

# Seasonal prediction skill of ECMWF System 4 and NCEP CFSv2 retrospective forecast for the Northern Hemisphere Winter

Hye-Mi Kim · Peter J. Webster · Judith A. Curry

Received: 16 September 2011 / Accepted: 12 April 2012 / Published online: 4 May 2012  
© The Author(s) 2012. This article is published with open access at Springerlink.com

**Abstract** The seasonal prediction skill for the Northern Hemisphere winter is assessed using retrospective predictions (1982–2010) from the ECMWF System 4 (Sys4) and National Center for Environmental Prediction (NCEP) CFS version 2 (CFSv2) coupled atmosphere–ocean seasonal climate prediction systems. Sys4 shows a cold bias in the equatorial Pacific but a warm bias is found in the North Pacific and part of the North Atlantic. The CFSv2 has strong warm bias from the cold tongue region of the eastern Pacific to the equatorial central Pacific and cold bias in broad areas over the North Pacific and the North Atlantic. A cold bias in the Southern Hemisphere is common in both reforecasts. In addition, excessive precipitation is found in the equatorial Pacific, the equatorial Indian Ocean and the western Pacific in Sys4, and in the South Pacific, the southern Indian Ocean and the western Pacific in CFSv2. A dry bias is found for both modeling systems over South America and northern Australia. The mean prediction skill of 2 meter temperature (2mT) and precipitation anomalies are greater over the tropics than the extra-tropics and also greater over ocean than land. The prediction skill of tropical 2mT and precipitation is greater in strong El Niño Southern Oscillation (ENSO) winters than in weak ENSO winters. Both models predict the year-to-year ENSO variation quite accurately, although sea surface temperature trend bias in CFSv2 over the tropical Pacific results in lower prediction skill for the CFSv2 relative to the Sys4. Both models capture the main ENSO teleconnection pattern of strong anomalies over the tropics, the North Pacific and the North America. However, both models have

difficulty in forecasting the year-to-year winter temperature variability over the US and northern Europe.

## 1 Introduction

Despite the chaotic internal dynamics of the atmosphere, the time average of atmospheric variables is predictable to some degree due to those components that have slow variations on time scales from months to seasons. The socioeconomic importance of accurate seasonal climate prediction has motivated development of better seasonal prediction systems. Recently, the development of coupled ocean–atmosphere dynamical model prediction systems has provided important advances in seasonal predictability (Kumar et al. 2005; Wang et al. 2005a; Kug et al. 2008). Several international projects have been undertaken to compare coupled climate predictions, including the Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER) (Palmer et al. 2004) and Asia–Pacific economic cooperation climate center (APCC)/climate prediction and its application to society (CliPAS) projects (Wang et al. 2009). Seasonal prediction skill and the model performance have been examined based on retrospective predictions of DEMETER and APCC/CliPAS (Jin et al. 2008; Kim et al. 2008; Kug et al. 2008; Wang et al. 2009; Lee et al. 2010).

Operational coupled seasonal forecast systems include Climate Forecast System from the National Center for Environmental Prediction (NCEP CFS) (Saha et al. 2006), the Australian POAMA (Wang et al. 2001), European Centre for Medium-Range Weather Forecasts (ECMWF), UK Meteorological Office and Meteo-France (Palmer et al. 2004). Operational climate forecast centers are now updating their seasonal prediction systems with improved

H.-M. Kim (✉) · P. J. Webster · J. A. Curry  
School of Earth and Atmospheric Science, Georgia Institute of Technology, 311 Ferst Dr., Atlanta, GA 30332, USA  
e-mail: hyemi.kim@eas.gatech.edu

physics and increased resolution. This study focuses on the ECMWF and NCEP CFS seasonal forecasting systems. ECMWF has been operating a seasonal forecast system since 1997 and the operational system, known as System 3, was introduced in March 2007. System 3 shows greater prediction skill for the sea surface temperature (SST) in the eastern Pacific and equatorial Indian Ocean than previous ECMWF operational systems (systems 1 and 2) (Stockdale et al. 2011). The ECMWF has now upgraded its operational seasonal forecasts from System 3 to System 4 with the later version being operational since late 2011. In the upgrade, it utilizes the use of the most recent atmospheric model version, higher resolution forecasts with a higher top of the atmosphere, more ensemble members and a larger reforecast data set (Molteni et al. 2011).

The NCEP CFS has been making coupled ocean–atmosphere forecasts since 2004. Skill of the CFS model has been examined in simulating and predicting El Niño–Southern Oscillation (ENSO) variability (Wang et al. 2005b), Asian–Australian/Indian monsoon (Yang et al. 2008; Wang et al. 2008; Pattanaik and Kumar 2010) and climatic variation in the US (Yang et al. 2009). The NCEP CFS version 2 (CFSv2, <http://cfs.ncep.noaa.gov/cfsv2.info/>) represents a substantial change to all aspects of the forecast system including model components, data assimilation system and ensemble configuration. The MJO simulation shows improvement in CFSv2 owing to a positive response to upgrades in the initial state compared to CFSv1 (Weaver et al. 2011).

The seasonal predictions of individual coupled seasonal forecast systems has been analyzed separately for various target of seasons, different time periods and regions with wide range of variables using regression and correlation analysis, composite analysis and principal component analysis (Wang et al. 2005b; Saha et al. 2006; Yang et al. 2008, 2009; Lee et al. 2010; Tompkins and Feudale 2010; Wang et al. 2010; Stockdale et al. 2011). However, the ECMWF and NCEP CFS seasonal forecast systems have not been compared with the same validation matrix. The choice of one model over the other, or the use of both models in a multi-model ensemble requires information that compares the predictions of both models and the determination of the bias of each model. We compare the simulation ability and seasonal prediction skill of the two systems using the same validation matrix. The results of this comparison may be useful for the community as a benchmark for future generations of seasonal prediction systems, and may provide valuable information for forecast providers and decision makers that use seasonal forecast products.

In this paper, we focus on the Northern Hemisphere (NH) winter when the magnitude of ENSO anomalies and teleconnections to the extratropics can be particularly high

(Peng et al. 2000). A companion paper for the NH summer has also been prepared. In particular, this study addresses how well the ECMWF System 4 and NCEP CFSv2 simulate the spatio-temporal climate variability for the NH winter. Section 2 introduces details of reforecast and observational data used in the present study. Section 3 examines the simulated climates and the seasonal prediction skill of surface temperature and precipitation. Section 4 examines the prediction of ENSO whilst Sect. 5 focuses on the prediction of the winter teleconnection patterns. A summary of the results and a general discussion are provided in Sect. 6.

## 2 Retrospective forecasts and observation data

The ECMWF System 4 (hereafter Sys4) and the NCEP CFSv2 (hereafter CFSv2) are fully coupled general circulation models (GCMs) that provide operational seasonal predictions. Both systems provide reforecast simulations for the purpose of evaluating and calibrating the model simulations. The ECMWF System 4 seasonal reforecasts, commencing in 1981, include 15 member ensembles consist of 7 month simulations initialized on the 1st day of every month. The atmospheric initial conditions come from ERA Interim reanalysis for the period 1981–2010. A new ocean model (NEMO) and ocean data assimilation system (NEMOVAR) is implemented, improving the mean state and SST forecast skill in the East Pacific and Tropical Atlantic oceans. Details for the ECMWF System 4 can be found in Molteni et al. (2011) and <http://www.ecmwf.int/products/forecasts/seasonal/documentation/system4>. The NCEP CFSv2 is an upgraded version of CFSv1 (Saha et al. 2006). CFSv2 produces a set of 9-month reforecast initiated from every 5th day with four ensemble members for the period 1982–2010. Initial conditions for the atmosphere and ocean come from NCEP Climate Forecast System Reanalysis (CFSR, Saha et al. 2010).

As prediction skill depends strongly on the ensemble size (Kumar and Hoerling 2000), we match the ensemble size, as well as lead-time for the comparison of the Sys4 and CFSv2 forecasts. The Sys4 reforecast consists of 15 ensembles initialized on November 1st and for CFSv2 16 member ensembles initialized from October 23rd to November 7th from the target variables and those from December to February (DJF), which we define as the NH winter. For example, 1997 winter is an average of December 1997 and January and February of 1998. A total of 28 boreal winters from 1982/1983 to 2009/2010 are examined in this study.

For the forecast evaluation, SST data is obtained from monthly NOAA Optimum Interpolation (OI) SST V2 (Reynolds et al. 2002). The air temperature at 2 meter

(2mT), mean sea level pressure (SLP), and geopotential height at 500 hPa data are obtained from the CFS reanalysis and ERA-Interim reanalysis products (Berrisford et al. 2009) from 1981. The CFSR is a major improvement over the first generation NCEP reanalyses (NCEP R1 and R2) as it is the product of a coupled ocean–atmosphere–land system at higher spatial resolution (Higgins et al. 2010; Saha et al. 2010). ERA-Interim (hereafter ERA) is the latest global atmospheric reanalysis produced by the ECMWF and shows improvements on ERA-40 (Uppala et al. 2005) due to the use of four-dimensional data assimilation (4D-Var), higher horizontal resolution, and bias correction of satellite radiance data (Dee and Uppala 2009; Dee et al. 2011). Global Precipitation Climatology Project (GPCP) version 2.1 combined precipitation dataset (Adler et al. 2003) is used as the validation dataset. It has to be noted that there are substantial differences in trends across different reanalyses (Ebisuzaki and Zhang 2011; Zhang et al. 2012).

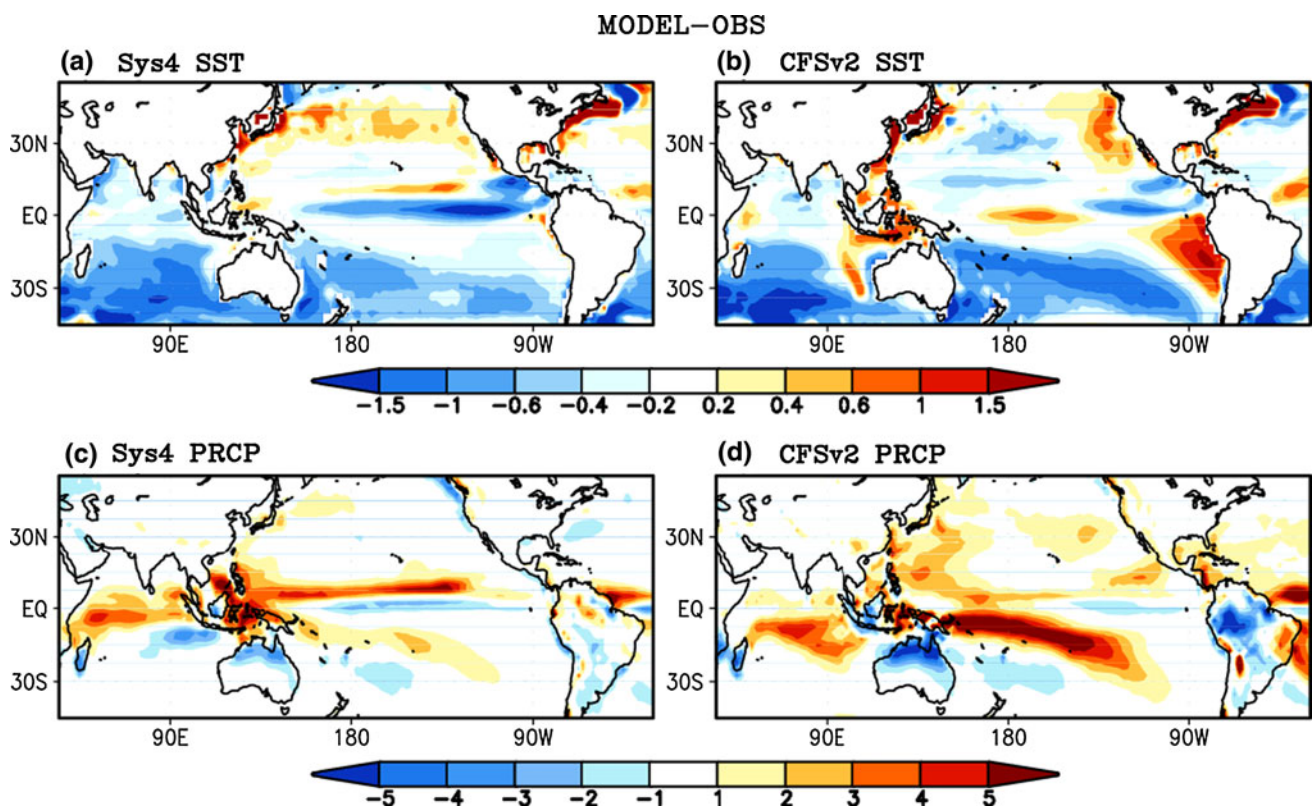
### 3 Seasonal prediction skill

Here, we examine the capability of the systems in simulating the spatial patterns of seasonal climatology and their

