthogonal set so that the reduced space can be obtained by the dot product of the APOPs with the original field.

The POPs and APOPs were obtained as described by Penland and Sardeshmukh (1995). We choose heat content anomalies (HCA)¹ to explore the prediction utility of the dynamical models because it is the primary

¹ Heat content is defined here as the integral of the temperature over the upper 250 m, calculated from $\text{HC} = (\sum_i h_i T_i / \sum_i h_i)$ where h_i and T_i are the thickness and temperature of level i , respectively.

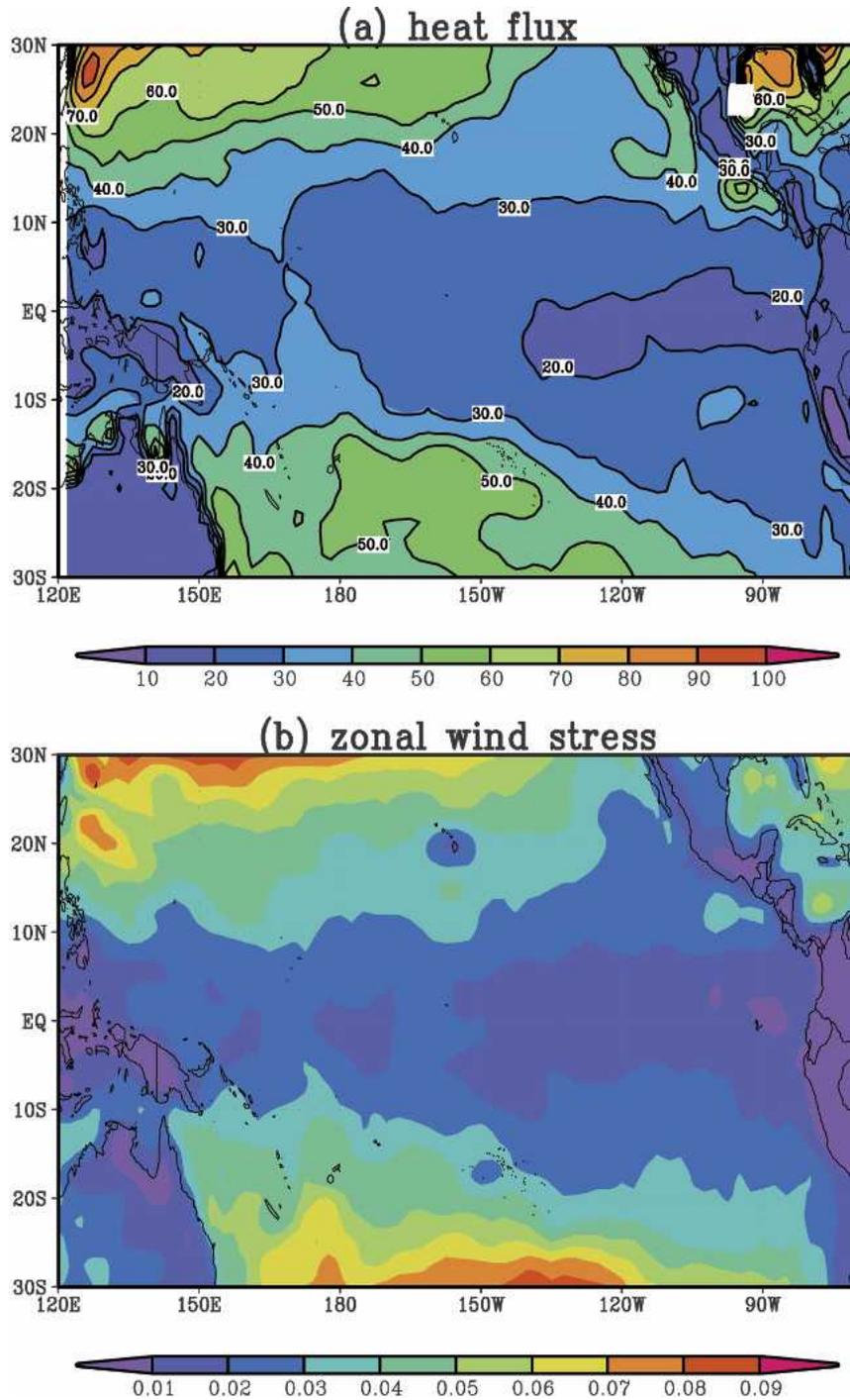


FIG. 2. Noise variance of heat flux and zonal wind stress for 1997, derived from NCEP reanalysis wind after the low-frequency signals with periods of 7 days and longer have been removed.

source of memory for the coupled system and is important for ENSO dynamics. Fluctuations in HCA are both systematic and significant in the evolution of ENSO (Tang and Hsieh 2003) and are thus an effective precursor of ENSO prediction. In addition, the HCAs are

also coincident anomalous features of the sea surface height and thermocline.

Figure 3 shows the real and imaginary components of the dominant POP and APOP (mode 1) of HCA, computed from the OGCM forced with observed wind

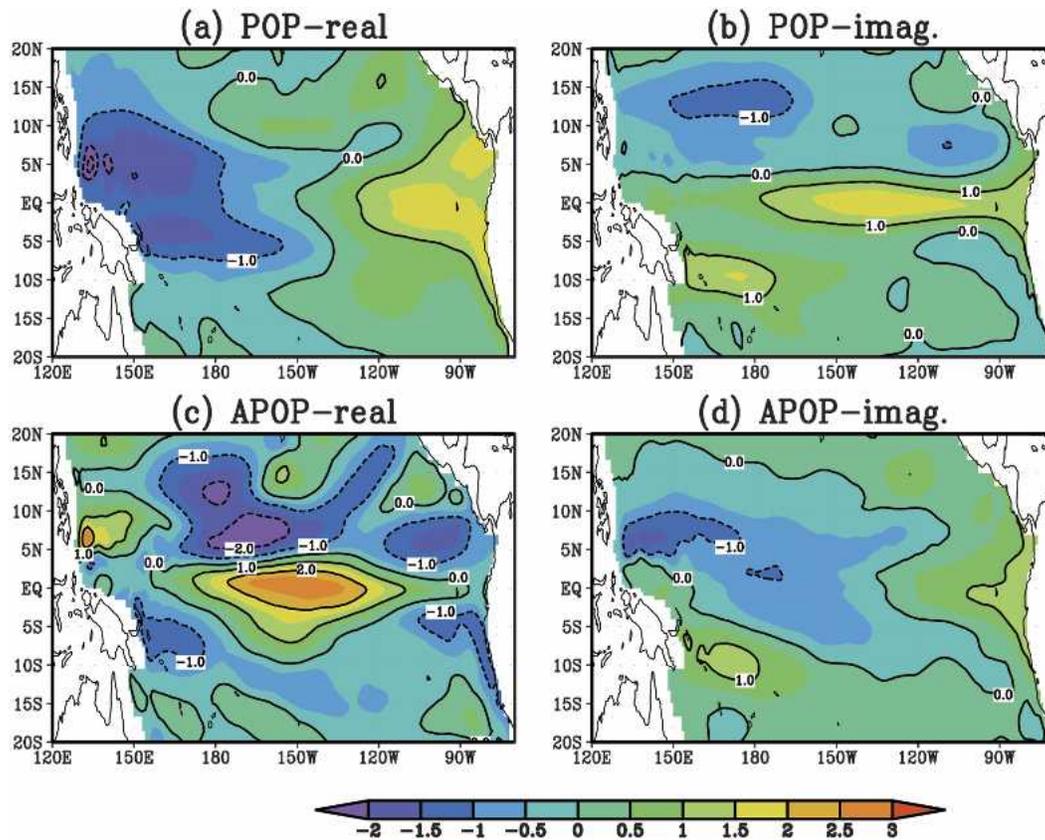


FIG. 3. Real and imaginary components of the first POP and APOP mode for HCA. Contour values are scaled by a factor of 10.

stress from 1961 to 1998. The dominant POPs and APOPs identify the mode having a period of 28 months and the decay time scale of 20 months relative to the period. As shown in Fig. 3, the characteristics of the POPs are very similar to those found in KM99: 1) Fig. 3a has a dipole zonal structure involving a western Pacific Rossby wave-like response of one sign and an eastern Pacific Kelvin wave-like response of the opposite sign; 2) Fig. 3b has a large amplitude signal of the one sign located mainly in the equatorial central/eastern Pacific. These patterns agree with the idea of a heat content buildup prior to El Niño as postulated by Wyrski (1975) and Jin (1997). Based on the traditional interpretation of POPs (Tang 2002), Figs. 3a,b are consistent with the delayed-action oscillator mechanism of Battisti (1988). They are also very similar to in structure the first two EOFs, which explain 80% of the heat content variability of the observations (not shown). Thus the reduced state space is effective in capturing a large amount of pointwise variance of the system.

Figures 3c,d show the spatial patterns of the APOP, which are used to obtain the reduced space. Compared to the POPs, the real component of the APOP resembles the imaginary part of POP mode, whereas the imaginary part of APOP is similar to the real part of the POP.

5. Prediction utility of the coupled models

a. The relation of prediction utility to prediction skill

Since (2) was derived based on a Gaussian assumption, we first examine the validity of this assumption prior to calculating the prediction utility. Figure 4 is an estimate of the pdf for Niño-3 HCA index predicted by HCM1 at lead times of 6 and 9 months. The pdf was obtained using ensemble predictions of 100 members with a randomly chosen initial condition. The first five SOs are used to perturb wind stress and heat flux forcings as described in section 3b. Figure 4 clearly indicates that the Gaussian assumption roughly holds for both cases. An examination of other variables such as SST and HCM2, produced similar results (not shown). However, as the lead time increases to 9 months, the Gaussian assumption no longer effectively holds and the pdfs become somewhat bimodal.

