Clim Dyn
DOI 10.1007/s00382-008-0397-3

Current status of ENSO prediction skill in coupled ocean–atmosphere models

Emilia K. Jin · James L. Kinter III · B. Wang · C.-K. Park · I.-S. Kang · B. P. Kirtman · J.-S. Kug · A. Kumar · J.-J. Luo · J. Schemm · J. Shukla · T. Yamagata

Received: 3 April 2007 / Accepted: 17 March 2008
© Springer-Verlag 2008

Abstract The overall skill of ENSO prediction in retrospective forecasts made with ten different coupled GCMs is investigated. The coupled GCM datasets of the APCC/CliPAS and DEMETER projects are used for four seasons in the common 22 years from 1980 to 2001. As a baseline, a dynamic-statistical SST forecast and persistence are compared. Our study focuses on the tropical Pacific SST, especially by analyzing the NINO34 index. In coupled models, the accuracy of the simulated variability is related to the accuracy of the simulated mean state. Almost all models have problems in simulating the mean and mean annual cycle of SST, in spite of the positive influence of realistic initial conditions. As a result, the simulation of the interannual SST variability is also far from perfect in most coupled models. With increasing lead time, this discrepancy gets worse. As one measure of forecast skill, the tier-1 multi-model ensemble (MME) forecasts of NINO3.4 SST have an anomaly correlation coefficient of 0.86 at the month 6. This is higher than that of any individual model as well as both forecasts based on persistence and those made with the dynamic-statistical model. The forecast skill of individual models and the MME depends strongly on season, ENSO phase, and ENSO intensity. A stronger El Niño is better predicted. The growth phases of both the warm and cold events are better predicted than the corresponding decaying phases. ENSO-neutral periods are far worse predicted than warm or cold events. The skill of forecasts that start in February or May drops faster than that of forecasts that start in August or November. This behavior, often termed the spring predictability barrier, is in part because predictions starting from February or May contain more events in the decaying phase of ENSO.

Keywords SST forecast · ENSO prediction · 10 CGCM intercomparison · Multi-model ensemble · APCC/CliPAS and DEMETER

E. K. Jin · J. L. Kinter III · J. Shukla
Department of Climate Dynamics, George Mason University, Fairfax, VA, USA

E. K. Jin (✉) · J. L. Kinter III · B. P. Kirtman · J. Shukla
Center for Ocean–Land–Atmosphere Studies, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705, USA
e-mail: kjin@cola.iges.org; kchin@gmu.edu

B. Wang
International Pacific Research Center, University of Hawaii, Honolulu, HI, USA

C.-K.Park
APEC Climate Center, Pusan, South Korea

I.-S.Kang · J.-S.Kug
Seoul National University, Seoul, South Korea

A. Kumar · J. Schemm
NCEP/NOAA Climate Prediction Center, Camp Spring, MD, USA

J.-J.Luo · T. Yamagata
FRCGC/JAMSTEC, Tokyo, Japan

1 Introduction

El Niño and the Southern Oscillation (ENSO), the most well-known coupled ocean–atmosphere phenomenon, is one of the most important seasonal to interannual climate variations (Philander 1990). Although ENSO originates and develops mainly in the tropical Pacific, the effects of ENSO are felt outside of the tropical Pacific Ocean (e.g., Trenberth et al. 1998; Rasmusson and Carpenter 1982; Bradley et al. 1987; Ropelewski and Halpert 1987, 1989)

with impacts on the global climate, the ecology of the tropical Pacific and the economies of many countries. Successful ENSO forecasts can offer decision makers an opportunity to take into account anticipated climate anomalies, potentially reducing the social and economic impacts of this natural phenomenon.

In the tropics, forecasting at the seasonal time-scale is linked to the availability of accurately predicted sea surface temperature (SST), since the tropical atmosphere responds to large-scale SST anomalies in a coherent and reproducible way (e.g., Stern and Miyakoda 1995; Shukla 1998). Because the characteristic time-scales of the ocean are longer than those of the atmosphere, the predictability limit is expected to be longer than in the atmosphere, possibly as much as several seasons. Based on this understanding, climate predictions at lead times of one to a few seasons are primarily a matter of predicting the SST (Goddard et al. 2001). The notion that the tropical atmosphere responds to SST forcing in a reproducible way, which forms the physical basis for the recent and present-day two-tier prediction approach (Bengtsson et al. 1993), has, however, been challenged and shown to be inadequate in the summer monsoon regions (e.g., Wang et al. 2005). Zheng et al. (2004) and Fu et al. (2004) have suggested that a two-tier approach has shortcomings for predicting intraseasonal oscillations as well.

There are a number of different strategies for forecasting SST in the tropical Pacific, including purely statistical techniques (e.g., Graham et al. 1987), combinations of dynamic and statistical models (e.g., Cane et al. 1986; Neelin and Dijkstra 1995), and purely dynamic models (e.g., Rosati et al. 1997; Stockdale et al. 1998; Schneider et al. 1999; Kirtman 2003). The models used in these forecasting strategies have varying degrees of sophistication and diverse initialization schemes. A reasonable level of prediction skill has been achieved with forecast models, and comparisons among different forecast methods have been made (Barnett et al. 1988; Barnston et al. 1994, 1999; Landsea and Knaff 2000). At present, these forecasting strategies are generally competitive, producing forecasts with roughly the same level of skill (Kirtman et al. 2001). Most of the major meteorological centers around the world have begun to produce dynamical seasonal forecasts with comprehensive ocean–atmosphere coupled general circulation models (CGCMs), using an ensemble approach (e.g., Mason et al. 1999; Kanamitsu et al. 2002).

The inherent but as yet unrealized advantage of CGCMs relative to other approaches is that improvements in prediction skill may be achieved with improved and more comprehensive representations of the relevant physical processes and improved utilization of the models using the admittedly imperfect initial conditions. A number of prediction efforts using CGCMs with various methods of data assimilation, initialization and component model coupling have been undertaken in the last decade to attempt to exploit the advantages of CGCMs (Ji et al. 1998; Rosati et al. 1997; Ji and Leetmaa 1997; Kirtman et al. 1997; Stockdale 1997, 1998; Leetmaa and Ji 1989; Schneider et al. 1999; Wang et al. 2002).

Although the representation of ENSO in coupled models has advanced considerably during the last decade, several aspects of the simulated climatology and ENSO are not well reproduced by the current generation of coupled models. The systematic errors in SST are often largest in the equatorial Pacific, and model representations of ENSO variability are often weak and/or incorrectly located (Neelin et al. 1992; Mechoso et al. 1995; Delecluse et al. 1998; Davey et al. 2002). These studies suggest that improvements in CGCMs are not yet sufficient to achieve realistic ENSO simulation (AchutaRao and Sperber 2006).

Because the memory of the system resides in the upper ocean, it is crucial to determine the oceanic initial state as accurately as possible. Accurate initialization of the forecast system does not, however, guarantee successful ENSO predictions (Gualdi et al. 2005). In addition to initial condition error, forecast error can also arise as a result of systematic errors in the coupled model, or errors in the representation of the evolution of the atmospheric or oceanic state. The systematic errors include errors in the time-mean basic state, the phase and amplitude of the seasonal cycle, and the statistics (amplitude, frequency, and location) of interannual anomalies. For example, in the tropical Pacific, most models simulate an equatorial cold tongue that is too prominent (Mechoso et al. 1995; Latif et al. 2001; Davey et al. 2002; AchutarRao and Sperber 2002). The models also differ in their representation of intraseasonal variability both due to weather noise and organized structures such as the Madden–Julian Oscillation (e.g., Lau and Waliser 2005).

The relationship between systematic errors in the climatological mean state and the simulation of interannual variability remains a subject of debate. Numerous authors have suggested the importance of properly simulating the mean climatology of the equatorial Pacific for achieving realistic interannual variability (Zebiak and Cane 1987; Schopf and Suarez 1988; Battisti and Hirst 1989). On the other hand, intercomparison studies by Neelin et al. (1992), Mechoso et al. (1995), and Latif et al. (2001) concluded that the simulation of ENSO variability in CGCMs seems to be relatively unaffected by the basic state of the model or by the behavior of the simulated seasonal cycle.

The systematic errors in coupled model forecasts also depend on the date of the initial conditions. Coupled models initialized from observed initial states tend to adjust toward their own climatological mean and variability, leading to forecast errors. The errors associated with such

adjustments tend to be more pronounced during boreal spring. This suggests that the so-called "spring predictability barrier" (e.g., Webster 1995) may be associated with a "spring variability barrier" in coupled models (Schneider et al. 2003). The spring barrier is most prominent during decades of relatively poor predictability (Balmaseda et al. 1995), but is not evident in all ENSO-prediction models, and so may not be an inherent feature of the ENSO phenomenon (Chen et al. 1995).

The main objective of this study is to quantitatively assess the current skill of seasonal forecast systems to predict ENSO anomalies. The ensemble forecasts of the tropical Pacific in ten global CGCMs developed within the last decade have been compared with each other and with observations. The comparable forecasts produced using a dynamic-statistical model and persistence serve as a baseline reference. The primary quantity analyzed is monthly mean tropical SST with regard to its annual mean, mean annual cycle, and interannual variability. We evaluate retrospective forecasts with respect to the relationship between ENSO and the phase of the annual cycle and global teleconnections associated with ENSO. We describe the common and distinct features of the various models' in order to explore possible causes for the varying quality of their ENSO forecasts. Throughout the remainder of this paper, "hindcast" (a term used to describe a forecast made retrospectively), "forecast" and "prediction" are used synonymously. Strictly speaking, all the results presented apply only to retrospective prediction.

The remainder of the paper is organized as follows. Section 2 briefly describes the models and experimental design of retrospective forecasts utilized in this study. The simulation of the annual mean and mean annual cycle in the equatorial Pacific SST is discussed in Sect. 3. Section 4 describes the simulation of the interannual variability in the equatorial Pacific, while Sect. 5 summarizes the forecast skill in terms of anomaly correlation coefficients. A brief summary and discussion of the major findings follows in Sect. 6.

2 Models and retrospective forecast experiments

The CGCM hindcast datasets that come from the APCC/CliPAS (APEC Climate Center/Climate Prediction and its Application to Society; Wang et al. 2007) and DEMETER (Development of a European Multimodel Ensemble system for seasonal to inTERannual prediction; Palmer et al. 2004) projects are used. The ten models used to produce the hindcast data are coupled ocean–land–atmosphere dynamical seasonal prediction systems. The hindcasts are 6- to 9-month model integrations starting from 6 to 15 different initial conditions in each ensemble. This hindcasts made for the 22 years from 1980 to 2001 are analyzed in this study, because they are common to all the hindcast data sets. Table 1 shows the CGCMs and the salient features of their component ocean and atmosphere models. In order to consider the possible impact of the annual cycle on the forecasts the retrospective forecasts were initialized in February, May, August, and November of each year (details are provided below). These models have diverse parameterizations of subgrid-scale physical processes and also differ considerably in their vertical and horizontal resolutions. It is important to note that some of these models were not specifically designed solely for forecasting ENSO.

