Clim Dyn
DOI 10.1007/s00382-013-1813-x

Global seasonal climate predictability in a two tiered forecast system. Part II: boreal winter and spring seasons

Haiqin Li · Vasubandhu Misra

Received: 1 November 2012 / Accepted: 20 May 2013

Abstract We examine the Florida Climate Institute–Florida State University Seasonal Hindcast (FISH50) skill at a relatively high (50 km grid) resolution two tiered Atmospheric General Circulation Model (AGCM) for boreal winter and spring seasons at zero and one season lead respectively. The AGCM in FISH50 is forced with bias corrected forecast sea surface temperature averaged from two dynamical coupled ocean–atmosphere models. The comparison of the hindcast skills of precipitation and surface temperature from FISH50 with the coupled ocean–atmosphere models reveals that the probabilistic skill is nearly comparable in the two types of forecast systems (with some improvements in FISH50 outside of the global tropics). Furthermore the drop in skill in going from zero lead (boreal winter) to one season lead (boreal spring) is also similar in FISH50 and the coupled ocean–atmosphere models. Both the forecast systems also show that surface temperature hindcasts have more skill than the precipitation hindcasts and that land based precipitation hindcasts have slightly lower skill than the corresponding hindcasts over the ocean.

H. Li · V. Misra (✉)
Center for Ocean-Atmospheric Prediction Studies,
Florida State University, 2035 E. Paul Dirac Dr.,
200 RM Johnson Bldg., Tallahassee, FL 32306-2840, USA
e-mail: vmisra@fsu.edu

V. Misra
Department of Earth, Ocean and Atmospheric Science,
Florida State University, P.O. Box 3064520, Tallahassee,
FL 32306-4520, USA

V. Misra
Florida Climate Institute, Florida State University,
2035 E. Paul Dirac Dr., 200 RM Johnson Bldg.,
Tallahassee, FL 32306-2840, USA

Published online: 13 June 2013

Keywords ENSO · Seasonal predictability · Forecast skill

1 Introduction

The basis for global seasonal climate prediction was initially best illustrated for the boreal winter climate when the El Niño and the Southern Oscillation (ENSO) Sea Surface Temperature (SST) anomalies in the equatorial anomalies are the largest (Bengtsson et al. 1993; Kumar and Hoerling 1995; Shukla 1998; Shukla et al. 2000). These 'success stories' spurred the climate modeling community that resulted in significant amount of literature on the impact of slowly varying surface boundary conditions on the genesis, sustenance and demise of several atmospheric climate anomalies (Barnston et al. 1994; Koster et al. 2000; Hoerling et al. 2001; Goddard et al. 2001). However these studies highlighted a somewhat arcane and possibly unattainable (in an operational environment) 'potential' predictability of the seasonal atmospheric anomalies as they were forced with observed SST. As a result the interest in diagnosing 'potential' predictability waned in the community, while efforts to develop and diagnose seasonal predictability of coupled ocean–atmosphere models increased (Stockdale et al. 1998; Kirtman et al. 2002; Kirtman 2003; DeWitt 2005). From these sustained efforts in the last decade or so, there has been a notable improvement in the dynamical ENSO prediction (Saha et al. 2006, 2010; Kirtman and Min 2009; Stockdale et al. 2011; Zhu et al. 2012). Saha et al. (2006) demonstrated for the first time that dynamical coupled ocean–atmosphere models were comparable if not better than statistical forecasts for ENSO. It has now culminated in a massive nation wide co-ordinated effort to develop the National Multi-Model Ensemble [NMME (Kirtman et al. 2013); http://www.cpc.ncep.noaa.gov/

products/ctb/nmme/] project to harvest the improvements made in the individual modeling centers towards improved seasonal prediction.

In this paper we seek to revisit the two-tiered seasonal forecast system for seasonal prediction. The motivation for this is several. One, the improvement in the forecasted SST anomalies from the dynamical prediction systems is worth leveraging. Second, coarseness of the coupled ocean–atmosphere models continues to be an issue. In this study we are investigating seasonal predictability with an Atmospheric General Circulation Model (AGCM) of $\sim$50 km grid resolution, which is two to four times higher resolution than the current coupled climate models in the NMME. To meet the growing needs of application studies say for example in hydrology, there is a push towards obtaining climate prediction products at a higher spatial and temporal scales (Bohn et al. 2010; Clark and Hay 2004; Shukla et al. 2010). Third, there are some recent studies suggesting that improvement of monsoon climate simulations in climate models is a result of the improved ENSO signal in the model (Delsole and Shukla 2012). Fourth, if the results from this study are promising then it opens the possibility of adding to the NMME effort at a comparatively lower encumbrance with potentially high pay off. Fifth, the basic premise of seasonal climate prediction of slowly varying boundary conditions preconditioning the atmospheric anomalies will always be valid and is worth revisiting periodically to at least assess the progress made in the prediction of the SST anomalies and other boundary conditions. In the next section we explain the experiment design and provide a brief model description. This is followed by the analysis of the results in Sect. 3 followed by summary and conclusions in Sect. 4.

2 Experiment design

The Florida Climate Institute-Florida State University Seasonal Hindcasts at 50 km grid resolution (FISH50) was implemented to initiate forecasts starting in the boreal winter and integrated through 6 months to end of May of the subsequent year. FISH50 was conducted for the period 1982–2008. Further details of the FISH50 experiments are provided in Table 1. FISH50 are two tiered hindcasts, meaning that the Atmospheric General Circulation Model (AGCM) was forced with forecasted SST from another prediction system. The forecasted monthly mean SST anomalies were averaged from two coupled ocean–atmosphere models (CFSv2 [Saha et al. 2010]; CCSM3.0 [Kirtman and Min 2009]), which are part of the family of the National Multi-Model Ensemble project (NMME [Kirtman et al. 2013]). The other coupled ocean–atmosphere models in the NMME project were not utilized as they were not available at the time of conceiving the FISH50 experiments. The multi-model average of the SST anomalies is found to have overall higher prediction skill than any single model (Kirtman and Min 2009). These multi-model averaged SST anomalies are overlaid on observed climatology that contains the seasonal cycle, secular changes and decadal variations. This bias correction of SST anomalies becomes necessary as the systematic errors in CFSv2 and CCSM3.0 in the equatorial Pacific and in the subtropical eastern oceans are grave (Fig. 1). For example, CCSM3.0 displays an equatorial central Pacific cold bias of $\sim$1 °C in both DJF and MAM seasons. Similarly CFSv2 shows a very large bias (of over 3 °C) in the southeastern equatorial Pacific in the MAM season.

Obviously in a two-tiered system as FISH50, we have greater flexibility of correcting these systematic errors. However, unlike other flux or bias correction attempts (LaRow 2013; Kirtman 2003; Kirtman et al. 2002; Drijfhout and Walsteijn 1998), care was taken to exclude the period of forecast (1982–2008) to develop this SST climatology. The importance and difficulty to adhere to this rule of excluding the FISH50 hindcast period to develop the observed SST climatology is highlighted in Fig. 2, which shows the differences in the SST climatology

Table 1 FISH50 experiment design

Seasonal hindcast feature	Detail					
Length of each seasonal hindcast integration	6 months					
Number of ensemble members for each seasonal hindcast	6 (E1, E, E3, E4, E5, E6)					
Seasonal hindcast period	1982–2008					
Seasonal hindcast start date	E1	E2	E3	E4	E5	E6
	28 Nov, 0000 UTC	29 Nov, 0000 UTC	30 Nov, 0000 UTC	01 Dec, 0000 UTC	02 Dec, 0000 UTC	03 Dec, 0000 UTC
Atmospheric initial conditions	Borrowed for subsequent days from 28 Nov to 03 Dec from the NCEP-DOE reanalysis (Kanamitsu et al. 2002a, b) for each ensemble member and interpolated to the FISH50 grid					
Land initial conditions	Interpolated from NCEP-DOE reanalysis and kept identical in all ensemble members for each season					