Displayed in Fig. 5 are the variations of prediction utility R for HCM1 and HCM2 during the period from 1981 to 1998 as a function of lead time and initial time. It is apparent that large prediction utility mainly resides in a few predictions such as those of the 1982–83 and 1997–98 ENSO events. For many other predictions, R is

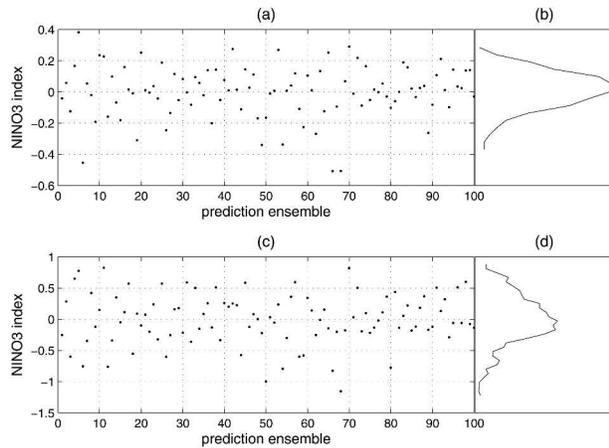


FIG. 4. (a) Niño-3 HCA index predicted by HCM1, as a function of prediction ensemble and (b) its pdf at a lead time of 6 months; (c), (d) As in (a) and (b) but for a lead time of 9 months.

small and exhibits significantly less variations with lead time. Since R measures the amount of extra information that resides in each prediction, a large value of R suggests a more informative and accurate prediction, whereas a small value of R often accompanies poor predictions. Figure 6 shows several predictions from HCM1, with both large and small values of R . It is obvious from Figs. 5 and 6 that the predictions with a large R are much better than those with a small R , compared with observations. The same is true for HCM2 (not shown).

To further explore the relation between prediction utility and prediction skill, we examine the contribution of each prediction to correlation skill r , traditionally defined as

$$r(t) = \frac{1}{N} \sum_i T_i^p(t) T_i^o(t), \quad (5)$$

where T denotes the normalized Niño-3 SSTA index with zero mean, p is for prediction, and o for observa-

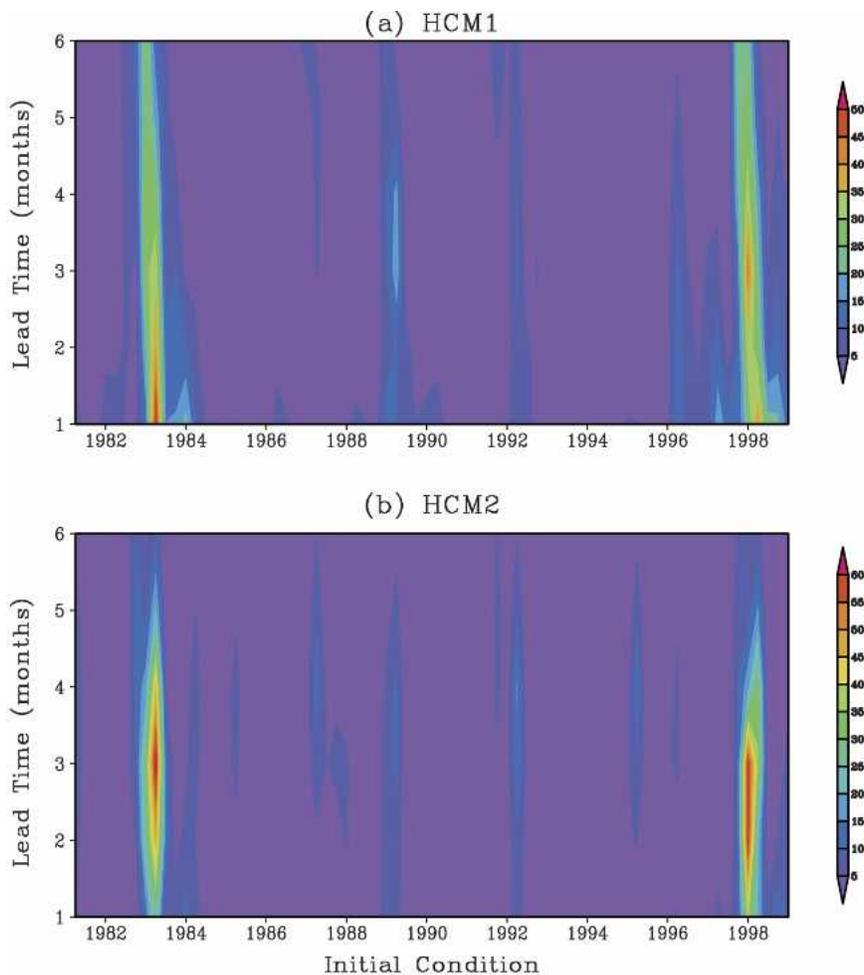


FIG. 5. Heat content (HC) prediction utility for HCM1 and HCM2 in POP reduced space.

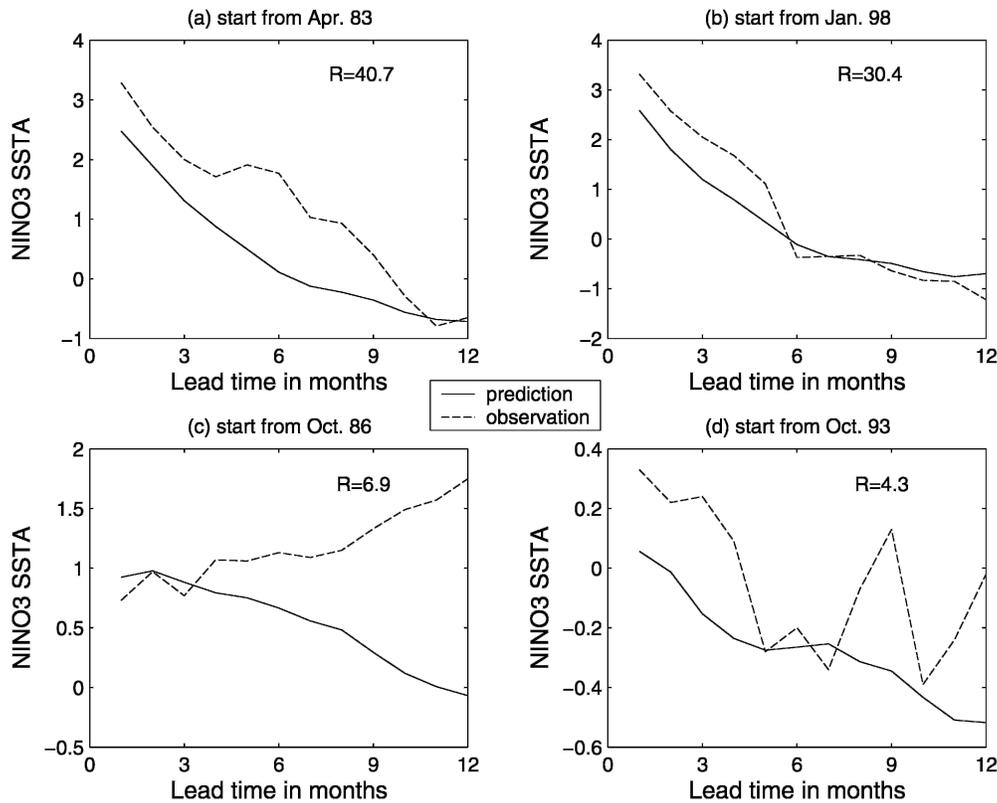


FIG. 6. HCM1 predicted and observed Niño-3 SSTA for several prediction cases. The value of R is for a lead time of 3 months. The prediction skill is evaluated for the single realization without the noise forcing.

tion; t denotes the lead time of the prediction and N is the number of samples used to calculate r .

The contribution of each prediction to $r(t)$, denoted as C , can be measured by

$$C_i(t) = \frac{N^{-1}T_i^p(t)T_i^o(t)}{r(t)} \times 100\%. \quad (6)$$

Figure 7 shows variations of C with lead time and initial time for HCM1 and HCM2. A striking feature shown in Fig. 7 is that there is a large variation of C with initial conditions. While some initial conditions lead to good predictions that account for significant contributions to r , most initial conditions correspond with a very small C . On the other hand, the variation of C with lead time is small, which seems reminiscent of the fact that the initial conditions play a critical role in ENSO prediction skill for all lead times.

Comparing Fig. 7 with Fig. 5 reveals that a large C generally corresponds to a large R . This is particularly true for HCM1, and for several typical ENSO events. For example, the utility R is far larger in the predictions initialized in 1983 and 1997–98 than at other times. Correspondingly, the accumulated contributions C to $r(t)$ from these predictions exceeds 30%. The relationship between R and C is further demonstrated in Figs. 8 and 9, which compare R and C for both models at two typi-

cal leading times (3 and 6 months). For HCM1, the correlation coefficient between R and C is 0.80 and 0.68 respectively at lead times of 3 and 6 months. The correlation between R and C is smaller in HCM2 than in HCM1. As shown in Fig. 9, R at 6-month lead time is significantly smaller than that at a 3-month lead time; in particular R is smaller in 1997–98 than in 1983. These are very different compared to HCM1 shown in Fig. 8, and seem unrealistic. It is not very clear why HCM2 displays these features that are absent in HCM1. One probable reason is that the POP reduced space used might be more suitable for HCM1. However, when the precise eigenmodes that were derived from HCM2's tangent-linear adjoint model are used, R still displays similar features (not shown). Nevertheless, the correlation coefficients still reach 0.71 and 0.53 at 3- and 6-month lead time for HCM2, which far exceed statistical significance. On the other hand, C in Figs. 9a,b also displays some features similar to Figs. 9c,d. For example, C at 6-month lead time is also slightly smaller in 1997–98 than in 1983. In this sense, the variations of R and C show rough consistency.

As demonstrated in Figs. 5 and 7, the large R and C values are mainly associated with predictions of large amplitude ENSO events. For most predictions, R and C are relatively small. This implies that the prediction

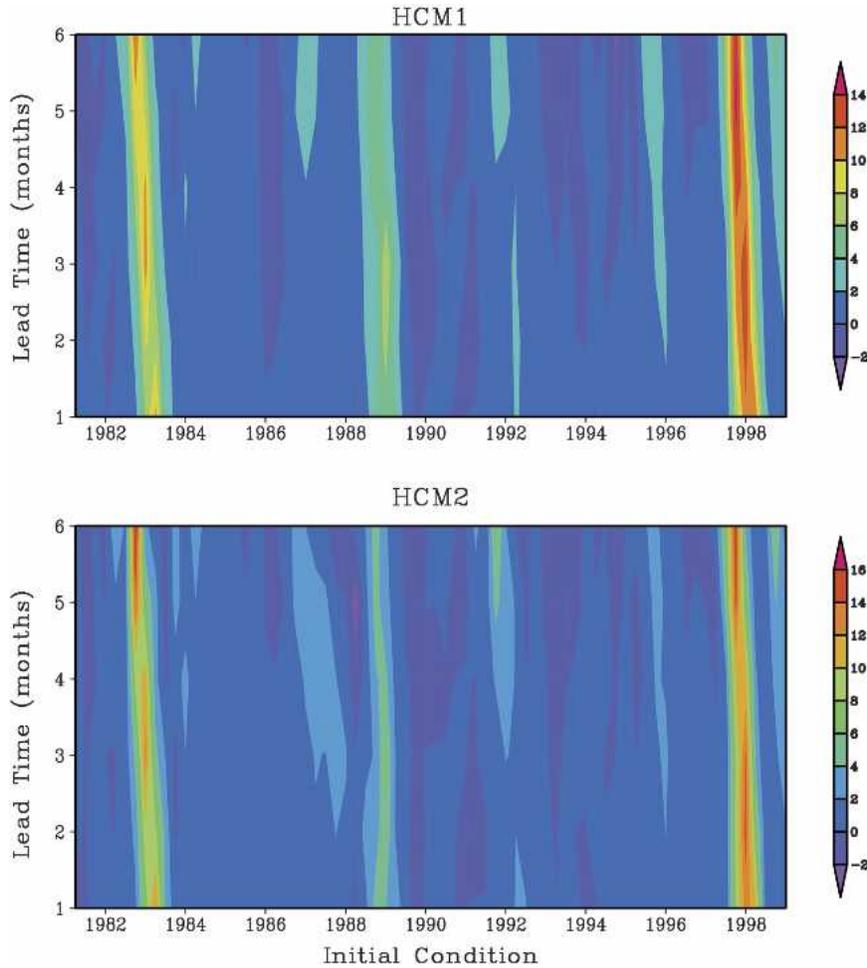


FIG. 7. Contribution of each prediction to correlation skill for (a) HCM1 and (b) HCM2.