To assess the skill of the CGCM hindcasts, a comparable set of hindcasts was produced using two much simpler and relatively inexpensive methods: persistence and the Dyn-Stat model. A persistence forecast, based purely on local observations of past and current climate (Goddard et al. 2001), is made simply by assuming that the observed monthly mean anomaly for the month prior to the initial time persists unaltered over the period of the forecast (referred to hereinafter as "PERSIST"). The Dyn-Stat model combines dynamical forecast information with statistics based on past observations (Kug et al. 2007). This global SST prediction system is developed for a boundary condition of AGCM forecast and four different SST predictions coming from dynamical El Niño prediction model, lagged linear regression model, pattern projection model, and persistence are simply composed. The SST analysis (referred to hereinafter as "observations") used to verify the SST forecasts is the HadISST1.1 (Rayner et al. 2003).

For each model, a climatological mean was computed from the 22 years of monthly mean hindcast values for each calendar month at each lead time. The anomaly in a given month for a given lead time is defined as the deviation from the climatological mean for that month and lead. The climatological means for the NCEP, and SINTEX-F models were calculated based on a slightly different period due to the limitations of data availability (Table 1). The same procedure was applied to the observations to compute the climatological mean and anomalies.

A 13th hindcast was constructed by forming a simple average of ten CGCM hindcast anomalies for each calendar month and lead time. This multi-model ensemble (hereafter referred to as the "MME") is compared to individual hindcasts and the observations.

Throughout the ENSO prediction literature there is some confusion regarding the appropriate definition of forecast lead time. For the purposes of this paper, forecast lead time is defined as in the following example. A forecast labeled as February 1982 initial conditions has its first full month of forecast in the month of February 1982. In the case of DEMETER data, these are forecasts initialized at 00Z on 1

Table 1 Models analyzed in this study

Modeling group	Model designation	AGCM	CGCM	Ensemble members	Climatology
CERFACS	CERFACS	ARPEGE	OPA 8.2	9	1980–2001
		T63 L31	2.0 × 2.0 L31		
		Deque (2001)	Delecluse and Madec (1999)		
ECMWF	ECMWF	IFS	HOPE-E	9	1980–2001
		T95 L40	1.4 × 0.3–1.429 L29		
		Gregory et al. (2000)	Wolff et al. (1997)		
INGV	INGV	ECHAM 4	OPA 8.1	9	1980–2001
		T42 L19	2.0 × 0.5–1.5 L31		
		Roeckner (1996)	Madec et al. (1998)		
LODYC	LODYC	IFS	OPA 8.2	9	1980–2001
		T95 L40	2.0 × 2.0 L31		
		Gregory et al. (2000)	Delecluse and Madec (1999)		
Meteo-France	Meteo-France	ARPEGE	OPA 8.0	9	1980–2001
		T63 L31	192–152, L31		
		Deque (2001)	Madec et al. (1997)		
MPI	MPI	ECHAM-5	MPI-IM1	9	1980–2001
		T42 L19	2.5 × 0.5–2.5 L23		
		Roeckner (1996)	Marsland et al. (2003)		
UK Met Office	UKMO	HadAM3	GloSea OGCM	9	1980–2001
		2.5 × 3.75 L19	1.25 × 0.3–125 L40		
		Pope et al. (2000)	Gordon et al. (2000)		
NCEP	NCEP	GFS	MOM 3	15	1981–2001
		T62 L64	1/3 × 5/8 L27		
		Saha et al. (2006)	Pacanowski and Griffies (1998)		
FRCGC SINTEX	SINTEX-F	ECHAM 4	OPA 8.2	9	1982–2001
		T106 L19	2 × 2 L31		
		Luo et al. (2005)	Madec et al. (1998)		
SNU	SNU	SNU	MOM 2.2	6	1981–2001
		T42 L21	1/3 × 1 L32		
		Kug et al. (2008)	Pacanowski (1995)		

February 1982; however, in the case of CFS forecasts, the ensemble members were initialized at various dates in January 1982 (see Saha et al. 2005 for details). The first monthly mean (i.e., the average of 1–28 February 1982) of the forecast is defined as the month 1. Similarly, the second monthly mean (i.e., 1–31 March 1982) is defined as the month 2. The remaining lead times are defined analogously.

3 Systematic errors in the SST forecasts

As described in the Introduction, a central question that remains under debate is how the quality of a given model's simulation of the annual mean and mean annual cycle affect the simulation and prediction of ENSO. Unlike previous studies that used long model runs without considering the impact of the initial conditions, retrospective forecasts are used here. The use of realistic initial conditions may contribute to the fidelity of model simulations at short lead times.

To compute the annual mean and mean annual cycle from the hindcast data, we used data at different lead times in two ways. First, short lead-time data were used as follows: we used the 1-month-lead data to calculate the means for February, May, August, and November, the 2-month-lead data for March, June, September, and December, and the 3-month-lead data for April, July, October, and January, averaging these 12 monthly means to form an annual mean. Second, we did the analogous calculation for longer leads using the 4-month-lead data for February, May, August, and November, and similarly for the 5-month-lead and 6-month-lead data. These two methods to estimate the annual mean and mean annual

cycle indicate how lead-time affects the representation of these quantities in hindcasts.

3.1 Annual mean

We show the observed and simulated equatorial SST (averaged over 2°N–2°S) as a function of longitude in Fig. 1. The coupled model simulations display a relatively narrow range of behavior that is quite good by comparison to earlier results (Mechoso et al. 1995; Latif et al. 2001; Davey et al. 2002; AchutaRao and Sperber 2002), which suggests that the influence of initial conditions contributes to the realistic reproduction of the climatology. Most of the models simulate a strong east-west SST gradient similar to the observed. However, the characteristics of the mean bias in coupled models are quite similar to those found in previous studies. In most models, the annual mean equatorial SST in the central Pacific is too cold. And substantial positive SST errors are common east of 100°W, where the observed mean SST rises approaching the South American coast. The warm SST errors adjacent to the South American coast in CGCMs are typically linked to errors in the simulation of stratus clouds: the deficient simulated cloud cover is associated with increased solar radiation into the ocean (Davey et al. 2002). This feature is more pronounced in the MME. Through the whole tropical band over the Pacific, the mean bias of the MME is less than 1°C. The Dyn-Stat and PERSIST forecasts have a larger warm bias east of 160°W.

The mean bias also increases with respect to increasing lead time (Fig. 1c, d). The gradient in the central part of the basin, however, is still simulated reasonably well by all models. There are two noticeable outliers: the simulated SST in the MPI CGCM is about 7°C too cold, and the SNU CGCM has a 4°C warm bias near the eastern boundary of the Pacific. The ECMWF and LODYC forecasts show a relatively large cooling with respect to lead time (compare Fig. 1a, b and 1c, d). The systematic error in all models develops very quickly.

3.2 Annual cycle

Time-longitude plots of the annual cycle of Pacific SST along the equator (with the annual mean removed) are shown in Fig. 2. The region of the cold tongue in the eastern equatorial Pacific is dominated by an annual period in the observations, while a semi-annual period is observed in the west. There is a pronounced westward propagation of the phase of the annual cycle of SST in the eastern and central equatorial Pacific in the observations. Many physical processes both in the ocean and atmosphere, as well as coupled feedbacks, contribute to the generation of the annual cycle in the east.

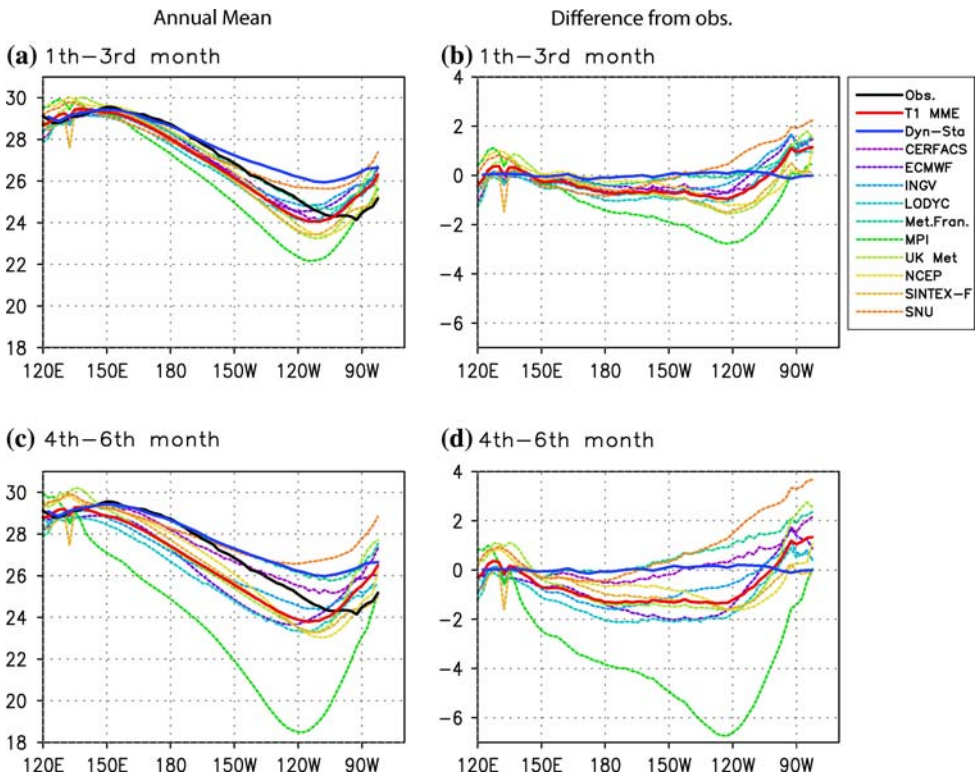

Fig. 1 *Right panels* show observed and simulated climatological mean, annual mean SST averaged over 2°S–2°N. *The black line* shows the observations, *red* for the MME forecasts, *blue* for the Stat-Dyn model forecasts, and *colored dots* for individual coupled models as shown in the legend, respectively. For models, the *upper panel* shows the 1- to 3-month-lead forecast mean (the annual mean is reconstructed based on Feb, May, Aug, and Nov data from the 1-month forecasts, Mar, Jun, Sep, and Dec data from the 2-month forecasts, and Apr, Jul, Oct, and Jan data from the 3-month forecasts, respectively), and the *lower panel* shows the same quantity for the 4- to 6-month-lead forecast mean with the same convention used to reconstruct the annual mean. *Left panels* show the differences from observation