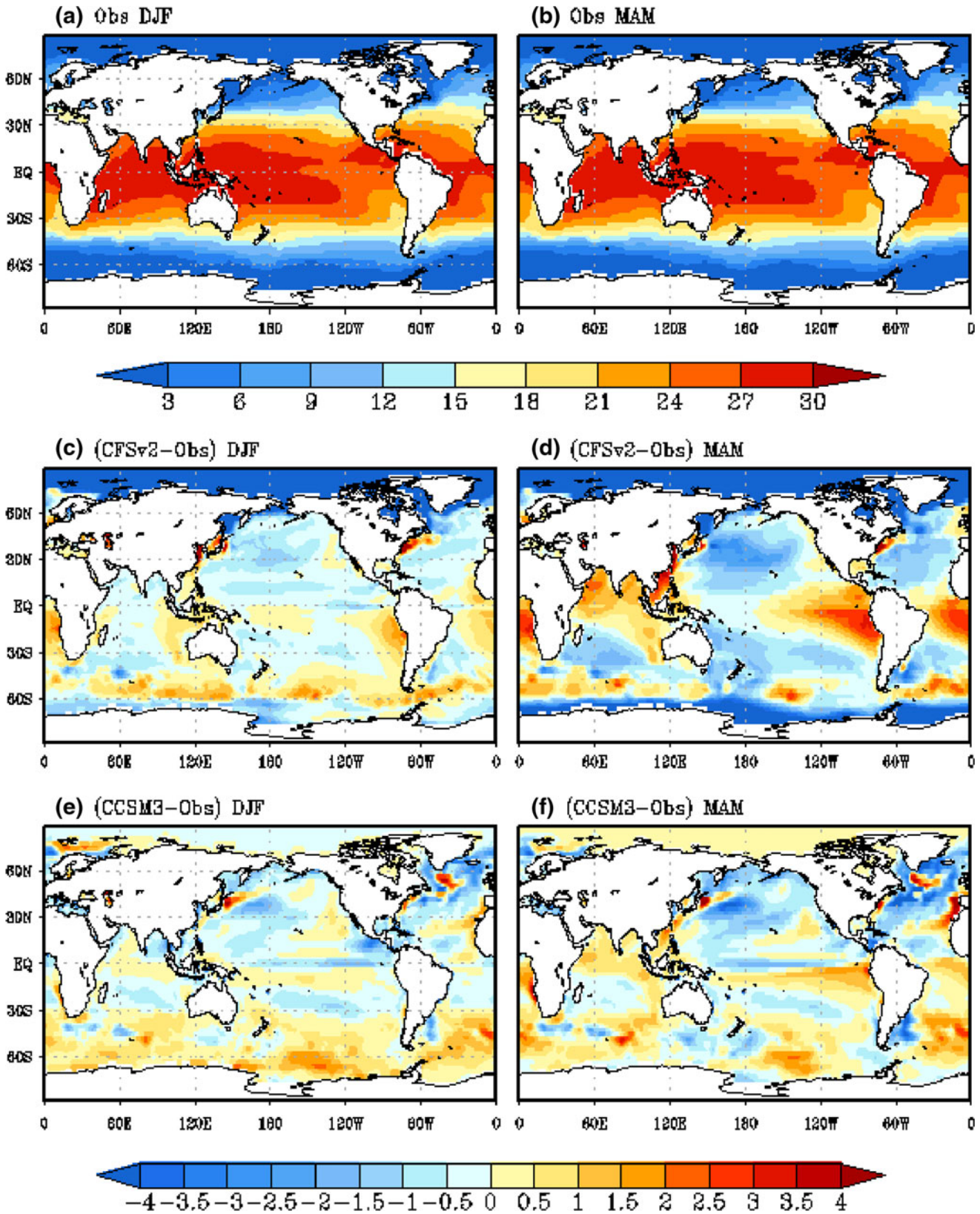

Fig. 1 The observed climatological SST for boreal **a** winter (DJF) season and **b** spring (MAM) season. The bias of hindcasted SST at zero lead for boreal winter season from **c** CFSv2, **e** CCSM3. Similarly, the bias of hindcasted SST at one season lead for boreal spring season from **d** CFSv2 and **f** CCSM3.0. The units are in °C

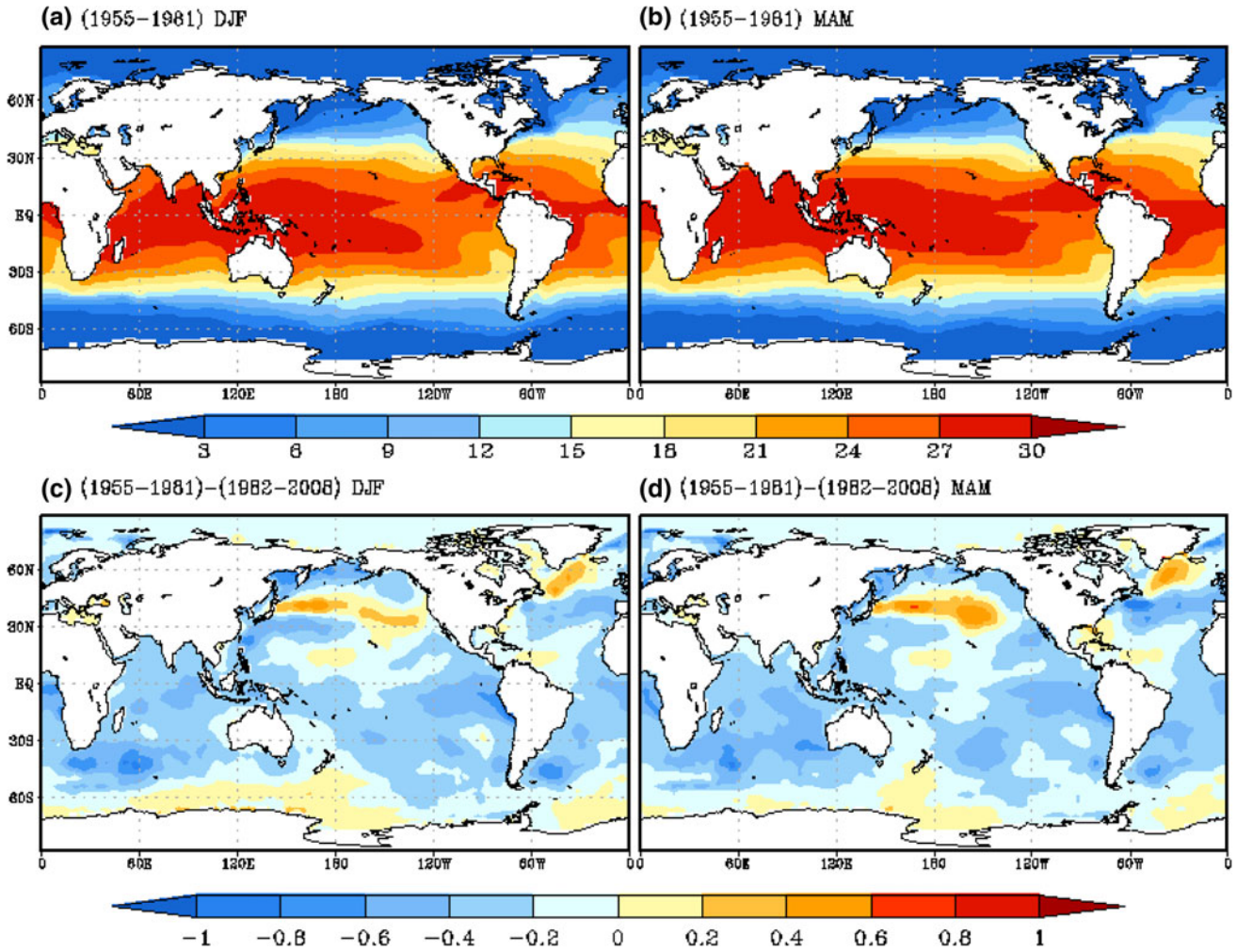

Fig. 2 The observed climatology SST computed over a period of 1955–1981 for **a** DJF season and **b** MAM season, and their corresponding differences with climatology computed over the period 1982–2008 for **c** DJF and **d** MAM. The units are in °C

computed in two adjacent 27 year periods of 1955–1981 and 1982–2008. In both the seasons of DJF and MAM, the systematic difference between the two periods in the equatorial oceans ranges from about 0.2 to 0.5 °C, which could be regarded as substantial. In other words, we cannot just use the previous 27 years of mean SST as the observed climatology for the FISH50 hindcast period. As an alternative we adopted a novel approach following Wu et al. 2009 to compute a time varying climatology that includes the secular change and decadal variations. Mathematically, this may be written as:

$$SST_F = SSTA_{CYCLE} + SSTA_{MME} + SST_{OLF} \quad (1)$$

where, SST_F is the forecast SST used to force FISH50 AGCM. $SSTA_{CYCLE}$ is the monthly climatology of ERSSTv3 (Smith et al. 2008) anomaly from 1901 to 1981. $SSTA_{MME}$ is the forecasted multi-model (CFSv2 and CCSM3) monthly mean averaged SST anomalies. SST_{OLF} is the observed low pass filtered SST obtained over the time period beginning from January 1880 to the start of seasonal hindcast. SST_{OLF} is updated at the start of each season and persisted through the integration period of the seasonal hindcast. To obtain SST_{OLF} we followed Misra et al. (2013), which involved first conducting a Multi-dimensional Ensemble Empirical Model Decomposition (MEEMD) analysis (Wu et al. 2009). MEEMD is a multi-dimensional (in space) data adaptive time series analysis of Ensemble Empirical Mode Decomposition (EEMD; Wu and Huang 2009). EEMD seeks to determine the intrinsic modes of oscillations in the data on the principle of local scale separation, which are called Intrinsic Mode Functions (IMFs). Although the decomposition in MEEMD does not make use of information on spatial coherence of the dataset, the obtained evolution of SST_{OLF} are both temporally and spatially coherent (true to its low frequency feature), which exhibits large spatial scale features when the SST_{OLF} is mapped on a global grid. The decomposition of the SST using MEEMD enables us to isolate the low frequency

modes ($\geq \sim 40$ years) and include them exclusively as part of SST_{OLF}. The monthly mean SST_F is interpolated to daily value following Taylor et al. (2000).

The AGCM used in FISH50 essentially follows from the formerly Experimental Climate Prediction Center's AGCM at Scripps Institute of Oceanography (Kanamitsu et al. 2002b; Shimpo et al. 2008) and now referred as the Florida Climate Institute-Florida State University Global Spectral Model (FGSM). A brief outline of the physics package used in the FGSM is presented in Table 2. It has 28 vertical (terrain following sigma) levels. We have however made some subtle but important changes of increasing the resolution to T248 spectral truncation (~ 50 km grid resolution) and replacing the convection scheme from Relaxed Arakawa Schubert (RAS; Moorthi and Suarez 1992) to Kain-Fritsch version 2 (KF2; Kain and Fritsch 1993; Kain 2004) scheme. The motivation for this change can easily be seen in the improvement of the AGCM's seasonal rainfall climatology for DJF and MAM seasons relative to the original version of the model (Fig. 3). For these test integrations displayed in Fig. 3, we ran a single ensemble member for 12 seasons (1982–1993) using the two different convection schemes at T248 spectral truncation forced with SST_F and compared the mean seasonal rainfall over the 12 seasons. In Fig. 3, it is clearly seen that KF2 improves the structure of the ITCZ globally relative to RAS. In the latter, the split ITCZ phenomenon is quite apparent especially in the tropical Indian and Pacific Oceans, which is greatly ameliorated in the KF2 version of the FGSM integrations. However, KF2 has a tendency to rain more relative to the RAS integration and observations.

Since we are going to be comparing FISH50 with the two coupled ocean–atmosphere models namely, CCSM3 and CFSv2, we briefly outline their physics in Table 2. The readers are referred to Collins et al. (2006a, b) and Saha et al. (2013) for further details on CCSM3 and CFSv2 respectively. In FISH50 we have 6 ensemble members per season (Table 1) and we therefore use as many from CCSM3 and CFSv2. The generation of the ensemble members in CCSM3 follows from Kirtman and Min (2009) and for CFSv2 from Saha et al. (2013).

3 Results

Since the seasonal hindcasts are global we will compare and validate the results on a larger scale (globally) and let region specific details for subsequent papers. We will in this paper hone in on surface air temperature and precipitation forecasts from FISH50, CCSM3.0 and CFSv2 seasonal hindcasts. The details of the validation datasets used in this section are provided in Table 3.

3.1 SST forcing

The bias in the seasonal mean SST_F (from Eq. 1; Fig. 4) for both DJF and MAM seasons is greatly reduced compared to the SST bias displayed by either CFSv2 or CCSM3.0 (Fig. 1). The bias in SST_F in Fig. 4 is uniformly in the range -0.5 to 0.5 °C, which is comparably far less than the large errors along the equatorial oceans, subtropical eastern oceans and in the higher latitude storm track regions of both hemispheres in CCSM3.0 and CFSv2 (Fig. 1). Similarly the standard deviation of the mean DJF SST_F (Fig. 5) shows that the variability along the equatorial Pacific Ocean and the Ecuadorian-Peruvian coast is comparable to the other two models. All three show slightly higher variability over the equatorial Pacific in the DJF season, while the mean DJF SST variation in the northern Pacific and in the northern Atlantic is marginally improved in SST_F compared to coupled seasonal hindcasts. In the MAM season, none of the SST forecast products capture the strong variations along the Ecuadorian-Peruvian coast. All of these forecasts contrary to observations exhibit strongest equatorial Pacific SST variations between ~ 90 and 160°W. The north Atlantic SST variations in

Table 2 A brief outline of the physics of the FISH50 AGCM, CFSv2 and CCSM3

Parameterization	FGSM	CFSv2	CCSM3
Cumulus parameterization	Kain-Fritsch (Kain and Fritsch 1993; Kain 2004)	Simplified Arakawa-Schubert (Hong and Pan 1998)	Zhang and McFarlane (1995)
Shallow convection	Tiektke scheme (Tiedtke 1983)	Tiektke scheme (Tiedtke 1983)	Hack (1994)
Boundary Layer	Nonlocal scheme (Hong and Pan 1996)	Nonlocal scheme (Hong and Pan 1996)	Holtsalag and Boville (1993)
Land surface	NOAH (Ek et al. 2003)	NOAH (Ek et al. 2003)	CLM3 (Oleson et al. 2004)
Gravity wave drag	Pierrehumbert (Alpert et al. 1988)	Chun and Baik (1998)	McFarlane (1987)
Shortwave radiation	M.-D. Chou (Chou and Lee 1996)	AER RRTM LW (Clough et al. 2005)	Berger (1978)
Longwave radiation	M.-D. Chou (Chou and Suarez 1994)	AER RRTM SW (Mlawer et al. 1997)	Ramanathan and Downey (1986)
CO_2	Time (annually) varying	Time (seasonally) varying	Fixed