predictive skill of seasonal anomalies. The prediction skill is calculated as an anomaly correlation based on the ensemble mean of each seasonal prediction and the target observations.

First, we examine the model bias for SST mean state. The long-term mean or climatology of the 28 year simulations of NH winter SST for each model is compared with observations. The SST climatology in both reforecasts generally matches the observed features of variability (not shown). The warm pool and the cold tongue in the equatorial eastern Pacific are well captured in both models. However, systematic biases are found in both simulations and are shown in Fig. 1a, b. In the Sys4 reforecast simulations, a cold bias is found from the equatorial western to eastern Pacific, whereas a warm bias is found in the North Pacific and part of the North Atlantic. The CFSv2, on the other hand, has strong warm bias from the cold tongue region to the equatorial central Pacific and cold bias in broad areas over the North Pacific and North Atlantic. A cold bias over the broad region in the Southern Hemisphere is common in both modeling systems.

Figure 1c, d shows the bias for winter mean precipitation (PRCP). The spatial pattern of the precipitation climatology in both Sys4 and CFSv2 are similar to the observation but include systematic biases. In Sys4, excessive precipitation is



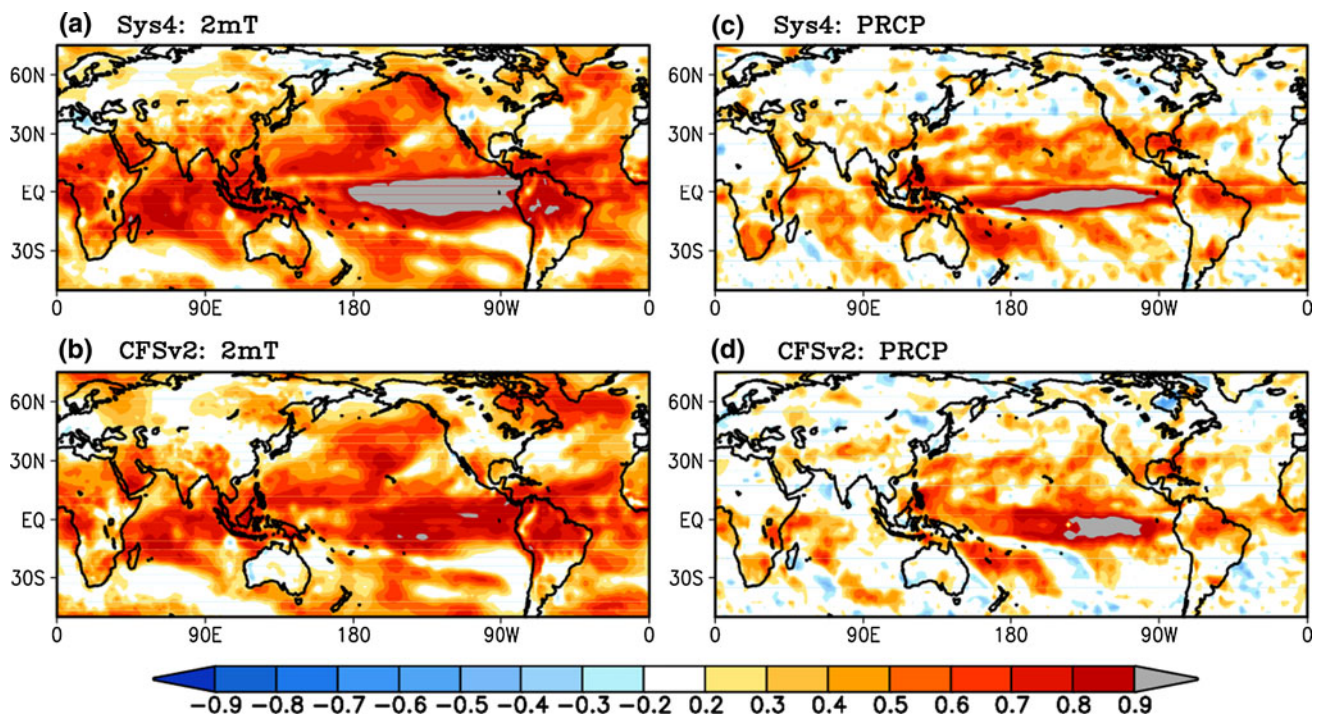
**Fig. 1** Climatological winter mean (DJF) bias (model-observation) of the SST (°C) for **a** Sys4, **b** CFSv2 and of PRCP (mm/day) for **c** Sys4 and **d** CFSv2



found along the Inter-Tropical Convergence Zone (ITCZ), equatorial Indian Ocean and western Pacific. In CFSv2, a strong wet bias is found along the South Pacific Convergence Zone (SPCZ) and the southern Indian Ocean as well as the western Pacific and dry biases are shown over the South America and the northern Australia consistent with Weaver et al. (2011). Wet bias in East Asia and the equatorial Atlantic is common in both systems.

To examine seasonal prediction skill, the correlation coefficients between reanalysis and reforecast anomalies are calculated for the ensemble mean determined from 28 winter seasons. Figure 2 shows the correlation coefficients for 2 meter temperature (2mT) and precipitation (PRCP) anomaly for each modeling system compared to ERA and GPCP. In both systems, the prediction skill for 2mT and PRCP is greater over the tropics than over the extra-tropics and greater over ocean than over land (Peng et al. 2000, 2011). 2mT has its greatest prediction skill in the tropical belt, especially in the ENSO region. The South Indian Ocean, the North Pacific and the equatorial North Atlantic also show high skill in both systems. There is almost no skill near the east coast of North America, a common problem in both systems (Fig. 2a, b). Prediction skill of precipitation in both reforecasts is generally lower than 2mT, but it also shows greatest skill over the equatorial Pacific which is influenced by ENSO (Fig. 2c, d).

A critical issue in evaluating the reforecast is the choice of the reanalysis dataset used for model evaluation. To examine the sensitivity of the prediction skill to different reanalysis datasets, we compare the 2mT prediction skill for each system with the CFSR which is used as initial conditions in CFSv2, and the ERA which is used as initial conditions in ECMWF System 4. Figure 3 shows the difference of 2mT prediction skill when ERA and CFSR is used as verification fields (ERA-CFSR) over 27 winters from 1982/1983 to 2008/2009. In the Sys4 reforecast, the skill decreases over part of North Atlantic and Indian Ocean when the model is compared with ERA than with CFSR. In CFSv2, most of the tropical ocean area shows a large decrease in prediction skill when ERA is used for verification. Compared to the an evaluation against the CFSR, the equatorial Indian Ocean, west coast of Africa, the equatorial Atlantic Ocean and the western Pacific show decrease in skill. To analyze the discrepancy on the two reanalysis, the correlation coefficient between ERA and CFSR 2mT anomaly is calculated over the 27 year DJF mean (Fig. 4). The two reanalysis data sets have weak correlation over the Indian Ocean, the equatorial western Pacific, the South America, over part of the equatorial Atlantic Ocean and over the Arctic. This comparison illustrates the uncertainty in the reanalysis datasets, which by extension contributes to uncertainty in the prediction



**Fig. 2** Correlation coefficients of (left) 2 meter temperature and (right) precipitation for (top) Sys4 and (bottom) CFSv2 for the period of 28 years from 1982 to 2009 winter

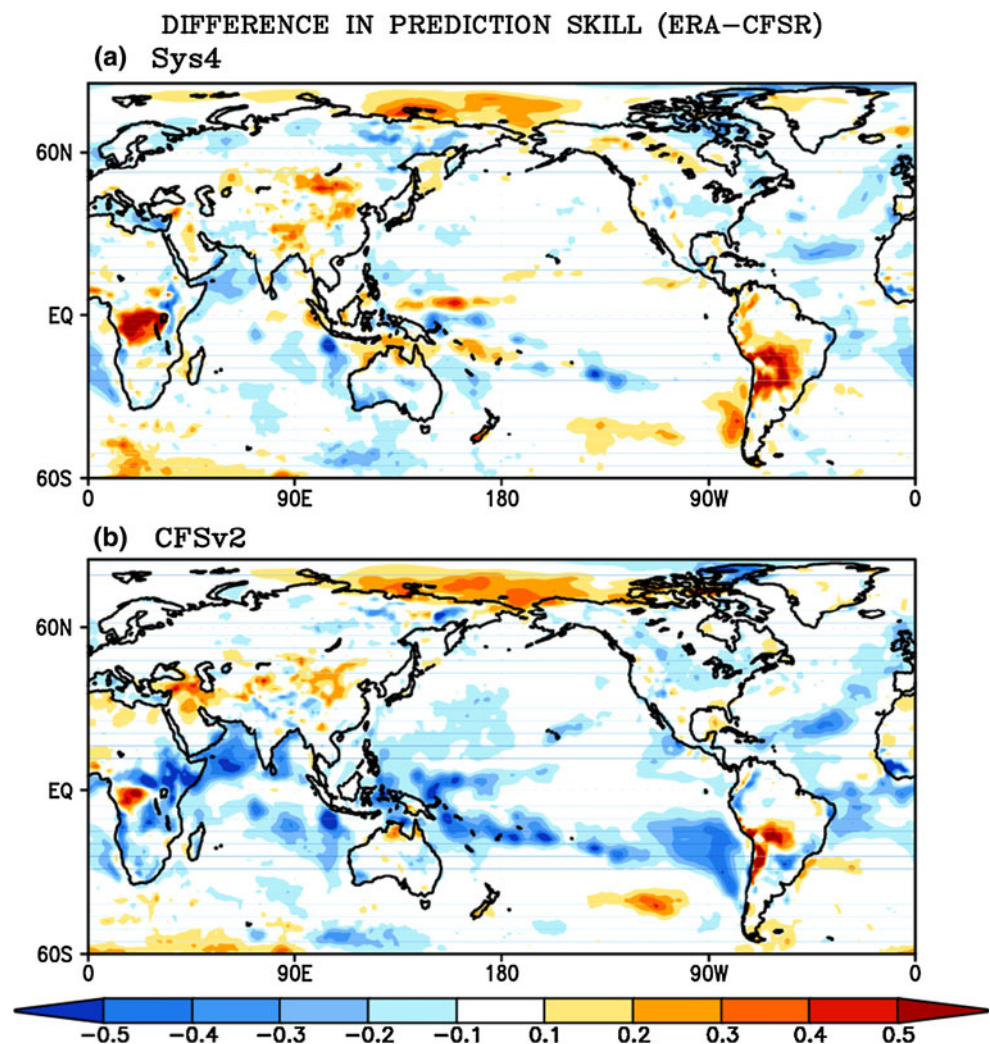
analysis. Therefore, the analyses in this study are conducted using both reanalysis datasets.