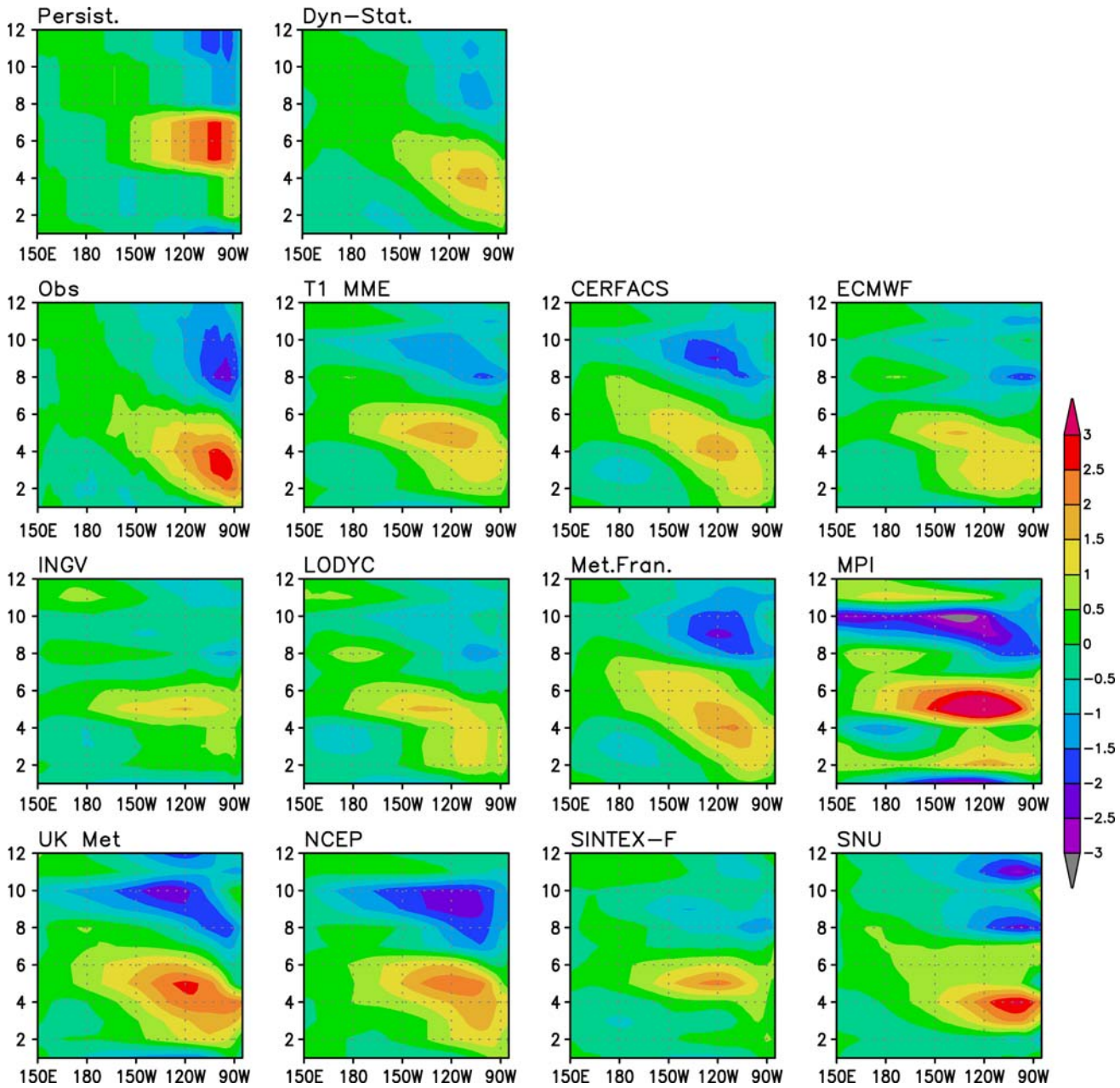

Fig. 2 Observed and simulated annual cycle of SST along the equator (average for 2°S–2°N). Shown are the deviations from the annual mean. For the models, the 1- to 3-month-lead forecast data are used

One of the main factors controlling the quality of CGCM simulation of the annual cycle is the amount of low-level stratus clouds in the atmospheric component models (Mechoso et al. 1995; Latif et al. 2001; Davey et al. 2002). However, dynamical processes in the ocean and atmosphere are also of great importance in generating the annual cycle. Thus, the annual cycle is a suitable benchmark for coupled models, since it involves complex dynamical and physical interactions among the components of the climate system. Based on this inference, Latif et al. (2001) argued that high meridional resolution in ocean component models is needed to improve the annual cycle in CGCMs. Other studies have shown that resolving the heat fluxes due either to penetrative radiation in the upper ocean (Murtugudde et al. 2002) or oceanic eddies (e.g., Jochum and Murtugudde 2004) is critically important for reduction of ocean model biases.

Many CGCMs appear to have problems in simulating the annual cycle of the SST in the eastern equatorial Pacific. Many models exhibit an annual cycle that is much too weak in the eastern Pacific. A phase shift in the annual cycle and a westward displacement of the annual cycle is

also found in most of the models. Other problems that occur are an annual cycle that is too strong (e.g., the MPI CGCM) or a predominantly semi-annual cycle instead of an annual cycle (e.g., the ECMWF, MPI, and SNU CGCMs). These discrepancies become more severe with increasing lead time (Fig. 3). In particular, there is an increasing phase shift with increasing lead time. Considering that these models have moderately high resolution in the tropical oceans, this suggests that high resolution is a necessary but not sufficient condition for a good annual cycle simulation. Note that persistent shows the discrete annual cycle due to reconstruction of data from four initial condition cases.

Due to the dominance of ENSO in global climate variability, the forecast skill for ENSO indices of equatorial SST anomalies such as NINO3, NINO3.4, and NINO4 are important measures of performance. The west–east gradient of SST over the tropics is also an important feature that helps to generate the mean circulation and atmospheric convection. To examine this, the climatological mean annual cycle of the NINO4 index (SST anomaly averaged over 160°E–150°W, 5°S–5°N) minus the NINO3 index

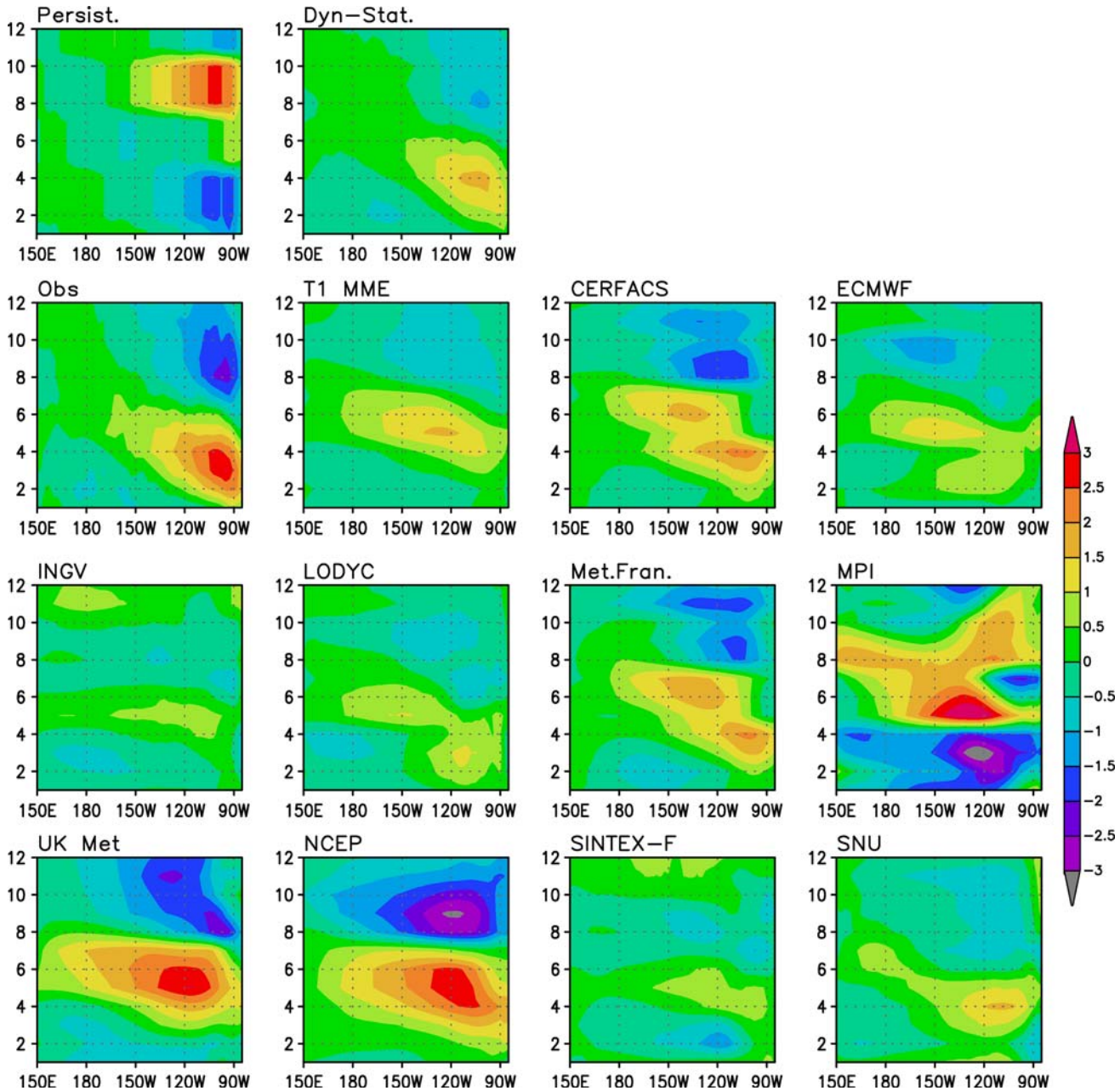

Fig. 3 Same as Fig. 2 except for the 4- to 6-month-lead forecasts

(SST anomaly averaged over 150°W–90°W, 5°S–5°N) is calculated (not shown). Most of the CGCMs underestimate the overall strength of this gradient throughout the year, in particular, in the May–January period. While most models produce a fairly reliable amplitude of the annual cycle of this measure of the SST gradient, the timing of the maximum and minimum values varies considerably. Most models correctly simulate the timing of the minimum value in April (except the MPI model), but the range of values is ±100% of the observed.

4 Interannual variability

In this section, we will examine how well the models represent the SST variability over the tropical Pacific in terms of amplitude and location. After removing the mean bias, the standard deviation of monthly mean SST anomalies was calculated (Fig. 4). Focusing on the Pacific east of the dateline (excluding the coast of South America), the observations show a clear seasonality with strong variance in winter and low variability from February through summer. This is a well-known linkage between the variability and the mean annual cycle of the eastern equatorial SST. There is also a minimum of SST variance in the western Pacific that is relatively constant throughout the year.

For lead times up to three months, the variability in most of the models is similar to the observed, with moderate seasonality. This feature is very clear in the MME. However, there are several features that distinguish the models from each other. The centers of action in summer and winter and the timing of the peak are quite different from one model to another. Many models simulate the center of action west of the observed location. For the Dyn-Stat and PERSIST models as well as many of the dynamic models, the variability is substantially lower than that observed. At the longer leads, more of the dynamic models have less variability than observed.

The intensity of the annual cycle shown in Figs. 2, and 3 seems to be related to the interannual variability, since the MPI, NCEP, and UKMO CGCMs have higher variability, even in the 4th to 6th lead months, unlike other models, both in the mean annual cycle and in the interannual variability. This is more evident for longer leads. Moreover, the interannual variability in the MPI and SNU CGCMs are quite different from the observed, as was already noted in the behavior of the mean annual cycle.

We also note that a phase shift of interannual anomalies with increasing lead time is found in several CGCMs (not shown), as in the mean annual cycle. This is likely to be associated with the fact that each model has its own intrinsic ENSO period, which is different from observations (Jin and Kinter 2007).