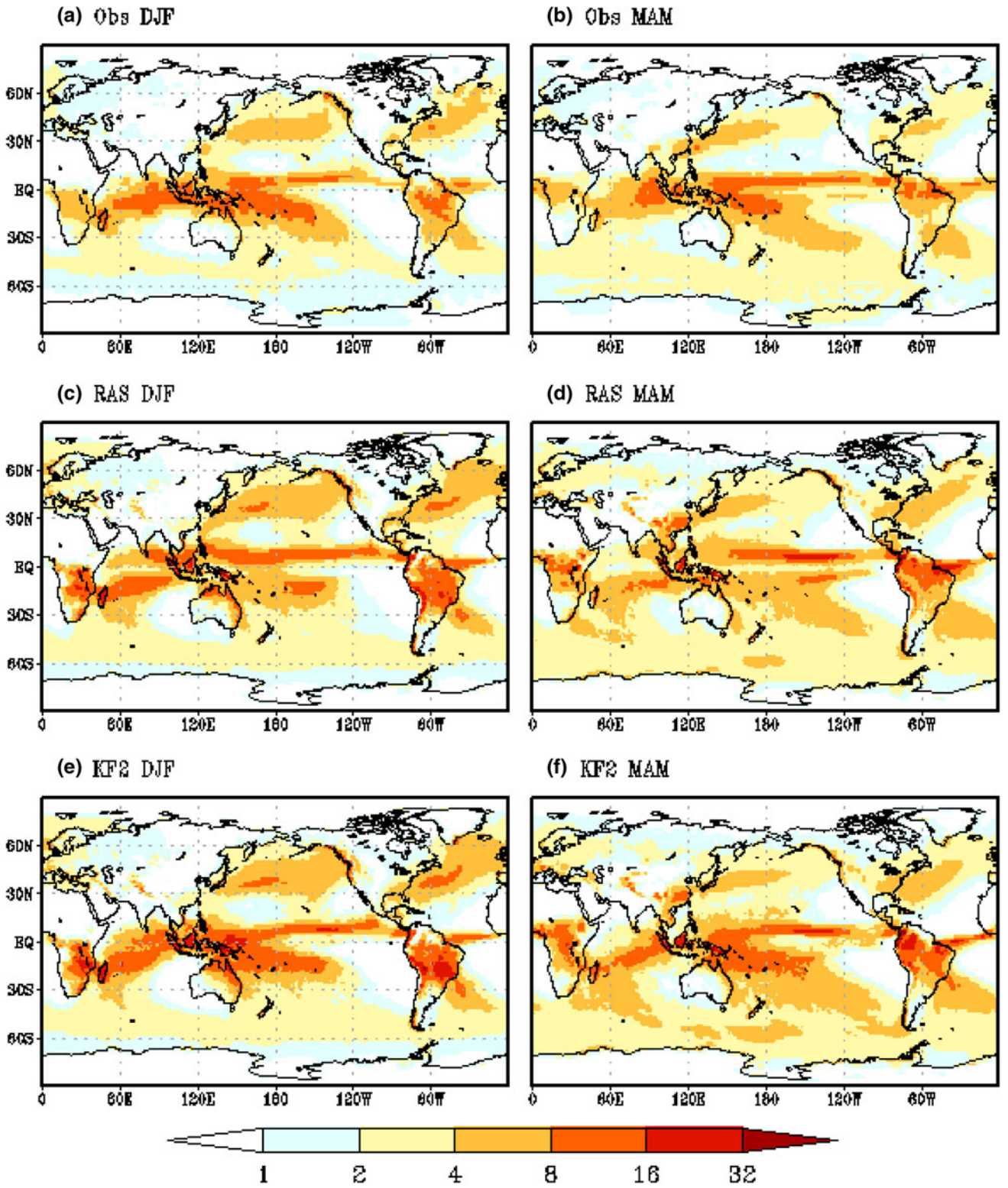

Fig. 3 The observed climatology of precipitation computed over a period of 1982–1993 for **a** DJF and **b** MAM seasons. The corresponding climatology of precipitation from a single member seasonal hindcast for the period of 1982–1993 using the RAS convection scheme for **c** DJF (at zero lead) and **d** MAM (one season lean) season. Likewise the climatology of precipitation from a single member seasonal hindcast for the period of 1982–1993 using the KF2 convection scheme for **e** DJF (at zero lead) and **f** MAM (one season lead). The units are in mm/day

Table 3 Details of the validation datasets used

Global dataset name	Variable	Reference	Resolution	Period available
Climate Prediction Center Merged Analysis of Precipitation (CMAP)	Precipitation	Xie and Arkin (1997)	2.5° × 2.5°	1979-present
Climate Research Unit version 3 (CRUv3)	Surface temperature	Mitchell and Jones (2005)	0.5° × 0.5°	1900-present

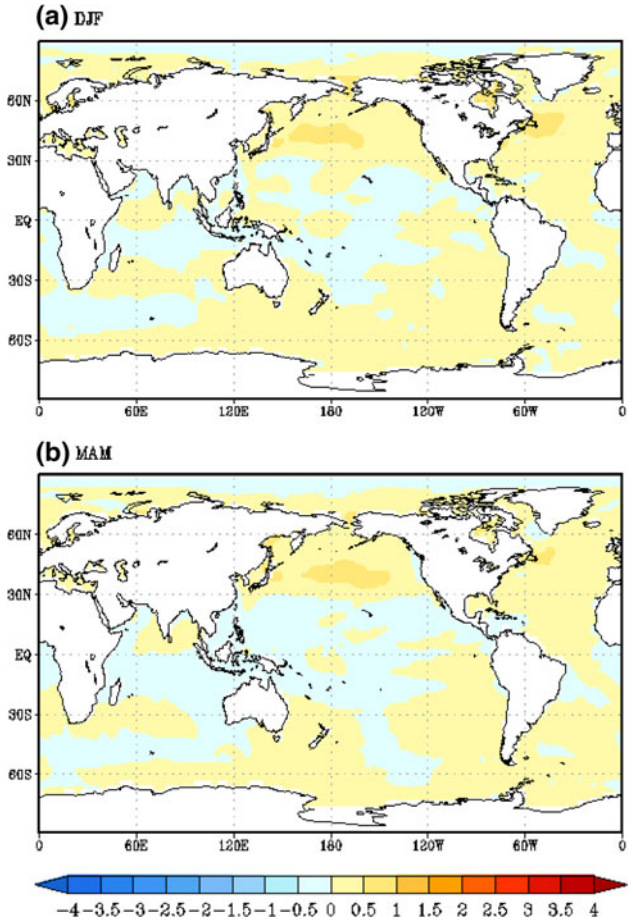

Fig. 4 The climatological SST bias computed for **a** DJF season (at zero lead) and **b** MAM season (at one season lead) from FISH50. The observed SST climatology was computed over the period 1982–2008 as shown in Fig. 1a, b. The units are in °C

SST_F are slightly improved over the corresponding variations in CFSv2 and CCSM3.0, while over north Pacific it is not so apparent.

3.2 FISH50 climatology

Figure 6 shows the observed seasonal climatology of precipitation for DJF and MAM seasons along with the corresponding RMSE for each of the three models. It is quite apparent that the RMSE of FISH50 is relatively much higher (as a result of erroneously higher precipitation rates) in the tropical latitudes than in either CCSM3.0 and CFSv2 hindcasts in both DJF and MAM seasons. In the DJF season, CCSM3.0 displays the least RMSE, while in FISH50 and in CFSv2 the RMSE are especially large over the western Pacific warm pool region, where they tend to rain more than the observations. In FISH50 the RMSE is also large over southeastern Africa and southeastern Brazil. In the MAM season the RSME are large in the southern tropics in CCSM3 relative to CFSv2 while still significantly less than that in FISH50. The RMSE in FISH50 in the MAM season continues to be large both over the tropical oceans and over land compared to either CFSv2 and CCSM3.0. The larger tropical RMSE in FISH50 may be highlighting the impact of the absence of the coupled air-sea interactions that could dampen the tropical rainfall activity.

The boreal winter and spring climatology of observed surface temperature and the corresponding RMSE from the three models are shown in Fig. 7. The RMSE over Sahara in northern Africa is rather unique in FISH50. Similarly, the errors in northern Russia and Canada are large in FISH50 compared to that in CCSM3.0 and CFSv2. From Fig. 7 it is seen that except for the semi-arid regions and high altitude region (e.g. Tibetan Plateau) the RMSE in FISH50 is comparatively small.

3.3 Deterministic predictability

Deterministic predictability in some ways is a fallacy of climate prediction if it is not complimented with probabilistic assessment of skill (Palmer et al. 2000; Kirtman 2003). It is argued that both weather and climate prediction are inherently non-deterministic because of uncertainty in the initial conditions, imperfect and non-linear model that result in chaotic evolution of the climate system. However deterministic skill analysis does provide some (but not complete) insight into the behavior of the forecast system as will be apparent by the conclusion of this paper, when we also compare the models for their probabilistic skill.