skills displayed in Fig. 1 may be mainly due to a few predictions. To explore this further, we recalculated the prediction skills for HCM1 and HCM2 after the 13 predictions with large R were removed. This is shown in Fig. 10. As can be seen, the prediction skill decreases dramatically and is even worse than persistence when the large R predictions are excluded. The vertical bar in Fig. 10 is an estimate of the correlation error bar after randomly removing any 13 predictions and was obtained by a bootstrap method (Tang et al. 2003). Obviously, the difference in the correlation skill shown in Fig. 10 significantly exceeds the correlation error due to the uncertainty of the finite sample size. This clearly indicates the critical importance of these predictions to model prediction skill.

b. The relations between prediction utility and the signal and dispersion components

In the previous subsection, we explored the relation between the prediction utility and model prediction

skill. Our results show that prediction utility R is a good indicator for prediction reliability. When R is large, the prediction is typically good, whereas when R is small, the prediction is often relatively poor. In this subsection, we will examine what determines variations in R .

The first two terms on the rhs of (2) are determined by the climatological variance and predictive variance. Since variance of climatology is time invariant, these terms represent a measurement of the dispersion or spread of the ensemble. The third term on the rhs of (2) is governed by the amplitude of the predicted mean field (L-2 norm), measuring the contribution of the predicted signal size to R . Kleeman (2002) refers to the first two terms minus n as the dispersion component (DC) and the third term as the signal component (SC). Therefore, we have $R = DC + SC$.

Figures 11 and 12 show the variation of SC and DC as a function of lead time and initial time for the period from 1981 to 1998 for HCM1 and HCM2. Both figures reveal that SC is significantly larger than DC, and that SC decreases with lead times. It is straightforward to

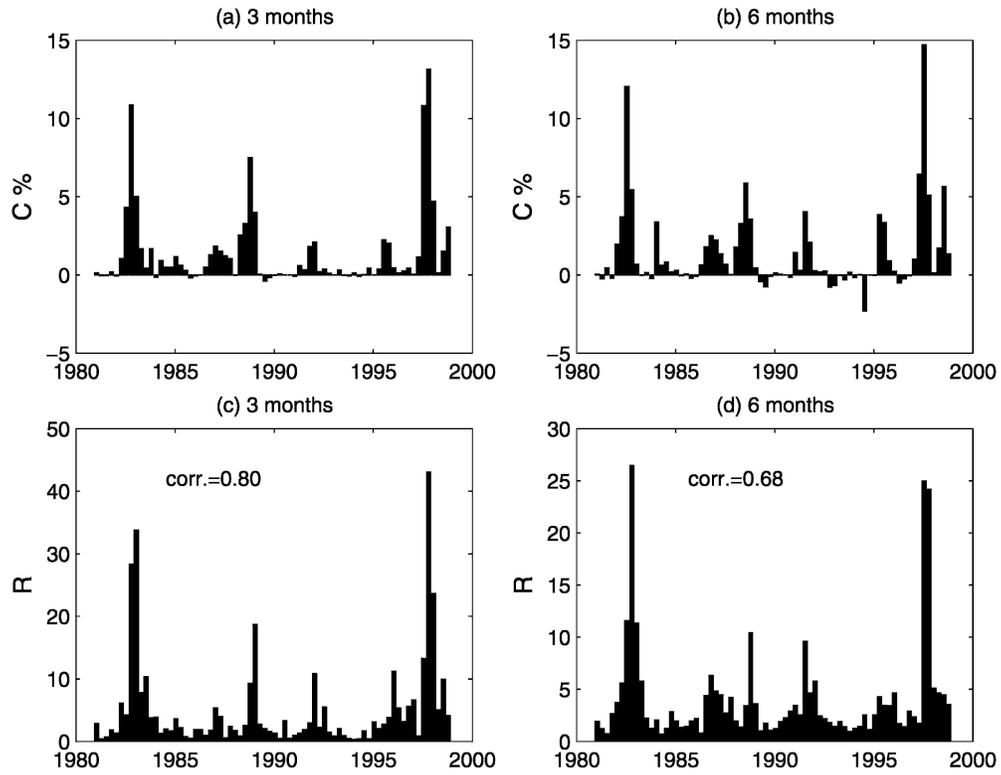


FIG. 8. Prediction utility R and the contribution of each prediction to correlation skill C for HCM1 at lead times of 3 and 6 months.

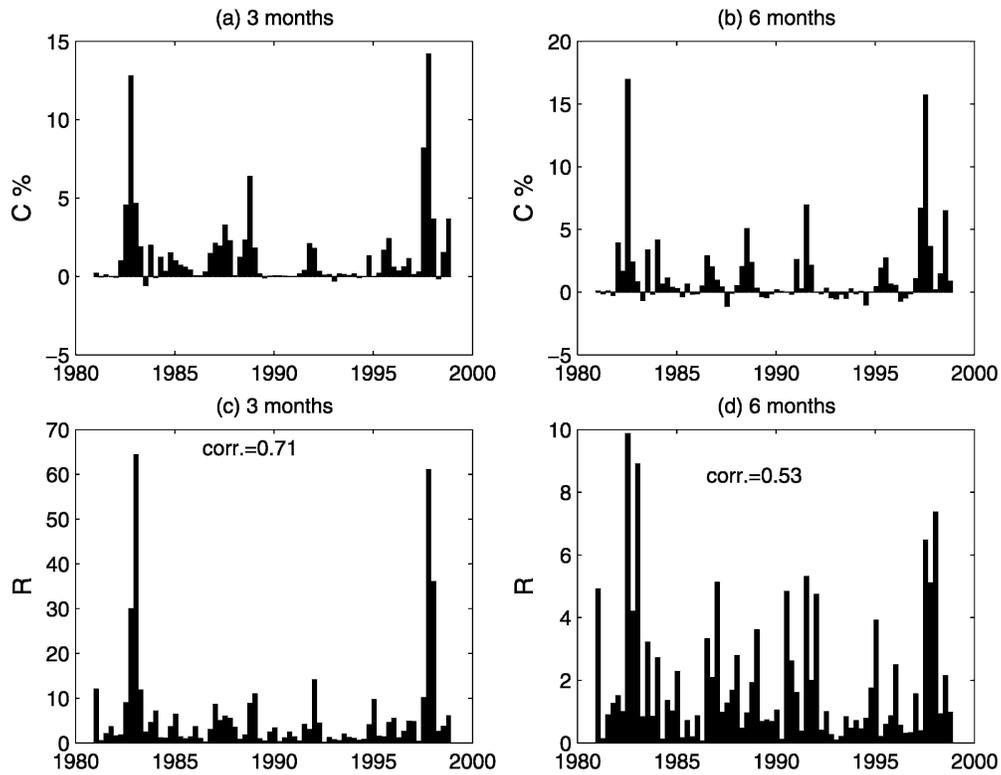


FIG. 9. As in Fig. 8 but for HCM2.

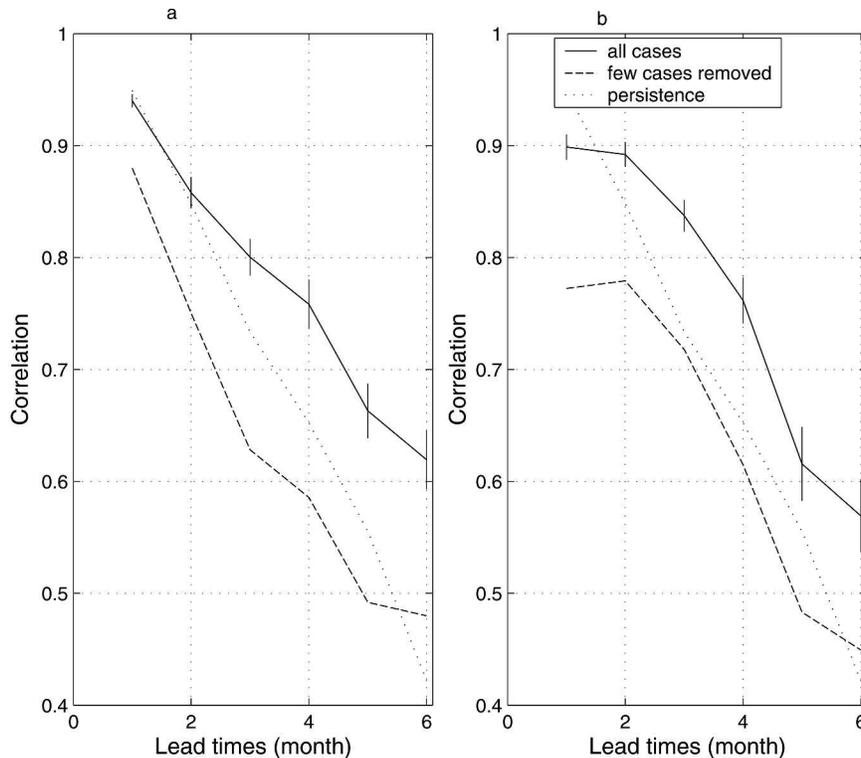


FIG. 10. Correlation skills for (a) HCM1 and (b) HCM2. The predictions are initialized from Jan 1981 to Oct 1998, three months apart. The solid line is the skill calculated using all prediction samples; the dashed line is the skill after removing the prediction cases Oct 1982, Jan 1983, Apr 1983, Jul 1988, Oct 1988, Jan 1989, Apr 1989, Jan 1997, Apr 1997, Jul 1997, Oct 1997, Jan 1998, and Apr 1998; dotted line is persistence skill. The vertical bar represents the error bar after randomly removing any 13 predictions.

understand the variation of SC with lead time since the system dissipates initial signals and reduces oscillation amplitudes at long lead times, leading to a small difference between the ensemble mean and the climatology. It is physically equivalent to the fact that the useful information contained in initial states will be gradually dissipated by stochastic processes with increasing lead time. In the next section, we will clearly show that SC is governed by the prediction initial state.