Figure 5 shows the longitudinal distribution of interannual variability along the equator. The standard deviation is calculated for 12 months during 1980–2001 periods. Right panels show the standard deviation and left show the difference from observed one. From west to east, the error of standard deviation gets larger and it is somewhat related with the intensity of variability. Most of models simulate westward center of action different from observed shown in 90°W. As a result, there are substantial error in the far eastern Pacific. Some of models simulate weak variability less than 50% of observed off the coast of Peru even for 1–3 lead. While, the simulated interannual variability over the central Pacific is moderately well. With increasing lead time, the interannual variability gets weaker over the central and eastern Pacific in models.

Interestingly, there exists a systematic relationship between the fidelity of the simulated annual cycle and the correct simulation of the phase locking between the ENSO and the annual cycle, in particular, the amplitude. Figure 6a shows the scatter diagram between interannual variability (x-axis) and intensity of annual cycle (y-axis) of NINO 3.4 index. The interannual variability is the standard deviation of SST anomalies of 12 calendar months during 22 years. The intensity of annual cycle is defined as a standard deviation from annual mean. All lead times from 1 to 6 are included for theses results. Models show linear relationship between annual cycle intensity and interannual variability suggesting that the accuracy of the simulated variability is related to the accuracy of the simulated mean state.

Figure 6b shows the relationship between the intra-ensemble and the interannual variability of 12 months during 1980–2001 with all lead times. It clearly shows the interannual variability is proportional to intra-ensemble variability. Note that SNU CGCM has 6 members and NCEP CGCM has 15 members different from other models having 9.

The relationship between errors of the annual cycle and those of anomalies is shown in Fig. 6c. The former is defined as a root-mean-square error of the deviation from annual mean. The latter is the root-mean-square error of anomalies after subtracting a 1980–2001 climatology. Even though the relationship is not clearly linear, models having smaller mean error show smaller error in the interannual variability, and vice versa. Interestingly, the MME shows much improved interannual predictability comparing to annual cycle.

5 ENSO forecast skill

In evaluating ENSO prediction models, two aspects can be considered: skill and usefulness (Landsea and Knaff 2000).

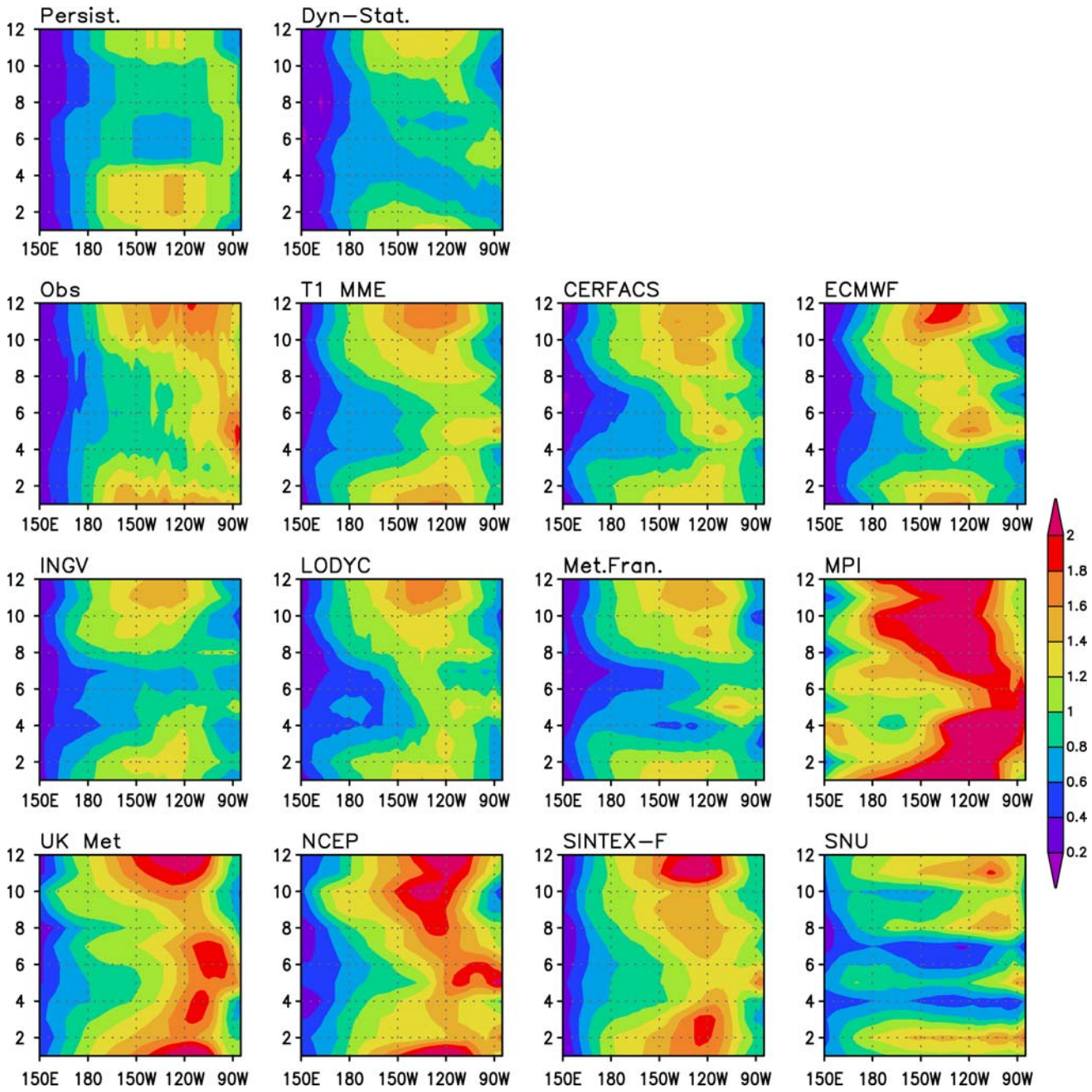

Fig. 4 Observed and simulated standard deviation of SST anomalies as a function of calendar month along the equator (2°S–2°N). For models, the 1- to 3-month-lead forecast data are used

Forecast skill is defined as the ability to improve upon some baseline prior estimate. In this study, we utilize both PERSIST and Dyn-Stat to serve as baselines. Usefulness is the degree to which predictions can differentiate between the phases (El Niño, La Niña, and neutral) and, when an El Niño or a La Niña is present, to determine its approximate magnitude.

For quantitative purposes, these aspects of forecasts are measured by the anomaly correlation coefficient and the root-mean-squared differences of the observed and forecast anomalies. The former (ACC) provides a measure of the quality of the forecast pattern, and the latter (RMSE) measures the agreement in amplitude of the anomalies. In general, NINO3.4 is a good index of ENSO variability (Barnston et al. 1997). Figure 7a shows the anomaly correlation coefficients and Fig. 7b shows the RMSE of NINO3.4 during 1980–2001 as a function of lead time, after removing the mean bias, for all four cases (February, May, August, and November initial conditions). As in Fig. 1, the thick red solid line is for the 10-CGCM MME,

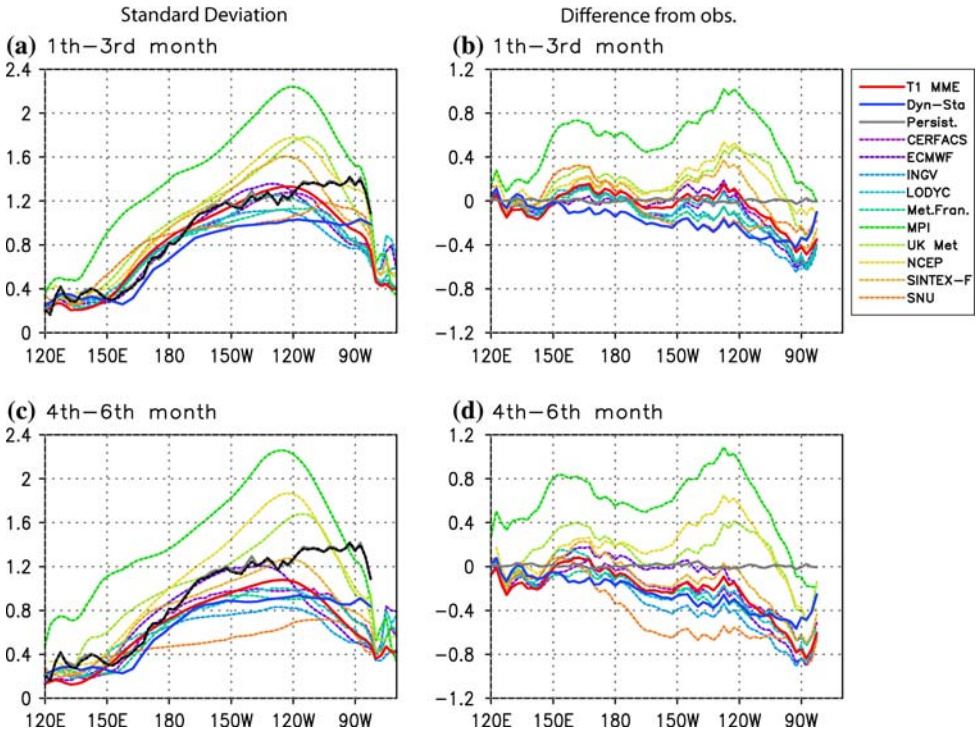

Fig. 5 *Right panels* show the longitudinal distribution of observed and simulated standard deviation of SST anomalies along the equator (2°S–2°N). For models, the *upper panel* shows the 1- to 3-month-lead forecast mean and the *lower panel* shows the same quantity for the 4- to 6-month-lead forecast mean. *Left panels* show the differences from observed standard deviation

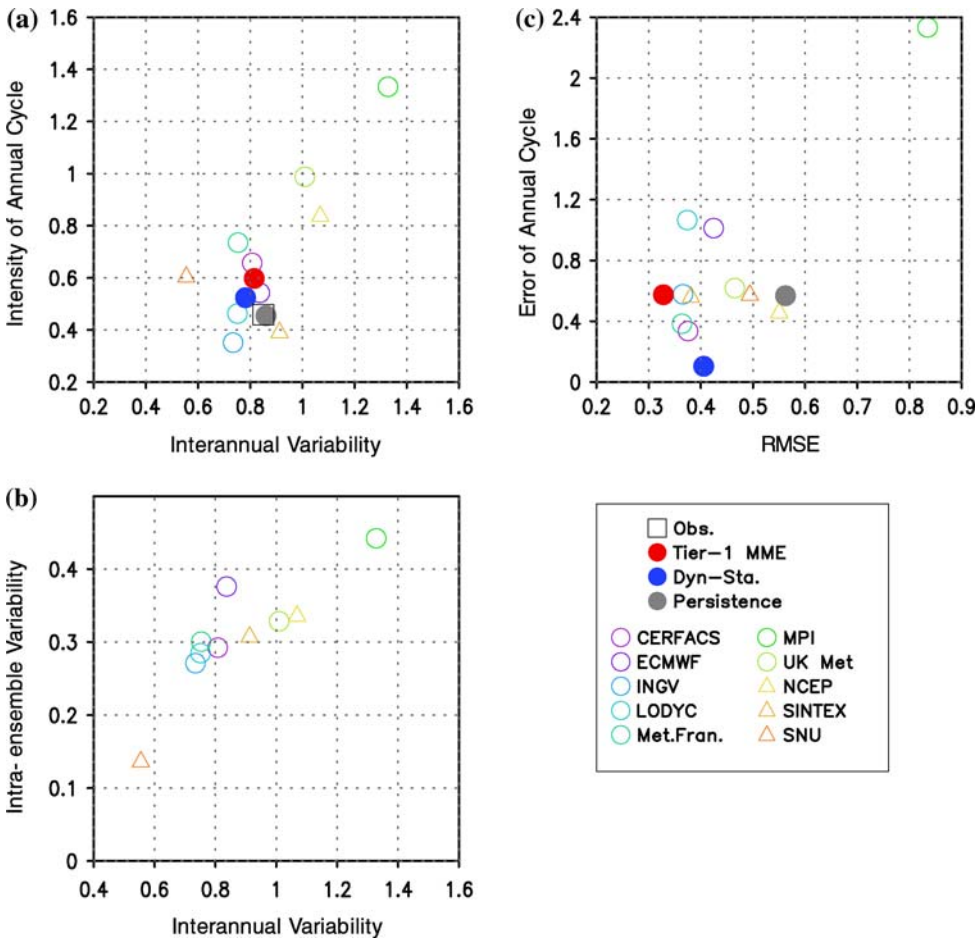

Fig. 6 Scatter diagram between **a** intensity of annual cycle and interannual variability, **b** intra-ensemble and interannual variability, and **c** error of annual cycle and interannual root-mean-square error after subtracting the climatology

blue is for Dyn-Stat, black is for PERSIST, and the dashed lines are for individual CGCMs, respectively. For all cases, the decline of skill (decrease of ACC and increase of RMSE) with respect to lead time is clear.