In Fig. 8 we show the correlation of the seasonal mean precipitation anomalies (of the ensemble mean) from the three model's seasonal hindcasts with the corresponding observed precipitation anomalies for both DJF (zero lead) and MAM (one season lead). In the DJF season, FISH50

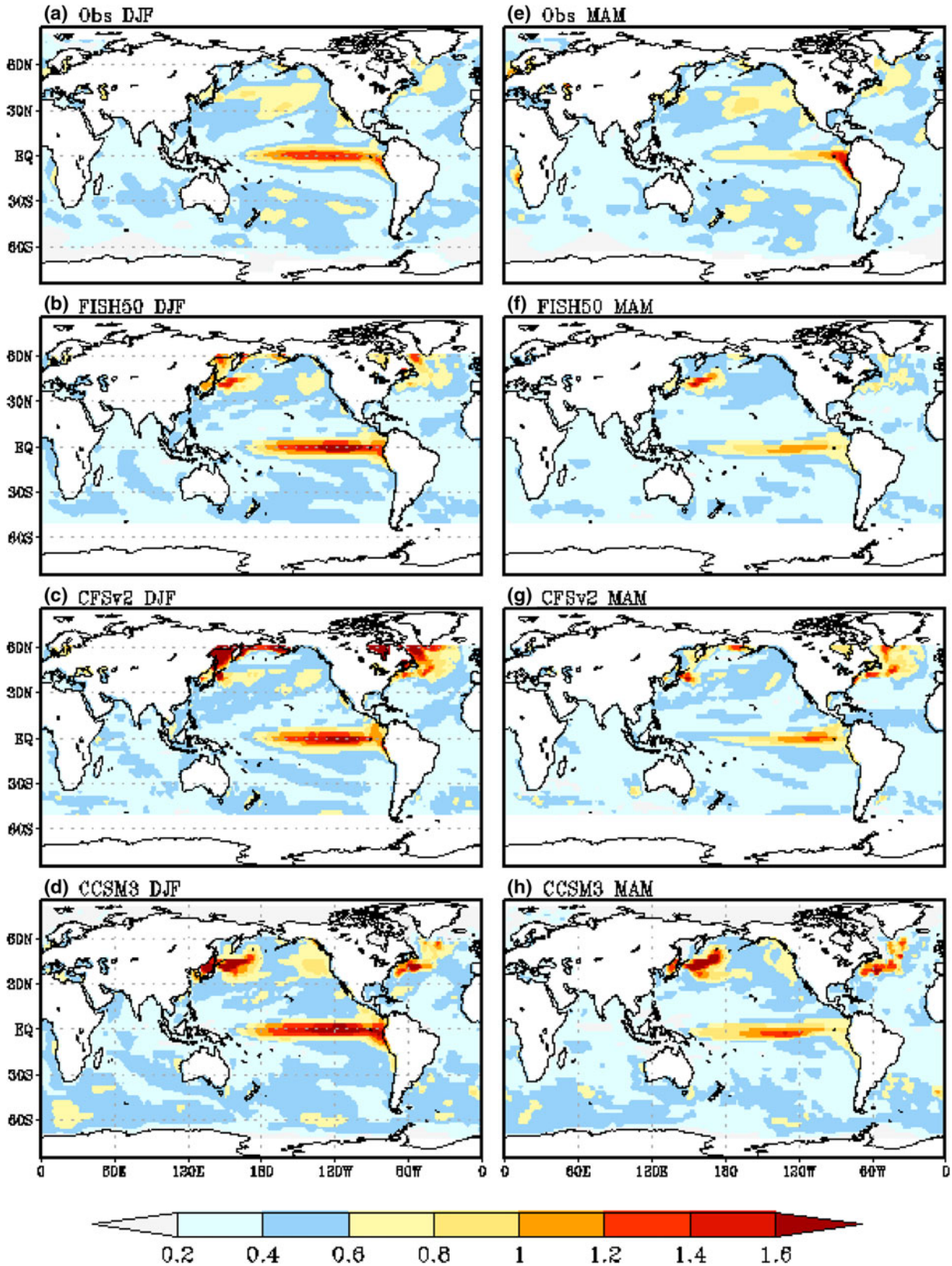

Fig. 5 The standard deviation of DJF seasonal mean SST from **a** observations, and seasonal hindcasts at zero lead from **b** FISH50, **c** CFSv2, and **d** CCSM3. Similarly the standard deviation of MAM seasonal mean SST from **e** observations, and seasonal hindcasts at one season lead from **f** FISH50, **g** CFSv2, and **h** CCSM3. The units are in °C

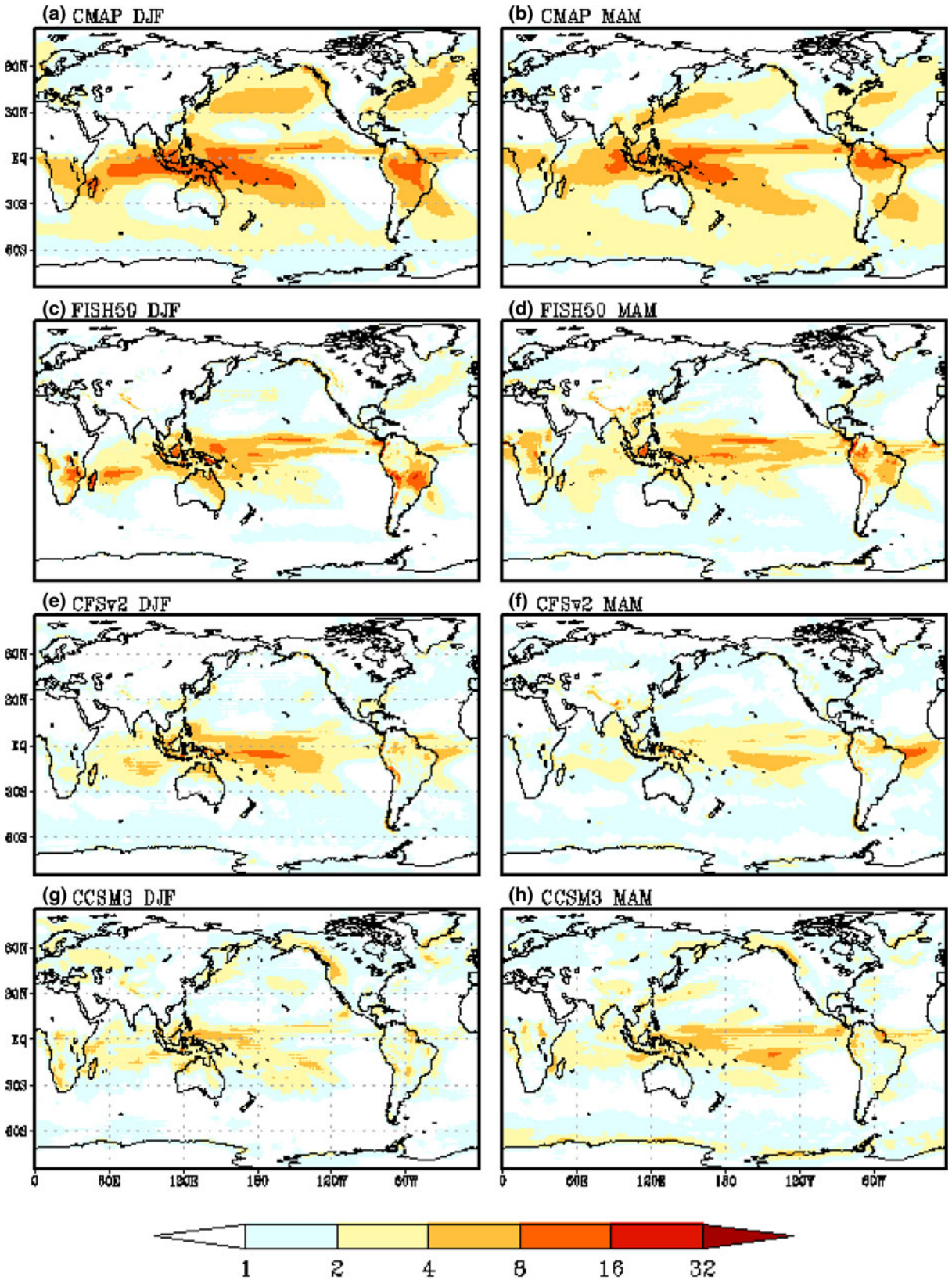

Fig. 6 The observed climatology of precipitation (1982–2008) in **a** DJF, and **b** MAM. The root mean square error of the ensemble mean precipitation for DJF (zero lead) for seasonal hindcasts from **c** FISH50, **e** CFSv2, and **g** CCSM3. Likewise, the root mean square error of the ensemble mean precipitation for MAM(one season lead) for seasonal hindcasts from **d** FISH50, **f** CFSv2, and **h** CCSM3. The units are in mm/day