Unlike SC, the differences of DC among different initial conditions are of little significant for short lead time of 1 to 2 months. This is especially obvious for HCM2 as shown in Fig. 12: DC decreases with lead times of 1–2 months. This is because (i) the ensemble spread σ_p is usually relatively small initially so that the first item of DC dominates its variations [see Eq. (2)] and (ii) the σ_p increases during the period. After the initial period, in most cases, the first term and the second term balance each other so DC stabilizes with lead time. For lead times of 4 to 6 months, the second term of DC outweighs the first item for some cases due to large ensemble spread σ_p , resulting in an increase of DC. Figure 13 shows the variation of DC and its two components for HCM1 for a typical case, arbitrarily chosen in October 1986.

Comparing Fig. 5 with Figs. 11 and 12 reveals that R is dominated by SC and that the DC contribution is small. Figures 14 and 15 show the relation of SC and DC to R at the lead time of 3 and 6 months for both models, respectively. As can be seen, R and SC vary linearly with a slope of unity. The correlation coefficients between R and SC are all over 0.95. In contrast to the good relation between SC and R , however, the relation between DC and R is much less significant.

Overall, the prediction utility R is mainly determined by the signal component SC. When the predictive mean signals are large, R is also large, suggesting that such predictions are reliable.

It should be noted that the good relation between SC and C is little sensitive to the choice of the variance of noise. As discussed in section 3b, the variance of noise is based on the high-pass filtered estimates, which might overestimate the amplitude of noise occurring at low frequency components. In order to examine the sensitivity of the noise amplitude to the results, we carried out a new set of ensemble predictions for a randomly chosen period (1981–85). In the new ensemble predictions, everything is kept the same as in the original ones presented in this paper but the variance of

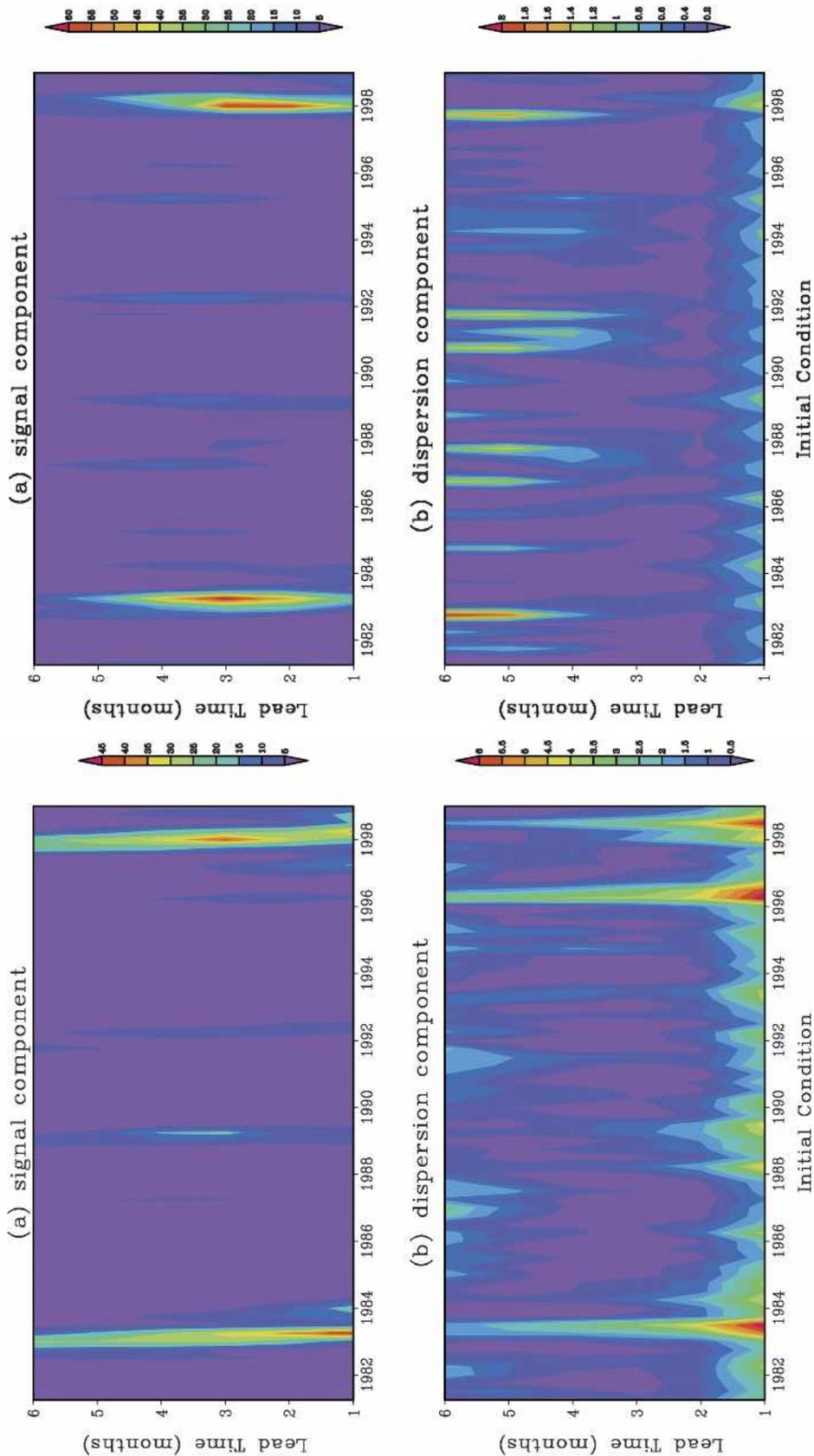


FIG. 12. As in Fig. 11 but for HCM2.

FIG. 11. Signal component (SC) and dispersion component (DC) of R for HCM1.

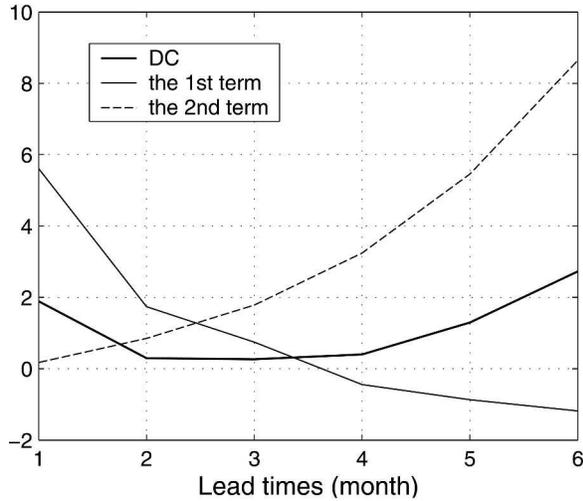


FIG. 13. The variation of DC and its two components with lead time for a typical case. The first term is $\ln[\det(\sigma_q^2)/\det(\sigma_p^2)]$ and the second term is $\text{istr}[(\sigma_p^2)(\sigma_q^{-2})]$.

stochastic forcing in Eq. (4) is doubled. The sensitivity experiments are shown in Fig. 16, indicating that the results presented above have very little sensitivity to the choice of the noise variance.

In the next section, we will find that SC is directly related to the eigenmode amplitudes present in the initial conditions. Using reasonable assumptions, KM99 proved that in theory the prediction correlation skills $r(t)$ can be measured by the predictive mean signals. This is consistent with the above result that DC has a small contribution to R . In KM99, the eigenmodes were represented by the leading POP modes. They argued that the periods during which slowly decaying eigenmodes are present with large amplitude should be intrinsically more predictable because such modes are able to resist dissipation by the more chaotic components of the system.

6. A simplified framework to measure prediction uncertainty

By applying relative entropy theory to two HCMs, we have shown that the prediction utility defined by (2) can effectively measure the reliability of ENSO predictions. However, evaluating (2) requires large ensembles of predictions, which greatly limits its application to more complex models, such as fully coupled GCMs. It is therefore in our best interest to explore more simple methods for estimating prediction utility. In this section, we will apply linear theory for this purpose.

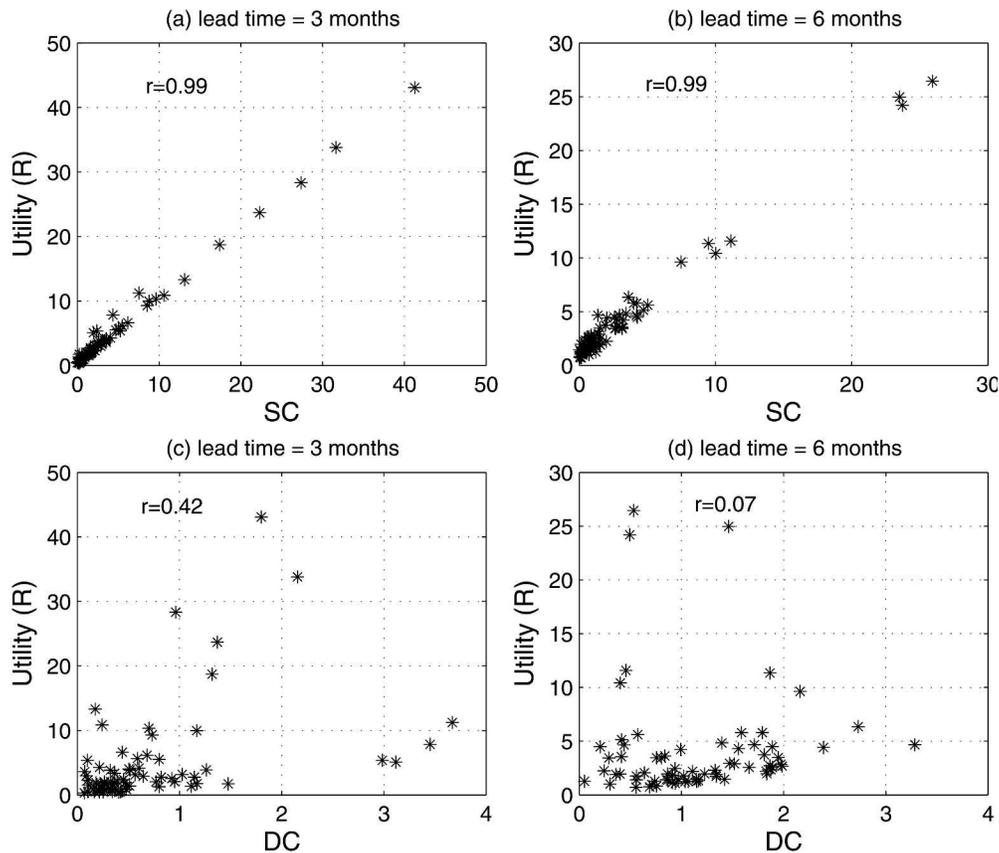


FIG. 14. The relation of SC and DC to R for HCM1 at several different lead times.