The results shown in Fig. 7 indicate that MME forecasts have better skill than those of any of the individual models (thin dashed lines). The ACC of the tier-1 MME forecast of NINO3.4 SST anomalies reaches 0.86 at month 6. This is in agreement with the results found with other coupled model forecast systems, and can be explained by the fact that the ensemble average reduces the noise present in the individual forecasts, increasing the correlations and reducing the systematic error (Krishnamurti et al. 2000; Palmer et al. 2000). MME forecasts appear to be generally better than not only PERSIST but also Dyn-Stat. In general, the statistical models provide a useful benchmark for evaluating the dynamical models. The skill of most of the models is better than that of PERSIST. Note that the skill for PERSIST, which has been the traditional standard for determining skill, is quite low. Moreover, half of the CGCMs outperform Dyn-Stat. Hence, it is readily apparent that most of the CGCMs provide skillful forecasts.

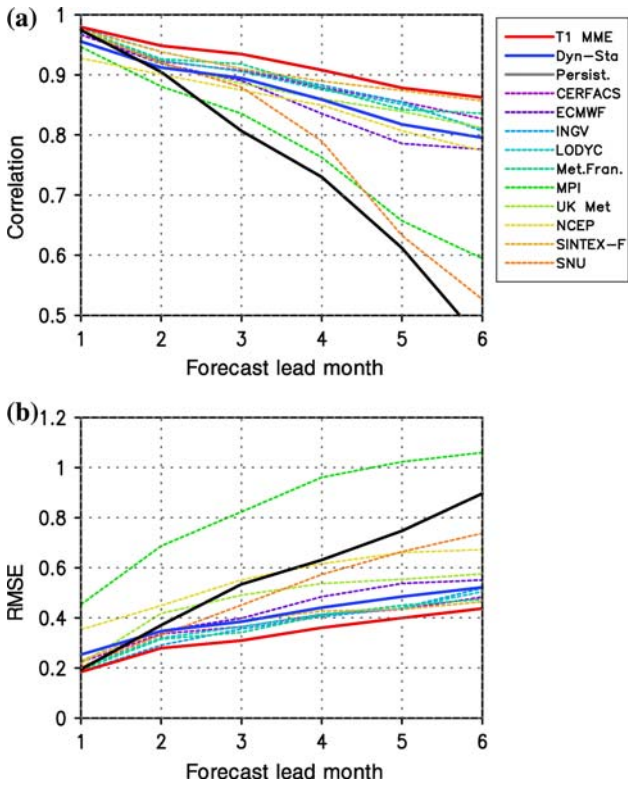

Fig. 7 **a** Anomaly correlation coefficients and **b** RMSE of NINO3.4 index during 1980–2001 with respect to lead time after removing the mean bias. The mean skill for all four cases including Feb, May, Aug, and Nov initial conditions is shown. *Black* for observation, *red* for 10 CGCM multi-model ensemble, *blue* for the Stat-Dyn forecast, *gray* for PERSIST, and *colored dots* for individual coupled models as shown in the legend, respectively

Compared with previous results using various coupled models (e.g., Kirtman et al. 2001), the performance of this set of models looks quite good in general.

Among the ten individual models in the MME, the SNU CGCM (brown dashed line) and the MPI CGCM (green dotted line) stand out as having worse skill than the other eight models. Since we have already seen that these two models are outliers in terms of their simulation of the mean annual cycle of SST in the tropical Pacific, this suggests that the erroneous climatology in specific models may have an influence on the anomaly forecast.

Breaking down the ACC by initial month, the skill clearly shows seasonality, with some suggestion of a "spring prediction barrier" (Fig. 8). The higher skill is found in the forecasts with short lead times (up to 5 months) that start in November and, even more so for longer lead forecasts, in August, compared to the forecasts that start during boreal winter and spring (February and May). The August cases have relatively low skill in the first lead time, but the decline of skill thereafter is small even through 6 months lead. Interestingly, the MME has better skill than individual models, especially for February and May cases, with a relatively slow drop in skill, despite the fast drop of skill among the individual models. The PERSIST and Dyn-Stat forecasts also have substantial seasonal dependency.

This seasonality of forecast skill has been the subject of many studies (e.g., Troup 1965; Wright 1979; Webster and Yang 1992; Xue et al. 1994; Latif et al. 1994). Many prediction schemes generally have a significant decline in skill in boreal spring with apparent skill recovery in subsequent seasons. The cause of this behavior has not yet been fully understood, and various hypotheses have been discussed. Some authors have suggested that it may be the result of relatively weak coupling between the ocean and atmosphere during boreal spring (Zebiak and Cane 1987; Battisti 1988; Goswami and Shukla 1991; Blumenthal 1991). More recent studies have emphasized the phase locking of the ENSO to the annual cycle (Balmaseda et al. 1995; Torrence and Webster 1998; An and Wang 2001), and the biennial component of ENSO (Clarke and van Gorder 1999; Yu 2005).

For a more detailed examination of forecast skill, several approaches to classify the cases are used here. First, forecast skill is assessed in terms of the intensity of the SST anomaly at the target time. In Fig. 9, cases are stratified according to the intensity of the observed SST anomaly at the target time. The warm cases are those for which the observed SST anomaly at the target time is more than half a standard deviation above the mean for all 6 lead months. For example, the NINO3.4 index exceeded 0.5 σ from April 1997 through May 1998, so the forecasts labeled May 1997, Aug 1997, and Nov 1997 would all be included in

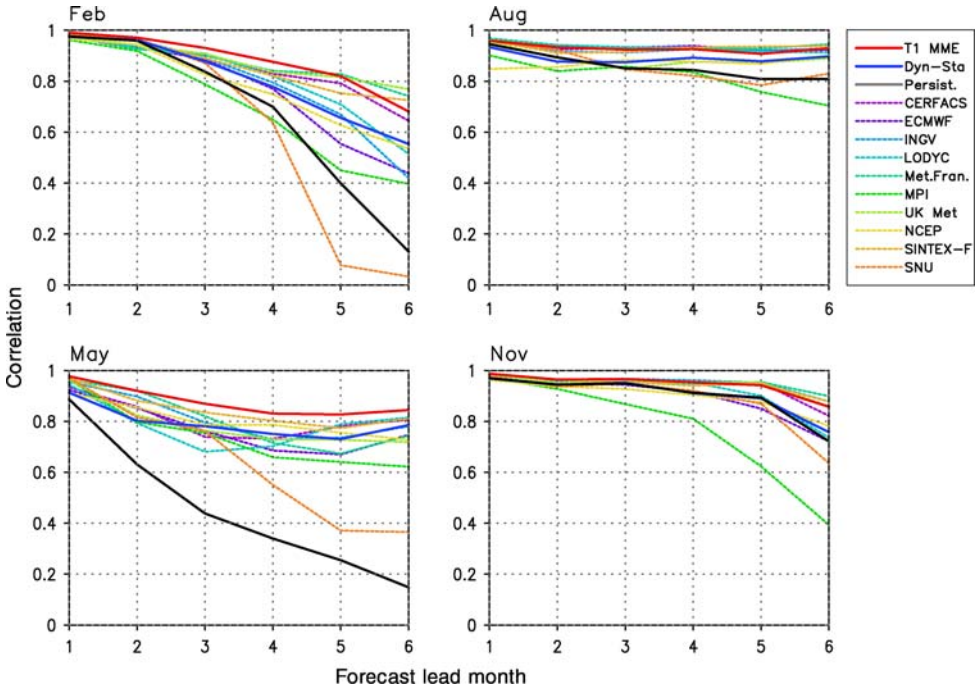

Fig. 8 Anomaly correlation coefficients of NINO3.4 index starting from Feb, May, Aug, and Nov initial conditions in the period 1980–2001

the group of warm cases, while neither the Feb 1997 nor the Feb 1998 forecasts would be. Cold cases are defined similarly. The normal cases include situations in which the absolute value of the observed SST anomaly is less than half a standard deviation. To remove the asymmetry between the amplitudes of warm and cold events that is present in the observations, the standard deviation is calculated separately for warm and cold anomalies.