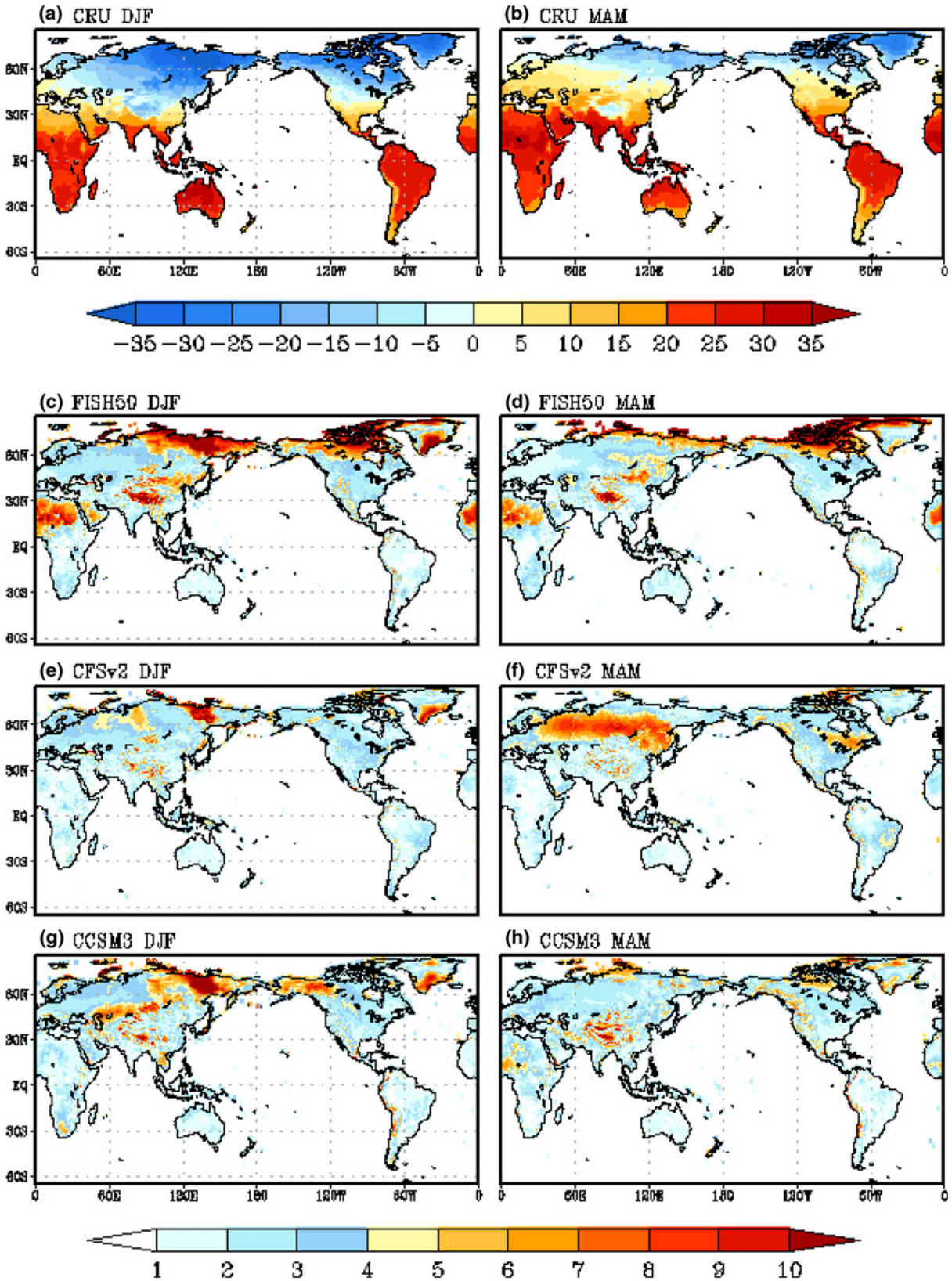

Fig. 7 The observed climatology of surface temperature (1982–2008) in **a** DJF, and **b** MAM. The root mean square error of the ensemble mean T2 m for DJF (zero lead) for seasonal hindcasts from **c** FISH50, **e** CFSv2, and **g** CCSM3. Likewise, the root mean square error of the ensemble mean T2 m for MAM (one season lead) for seasonal hindcasts from **d** FISH50, **f** CFSv2, and **h** CCSM3. The units are in °C

displays a strong positive correlation over southwestern and southeastern United States, eastern Africa, and over northeastern South America, which are some well known regions for ENSO teleconnections (Ropelewski and Halpert 1986, 1987; AchutaRao and Sperber 2006). In the CFSv2 and CCSM3 seasonal hindcasts, the DJF seasonal precipitation anomalies displays similar but somewhat weaker correlations over these continental regions. As noted earlier, the MAM season comparatively shows weaker correlations than in the DJF season in all three models. In case of FISH50, the positive correlations are shifted to northwestern US and northeast Brazil that are well known again for ENSO teleconnections in the boreal spring season (Ropelewski and Halpert 1986; Moura and Hastenrath 2004). Similarly CFSv2 display positive correlations, albeit weaker than FISH50 over both of these continental regions. CCSM3.0 seasonal hindcasts, however display the weakest positive correlations in the MAM season compared to the other two models over northeast Brazil and over northwestern US.

A similar correlation of seasonal mean surface temperature anomalies from all three models with corresponding observations is shown in Fig. 9. In the DJF season, FISH50 displays a more extensive positive correlation over Africa, Australia, South America, and North America than either CFSv2 and CCSM3.0. It could be argued in the DJF season that CFSv2 exhibits a larger positive correlation than any other model over equatorial Africa, Arabian Peninsula, and tropical South America. The corresponding correlations in CCSM3.0 display the least extensive and weakest positive correlations over majority of the continental regions. This could be an artifact of the concentration of CO_2 being fixed in the CCSM3.0 seasonal hindcasts unlike in the other two models in which it is varied. In the MAM season, FISH50 displays a less extensive positive correlation over Africa (with notable reduction over Sahara), reduction of correlation over Australia and less extensive correlations over United States (with notable positive correlations over southwestern and northwestern United States). The corresponding correlations in CFSv2 are nearly comparable to that in FISH50. CCSM3.0 is able to maintain the correlations from the previous season.

3.4 Probabilistic prediction skill

Following Mason and Graham (1999, 2002) we compute the Area under the Relative Operating Characteristic Curve (AROC) to assess the probabilistic skill of the seasonal hindcasts. In Appendix we compare the seasonal hindcasts of the three models (FISH50, CFSv2, and CCSM3) with the rest of the National Multi-Model seasonal hindcasts. We have analyzed these skills for the lower, middle and upper terciles for both seasonal mean precipitation and surface land temperature for both DJF and MAM seasons. Unlike the deterministic skill that evaluates the ensemble mean anomalies, AROC is a conditional probability metric that provides the forecast probability for events defined by the user (in this case terciles). The thresholds for the terciles were based on the respective model hindcasts and observations separately. In an operational environment, AROC serves as a useful way to assess a priori an optimal strategy to issue warnings for specific events based on the hindcast performance of the forecast system.

In Fig. 10 we show the AROC for FISH50 seasonal precipitation anomalies in DJF and MAM seasons for lower, middle, and upper terciles. In comparison to Fig. 8 it is immediately apparent that more hindcast skill can be harvested through this approach. In DJF season, FISH50 displays much higher skill than climatology for low and upper tercile events over a vast transect of the global tropics including both land and ocean points. Even higher latitude regions display higher skill than climatology in the first season (DJF) of the FISH50 hindcast, which does not seem to be so obvious from Fig. 8. In the subsequent season of MAM (one season lead) the AROC values diminish in their magnitude, but are still sustained (higher than climatology) from the previous season. It is interesting to note that in both seasons, the middle tercile has lesser AROC values than extreme terciles and are also less spatially coherent.