Many studies suggest that ENSO can be described as a linear, damped, stochastically forced system (e.g., KM99; Kleeman and Moore 1997; Moore and Kleeman 1999; Xue et al. 1994, 1997; Penland and Sardeshmukh 1995). Following these works, ENSO can be modeled by the following discrete, linear, stochastic equation,

$$\mathbf{X}_k = \Phi_{k-1}\mathbf{X}_{k-1} + w_{k-1}, \quad (7)$$

where \mathbf{X}_{k-1} is the system state vector at the time $k-1$, \mathbf{X}_k is its value at the time k , Φ_{k-1} is the state transition matrix for the system at the time k , w_{k-1} is a white noise, that is, $E\langle w_k \rangle = 0$, $E\langle w_{t_1} w_{t_2} \rangle = \delta(t_1 - t_2)\mathbf{Q}$ (\mathbf{Q} is a time invariant matrix), and $E\langle \cdot \rangle$ denotes the expectation operator.

The state transition matrix Φ_{k-1} is usually real and symmetric. Denoting the eigenvectors and eigenvalues of Φ_{k-1} by \mathbf{P}_{k-1} and Λ_{k-1} , we have $\mathbf{P}_{k-1}^{-1}\Phi_{k-1}\mathbf{P}_{k-1} = \Lambda_{k-1}$. Here Λ_{k-1} and \mathbf{P}_{k-1} may be complex and Λ_{k-1} is a diagonal matrix.

Multiplying (7) by \mathbf{P}_{k-1}^{-1} on both sides of Eq. (7) yields

$$\mathbf{P}_{k-1}^{-1}\mathbf{X}_k = (\mathbf{P}_{k-1}^{-1}\Phi_{k-1}\mathbf{P}_{k-1})\mathbf{P}_{k-1}^{-1}\mathbf{X}_{k-1} + \mathbf{P}_{k-1}^{-1}w_{k-1}. \quad (8)$$

We will denote a new white noise $\tilde{w}_{k-1} = \mathbf{P}_{k-1}^{-1}w_{k-1}$ and assume it has zero mean and white covariance, namely

$$E\langle \tilde{w}_{t_1} \tilde{w}_{t_2} \rangle = \delta(t_1 - t_2)\tilde{\mathbf{Q}}.$$

Equation (8) can then be written

$$\mathbf{P}_{k-1}^{-1}\mathbf{P}_k\mathbf{Z}_k = \Lambda_{k-1}\mathbf{Z}_{k-1} + \tilde{w}_{k-1}, \quad (9)$$

where $\mathbf{Z}_{k-1} = \mathbf{P}_{k-1}^{-1}\mathbf{X}_{k-1}$. For a large-scale and slowly varying climate system, such as ENSO, the difference between the transition matrix Φ at two adjacent times $k-1$ and k (and in particular the difference in the few leading eigenmodes of Φ at $k-1$ and k) \mathbf{P}_{k-1} and \mathbf{P}_k , will be small and is assumed to be negligible. Equation (9) is thus simplified:

$$\mathbf{Z}_k = \Lambda_{k-1}\mathbf{Z}_{k-1} + \tilde{w}_{k-1}. \quad (10)$$

Equation (10) describes the trajectory of the dynamical system (7) on its leading eigenmode space ($\mathbf{Z} = \mathbf{P}^{-1}\mathbf{X}$). From (10), an element z_k^j of vector \mathbf{Z}_k will satisfy

$$z_k^j = \lambda_{k-1}^j z_{k-1}^j + \tilde{w}_{k-1}^j. \quad (11)$$

Comparing with the reduced space discussed in section 5 and the classical POP analysis (von Storch and Zwiers 1999), one finds that (10), in fact, is a projection of the original dynamical system (7) on POP space, and (11) describes the variation of POP amplitude coefficients.

Therefore, we can evaluate the prediction utility for (10) using Eq. (2). The predictive mean vector $\boldsymbol{\mu}^p$ and covariance vector $\boldsymbol{\sigma}_p$ can be written as (see appendix)

$$\boldsymbol{\mu}_k^p = \Lambda_{k-1}\boldsymbol{\mu}_{k-1}^p \quad (12)$$

$$\boldsymbol{\sigma}_p(k) = \Lambda_{k-1}\boldsymbol{\sigma}_p(k-1)\Lambda_{k-1}^T + \tilde{\mathbf{Q}}. \quad (13)$$

We denote by $\boldsymbol{\mu}^p(0)$ and $\boldsymbol{\sigma}_p(0)$ the values of $\boldsymbol{\mu}^p$ and $\boldsymbol{\sigma}_p$ at the initial time so that

$$\boldsymbol{\mu}_k^p = \prod_{i=0}^{k-1} \Lambda_i \boldsymbol{\mu}^p(0) \quad (14)$$

$$\boldsymbol{\sigma}_p(k) = \prod_{i=0}^{k-1} \Lambda_i \boldsymbol{\sigma}_p(0) \Lambda_i^T + \Pi \tilde{\mathbf{Q}}, \quad (15)$$

where $\Pi = \prod_{i=0}^{k-1} \Lambda_i + \prod_{i=1}^{k-1} \Lambda_i + \dots + \prod_{i=k-1}^{k-1} \Lambda_i + \mathbf{I}$, and \mathbf{I} is an identity matrix.

With (14) and (15) the predicted mean and covariance vector can be directly computed without the requirement of ensemble runs if the eigenvalues of the state transition matrix are known. Compared with costly ensemble runs, the computation of the eigenvalues is relatively cheap, even for sophisticated GCMs. The calculation of the prediction utility R based on (14) and (15) will be discussed in detail in another paper.

In the remainder of this section, we will mainly explore the signal component SC, which contributes most to the prediction utility R as shown in section 5. For this purpose, we assume that the dynamical process (10) or (11) is stationary, which is widely assumed in POP theory (von Storch and Zwiers 1999). The condition of stationarity of (10) is that $\det(\Lambda_{k-1}) < 1$. As such, the climatological mean $\boldsymbol{\mu}^q = 0$ and the climatological variance $\boldsymbol{\sigma}_q = \tilde{\mathbf{Q}}$ at $k \rightarrow +\infty$. Also $\tilde{\mathbf{Q}}$ is a constant square matrix, in which case SC can be written as

$$\begin{aligned} \text{SC}(k) &= (\boldsymbol{\mu}_k^p)^T \boldsymbol{\sigma}_q^{-1} \boldsymbol{\mu}_k^p = (\boldsymbol{\mu}^p(0))^T \tilde{\mathbf{Q}}^{-1} \boldsymbol{\mu}^p(0) \\ &= \boldsymbol{\mu}^p(0)^T \boldsymbol{\Gamma}^T \tilde{\mathbf{Q}}^{-1} \boldsymbol{\Gamma} \boldsymbol{\mu}^p(0), \end{aligned} \quad (16)$$

where $\boldsymbol{\Gamma} = \prod_{i=0}^{k-1} \Lambda_i$, and is a square (diagonal) matrix consisting of the eigenvalues of the state transition matrix.

Equation (16) indicates that SC is inversely proportional to the noise variance $\tilde{\mathbf{Q}}$ and proportional to the initial signal variance (ISV) projected onto eigenmode space $\boldsymbol{\mu}^p(0)^T \boldsymbol{\mu}^p(0)$, modified by $\boldsymbol{\Gamma}^T \boldsymbol{\Gamma}$. When the initial signal variance is large, the eigenmodes are able to resist dissipation by the noise components of the system, leading to a large SC and a good prediction. On the other hand, if the signal variance is small, the noise components will quickly dominate the dynamical system and result in a low SC and a poor prediction. Thus ISV plays a critical role in determining prediction performance.

Compared with the earlier work of KM99, (16) is very consistent with their Eq. (2.8), which was derived using a completely different theoretical approach to that used here. In KM99, the prediction correlation function $r(t)$ is used to examine the dominant factors affecting its variations. It was found that $r(t)$ depends on the noise forcing and the eigenmode amplitudes present in the initial conditions, that is, ISV. Such a consistent result from two different theoretical frameworks indicates that (i) the prediction utility R can

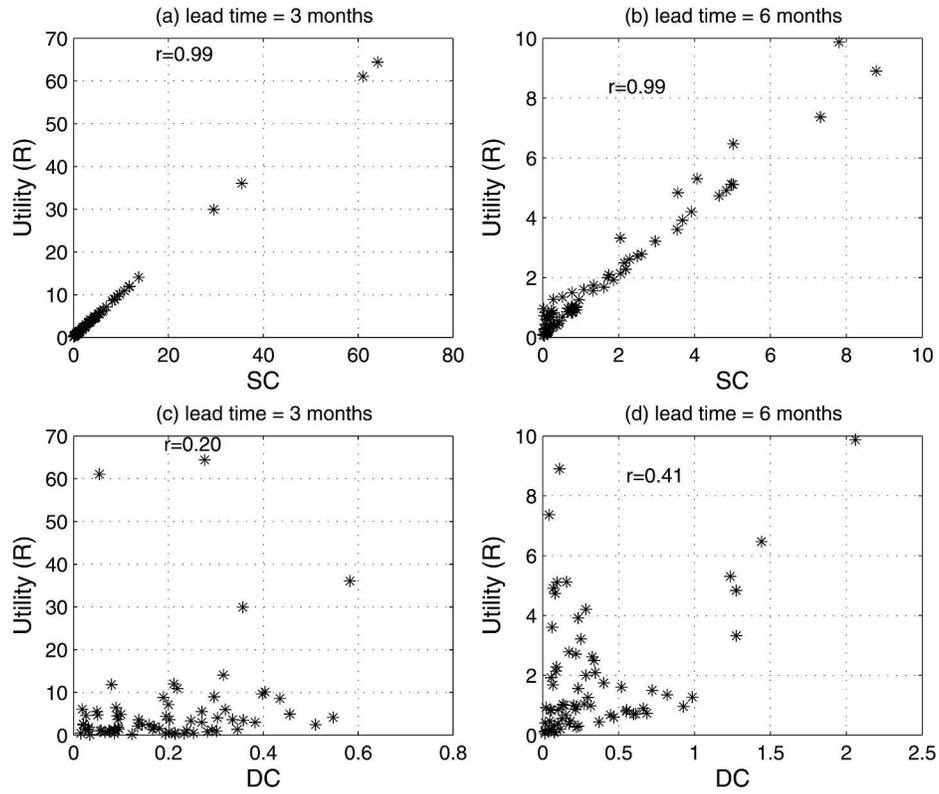


FIG. 15. As in Fig. 14 but for HCM2.