Forecasts of warm and cold cases significantly outperform those for normal conditions, based on ACC. Because ENSO events are typically confined to boreal winter, this is consistent with the earlier finding that forecast skill is

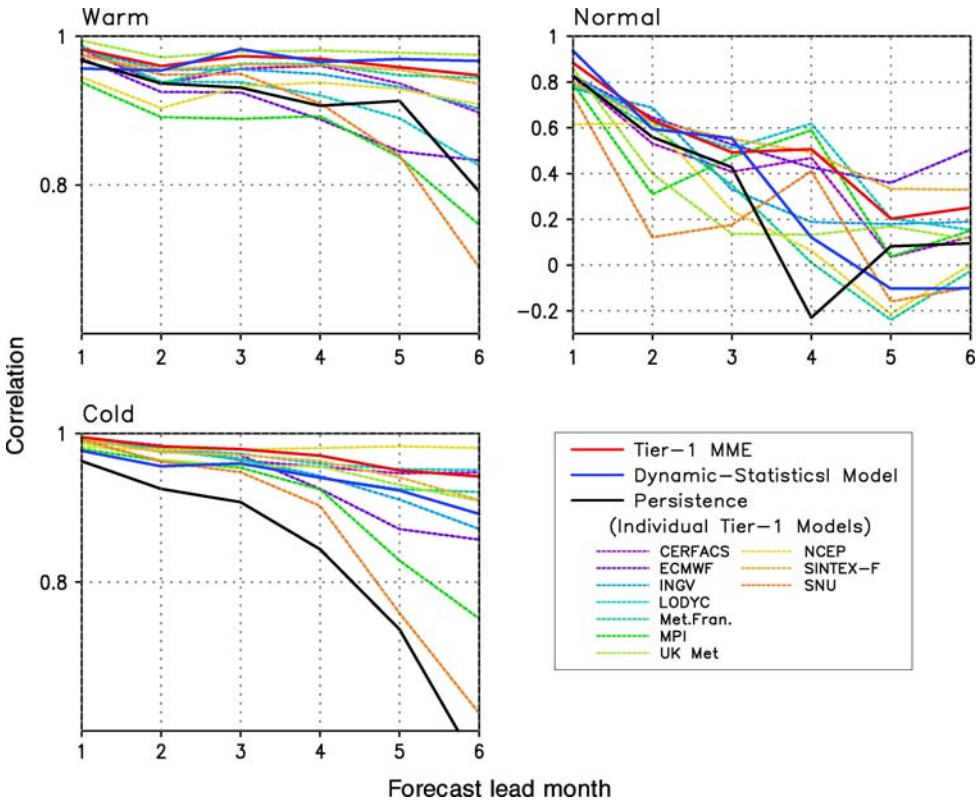

Fig. 9 Anomaly correlation coefficients of NINO3.4 index with respect to SST intensity. Three cases are classified following to intensity of the SST anomalies of target month. Warm case (*upper left*) denotes the case having observed SST anomalies with more than half standard deviation of warm SST anomalies, cold case (*lower left*) denotes half standard deviation of cold SST anomalies, and normal (*right*) denotes SST anomalies with less than half standard deviation

higher in that season. There is also some indication that forecasts of warm events have higher skill than those for cold events. It should be noted that, in general, ACC is higher for anomalous events than normal conditions since small errors will have a larger deleterious effect on ACC when the amplitude of the anomalies is relatively smaller.

As another way of examining the spring predictability barrier, we consider the phase locking between ENSO and the annual cycle. Comparisons have been made to determine how well the models forecast the various phases of ENSO. Based on monthly observed SST anomalies, cases were stratified with respect to ENSO phase at the initial time in three categories: ENSO growth, decay, and small variation. The cases that fall into these three categories are enumerated in Table 2. Because ENSO is often phase-locked to the annual cycle, the growth phase occurs most frequently for August and November initial conditions cases, and the decaying phase is most often found in the February and May cases. In Fig. 10, the skill is shown as a function of the ENSO phase at the initial time of the forecast. The growth phase of both warm and cold events is better predicted than the decay phase. Normal conditions are far worse predicted than either growing or decaying warm and cold events. Again, this is consistent with the fast drop of skill in February and May cases since these cases include more events in their decaying phase, having lower skill than those in the growth phase. The August and November forecasts have relatively better skill, when there are fewer decaying phase events.

Figure 11 shows the results for dependence of forecast skill on the month of the initial conditions, initial ENSO phase and ENSO amplitude for the MME, providing an overall perspective of forecast skill. Similar to the correlation coefficient, we also plot the normalized RMSE, which was calculated by dividing the actual RMSE by the observed standard deviation of SST for the cases in the particular category. Thus the normalized RMSE for "warm SST" events is the RMSE relative to large amplitude anomalies, while for "normal SST" events the normalized RMSE is relative to small amplitude anomalies. The MME results are by and large similar to the findings for individual models. The ACC is high for the August and November forecasts, and poorer for the February and May forecasts, although the MME does better even in the poor cases than individual models. Similarly, the normalized RMSE of the MME forecasts is lowest in August and November cases. When a warm or cold event is growing in the initial month of a forecast, the skill (ACC) is generally higher than when an event is decaying and much higher than forecasts when neither growth nor decay is occurring at the initial time. Similarly, warm and cold event growth cases are best in the RMSE measure, except at lead times of 1 and 2 months. Also, very high correlation is found for forecasts of large amplitude anomalies and these have the lowest normalized RMSE. Note that without normalization, warm/cold cases have large RMSE, i.e., the absolute errors are large. The coincidence between high correlation and large RMSE of SST anomalies during intense El Niño events suggests that the MME successfully reproduces the evolution of the SST anomaly patterns in the NINO3.4 region, but it has problems with the amplitude of the anomalies. While several studies argue that the tendency to simulate weak ENSO anomalies is a feature common to many CGCMs (e.g., AchutaRao and Sperber 2002; Fischer 2002), it is not evident in this study, because the intensity of anomalies shown in the MME is quite comparable to that of the observed by using composite analysis (not shown).

6 Summary and conclusion

An evaluation of the skill of equatorial SST forecasts made with state-of-the-art CGCMs, both in comparison with standard benchmark forecasts and MME forecasts was undertaken to determine the current level of achievable forecast skill and to quantitatively assess the efficacy of the MME approach. The hypothesis that the quality of the models forecasts of the evolution of interannual anomalies is related to their simulation of the annual mean and mean annual cycle, as a function of lead time, was also tested. To assess the forecast skill and usefulness, anomaly

Table 2 Specific cases distinguished by ENSO phase of initial time

Phase	El Niño		La Niña		Normal
	Growth	Decay	Growth	Decay	
No.	13	10	13	19	33
Jan	87	82,92,95,98	84,99	81,85,86,89,96,00,01	80,82,88,90,91,93,94,97
Apr	82,87,97	83,92,93,98	84,88,99	85,86,89,96,00	80,81,90,91,94,95,01
Jul	82,87,91,94,96		84,88,98,99	85,89,00	80,81,83,86,90,92,93,95,96,01
Oct	82,86,91,94	87,97	83,84,98,99	85,88,95,00	80,81,89,90,92,93,96,01
Initial month	Specific year				

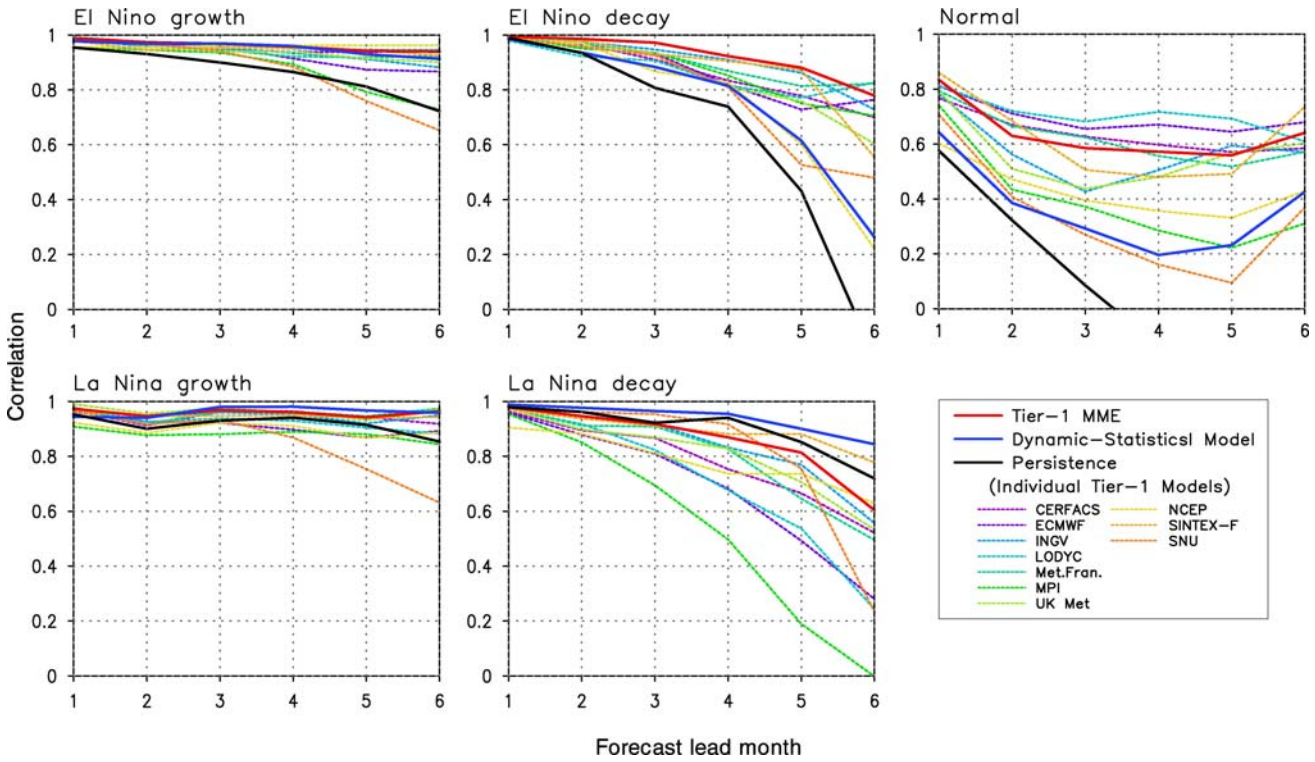

Fig. 10 Anomaly correlation coefficients of NINO3.4 index distinguished by ENSO phase of initial time. ENSO growth, decay, and Normal cases are distinguished

correlation coefficients and root-mean-squared errors were calculated in terms of ENSO intensity and phase.

Even though CGCMs tend to accurately simulate the annual mean SST climatology in the tropics (equatorial Pacific?) when realistic initial conditions are imposed, the simulated annual cycle of most models is far from perfect. The CGCMs in this study have an annual cycle in the eastern Pacific that is too weak, except for the MPI, NCEP, and UKMO models. The phase and peak amplitude of westward propagation of the annual cycle in the eastern and central equatorial Pacific are different from those of observed. The phase error tends to increase with lead time in most of models.

In coupled models, the realism of interannual variability seems to be associated with the accuracy of the simulation of the climatological annual cycle in terms of amplitude. In particular, the SNU and MPI models, which have erroneous interannual variability, are also outliers in their simulation of the climatology, including both the annual mean and the annual cycle. These two models also exhibit the worst forecast skill. However, none of the models we examined attain good performance in simulating the mean annual cycle of SST, even with the advantage of starting from realistic initial conditions. The error in models' simulation of the interannual SST variability is quite noticeable in most coupled models. The interannual variability of the coupled system is directly related to the phase locking of ENSO to the annual cycle and also to ENSO forecast skill. Consequently, there is substantial potential for model improvement. Finally, all models have a noticeable degradation of skill with lead time, including a prevalent phase shift of the center of action.

To assess the forecast skill, anomaly correlations and RMSE were calculated for each of the models and the MME. The MME has several distinctive features. The correlation skill of the tier-1 MME forecast of NINO3.4 SST anomalies reaches 0.86 at 6-months lead. The MME outperforms not only the individual coupled models but also both the persistence and the dynamic-statistical model. Most state-of-the-art CGCMs beat persistence. The skill difference, while not huge, is quite noticeable.

The forecast skill depends strongly on season. The skill of forecasts starting in February and May drops faster than that of those starting in August and November. This is consistent with the spring predictability barrier also described in many studies (e.g., Latif et al. 1994). Interestingly, the forecast skill advantage of the MME is more pronounced in the February and May initial conditions cases, with a wide range of skill in the individual models.