Figure 11 compares the AROC for seasonal precipitation anomalies over global oceans, global tropical oceans, global land, and global tropical land regions between the three models. The following may be noted from Fig. 11:

- All three models over all these four regions show a reduction of AROC from going from DJF to MAM season for all terciles.
- All three models display a higher AROC for the extreme terciles than the middle tercile.
- All three models display higher skill over the oceans than over land in both seasons and for the extreme tercile events.
- Outside of the global tropics, FISH50 has marginally higher AROC than either of the two models for the high tercile events in the DJF and MAM seasons.
- For the low tercile events outside of the global tropics, the AROC's are comparable between FISH50 and CFSv2 with significantly lower values for CCSM3.0.
- For extreme terciles in the global tropics for the DJF season CFSv2 seems to have the highest AROC except for high tercile events over the global tropical land where FISH50 displays the highest values.
- In the MAM season over global tropical oceans, the AROC values in FISH50 and CFSv2 are comparable while in CCSM3.0 it is significantly less but still above climatology. However over global tropical land

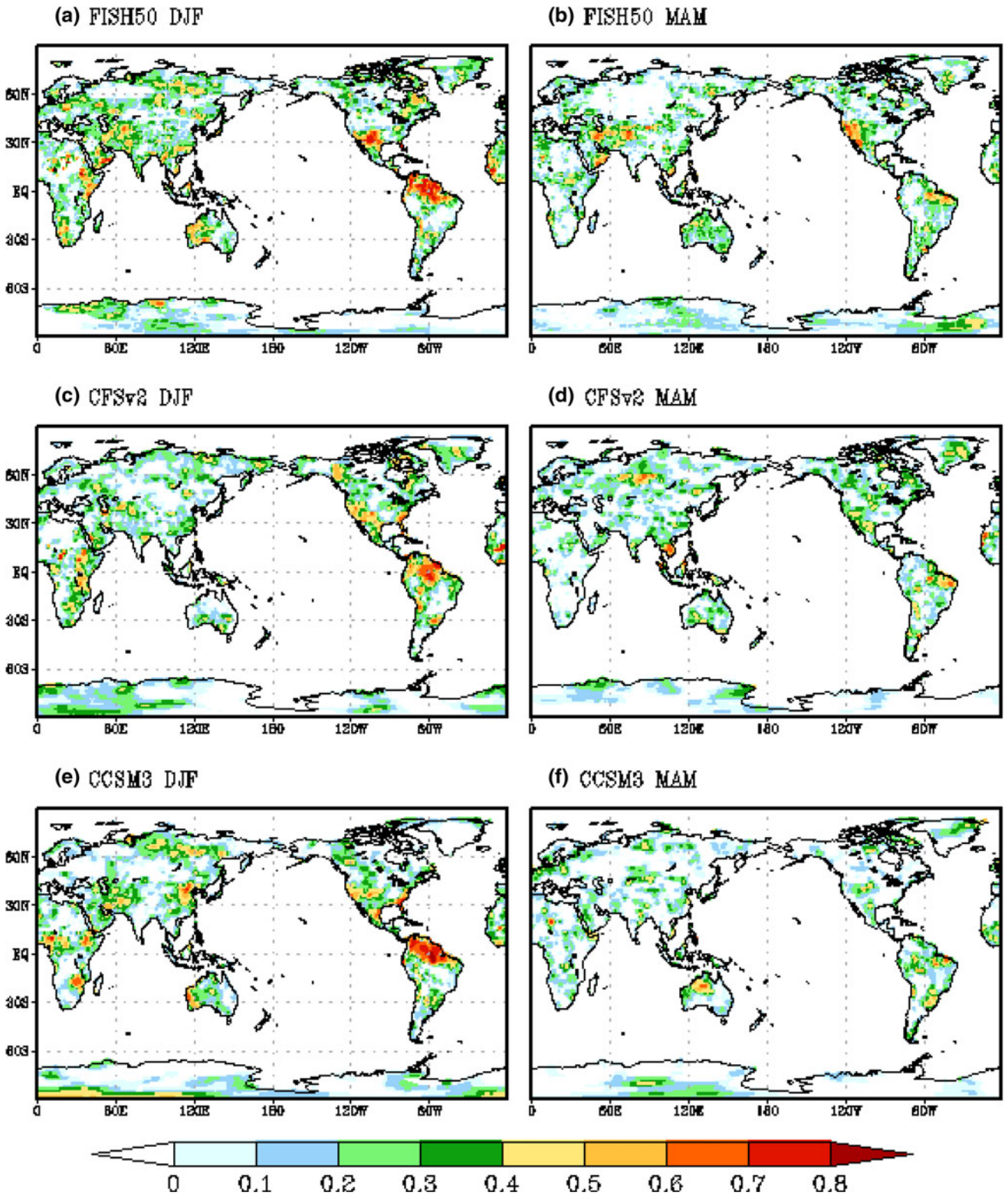

Fig. 8 The correlation of the ensemble mean precipitation for DJF (zero lead) from **a** FISH50, **c** CFSv2, and **e** CCSM3. Similarly, the correlation of the ensemble mean precipitation for MAM (one season lead) from **b** FISH50, **d** CFSv2, and **f** CCSM3

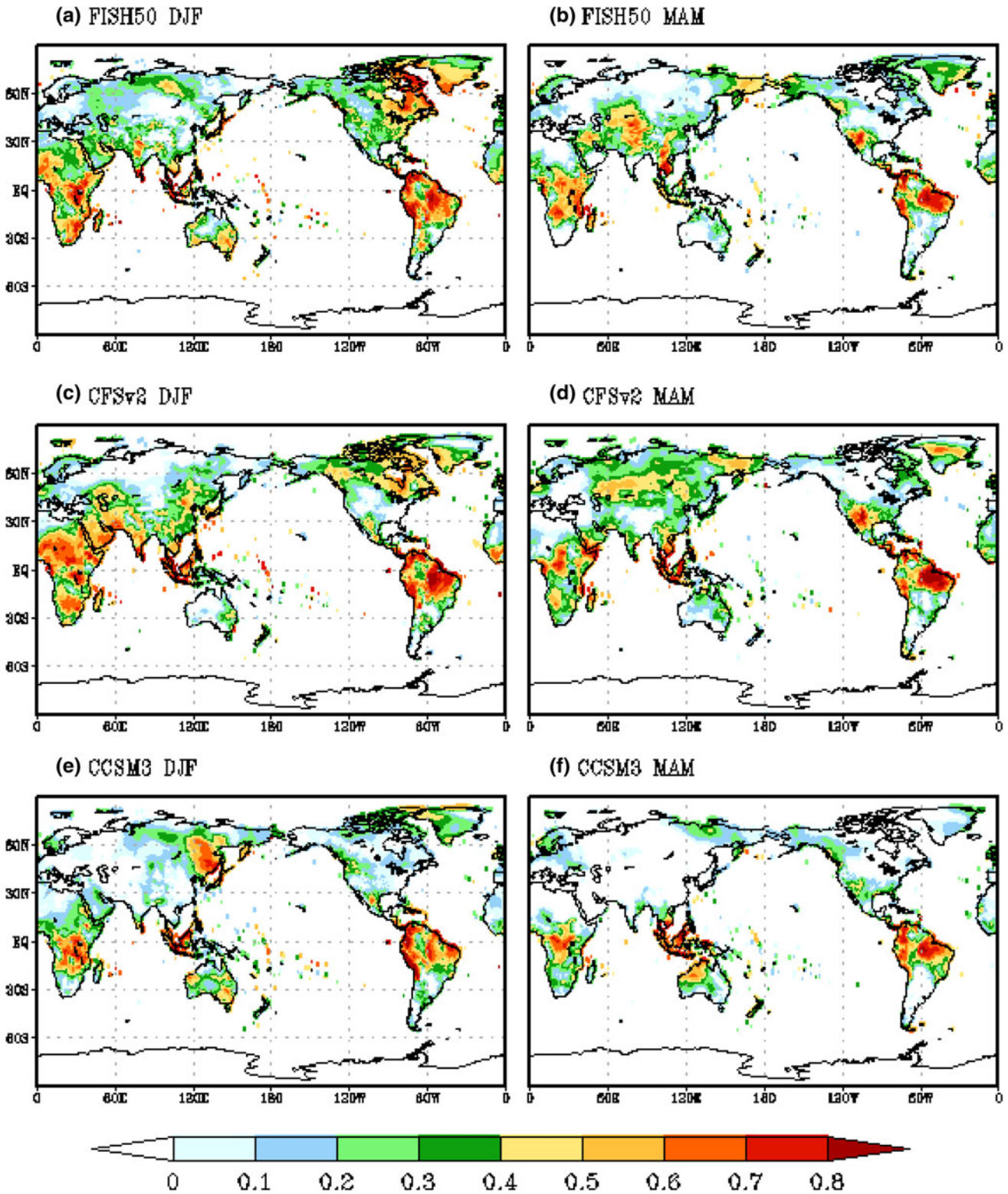

Fig. 9 The correlation of the ensemble mean T2 m for DJF (zero lead) from **a** FISH50, **c** CFSv2, and **e** CCSM3. Similarly, the correlation of the ensemble mean precipitation for MAM (one season lead) from **b** FISH50, **d** CFSv2, and **f** CCSM3