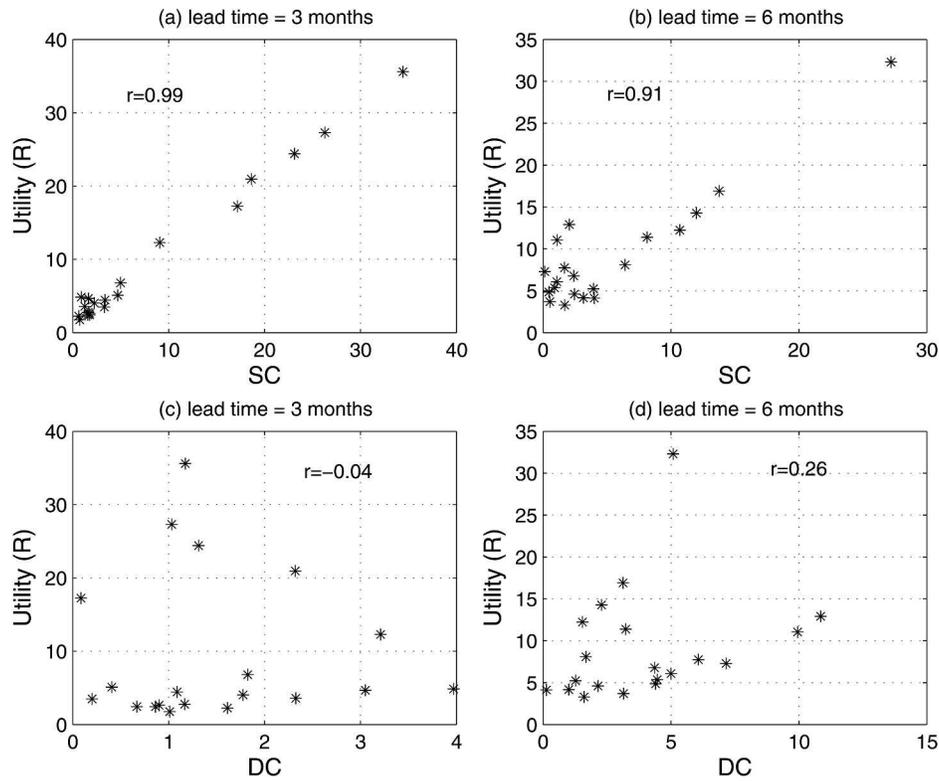


FIG. 16. As in Fig. 14 but the variance of stochastic forcing is doubled.

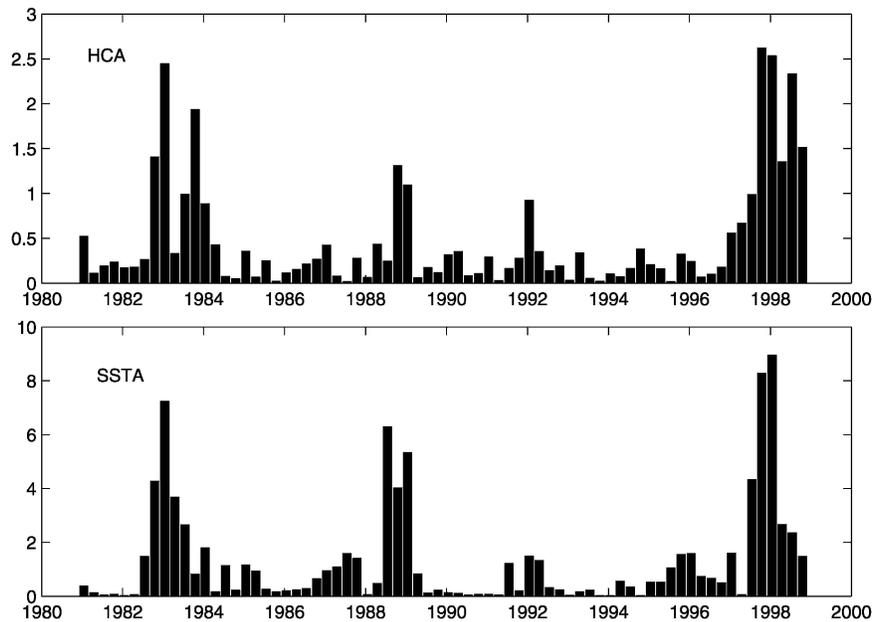


FIG. 17. POP amplitude for HCA and SSTA of prediction initial oceanic conditions generated by the data assimilation of subsurface temperature.

quantify the predictability of a skill measure such as $r(t)$ and (ii) the eigenmode amplitudes present in the initial conditions are a dominant factor affecting the model prediction skill.

Figure 17 shows the first eigenmode amplitude ISV of HCM1 (or HCM2) present in the initial conditions. This eigenmode was estimated by a POP analysis with the initial fields of HCA and SSTA, as discussed in previous sections. As can be seen, both the HCA and SSTA ISV have a large amplitude on 1982–83, 1988–89, and 1997–98, consistent with the timing of large R and C . Figure 18 compares the variation of HCA and SSTA ISV with that of SC at two lead times of prediction and shows a good relationship between them. Figure 19 shows the correlation coefficient between HCA ISV and SC for both HCM1 and HCM2. As can be seen, the correlation coefficients decrease with the lead times for both models, but exceed 0.6 for HCM1 and 0.55 for HCM2 at all lead times and are statistically significant.

From the discussions above, we conclude that we do not always have to perform expensive ensemble runs for measuring the reliability of ENSO predictions. Instead, we might be able to explore the ISV of the initial conditions to estimate the prediction reliability. This is very inexpensive and practical. If ISV is found to be large, the model prediction will most likely be reliable, and vice versa. The relationship between ISV and model predictability can further be demonstrated in Figs. 20 and 21. Shown in Fig. 20 is the comparison of model predictive skill, initialized with and without data assimilation. In the latter case, the model initial conditions are generated by spinning up the OGCM with FSU wind stress. As can be seen, the prediction skills

are significantly better with data assimilation than without data assimilation, especially for HCM2. Figure 21 presents the ISV from the initial fields without data assimilation. Compared with Fig. 17, the ISV without data assimilation is smaller than its counterpart with data assimilation, which is consistent with the difference in the model prediction skills shown in Fig. 20, indicating the key role of ISV in determining the model predictability. This also suggests that an important contribution of data assimilation might be to increase the ISV of the initial fields.

It should be noted that ISV only approximately estimates SC and represents the information that the prediction utility possesses. A complete and accurate measure of the reliability of a dynamical prediction should be directly obtained from (2). The ISV is only a good and economic alternative to ensemble prediction to evaluate model predictability. However, some proprieties and advantages that ensemble predictions possess are absent in the ISV. For example, the best use of ensembles is to make the probabilistic forecasts that include the uncertainty estimate, which is not available for the application of ISV.

7. Discussion and summary

A central task of ENSO predictability studies is to measure the reliability of the prediction and determine the dominant factors that affect the prediction accuracy. By applying a new theoretical framework introduced by Kleeman (2002) we have explored this issue using two realistic hybrid coupled models. It was found

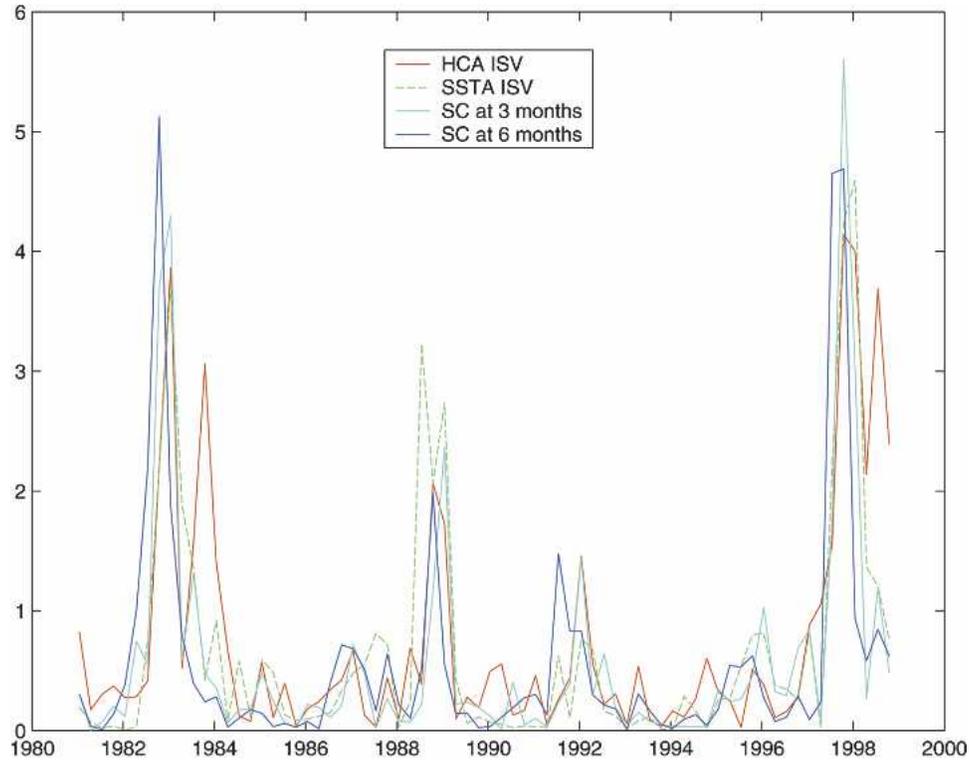


FIG. 18. The variations of HCA ISV, SSTA ISV, and SC at 3 and 6 months for predictions using HCM1. Each variable is normalized by its own standard deviation.

that the prediction utility R , defined by relative entropy, can measure well the reliability of the predictions of these models. In general, when R is large, the corresponding prediction was found to be reliable whereas, when R is small, the prediction is found to be less reliable.