To compare the year-to-year variability of seasonal prediction skill, the pattern correlation between the predictions and reanalysis is calculated over the entire globe (0–360°E, 60°S–60°N) and the tropical pacific (40°E–300°E, 20°S–20°N) over the 28 winters. Figure 5 shows the correlation coefficient for 2mT and PRCP for each region for both modeling systems compared to the ERA. The global 2mT prediction skill shows strong interannual variation over 28 winters (Fig. 5a). The 28 year mean correlation coefficient for the global 2mT is similar for both modeling systems, showing little dependence on the reanalysis data set. For the tropics (Fig. 5b), Sys4 shows the greatest prediction skill in 1997 winter and lowest in 1990. In CFSv2, the highest skill is also shown in 1997, but the lowest skill occurs in 1987 winter. The 28-year mean prediction skill for tropical 2mT is 0.54 for Sys4 and 0.42 for CFSv2. Figure 18 (to be discussed later) shows the summary of the mean prediction skill for each variables

compared with ERA (dark shading) and CFSR (light shading). Precipitation over the tropics shows strong interannual variation (Fig. 5c) and mean prediction skill for the PRCP is 0.47 and 0.41 for Sys4 and CFSv2, respectively (Fig. 18).

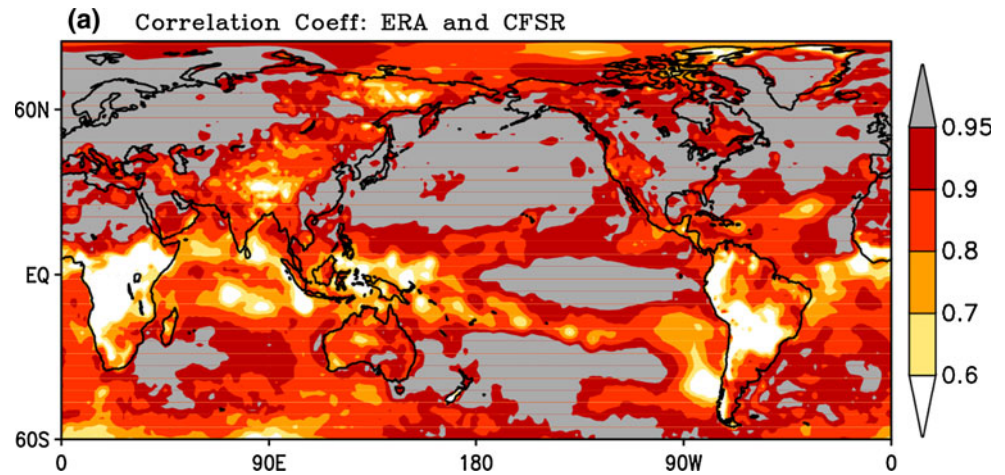
Both systems have the highest predictive skill for tropical 2mT in winters with strong ENSO amplitudes, specifically 1982, 1988, 1997 and 2007. To compare the relationship between the seasonal prediction skill and ENSO, the temporal correlation coefficient between the year-to-year tropical 2mT prediction skill and ENSO amplitude is calculated. The ENSO amplitude is defined as a standard deviation of NH winter Nino 3 index (Fig. 5, gray bar). The correlation coefficient between the 2mT prediction skill over the tropics and ENSO amplitude is 0.63 and 0.57 for Sys4 and CFSv2, respectively. The correlations between the PRCP prediction skill and ENSO amplitude is 0.46 and 0.60 in Sys4 and CFSv2, respectively. Hence, during strong ENSO winters the prediction skill of tropical 2mT and PRCP is higher than for weak

**Fig. 3** Difference in 2mT prediction skill between ERA interim and CFSR (ERA-CFSR) used for verification in **a** Sys4 and **b** CFSv2

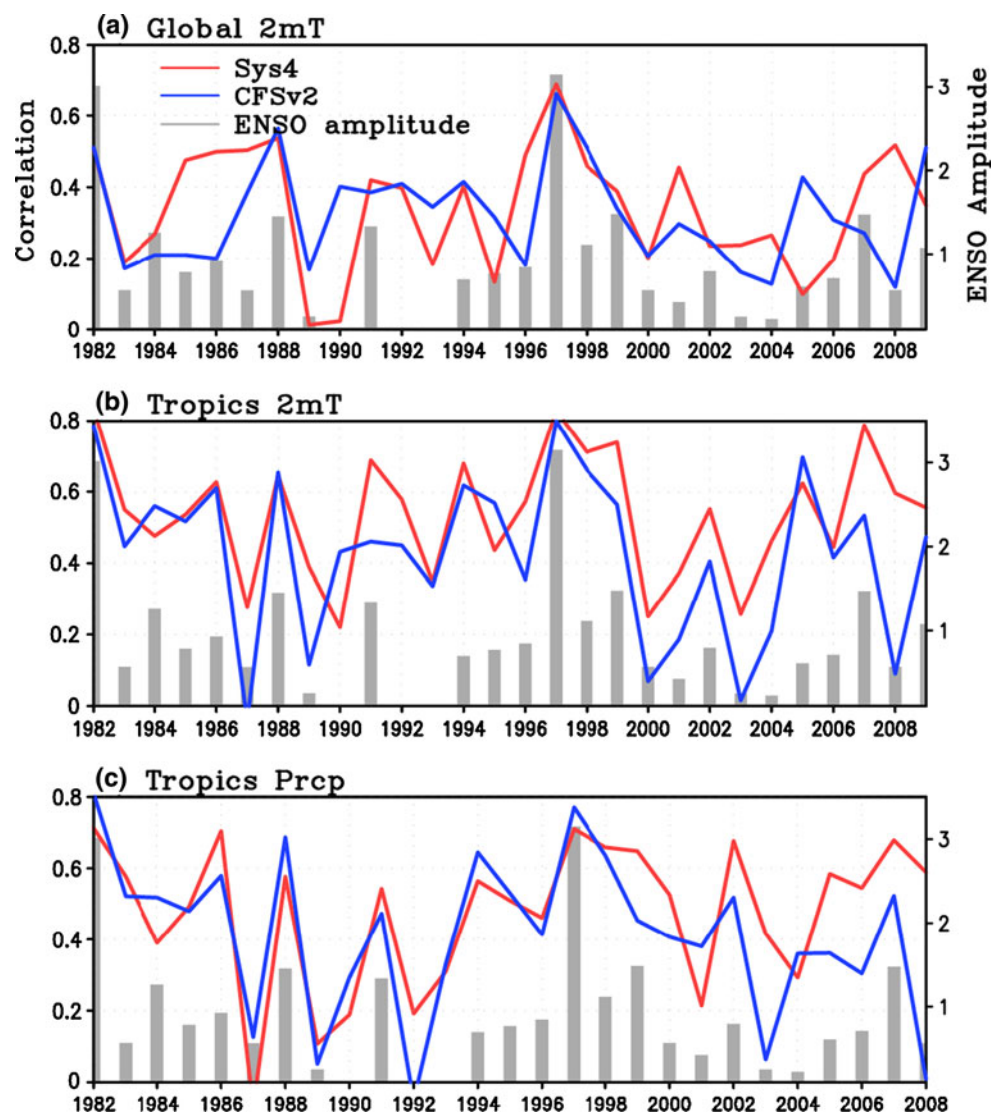




**Fig. 4** Correlation coefficients for DJF 2mT between ERA interim and CFSR over the period from 1982/1983 to 2008/2009



**Fig. 5** Anomaly pattern correlation for **a** global area and **b, c** tropical Pacific area for **b** 2 meter temperature and **c** precipitation. Gray bar is the ENSO amplitude. Mean correlation coefficients are displayed in Fig. 18



ENSO winters. Figure 6 shows the mean prediction skill of tropical 2mT from Fig. 5b plotted in descending order of the amplitude of ENSO arranged according to the absolute

value of the ENSO amplitude. The ENSO amplitude and skill are the moving average for 7 years from the largest ENSO amplitude year to the smallest. For example, the

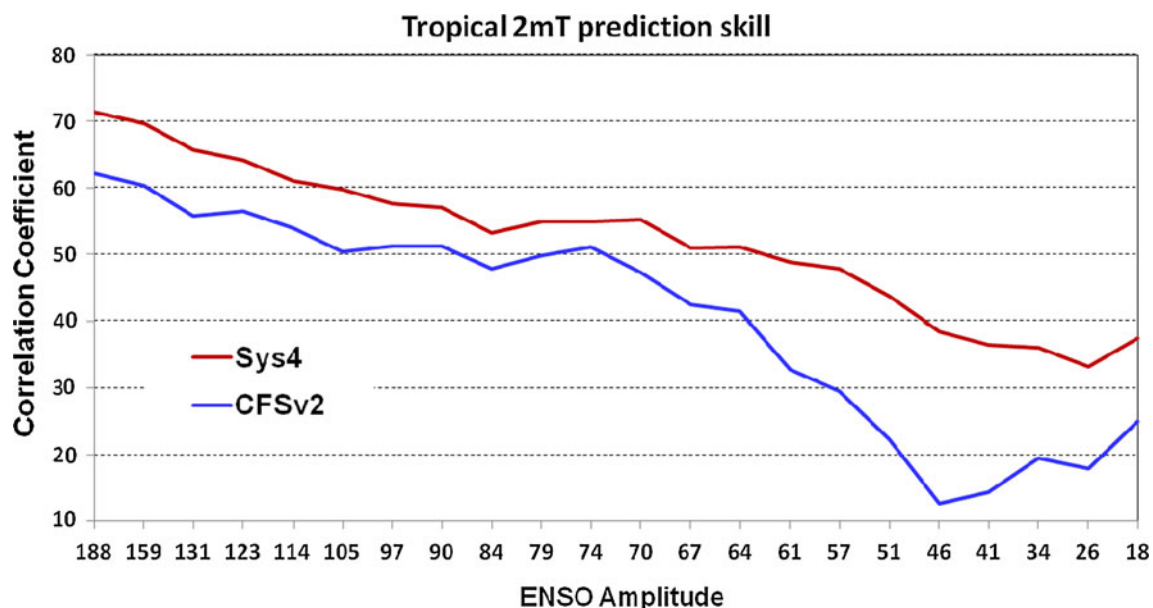
mean prediction skill from the largest ENSO amplitude years is 0.71 which is the average of seven strongest ENSO years (1997, 1982, 1999, 2007, 1988, 1991, and 1984). It is consistent in both modeling systems that the prediction skill increases with ENSO amplitude (Peng et al. 2000, 2011).