This seasonality is associated with the phase of ENSO and the magnitude of the SST anomaly. The stronger warm ENSO (El Niño) cases are better predicted in terms of

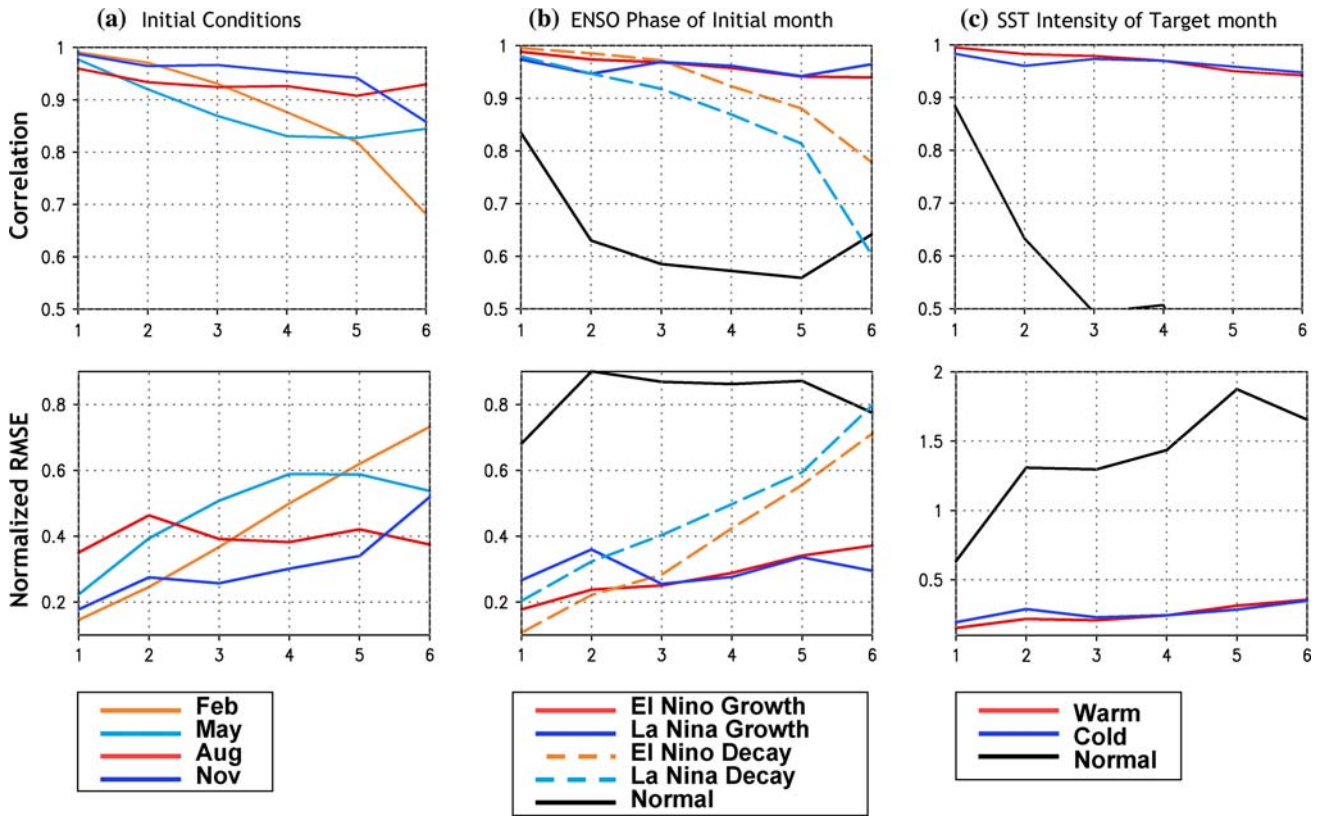

Fig. 11 Anomaly correlation coefficients (*top*) and normalized root-mean-square error (*bottom*) of Multi-model ensemble of NINO3.4 index during 1980–2001 with respect to lead time. Collection by **a** initial conditions, **b** ENSO phase of initial time, and **c** SST intensity of target time are shown

higher anomaly correlation with the observed. There are substantial differences in forecast skill between normal conditions and ENSO cases. Interestingly, the advantage in terms of anomaly correlation does not appear to carry over the RMSE; however, this is not necessarily a result of deficient amplitude ENSO events in the forecasts. The growth phases of both the warm and cold events are better predicted than the corresponding decaying phases. ENSO-neutral years are far worse predicted than growing warm and cold events. Predictions starting from February and May tend to have more events in the decaying phase than those in August and November, which is consistent with the fact that forecasts in these months yield much lower skill.

Based on these results, we conclude that accurately predicting the strength and timing of ENSO events continues to be a critical challenge for dynamical models of all levels of complexity. Improvement in the ability of couple models to simulate the mean climate, including the mean annual cycle, and interannual variability is very much required to improve forecasts. Improved models, data and initialization strategies are required to address the problem of predicting tropical eastern Pacific SST anomalies. Prediction of regional precipitation and circulation will not be possible without accurate predictions of SST anomalies.

Acknowledgments This research was supported by APEC Climate Center (APCC) as a part of APCC International research project. The second author was supported by grants from the National Science Foundation (ATM-0332910), the National Oceanic and Atmospheric Administration (NA04OAR4310034) and the National Aeronautics and Space Administration (NNG04GG46G). The 5th and 7th authors were supported by the SRC program of the Korean Science and Engineering Foundation and the Brain Korea 21 project. We would like to thank Duane E. Waliser and one anonymous reviewer for their constructive comments on the earlier version of this manuscript.

References

AchutaRao K, Sperber KR (2002) Simulation of the El Niño-Southern Oscillation: results from the coupled model intercomparison project. Clim Dyn 19:191–209

AchutaRao K, Sperber KR (2006) ENSO Simulation in coupled ocean–atmosphere models: are the current models better? Clim Dyn 27:1–15

An SI, Wang B (2001) Mechanisms of locking the El Nino and La Nina mature phases to boreal winter. J Clim 14:2164–2176

Balmaseda MA, Davey MK, Anderson DLT (1995) Decadal and seasonal dependence of ENSO prediction skill. J Clim 8:2705–2715

Barnett TP, Graham NE, Cane MA, Zebiak SE, Dolan SC, O'Brien J, Legler D (1988) On the prediction of the El Niño of 1986–1987. Science 241:192–196

Barnston AG, Glantz M, He Y (1999) Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997–98 El Niño and the 1998 La Niña onset. Bull Amer Met Soc 80:217–243

Barnston AG, van den Dool HM, Zebiak SE, Barnett TP, Ji M, Rodenhuis DR, Cane MA, Leetmaa A, Graham NE, Ropelewski CR, Kousky VE, O'Lenic EA, Livezey RE (1994) Long-lead seasonal forecasts—Where do we stand? Bull Am Met Soc 75:2097–2114

Barnston AG, Chelliah M, Goldenberg SB (1997) Documentation of a highly ENSO-related SST region in the equatorial Pacific. Atmos–Ocean 35:367–383

Battisti DS, Hirst AC (1989) Interannual variability in the tropical atmosphere–ocean system: influence of the basic stated, ocean geometry, and non-linearity. J Atmos Sci 46:1687–1712

Battisti DS (1988) Dynamics and thermodynamics of a warming event in a coupled tropical atmosphere–ocean model. J Atm Sci 45:2889–2919

Bengtsson L, Schlese U, Roeckner E, Latif M, Barnett TP, Graham NE (1993) A two-tiered approach to long-range climate forecasting. Science 261:1027–1029

Blumenthal MB (1991) Predictability of a coupled ocean–atmosphere model. J Clim 4:766–784

Bradley RS, Diaz HF, Kiladis GN, Eishcheid JK (1987) ENSO signal in continental temperature and precipitation records. Nature 327:497–501

Cane MA, Zebiak SE, Dolan SC (1986) Experimental forecasts of El Niño. Nature 321:827–832

Chen D, Zeiak SE, Busalacchi AJ, Cane MA (1995) An improved procedure for El Niño forecasting: implications for predictability. Science 269:1699–1702

Clarke AJ, Van Gorder S (1999) The correlation between the boreal spring Southern Oscillation persistence barrier and biennial variability. J Clim 12:610–620

Davey M, Huddleston M, Sperber KR, Braconnot P, Bryan F, Chen D, Colman RA, Cooper C, Cubasch U, Delecluse P, DeWitt D, Fairhead L, Flato G, Gordon C, Hogan T, Ji M, Kimoto M, Kitoh A, Knutson TR, Latif M, Le Treut H, Li T, Manabe S, Mechoso CR, Meehl GA, Oberhuber J, Power SB, Roeckner E, Terray L, Vintzileos A, Voss R, Wang B, Washington WM, Yoshikawa I, Yu JY, Yukimoto S, Zebiak SE (2002) A study of coupled model climatology and variability in tropical ocean regions. Clim Dyn 18:403–420

Delecluse P, Davey M, Kitamura Y, Philander S, Suarez M, Bengtsson L (1998) TOGA review paper: coupled general circulation modeling of the tropical Pacific. J Geophys Res 103:14357–14373

Delecluse P, Madec G (1999) Ocean modeling and the role of the ocean in the climate system. In: Holland WR, Joussaume S, David F (eds) Modeling the earth's climate and its variability. Elsevier, Amsterdam, pp 237–313

Deque M (2001) Seasonal predictability of tropical rainfall: Probabilistic formulation and validation. Tellus 53A:500–512

Fischer M (2002) ENSO predictions with coupled ocean–atmosphere models. In: Pinardi N, Woods J (eds) Ocean forecasting. Springer, Berlin, pp 307–338

Fu X, Wang B, Li T, McCreary J (2004) Coupling between northward-propagating, intraseasonal oscillations and sea surface temperature in the Indian Ocean. J Atmos Sci 60:1733–1753

Goddard L, Mason SJ, Zebiak SE, Ropelewski CF, Basher R, Cane MA (2001) Current approaches to seasonal-to-interannual climate predictions. Int J Climatol 21:1111–1152

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. Clim Dyn 16:147–168

Goswami BN, Shukla J (1991) Predictability of the coupled ocean–atmosphere model. J Clim 4:3–22

Graham NE, Michaelson J, Barnett TP (1987) An investigation of the El Niño-Southern Oscillation cycle with statistical models: II. Model results. J Geophys Res 92:14271–14289

Gregory D, Morcrette JJ, Jakob C, Beljaars ACM, Stockdale T (2000) Revision of convection, radiation and cloud schemes in the ECMWF Integrated Forecasting System. Q J R Meteor Soc 126:1685–1710

Gualdi S, Alessandri A, Navarra A (2005) Impact of atmospheric horizontal resolution on El Niño Southern Oscillation forecasts. Tellus 57A:357–374

Ji M, Behringer DW, Leetmaa A (1998) An improved coupled model for ENSO prediction and implications for ocean initialization. Part II: The coupled model. Mon Weather Rev 126:1022–1034

Ji M, Leetmaa A (1997) Impact of data assimilation on ocean initialization and El Niño prediction. Mon Weather Rev 125:742–753

Jin EK, Kinter JL III (2007) Characteristics of tropical pacific SST predictability in coupled GCM forecasts using the NCEP CFS. J Clim (Submitted)

Jochum M, Murtugudde R (2004) Internal variability of the tropical pacific ocean. Geophys Res Lett 31:L14309. doi:10.1029/2004GL020488