A direct strategy for estimating R is via (2) using ensemble predictions. Several important issues should

be addressed prior to using (2). The first is how to generate forecast ensembles. Obviously, the larger the size of the ensemble, the better an estimate of the evolving pdf is likely to be. In theory, each component of the model state vector should be perturbed independently. However, the number of state variables of a realistic numerical model (10^6) far exceeds the maximum affordable ensemble size (10–50), which requires that one must choose perturbations wisely. In this work, we used stochastic optimals for this purpose. The stochastic optimals represent the spatial patterns onto which stochastic forcing must project in order to maximize error growth over a given time interval.

A second issue is the choice of a suitable variable and reduced space to which (2) can be applied and evaluated. The large model dimension limits the application of (2). Any reduced state space should be one that characterizes the model dynamics and at the same time represents the most significant physical features present in the original space. In this study, we choose subsurface heat content as the argument for the prediction utility R , as it represents the most useful information in ENSO prediction. It has been found that subsurface heat content is a good precursor for ENSO evolution, and that assimilation of heat content can significantly improve model prediction skill (Tang and Hsieh 2003). We also used the first POP mode to construct the reduce space since it describes very well ENSO oscillations.

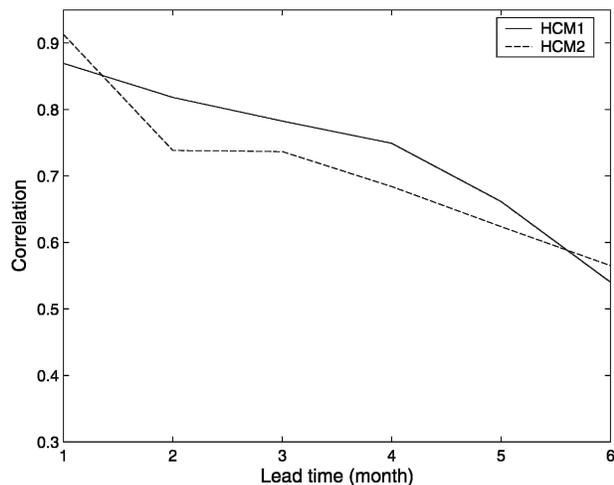


FIG. 19. Correlation of HCA ISV with HCM1 and HCM2 SC as a function of lead time.

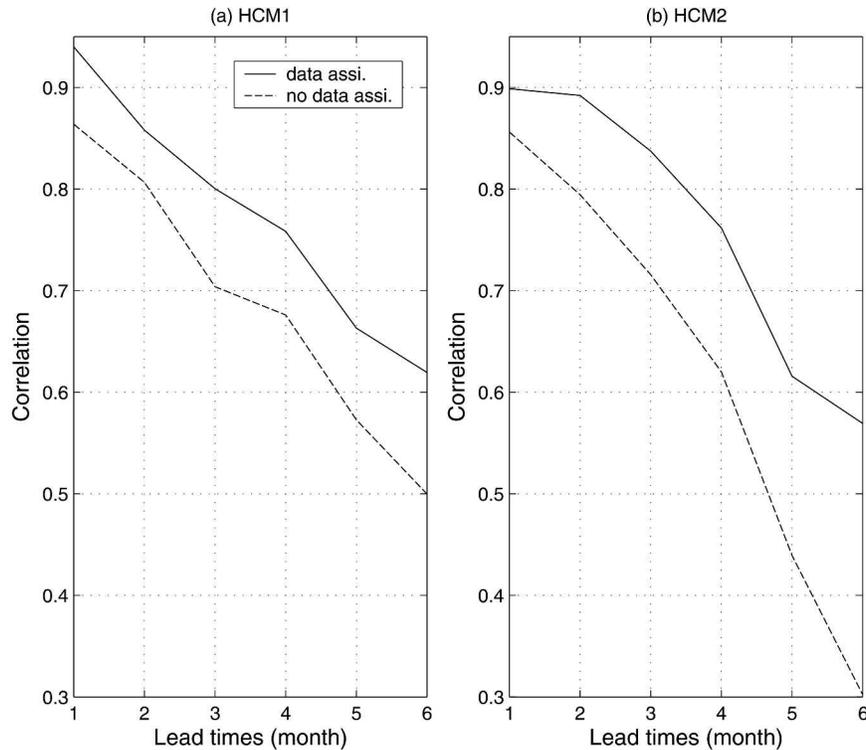


FIG. 20. Predictive correlation skill of HCM1 and HCM2, initialized with and without data assimilation.

tory behavior and approximates the leading eigenmode of the dynamical system. Sensitivity experiments revealed that the prediction utility R of HCM1 and HCM2 depend, to some extent, on the variable and the reduced space used. To illustrate, Fig. 22 shows SSTA prediction utility R for HCM1 estimated using a reduced state space that was constructed by the leading six EOF modes. As can be seen, SSTA prediction utility has a maximum value during 1989. This is different from Fig. 5 discussed in section 6.

One interesting finding in this paper is that there is a good relationship between the prediction utility R and the initial signal variance (ISV) in eigenmode space. By considering a linear stochastic dynamical system, we identified the signal component that dominates R with ISV. In general, it was found that when ISV is small (large), then R is also small (large). The ensemble results from HCM1 and HCM2 confirmed the analytical analysis. This finding has practical significance since it suggests that the reliability of ENSO predictions can be estimated very inexpensively using a simple method without the need for expensive ensemble of forecasts. The ideas presented here are also consistent with a similar work reported in KM99 using a different theoretical approach.

It is of interest to further explore the underlying physical interpretation of the relationship between ISV and model predictability. Since ISV is the amplitude of

leading POP mode at initial time, a large ISV means a strong spatial pattern similar to Figs. 3a,b present in the initial field. As argued in the literature (e.g., Tang 2002; KM99), Figs. 3a,b are consistent with the delayed-action oscillator mechanism for ENSO. For a freely evolving POP oscillation, the patterns appear in the sequence $\text{POP}^{\text{imag}} \rightarrow \text{POP}^{\text{real}} \rightarrow -\text{POP}^{\text{imag}} \rightarrow -\text{POP}^{\text{real}} \rightarrow \text{POP}^{\text{imag}}$. Thus the warm water present in the central equatorial and eastern Pacific Ocean yields the warm SST and heat content (HC) anomalies in this region (Fig. 3b) prior to the peak phase of an ENSO warm event (Fig. 3a). A strong zonal HC gradient at the central equatorial Pacific weakens the upwelling there and intensifies the warm Kelvin waves propagating eastward. The pattern in Fig. 3b can be regarded as a precursor pattern that is observed during the onset phase of El Niño, which is consistent with the “pile up” hypothesis of Wyrki (1975) and the recharge oscillator of Jin (1997). In this sense, a large ISV in fact corresponds to a strong delayed-action oscillator signal residing in the initial fields, leading to a reliable prediction.

It has been recognized for decades that model initial conditions exert a strong influence on model prediction skill. However, this recognition has been general and somewhat limited in a quantitative sense. These are still numerous outstanding issues such as 1) what are the dominant components present in initial conditions that impact model skill and 2) how can we quantitatively

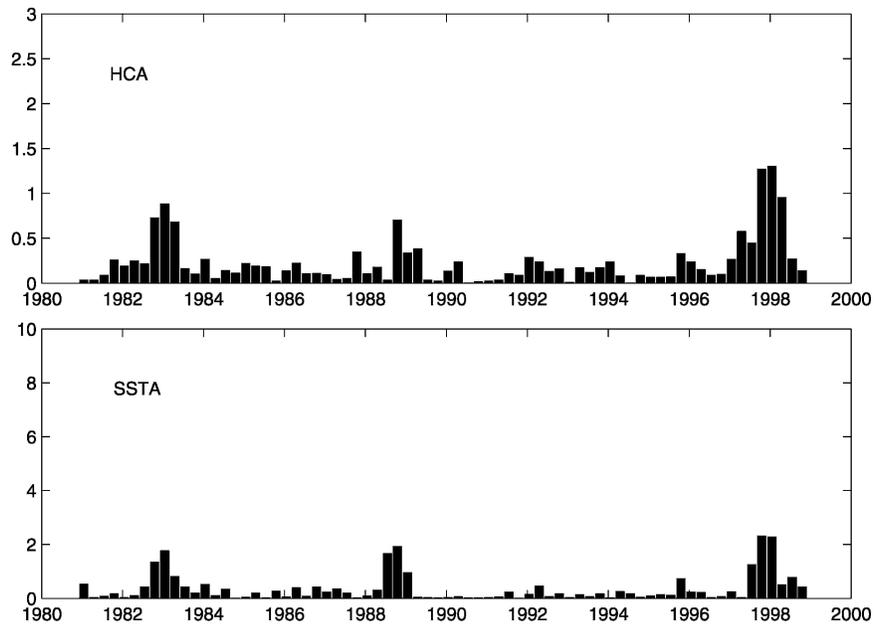


FIG. 21. POP amplitude of prediction initial conditions, generated without data assimilation, for HCA and SSTA. Compared with Fig. 14, the POP amplitudes are considerably smaller in this case.

measure the importance of initial conditions on prediction skill. These questions are interesting but difficult to answer. In this study, we explored these questions using information theory, and have presented some important findings and results relevant to ENSO prediction.

In this study, we confine our attentions to correlation skill for the measure of model predictability. Besides the correlation skill, the other skill often used is the root-mean-square error (rmse). We also examined the relation of the relative entropy to rmse. The results show that their relationship is less significant than that

of the relative entropy to correlation skill. This is because a small rmse does not have to indicate a good prediction. rmse variation greatly depends on the anomalous magnitude of SSTA variations. One prediction of anomalous event is usually associated with a larger rmse than that of a normal event. For example, the predictions shown in Figs. 6a,b have high correlation skills but a large rmse, whereas the prediction in Fig. 6d has a very poor correlation skill but a very small rmse. That is why we mainly focus on the analysis of correlation skill in this paper. Our results indicate that

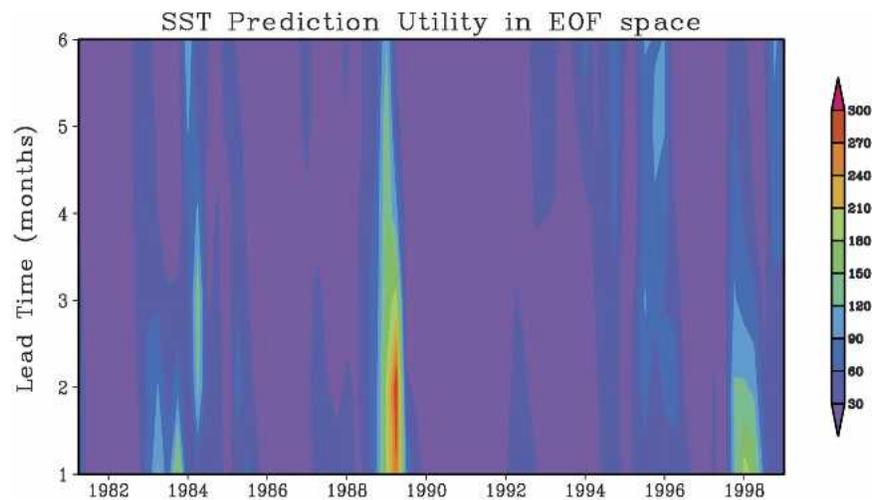


FIG. 22. SST prediction utility for HCM1 in an EOF reduced state space using SSTA.

a good prediction of the anomalous event (El Niño or La Niña) usually corresponds with large signal components such as for 1982–82, 1997–98 events.