#### 4 ENSO prediction

As described above, the amplitude of ENSO dominates the winter seasonal prediction skill. Jin et al. (2008) examined the current status of ENSO prediction using retrospective forecasts made with ten different coupled GCMs from DEMETER and CliPAS/APCC model sets and found that the ENSO prediction skill in the state-of-the-art dynamical predictions depends on the ENSO phase and amplitude. Generally, dynamical models tend to have better prediction skill when initialized at NH winter than spring due to the ‘spring predictability barrier’ (Webster and Yang 1992; Webster 1995; Torrence and Webster 1998; Jin et al. 2008; Kim et al. 2009; Hendon et al. 2009). This study focuses on the boreal winter prediction when the initial condition already contains a strong ENSO signal. The ECMWF forecast model has been found to be better than statistical models at forecasting ENSO events (Van Oldenborgh et al. 2005) and NCEP CFS is shown to be competitive with other statistical models in predicting tropical SST variability (Saha et al. 2006). Here we compare ECMWF System 4 and CFSv2 in terms of winter ENSO prediction.

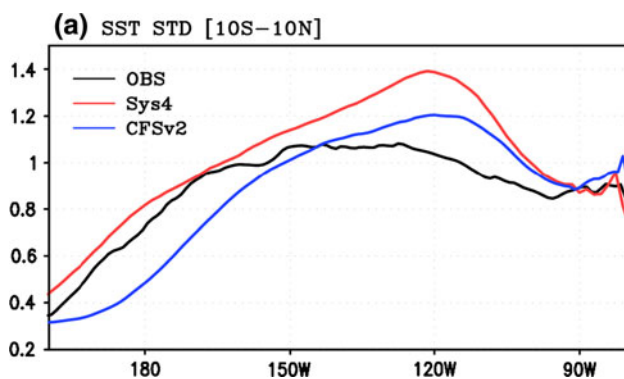
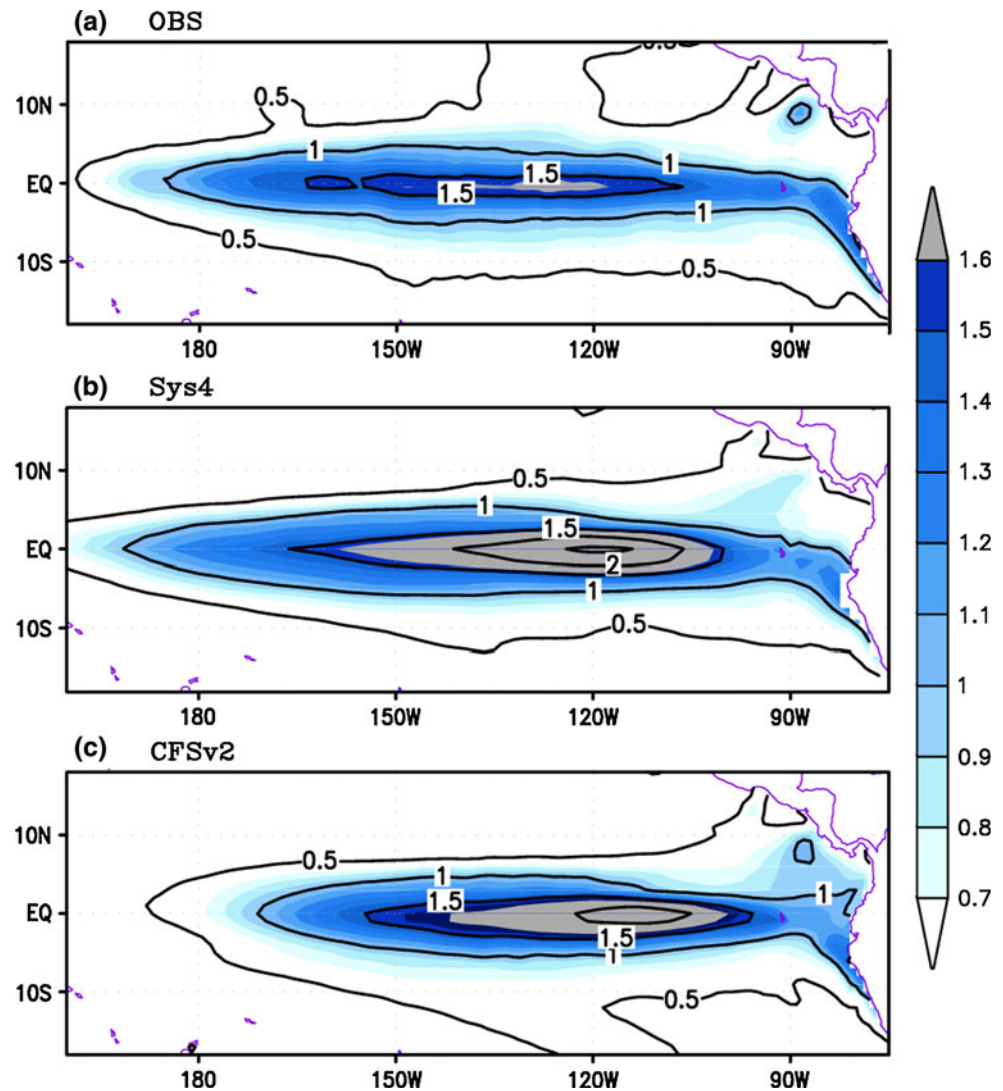
Figure 7 compares the predicted SST with OISSTv2 variability over the tropical Pacific for each forecast system. The SST variability is calculated by the standard deviation of NH winter SST anomalies over the 28 year period. Both modeling systems show similar patterns to the observations with maximum variability over the central to eastern Pacific, but with stronger magnitudes (Fig. 7). It has been previously noted that NCEP CFSv1 and v2 consistently tends to forecast larger ENSO amplitude (Wang et al. 2010). Figure 8 shows the latitudinal average of SST standard deviation (Fig. 7) over the tropics ( $10^{\circ}\text{S}$ – $10^{\circ}\text{N}$ ). Sys4 overestimates the amplitude of SST variability over the entire Tropics and CFSv2 overestimates the amplitude especially from  $150^{\circ}\text{W}$  to the eastern Pacific and underestimates it in the western Pacific.

To analyze the SST variance and systematic bias in both modeling systems, an empirical orthogonal function (EOF) analysis is applied to both the observed and predicted NH winter SST anomaly fields. To examine the simulation ability for increasing forecast lead times, we applied the EOF analysis for the winter mean (DJF) SST predictions initialized in November (0-month lead), October (1-month lead) and September (2-month lead), respectively. Each DJF predictions with 0- to 2-month lead include 16 ensemble members from August to November. The EOF analysis is applied to the predicted SST of individual ensemble members and then averaged. Figure 9 shows the eigenvector of the first EOF mode for the observation and for two prediction systems. Figure 10 compares the latitudinal and longitudinal mean of the first eigenvector for each lead time. The leading



**Fig. 6** Tropical 2mT prediction skill as a function of ENSO amplitude from Fig. 5b. ENSO amplitude and correlation coefficients are multiplied by 100. Years are arranged in the ascending order of amplitude of the ENSO

**Fig. 7** Standard deviation of winter mean SST anomalies for **a** observation, **b** Sys4 and **c** CFSv2



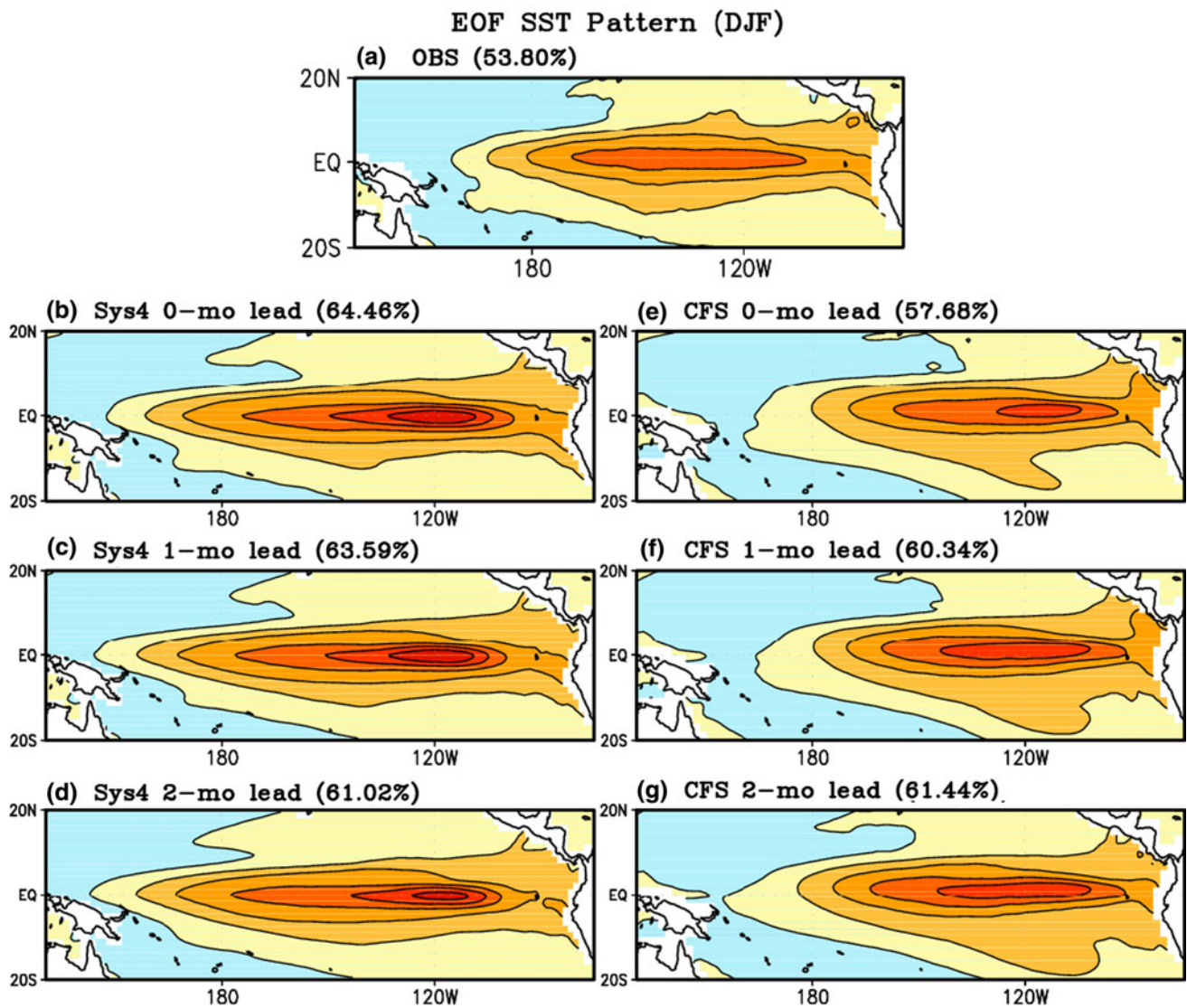
**Fig. 8** Latitudinal mean of SST standard deviation (Fig. 7) between 10°S to 10°N for observation (*black*), Sys4 (*red*) and CFSv2 (*blue*)

EOF mode for the observation explains 54 % of the total variance and the eigenvector is characterized by large positive components over the central to equatorial eastern Pacific (Fig. 9a). The spatial pattern of SST in the model

counterpart differs from observation which can be expected from Fig. 7. Both systems overestimate the amplitude in the eastern Pacific compared to observations. In the Sys4 results, the positive maximum value is concentrated to the region around 120°W and is shifted to south relative to the observation (Figs. 9, 10). The patterns do not change much as the forecast lead time increases. In the CFSv2, the center of maximum variability matches the observations well but is slightly shifted to the east.