Kanamitsu M, Kumar A, Juang HM, Schemm JK, Wang W, Yang F, Hong SY, Peng P, Chen W, Moorthi S, Ji M (2002) NCEP dynamical seasonal forecast system 2000. Bull Amer Met Soc 83:1019–1037

Kirtman BP (2003) The COLA anomaly coupled model: ensemble ENSOprediction. Mon Weather Rev 131:2324–2341

Kirtman BP, Shukla J, Balmaseda M, Graham N, Penland C, Xue Y, Zebiak SE (2001) Current status of ENSO forecast skill: A report to the Climate Variability and Predictability (CLIVAR) Working Group on Seasonal to Interannual Prediction. WCRP Informal Report No 23/01, 31pp

Kirtman BP, Shukla J, Huang B, Zhu Z, Schneider EK (1997) Multiseasonal prediction with a coupled tropical ocean global atmosphere system. Mon Weather Rev 125:89–808

Krishnamurti TN, Kishitawal CM, Zhang Z, Larow T, Bachiochi D, Williford E (2000) Multimodel ensemble forecasts for weather and seasonal climate. J Clim 13:4196–4216

Kug JS, Kang IS, Choi DH (2008) Seasonal climate predictability with tier-one and tier-two prediction Systems. Clim Dyn. doi:10.1007/s00382-007- 0264-7

Kug JS, Lee JY, Kang IS (2007) Global sea surface temperature prediction using a multi-model ensemble. Mon Wea Rev 135:3239–3247

Landsea CW, Knaff JA (2000) How much skill was there in forecasting the very strong 1997–98 El Niño? Bull Amer Met Soc 81:2107–2120

Latif M, Barnett TP, Cane MA, Flügel M, Graham NE, Von Storch H, Xu JS, Zebiak SE (1994) A review of ENSO prediction studies. Clim Dyn 9:167–179

Latif M, Sperber K, Arblaster J, Braconnot P, Chen D, Colman A, Cubasch U, Cooper C, Delecluse P, DeWitt D, Fairhead L, Flato G, Hogan T, Ji M, Kimoto M, Kitoh A, Knutson T, Le Treut H, Li T, Manabe S, Marti O, Mechoso C, Meehl G, Power S, Roeckner E, Sirven J, Terray L, Vintzileos A, Vob R, Wang B, Wasington W, Yoshikawa I, Yu JY, Zebiak SE (2001) ENSIP: the El Niño simulation intercomparison project. Clim Dyn 18:255–276

Lau KM, Waliser DE (2005) Intraseasonal variability in the atmosphere–ocean climate system. Springer, Berlin

Leetmaa A, Ji M (1989) Operational hindcasting of the tropical Pacific. Dyn Atmos Oceans 13:465–490

Luo JJ, Masson S, Behera S, Shingu S, Yamagata T (2005) Seasonal climate predictability in a coupled OAGCM using a different approach for ensemble Forecasts. J Clim 18:4474–4497

Madec G, Delecluse P, Imbard M, Levy C (1997) OPA release 8. Ocean general circulation model reference manual. LODYC Internal Rep., Paris, France, 200pp

Madec G, Delecluse P, Imbard M, Levy C (1998) OPA version 8.1 Ocean general circulation model reference manual. LODYC Tech Rep. 11, Paris, France, 91pp

Marsland SJ, Haak H, Jungclaus JH, Latif M, Roske F (2003) The Max-Plank-Institute global ocean/sea ice model with orthogonal curvilinear coordinates. Ocean Modell 5:91–127

Mason SJ, Goddard L, Graham NE, Yulaeva E, Sun L, Arkin PA (1999) The IRI seasonal climate prediction system and the 1997/98 El Niño event. Bull Amer Met Soc 80:1853–1873

Mechoso CR, Robertson AW, Barth N, Davey MK, Delecluse P, Gent PR, Ineson S, Kirtman BP, Latif M, Le Treut H, Nagai T, Neelin JD, Philander SGH, Polcher J, Schopf PS, Stockdale TN, Suarez MJ, Terray L, Thual O, Tribbia JJ (1995) The seasonal cycle over the tropical Pacific in coupled atmosphere–ocean general-circulation models. Mon Weather Rev 123:2825–2838

Murtugudde R, Beauchamp J, McClain CR, Lewis M, Busalacchi AJ (2002) Effects of penetrative radiation on the upper tropical ocean circulation. J Clim 15:470–486

Neelin JD, Dijkstra HA (1995) Ocean–atmosphere interaction and the tropical climatology. Part I: the dangers of flux correction. J Clim 8:1325–1342

Neelin JD, Latif M, Allaart MAF, Cane MA, Cubasch U, Gates WL, Gent PR, Ghil M, Gordon C, Lau NC, Mechoso CR, Meehl GA, Oberhuber JM, Philander SGH, Schopf PS, Sperber KR, Sterl A, Tokioka T, Tribbia J, Zebiak SE (1992) Tropical air-sea interaction in general circulation models. Clim Dyn 7:73–104

Pacanowski RC (1995) MOM 2.2 manual. NOAA/GFDL

Pacanowski RC, Griffies SM (1998) MOM 3.0 manual. NOAA/GFDL. http://www.gfdl.noaa.gov/_smg/MOM/web/guide_parent/guide_parent.html

Palmer TN, Alessandri A, Andersen U, Cantelaube P, Davey M, Delecluse P, Deque M, Diez E, Doblas-Reyes FJ, Feddersen H, Graham R, Gualdi S, Gueremy JF, Hagedorn R, Hoshen M, Keenlyside N, Latie M, Lazar A, Maisonnave E, Marletto V, Morse AP, Orfila B, Rogel P, Terres JM, Thomson MC (2004) Development of a European multi-model ensemble system for seasonal to interannual prediction (DEMETER). Bull Amer Met Soc 85:853–872

Palmer TN, Brankovic C, Richardson DS (2000) A probability and decision-model analysis of PROBOST seasonal multi-model ensemble integrations. Q J R Meteor Soc 126:2013–2034

Philander SGH (1990) El Niño, La Niña and the Southern Oscillation. Academic Press, San Diego, pp 293

Pope VD, Gallani ML, Rowntree PR, Stratton RA (2000) The impact of new physical parameterizations in the Hadley Centre climate model: HadAM3. Clim Dyn 16:123–146

Rasmusson EM, Carpenter TH (1982) Variations in tropical sea surface temperature and surface wind fields associated with Southern Oscillation/El Niño. Mon Weather Rev 110:354–384

Rayner NA, Parker DE, Horton EB, Folland CK, AlenxanderLV, Rowell DP, Kent EC, Kaplan A (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J Geophys Res 108:4407–4443

Roeckner E (1996) The atmospheric general circulation model ECHAM-4: Model description and simulation of present-day climate. Max-Planck-Institut fur Meteorologie Tech. Rep. 218, Hamburg, Germany, 90pp. Available from Max-Planck Institut fur Meteorologie, Bundesstr. 55, D-20146 Hamburg, Germany

Ropelewski CF, Halpert MS (1987) Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. Mon Weather Rev 115:1606–1626

Ropelewski CF, Halpert MS (1989) Precipitation patterns associated with the high index phase of the Southern Oscillation. J Clim 2:268–284

Rosati A, Miyakoda K, Gudgel R (1997) The impact of ocean initial conditions on ENSO forecasting with a coupled model. Mon Weather Rev 125:754–772

Saha S, Nadiga S, Thiaw C, Wang J, Wang W, Zhang Q, van den Dool HM, Pan HL, Moorthi S, Behringer D, Stokes D, White G, Lord S, Ebisuzaki W, Peng P, Xie P (2005) The NCEP Cliamte Forecast System. J Clim 15:3483–3517

Schneider EK, DeWitt DG, Rosati A, Kirtman BP, Ji L, Tribbia JJ (2003) Retrospective ENSO forecasts: sensitivity to atmospheric model and ocean resolution. Mon Weather Rev 131:3038–3060

Schneider EK, Huang B, Zhu Z, DeWitt DG, Kinter III JL, Kirtman BP, Shukla J (1999) Ocean data assimilation, initialization and prediction of ENSO with a coupled GCM. Mon Weather Rev 127:1187–1207

Schopf PS, Suarez MJ (1988) Vacillations in a coupled ocean–atmosphere model. J Atmos Sci 45:549–566

Shukla J (1998) Predictability in the Midst of Chaos: A Scientific Basis for Climate Forecasting. Science 282:728–731

Stern W, Miyakoda K (1995) Feasibility of seasonal forecasts inferred from multiple GCM simulations. J Clim 8:1071–1085

Stockdale TN (1997) Coupled ocean–atmosphere forecasts in the presence of climate drift. J Clim 10:809–818

Stockdale TN, Anderson DLT, Alves JOS, Balmaseda MA (1998) Global seasonal rainfall forecasts using a coupled ocean–atmosphere model. Nature 392:370–373

Torrence C, Webster PJ (1998) The annual cycle of persistence in the El Niño /Southern Oscillation, Q J R Meteor Soc 124:1985–2004

Trenberth KE, Branstator GW, Karoly D, Kumar A, Lau NC, Ropelewski C (1998) Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. J Geophys Res 103:14291–14324

Troup AJ (1965) The "southern oscillation". Q J R Meteor Soc 91:490–506

Wang B, Ding QH, Fu XH, Kang IS, Jin K, Shukla J, Doblas-Reyes F (2005) Fundamental challenge in simulation and prediction of summer monsoon rainfall. Geophys Res Lett 32:L15711. doi: 10.1029/2005GL022734

Wang B, Lee JY, Kang IS, Shukla J, Kug JS, Kumer A, Schemm J, Luo JJ, Yamagata T, Park CK (2007) How accurately do coupled climate models predict the Asian-Australian Monsoon interannual variability? Clim Dyn (Submitted)

Wang G, Kleeman R, Smith N, Tseitkin F (2002) The BMRC coupled general circulation model ENSO forecast system. Mon Weather Rev 130:975–991

Webster PJ (1995) The annual cycle and the predictability of the tropical coupled ocean–atmosphere system, Meteor Atmos Phys 56:33–55

Webster PJ, Yang S (1992) Monsoon and ENSO: Selectively interactive systems. Q J R Meteor Soc 118:877–925

Wolff JE, Maier-Reimer E, Legutke S (1997) The Hamburg Ocean primitive equation model. Deutsches Klimarechenzentrum Tech. Rep. 13, Hamburg, Germany, 13pp. Available from Model and Data Group c/o Max-Planck Institut fur Meteorologie, Bundesstr. 55, D-20146 Hamburg, Germany

Wright PB (1979) Persistence of rainfall anomalies in the central Pacific. Nature 277:371–374

Xue Y, Cane MA, Zebiak SE, Blumenthal MB (1994) On the prediction of ENSO: a study with a low-order Markov model. Tellus 46A:512–528

Yu JY (2005) Enhancement of ENSO's persistence barrier by biennial variability in a coupled atmosphere–ocean general circulation model. J Geophys Res 32:L13707. doi:10.1029/2005GL023406

Zebiak SE, Cane MA (1987) A model El Niño-Scouthern Oscillation. Mon Weather Rev 115:2262–2278

Zheng Y, Walier DE, Stern WF, Jones C (2004) The role of coupled sea surface temperatures in the simulation of the tropical intraseasonal oscillation. J Clim 21:4109–4134