It has been found that the signal component has much more important contribution to prediction utility than the ensemble spread in this study. This seems a striking counterexample to the widespread perception that ensemble spread is the main determinant of potential forecast skill for numerical weather prediction models. This is most probably because the ENSO system could be viewed as a stochastically forced damped linear system which leads to the covariance of transient distributions being independent of the initial conditions for a particular prediction, namely that the signal component of the prediction utility R shows any variation with initial conditions. Kleeman (2002) has theoretically proved this conclusion by a simplified system analogy to ENSO. He also showed that for other systems, for example, a Lorenz system analogy to weather, the en-

semble spread could play a dominant role in prediction utility. DelSole (2001) and Kleeman and Moore (1999) also found that the subdominance of the ensemble spread is probably due mainly to the optimally persistent patterns or signal amplitude present in the initial fields, which determines the model predictability for a large-scale and slowly varying climate system like ENSO.

Acknowledgments. This work is supported by NSF Grant ATM-0071342. We thank Jérôme Vialard and Anthony Weaver for their help in configuring the ocean model.

APPENDIX

Definition of σ_p

For dynamical system (10), $\sigma_p(k)$ is defined as

$$\begin{aligned}\sigma_p(k) &= E\langle[\mathbf{Z}_k - E\langle\mathbf{Z}_k\rangle][\mathbf{Z}_k - E\langle\mathbf{Z}_k\rangle]^T\rangle = E\langle[\Lambda_{k-1}(\mathbf{Z}_{k-1} - E\langle\mathbf{Z}_{k-1}\rangle) + \tilde{w}_{k-1}][\Lambda_{k-1}(\mathbf{Z}_{k-1} - E\langle\mathbf{Z}_{k-1}\rangle) + \tilde{w}_{k-1}]^T\rangle \\ &= E\langle\Lambda_{k-1}[\mathbf{Z}_{k-1} - E\langle\mathbf{Z}_{k-1}\rangle][\mathbf{Z}_{k-1} - E\langle\mathbf{Z}_{k-1}\rangle]^T\Lambda_{k-1}^T + \Lambda_{k-1}[\mathbf{Z}_{k-1} - E\langle\mathbf{Z}_{k-1}\rangle]\tilde{w}_{k-1}^T \\ &\quad + \tilde{w}_{k-1}[\mathbf{Z}_{k-1} - E\langle\mathbf{Z}_{k-1}\rangle]^T\Lambda_{k-1}^T + \tilde{w}_{k-1}\tilde{w}_{k-1}^T\rangle = \Lambda_{k-1}E\langle[\mathbf{Z}_{k-1} - E\langle\mathbf{Z}_{k-1}\rangle][\mathbf{Z}_{k-1} - E\langle\mathbf{Z}_{k-1}\rangle]^T\Lambda_{k-1}^T \\ &\quad + \Lambda_{k-1}E\langle[\mathbf{Z}_{k-1} - E\langle\mathbf{Z}_{k-1}\rangle]\tilde{w}_{k-1}^T\rangle + E\langle\tilde{w}_{k-1}[\mathbf{Z}_{k-1} - E\langle\mathbf{Z}_{k-1}\rangle]^T\Lambda_{k-1}^T + E\langle\tilde{w}_{k-1}\tilde{w}_{k-1}^T\rangle \\ &= \Lambda_{k-1}\sigma_p(k-1)\Lambda_{k-1}^T + \tilde{Q}.\end{aligned}$$

REFERENCES

- Battisti, D. S., 1988: Dynamics and thermodynamics of a warming event in a coupled tropical atmosphere–ocean model. *J. Atmos. Sci.*, **45**, 2889–2919.
- Blanke, B., and P. Delecluse, 1993: Variability of the tropical Atlantic Ocean simulated by a general circulation model with two different mixed layer physics. *J. Phys. Oceanogr.*, **23**, 1363–1388.
- Buizza, R., and T. N. Palmer, 1998: Impact of ensemble size on ensemble prediction. *Mon. Wea. Rev.*, **126**, 2503–2518.
- Chen, Y.-Q., D. S. Battisti, T. N. Palmer, J. Barsugli, and E. S. Sarachik, 1997: A study of the predictability of tropical Pacific SST in a coupled atmosphere–ocean model using singular vector analysis: The role of the annual cycle and the ENSO cycle. *Mon. Wea. Rev.*, **125**, 831–845.
- Cover, T. M., and J. A. Thomas, 1991: *Elements of Information Theory*. Wiley, 576 pp.
- DelSole, T., 2001: Optimally persistent patterns in time-varying fields. *J. Atmos. Sci.*, **58**, 1341–1356.
- Epstein, E. S., 1969: Stochastic dynamic predictions. *Tellus*, **21**, 388–407.
- Farrell, B. F., P. J. Ioannou, and J. Petros, 1993: Stochastic dynamics of baroclinic waves. *J. Atmos. Sci.*, **50**, 4044–4057.
- Hasselmann, K., 1988: PIPs and POPs: The reduction of complex dynamical systems using principal interaction and oscillation patterns. *J. Geophys. Res.*, **93**, 11 015–11 021.
- Ji, M., R. W. Reynolds, and D. W. Behringer, 2000: Use of TOPEX/Poseidon sea level data for ocean analyses and ENSO prediction: Some early results. *J. Climate*, **13**, 216–231.
- Jin, F.-F., 1997: An equatorial ocean recharge paradigm for ENSO. Part I: Conceptual model. *J. Atmos. Sci.*, **54**, 811–829.
- Kirtman, B. P., and J. Shukla, 1998: Current status of ENSO forecast skill. Climate Variability and Predictability (CLIVAR) Numerical Experimental Group Rep. [Available online at http://www.clivar.org/publications/wg_reports/wgsip/nino3/report.htm].
- Kleeman, R., 1991: A simple model of the atmospheric response to ENSO sea surface temperature anomalies. *J. Atmos. Sci.*, **48**, 3–18.
- , 2002: Measuring dynamical prediction utility using relative entropy. *J. Atmos. Sci.*, **59**, 2057–2072.
- , and A. M. Moore, 1997: A theory for the limitation of ENSO predictability due to stochastic atmospheric transients. *J. Atmos. Sci.*, **54**, 753–767.
- , and —, 1999: A new method for determining the reliability of dynamical ENSO predictions. *Mon. Wea. Rev.*, **127**, 694–705.
- , and A. J. Majda, 2005: Predictability in a model of geophysical turbulence. *J. Atmos. Sci.*, in press.
- , Y. Tang, and A. Moore, 2003: The calculation of climatically relevant singular vectors in the presence of weather noise. *J. Atmos. Sci.*, **60**, 2856–2867.
- Madec, G., P. Delecluse, M. Imbard, and C. Levy, 1998: OPA 8.1 ocean general circulation model reference manual. Institut Pierre Simon Laplace (IPSL), 91 pp.
- Molteni, R., and T. N. Palmer, 1993: Predictability and finite-time instability of the northern winter circulation. *Quart. J. Roy. Meteor. Soc.*, **119**, 269–298.
- Moore, A. M., and R. Kleeman, 1998: Skill assessment for ENSO using ensemble prediction. *Quart. J. Roy. Meteor. Soc.*, **124**, 557–584.
- , and —, 1999: Stochastic forcing of ENSO by the intraseasonal oscillation. *J. Climate*, **12**, 1199–1220.
- Palmer, T. N., 1999: Predicting uncertainty in forecast of weather and climate. ECMWF Tech. Memo. 294, 93 pp.

- Penland, C., and P. D. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, **8**, 1999–2024.
- Segschneider, J., D. L. T. Anderson, J. Vialard, M. Balmaseda, T. N. Stockdale, A. Troccoli, and K. Haines, 2001: Initialization of seasonal forecasts assimilating sea level and temperature observations. *J. Climate*, **14**, 4292–4307.
- Smith, T. M., 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.
- Tang, Y., 2002: Hybrid coupled models of the tropical Pacific: I. Interannual variability. *Climate Dyn.*, **19**, 331–342.
- , and W. W. Hsieh, 2003: ENSO simulation and prediction in a hybrid coupled model with data assimilation. *J. Meteor. Soc. Japan*, **81**, 1–19.
- , R. Kleeman, A. M. Moore, A. Weaver, and J. Vialard, 2003: The use of ocean reanalysis products to initialize ENSO predictions. *Geophys. Res. Lett.*, **30**, 1694, doi:10.1029/2003GL017664.
- , —, —, J. Vialard, and A. Weaver, 2004: An off-line, numerically efficient initialization scheme in an oceanic general circulation model for El Niño–Southern Oscillation prediction. *J. Geophys. Res.*, **109**, C05014, doi:10.1029/2003JC002159.
- Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330.
- Trevisan, A., F. Pancotti, and F. Molteni, 2001: Ensemble prediction in a model with flow regimes. *Quart. J. Roy. Meteor. Soc.*, **127**, 343–358.
- Vialard, J., P. Delecluse, and C. Menkes, 2002: A modeling study of salinity variability and its effects in the tropical Pacific Ocean during the 1993–1999 period. *J. Geophys. Res.*, **107**, 8005, doi:10.1029/2000JC000758.
- von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.
- Wyrski, K., 1975: Fluctuations of the dynamic topography in the Pacific Ocean. *J. Phys. Oceanogr.*, **5**, 450–459.
- Xu, J. S., and H. von Storch, 1990: Predicting the state of the Southern Oscillation using principal oscillation pattern analysis. *J. Climate*, **3**, 1316–1329.
- 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.
- , —, —, and T. N. Palmer, 1997: Predictability of a coupled model of ENSO using singular vector analysis. Part II: Optimal growth and forecast skill. *Mon. Wea. Rev.*, **125**, 2057–2073.
- Zavala-Garay, J., A. M. Moore, C. L. Perez, and R. Kleeman, 2003: The response of a coupled model of ENSO to observed estimates of stochastic forcing. *J. Climate*, **16**, 2827–2842.