The eigenvectors and their corresponding normalized time series of principal components (PC) of the EOF 1st mode are related to ENSO variability. The PC time series for observation and model with different lead times capture the dominant ENSO variability (Fig. 11), although the model eigenvectors show bias in their spatial pattern. The similarity between the observed and predicted PC time series provides possibilities for model error correction using a statistical approach (Kang et al. 2004; Kim et al. 2008). In both systems, the percentage of total variance of





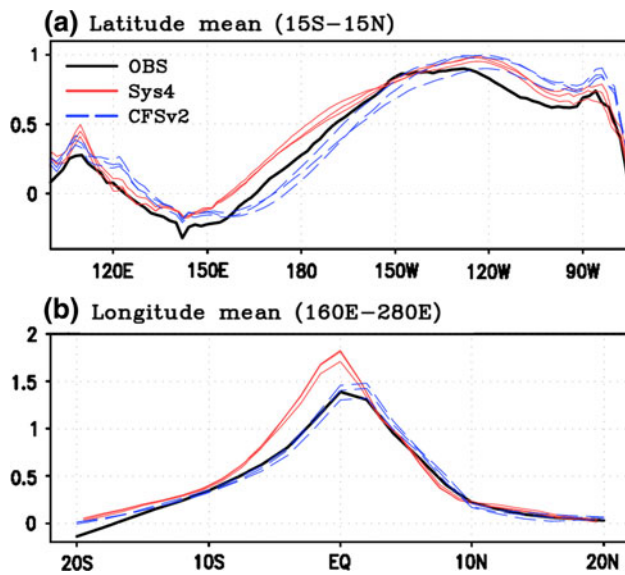
**Fig. 9** Eigenvectors of the first EOF mode for DJF SST anomaly for **a** observation, *left* Sys4 and *right* CFSv2 initiated at **b, e** November, **c, f** October, **d, g** September

the SST anomaly is larger than observed and differs in each lead time (Fig. 9).

The year-to-year ENSO prediction skill is assessed by using the Nino 3.4 index, defined as a mean SST anomaly averaged over the region from 190°E to 240°E and from 5°S to 5°N. The index possesses a strong interannual variability (Fig. 12) and both prediction systems capture the year-to-year ENSO variability very well. The correlation coefficient between the reforecasts and observations for Sys4 is 0.97 with root-mean-square error (RMSE) of 0.37, and for CFSv2 is 0.85 with RMSE of 0.67. Although the ENSO phase is well predicted in CFSv2, the magnitude of ENSO is overestimated in the system as noted earlier. Relatively low prediction skill and large RMSE in CFSv2 result from larger SST variability over the tropics. For example, the observed Nino 3.4 index in 1988 winter is around  $-2$  K while CFSv2

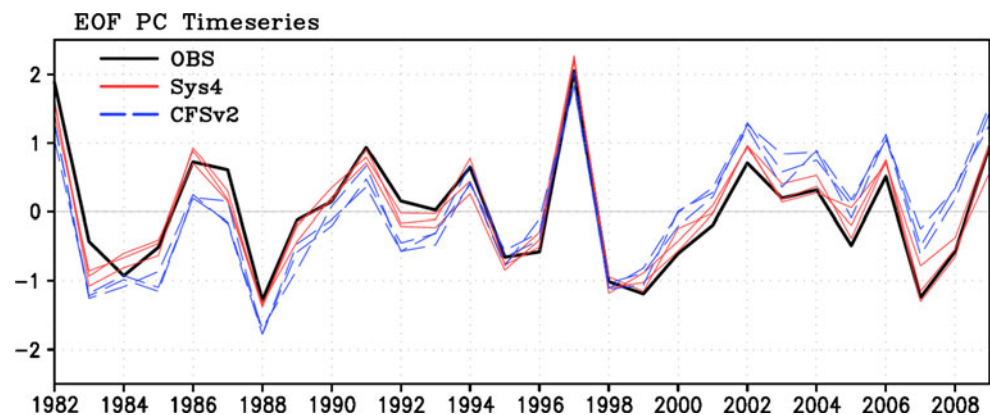
predicts a value almost 1 K lower than the observation. Before 1993, CFSv2 underestimates the Nino 3.4 values, but after 1998 CFSv2 overestimates the Nino 3.4 continuously, about 0.5 K higher than the observation (Fig. 12). A clear upward trend in the predicted winter Nino 3.4 index is found in CFSv2 (Xue et al. 2011).

Figure 13 compares the winter SST trend [K/year] of both modeling systems with observations. The observations show an upward trend over the most of the globe, while it shows a negative trend in the eastern Pacific and part of the North Pacific. Sys4 captures the trend very well, except with weaker amplitude over the globe. However, the CFSv2 has a very strong warming trend in winter SST even over the equatorial eastern Pacific, whereas the observations and Sys4 show a negative trend (Fig. 13c). The earlier version of CFS (Saha et al. 2006) shows a weaker warming trend perhaps due to the

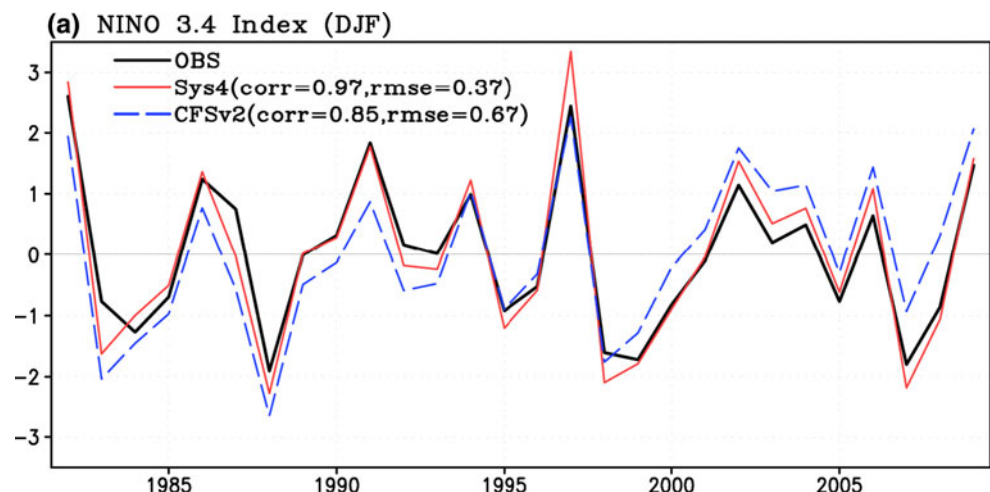


**Fig. 10** The eigenvector of the first EOF mode for **a** latitudinal mean (15°S–15°N) and for **b** longitudinal mean (160°E–280°E). Black line indicates the observation and thin blue and red lines indicate Sys4 and CFSv2 for lags from 0 to 2 month

**Fig. 11** Normalized timeseries of PCs of the first EOF modes from observation (black), Sys4 (red) and CFSv2 (blue)



**Fig. 12** Nino 3.4 index for observation (black), Sys4 (red) and CFSv2 (blue) from 1982 to 2009. Correlation coefficient and root-mean-square error between observation and hindcasts are indicated together



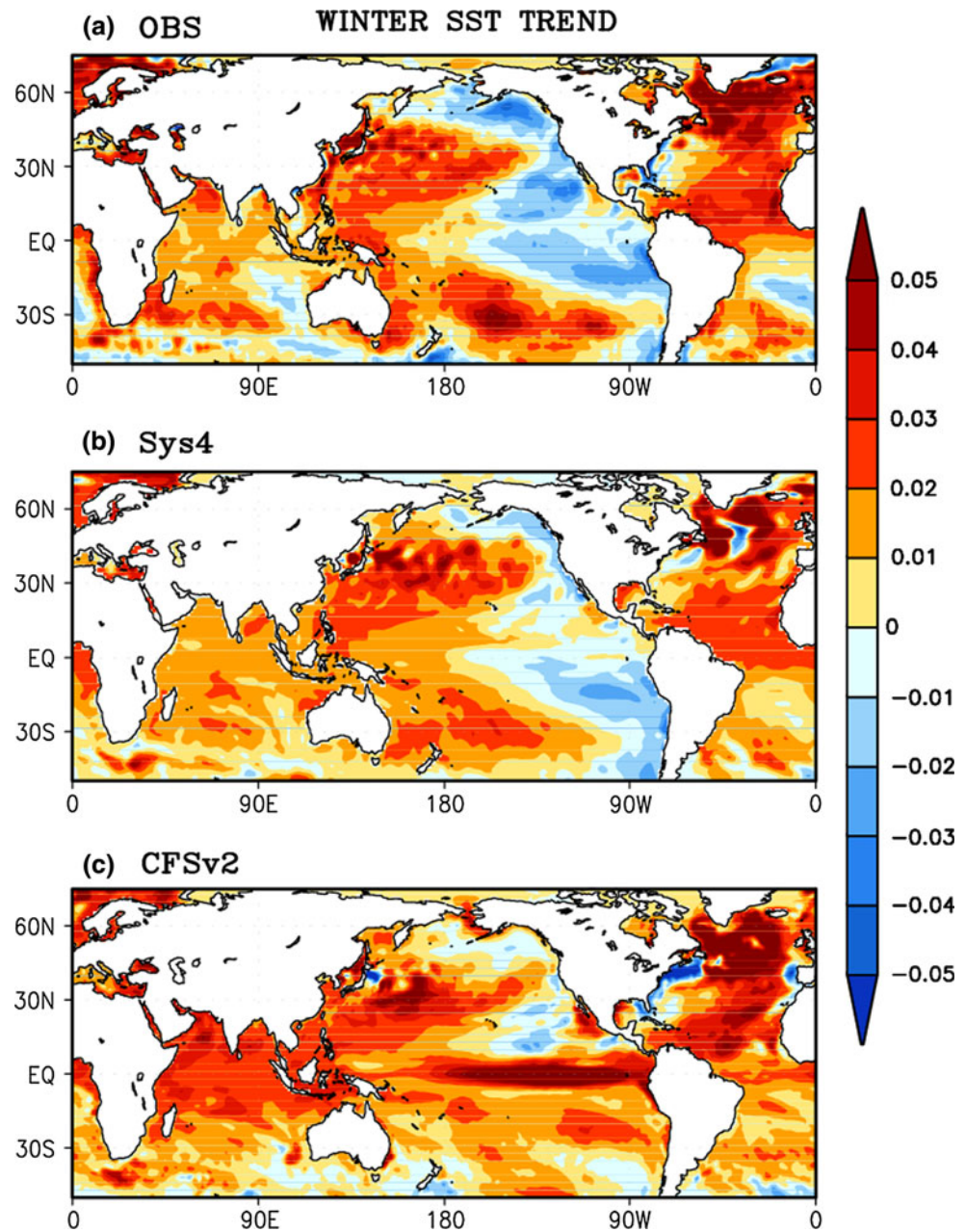
use of fixed greenhouse gas concentration. The CFSv2, on the other hand, uses prescribed  $\text{CO}_2$  concentrations as a function of time in its atmospheric initial condition (Cai et al. 2009). The large warming trend in the eastern Pacific SST is primarily associated with changes in satellite observing system that occurred in 1998/1999 period that were assimilated in the CFSR (Xue et al. 2011; Wang et al. 2011). An assessment of the trend is beyond the scope of this study, but it certainly needs further examination.

## 5 Teleconnection patterns in the extratropics

### 5.1 ENSO teleconnection

We now examine how the models predict winter teleconnection patterns in relation to the ENSO phase. Clearly, the NH winter is strongly influenced by the warm and cold phases of ENSO, especially the North Pacific and North America. Figures 14 and 15 shows the composite map of the ERA 2mT, the 500 hPa geopotential height and the PRCP anomaly in

**Fig. 13** Temperature changes (K/year) for **a** observation, **b** Sys4 and **c** CFSv2 from 1982 to 2009 NH winter



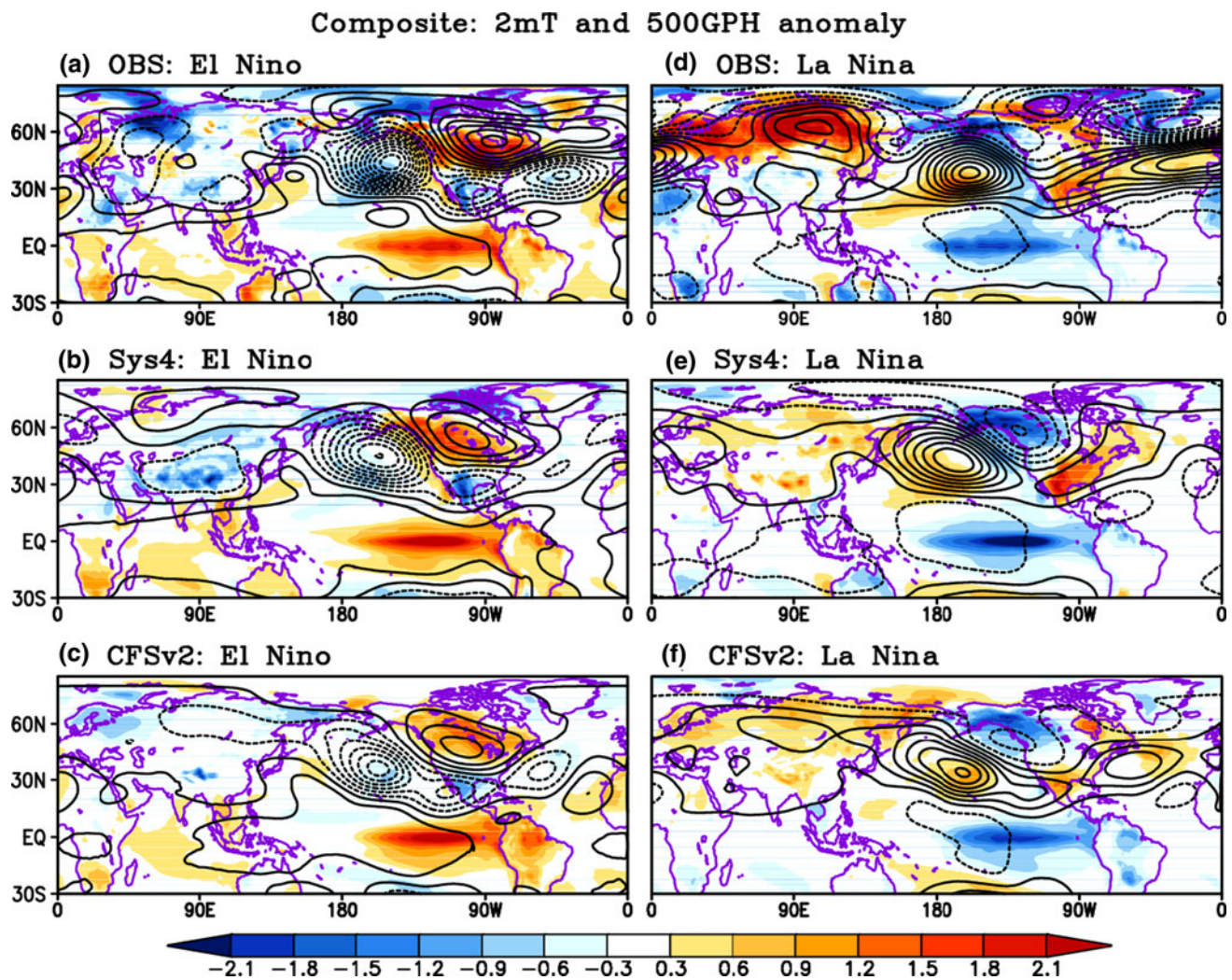
four strong El Nino (1982, 1991, 1997 and 2009) and La Nina (1988, 1998, 1999 and 2007) winters.

The composite patterns in CFSR are similar to the ERA analyses (not shown). The conventional El Nino pattern is apparent, with warm/wet anomaly across the equatorial central to eastern Pacific produced by the shifting pattern of the Walker circulation (Figs. 14, 15). A boomerang pattern of cold and dry anomaly appears to the north and south of the equatorial western Pacific. Although the La Nina pattern is not exactly the mirror image of El Nino, it is almost the opposite from El Nino in the extratropics. Both prediction systems simulate well the general pattern of ENSO response over the tropics, although the boomerang pattern

in the western Pacific is not well simulated by either system. The magnitude of the SST anomaly in both prediction systems is larger than the observed anomaly. The warm anomaly over the South Indian Ocean during El Nino and the warm/cold anomaly over the northern part of Australia in El Nino/La Nina are well captured in Sys4 (Fig. 14b, e).

The ENSO forcing of the Polar Jet over the North Pacific and North America is known to be responsible for ENSO teleconnections such as Pacific North America (PNA) (Wallace and Gutzler 1981). The southern part of North America experiences a cold and wet winter during El Nino and a warm and dry winter during La Nina (Figs. 14, 15). The northwestern part of North America experiences





**Fig. 14** Composite map of 2 meter temperature (K, shading) and 500 hPa geopotential height anomaly (m, contour) for *top* ERA interim, *middle* Sys4 and *bottom* CFSv2 for *left* El Niño and *right* La Niña winter

milder winter during the El Niño and colder winter during the La Niña phase. Both modeling systems capture the gross global patterns in strong ENSO winters. The 500 hPa high pressure area over the North America in El Niño winter is well captured in Sys4 but with weaker magnitude, and it is shifted to the west in CFSv2. The strong low pressure area in the North Pacific is well captured in both models, but slightly shifted to the south in CFSv2 (Figs. 14, 15). The other low pressure area in the southern part of US and the Atlantic Ocean is not well simulated in Sys4. In La Niña winters, the models have a tendency that is similar but slightly asymmetric to El Niño winters (Figs. 14, 15).

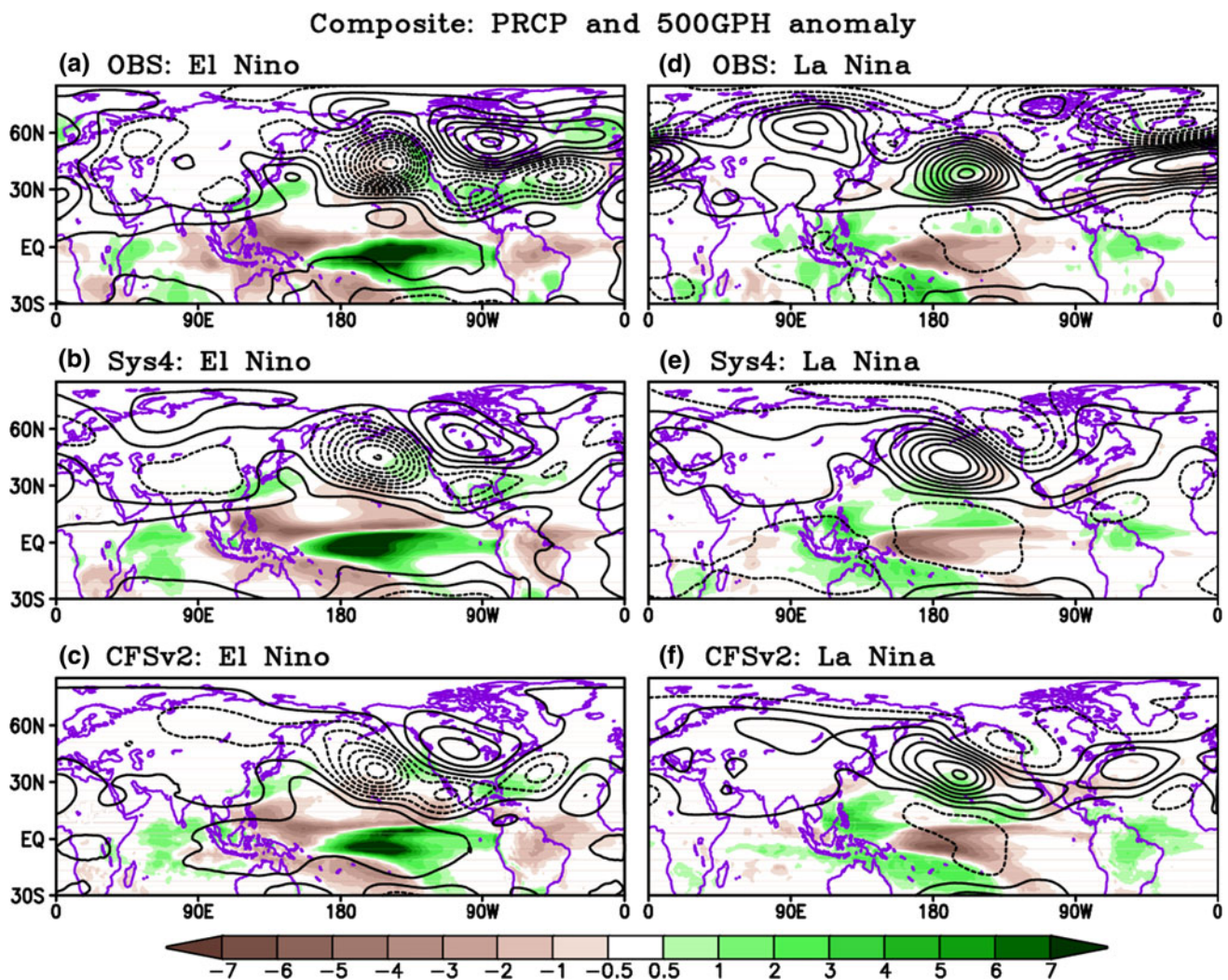
## 5.2 PNA and NAO

We have shown that the ENSO teleconnection pattern over the North Pacific and the North America is generally well predicted for strong ENSO winters. However, the year-

to-year winter climate variability in extratropics is influenced not only by tropical forcing but by oscillations of atmospheric mass between mid- and high-latitudes, such as PNA or North Atlantic Oscillation (NAO; Wallace and Gutzler 1981; Barnston and Livezey 1987). The NAO and PNA patterns are the two most important modes of variability in the NH mid- and high-latitudes, thus the prediction skill of the NH extratropics is related to the skill of predicting these patterns. In this section, we examine how well the models predict the dominant winter climate oscillations.

The NAO is one of the most prominent wintertime teleconnection patterns that modulate climate over the North America to the northern Europe (e.g., Hurrell 1995). The NAO index is defined as a difference between normalized DJF mean SLP anomaly from 80°W to 30°E and at 35°N and 65°N (Li and Wang 2003). The NAO has exhibited considerable variability over the past 28 years in both the ERA interim and CFSR data sets (Fig. 16a).





**Fig. 15** As in Fig. 14, but for the precipitation anomaly (mm/day, shading)

The correlation coefficients between the ERA interim and CFSR is 0.78. Neither prediction system captures the year-to-year NAO variability during DJF. Coefficients between the ERA reanalysis and predicted NAO index are 0.16 and 0.25 for Sys4 and CFSv2, respectively (Figs. 16a, 18). The correlation coefficients between CFSR and the predicted NAO index are 0.11 and 0.21 for Sys4 and CFSv2, respectively (Figs. 16a, 18).

The PNA is also a dominant low frequency mode of climate variability over the NH winter. The PNA index is determined following Wallace and Gutzler (1981):

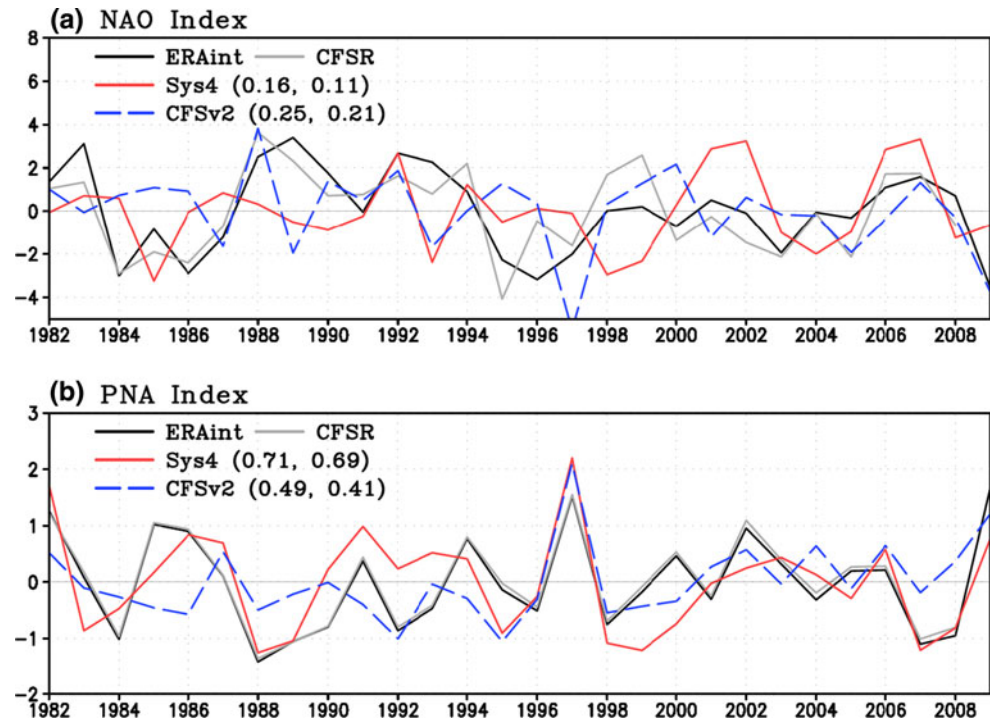
$$\text{PNA} = 0.25 \times [Z(20^\circ\text{N}, 160^\circ\text{W}) - Z(45^\circ\text{N}, 165^\circ\text{W}) + Z(55^\circ\text{N}, 15^\circ\text{W}) - Z(30^\circ\text{N}, 85^\circ\text{W})]$$

where  $Z$  is standardized value of the 500 hPa geopotential height. Figure 16b shows interannual variability of the PNA index for 28 winters from ERA interim and CFSR. Although the PNA pattern is a natural internal mode of climate variability, it is also modulated by the ENSO. The

correlation coefficient between observed NH winter PNA and Nino 3.4 index is highly correlated at 0.7. The correlation coefficients between ERA interim and CFSR is 0.99. The two modeling systems predict the PNA quite well, with correlation coefficients between 0.4 and 0.7 with ERA interim and CFSR (Figs. 16b, 18). The Sys4 system predicts the PNA better than the CFSv2 system, especially in strong ENSO winters (particularly for the winters of 1982, 1988, 1991, 1997 or 2007: not shown). Due to the association of the PNA and its low-frequency variability and the influence of ENSO forcing, relatively higher prediction skill occurs for the PNA than for the NAO in general agreement with dynamical predictions (Johansson 2007; Müller et al. 2005).

We have examined the prediction skill of the NAO and PNA, both of which influence North American and northern Europe climate variability. How well do the models predict the winter climate over the North America and the northern Europe? Figure 17 shows the year-to-year area

**Fig. 16** **a** NAO and **b** PNA index for ERA interim (black), CFSR (gray), Sys4 (red) and CFSv2 (blue) from 1982 to 2009 winter. Numbers indicate the temporal correlation coefficient compared with ERA interim and CFSR



averaged 2mT for the North America (Fig. 17a) and for the northern Europe (Fig. 17b) compared with both the ERA and CFSR. The average skill over the North America is 0.14 for Sys4 and 0.30 for CFSv2 (Fig. 17a). The skill changes to 0.29 for Sys4 and 0.42 for CFSv2 when evaluated against the CFSR. The skill over the northern Europe is 0.39 for Sys4 and 0.33 for CFSv2 when compared with ERA, and 0.40 and 0.41, respectively, when compared with the CFSR (Fig. 17b). No relationship between the prediction skill of the North American and European regions and NAO/PNA has been found. Similar difficulty occurs in finding coherence in the prediction skill of both models.

## 6 Summary and discussion

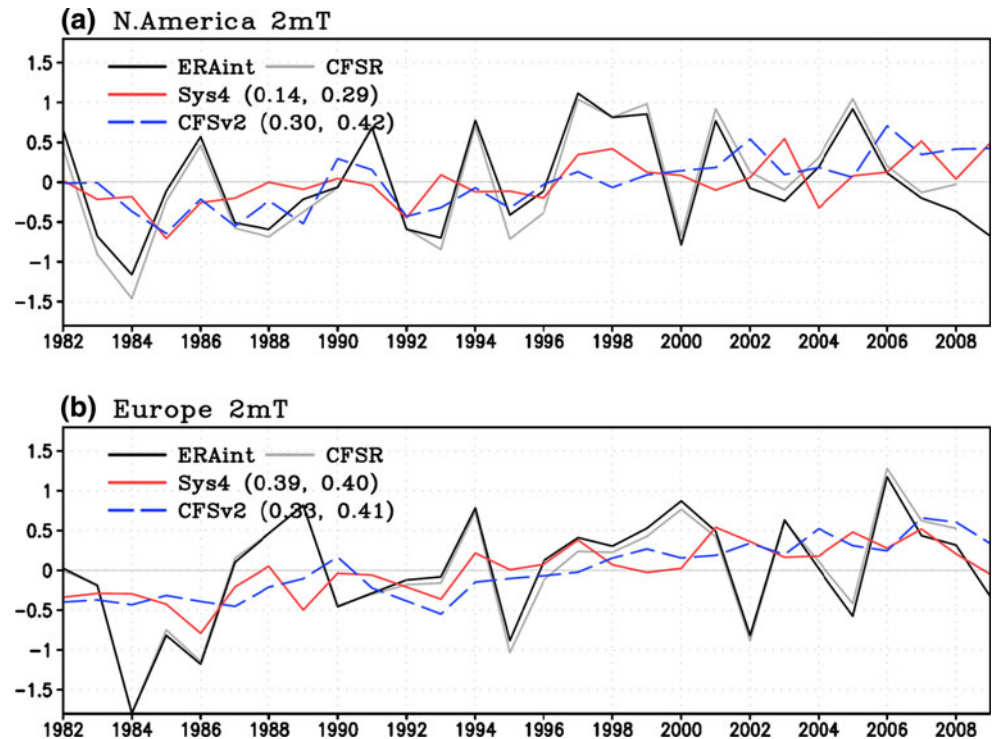
This study has examined the seasonal prediction skill for NH winter using retrospective predictions (reforecasts) by the ECMWF System 4 and NCEP CFSv2. The temperature, precipitation and geopotential height from the reforecast for the period 1982–2010 were compared with two reanalysis products: the ERA interim and the CFSR. The simulation ability of long-term mean climatology and the year-to-year variation were assessed. Both Sys4 and CFSv2 reproduce realistically the observed climatology pattern. However, systematic biases are found in both simulations. For the Sys4, a cold bias is found across the equatorial Pacific although a warm bias is found in the North Pacific and part of the North Atlantic. The CFSv2 has strong warm bias from the cold

tongue region of the Pacific to the equatorial central Pacific and cold bias in broad areas of the North Pacific and the North Atlantic. A cold bias over large regions of the Southern Hemisphere is a common property of both reforecasts. With respect to precipitation, the Sys4 produced excesses along the ITCZ, the equatorial Indian Ocean and the western Pacific in Sys4. In the CFSv2, a strong wet bias is found along the SPCZ and the southern Indian Ocean as well as in the western Pacific. A dry bias is found for both modeling systems over South America and northern Australia and wet bias in East Asia and the equatorial Atlantic.

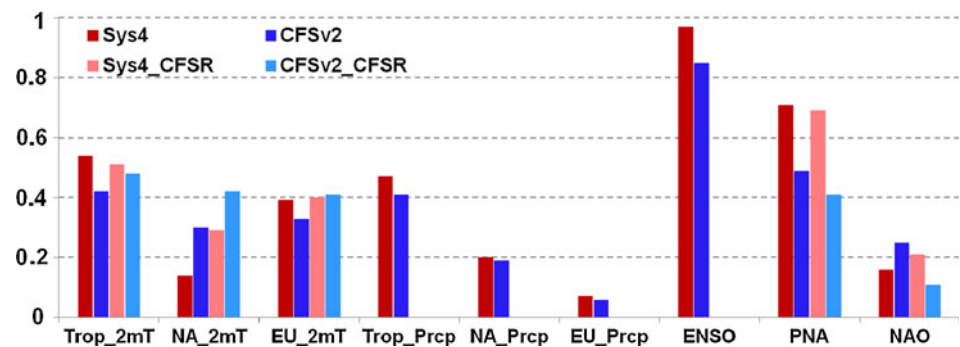
For both the Sys4 and CFSv2 systems, the mean prediction skill of 2mT and precipitation is higher over the tropics than the extra-tropics and higher over ocean than land. The 2mT over the South Indian Ocean, the North Pacific and equatorial North Atlantic shows high predictive skill in both reforecasts. The actual prediction skill of the 2mT depends on the reanalysis data set which is used as verification field. The discrepancy in two reanalysis (ERA interim and CFSR) is clear over the Indian Ocean, the equatorial western Pacific, the South America, over part of the equatorial Atlantic Ocean and over the Arctic. Therefore, the analyses are conducted using both reanalysis datasets. The 2mT and precipitation show the greatest skill in the tropical belt, especially in ENSO region when it is verified with both ERA interim and CFSR. In both modeling systems, the prediction skill of both tropical 2mT and precipitation is higher during strong ENSO winters than during weak ENSO winters.



**Fig. 17** As in Fig. 16, but for the area averaged 2 meter temperature anomaly over the **a** North America and **b** Europe



**Fig. 18** Mean prediction skill (correlation coefficient) in different variables and regions for Sys4 (red) and CFSv2 (blue). For 2mT, PNA and NAO, dark (light) shadings indicate the mean prediction skill compared to ERA interim (CFSR)



In both systems, the standard deviation of winter mean SST anomaly shows similar patterns to observations with maximum variability over the central to eastern Pacific with a stronger magnitude than observed. Although the ENSO SST variability is spatially biased in the models, both models predict the year-to-year ENSO variation accurately. Bias in winter SST trend over the ENSO region in CFSv2 results in relatively low ENSO prediction skill and high RMS error compared to Sys4. Both models capture the main ENSO teleconnection pattern of strong anomalies over the tropics, the North Pacific, the North America and for PNA. However, both models have difficulty in forecasting the NAO and the year-to-year winter temperature variability over the North America and northern Europe. Figure 18 shows the summary of the mean prediction skill for different variables and regions in Sys4 and CFSv2.

This study has examined the prediction skill of the NH winter from the most recently upgraded seasonal forecast systems from ECMWF and NCEP. However, to provide physical insights to differences in prediction skill regarding to the set up of forecast systems, it would be useful to compare the skill between CFSv1 and CFSv2 and between ECMWF System 3 and system 4. This will be the subject of future research.

**Acknowledgments** We would like to thank the reviewers for thoughtful and helpful comments. The ECMWF System 4 reforecasts were obtained by the authors through a commercial agreement with ECMWF. We would like to acknowledge the continued support of ECMWF especially, with regard to the paper, the comments of Drs. Franco Molteni and Tim Stockdale. The Climate Dynamics Division of the National Science Foundation under grant NSF-AGS 0965610 provided funding support for this research.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Adler RF et al (2003) The version 2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–Present). *J Hydrometeorol* 4:1147–1167
- Barnston AG, Livezey RE (1987) Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon Weather Rev* 115:1083–1126
- Berrisford P, Dee D, Fielding K, Fuentes M, Kallberg P, Kobayashi S, Uppala S (2009) The ERA-interim archive. ERA report series. No. 1. ECMWF: Reading, UK
- Cai M, Shin CS, van den Dool HM, Wang W, Saha S, Kumar A (2009) The role of long-term trends in seasonal predictions: implication of global warming in the NCEP CFS. *Weather Forecast* 24:965–973
- Dee DP, Uppala S (2009) Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Q J R Meteorol Soc* 135:1830–1841
- Dee DP et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137:553–597
- Ebisuzaki W, Zhang L (2011) Assessing the performance of the CFSR by an ensemble of analyses. *Clim Dyn* 37:2541–2550
- Hendon HH, Lim E, Wang G, Alves O, Hudson D (2009) Prospects for predicting two flavors of El Niño. *Geophys Res Lett* 36:L19713. doi:[10.1029/2009GL040100](https://doi.org/10.1029/2009GL040100)
- Higgins RW, Kousky VE, Silva VBS, Becker E, Xie P (2010) Intercomparison of daily precipitation statistics over the United States in observations and in NCEP reanalysis products. *J Clim* 23:4637–4650
- Hurrell JW (1995) Decadal trends in the North Atlantic oscillation: regional temperatures and precipitation. *Science* 269:303–313
- Jin EK et al (2008) Current status of ENSO prediction skill in coupled ocean-atmosphere models. *Clim Dyn* 31:647–664
- Johansson A (2007) Prediction skill of the NAO and PNA from daily to seasonal time scales. *J Clim* 20:1957–1975
- Kang IS, Lee JY, Park CK (2004) Potential predictability of summer mean precipitation in a dynamical seasonal prediction system with systematic error correction. *J Clim* 17:834–844
- Kim HM, Kang IS, Wang B, Lee JY (2008) Interannual variations of the boreal summer intraseasonal variability predicted by ten atmosphere-ocean coupled models. *Clim Dyn* 30:485–496
- Kim HM, Webster PJ, Curry JA (2009) Impact of shifting patterns of Pacific Ocean warming events on North Atlantic tropical cyclone. *Science* 325:77–80
- Kug JS, Kang IS, Choi DH (2008) Seasonal climate predictability with tier-one and tier-two prediction systems. *Clim Dyn* 31:403–416
- Kumar A, Hoerling MP (2000) Analysis of a conceptual model of seasonal climate variability and implications for seasonal predictions. *Bull Am Meteorol Soc* 81:255–264
- Kumar KK, Hoerling M, Rajagopalan B (2005) Advancing Indian monsoon rainfall predictions. *Geophys Res Lett* 32:L08704. doi:[10.1029/2004GL021979](https://doi.org/10.1029/2004GL021979)
- Lee JY et al (2010) How are seasonal prediction skills related to models' performance on mean state and annual cycle? *Clim Dyn* 35:267–283
- Li J, Wang J (2003) A new North Atlantic oscillation index and its variability. *Adv Atmos Sci* 20(5):661–676
- Molteni F, Stockdale T, Balmaseda M, Balsamo G, Buizza R, Ferranti L, Magnusson L, Mogensen K, Palmer T, Vitart F (2011) The new ECMWF seasonal forecast system (System 4). ECMWF Technical Memorandum 656
- Müller WA, Appenzeller C, Schär C (2005) Probabilistic seasonal prediction of the winter North Atlantic oscillation and its impact on near surface temperature. *Clim Dyn* 24:213–226
- Oldenborgh GJV, Balmaseda MA, Ferranti L, Stockdale TN, Anderson DLT (2005) Did the ECMWF seasonal forecast model outperform statistical ENSO forecast models over the last 15 years? *J Clim* 18(16):3240–3249
- Palmer T et al (2004) Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull Am Meteorol Soc* 85:853–872
- Pattanaik DR, Kumar A (2010) Prediction of summer monsoon rainfall over India using the NCEP climate forecast system. *Clim Dyn* 34:557–572
- Peng P, Kumar A, Barnston AG, Goddard L (2000) Simulation skills of the SST-forced global climate variability of the NCEP-MRF9 and the Scripps-MPI ECHAM3 models. *J Clim* 13:3657–3679
- Peng P, Kumar A, Wang W (2011) An analysis of seasonal predictability in coupled model forecasts. *Clim Dyn* 36:419–430
- Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W (2002) An improved in situ and satellite SST analysis for climate. *J Clim* 15:1609–1625
- Saha S et al (2006) The NCEP climate forecast system. *J Clim* 19:3483–3517
- Saha S et al (2010) The NCEP climate forecast system reanalysis. *Bull Am Meteorol Soc* 91(8):1015–1057
- Stockdale TN, Anderson DLT, Balmaseda MA, Doblas-Reyes FJ, Ferranti L, Mogensen K, Palmer TN, Molteni F, Vitart F (2011) ECMWF seasonal forecast system 3 and its prediction of sea surface temperature. *Clim Dyn*. doi:[10.1007/s00382-010-0947-3](https://doi.org/10.1007/s00382-010-0947-3)
- Tompkins AM, Feudale L (2010) Seasonal ensemble predictions of West African monsoon precipitation in the ECMWF system 3 with a focus on the AMMA special observing period in 2006. *Weather Forecast* 25:768–788
- Torrence C, Webster PJ (1998) The annual cycle of persistence in the El Niño-southern oscillation statistics. *Q J R Meteorol Soc* 124:1985–2004
- Uppala S et al (2005) The ERA-40 re-analysis. *Q J R Meteorol Soc* 131:2961–3012
- Wallace JM, Gutzler DS (1981) Teleconnections in the geopotential height field during the northern hemisphere winter. *Mon Weather Rev* 109:784–812
- Wang G, Kleeman R, Smith N, Tseitin F (2001) The BMRC coupled general circulation model ENSO forecast system. *Mon Weather Rev* 130:975–991
- Wang B, Ding QH, Fu XH, Kang IS, Jin K, Shukla J, Doblas-Reyes F (2005a) Fundamental challenge in simulation and prediction of summer monsoon rainfall. *Geophys Res Lett* 32:L15711
- Wang W, Saha S, Pan HL, Nadiga S, White G (2005b) Simulation of ENSO in the new NCEP coupled forecast system model. *Mon Weather Rev* 133:1574–1593
- Wang B, Lee JY, Kang IS, Shukla J, Kug JS, Kumar A, Schemm J, Luo JJ, Yamagata T, Park CK (2008) How accurately do coupled climate models predict the leading modes of Asian-Australian monsoon interannual variability? *Clim Dyn* 30:605–619
- Wang B et al. (2009) Advance and prospectus of seasonal prediction: assessment of the APCC/CLIPAS 14-model ensemble retrospective seasonal prediction (1980–2004). *Clim Dyn* 33. doi:[10.1007/s00382-008-0460-0](https://doi.org/10.1007/s00382-008-0460-0)
- Wang W, Chen M, Kumar A (2010) An assessment of the CFS real-time seasonal forecasts. *Weather Forecast* 25:950–969

- Wang W, Xie P, Yoo SH, Xue Y, Kumar A, Wu X (2011) An assessment of the surface climate in the NCEP climate forecast system reanalysis. *Clim Dyn* 37:1601–1620
- Weaver SJ, Wang W, Chen M, Kumar A (2011) Representation of MJO variability in the NCEP climate forecast system. *J Clim*. doi:[10.1175/2011JCLI4188.1](https://doi.org/10.1175/2011JCLI4188.1)
- Webster PJ (1995) The annual cycle and the predictability of the tropical coupled ocean-atmosphere system. *Meteorol Atmos Phys* 56:33–55
- Webster PJ, Yang S (1992) Monsoon and ENSO: selectively interactive systems. *Q J R Meteorol Soc* 118:877–926
- Xue Y, Huang B, Hu ZZ, Kumar A, Wen C, Behringer D (2011) An assessment of oceanic variability in the NCEP climate forecast system reanalysis. *Clim Dyn* 37:2511–2539
- Yang S, Zhang Z, Kousky VE, Higgins RW, Yoo SH, Liang J, Fan Y (2008) Simulations and seasonal prediction of the Asian summer monsoon in the NCEP climate forecast system. *J Clim* 21:3755–3775
- Yang S, Jiang Y, Zheng D, Higgins RW, Zhang Q, Kousky VE, Wen M (2009) Variations of US regional precipitation and simulations by the NCEP CFS: focus on the Southwest. *J Clim* 22:3211–3231
- Zhang L, Kumar A, Wang W (2012) Influence of changes in observations on precipitation: a case study for the climate forecast system reanalysis (CFSR). *J Geophys Res*. doi:[10.1029/2011JD017347](https://doi.org/10.1029/2011JD017347)