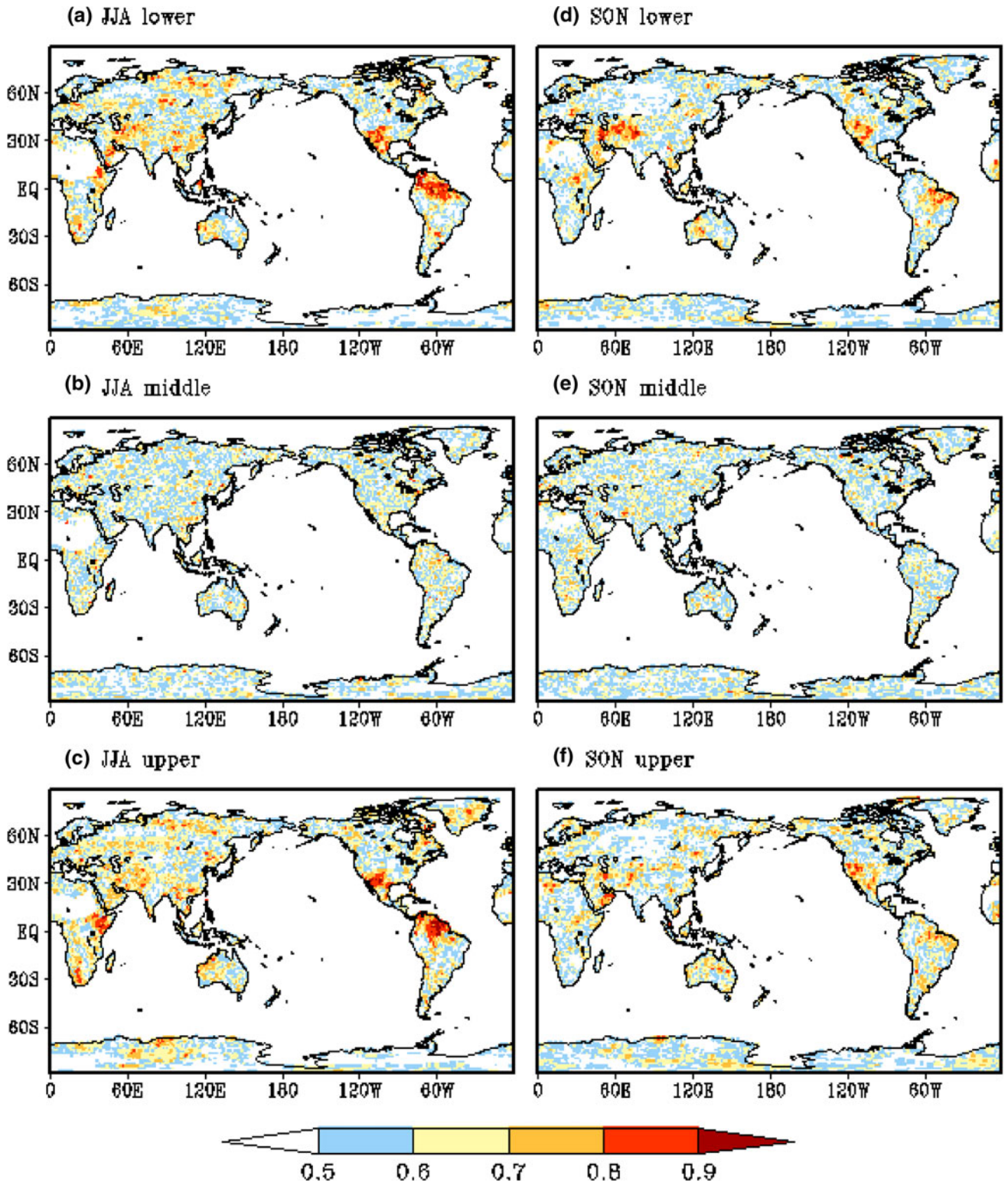

Fig. 10 The area under the relative operation characteristic curve (AROC) for **a** *lower*, **b** *middle*, and **c** *upper* tercile for DJF (zero season lead) from FISH50 precipitation, Similarly, the area under the ROC for **d** *lower*, **e** *middle*, and **f** *upper* tercile for MAM (one season lead) from FISH50 precipitation. Area over 0.5 is colored and signifies higher skill than climatology

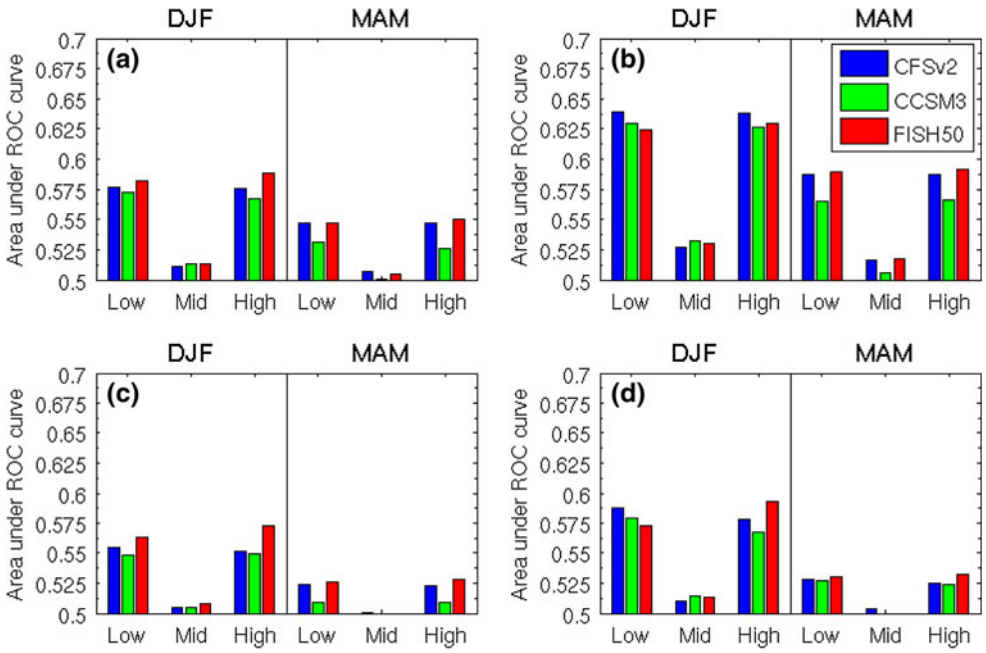

Fig. 11 Area under the ROC averaged over **a** global oceans, **b** tropical oceans, **c** global land, and **d** tropical land for *low*, *middle*, and *upper* terciles of CFSv2, CCSM3, and FISH50 precipitation in DJF and MAM

regions, all three models display comparable AROC values, with notably less skill than climatology for the middle tercile.

Figure 12 similar to Fig. 10 shows the spatial distribution of the AROC for seasonal mean surface land temperature anomalies from FISH50 for the three terciles. In contrast to Fig. 9, we again see evidence of more probabilistic prediction skill that can be gleaned from FISH50. However, unlike seasonal precipitation anomalies in Fig. 10, the AROC for extreme terciles of surface temperature are higher and more extensive over land especially over Africa, South and North America in both DJF and MAM seasons. Once again we notice that for middle tercile events, the hindcast skill of FISH50 is relatively less than that for the extreme tercile events. Comparing the AROC's across the three models for surface land temperature for global land and tropical land areas (Fig. 13) we notice that in the DJF season the high skills for the extreme tercile events are comparable between CFSv2 and FISH50 with CCSM3.0 showing slightly less skill. However in the MAM season, CFSv2 shows its superiority over the other two models both over global land and global tropical land areas, while AROC in FISH50 remains higher than CCSM3.0 especially outside of the global tropics.

4 Summary and conclusions

In part I of this paper we presented the results of the seasonal hindcasts for boreal summer and fall seasons. In this part II of the paper the analysis of seasonal hindcasts for boreal winter and spring seasons are analyzed. At the outset it seems that the two tiered system of FISH50 forced with the bias corrected forecasted SST from single tiered system offers complementary seasonal prediction skill to the coupled ocean–atmosphere forecast systems. A systematic comparison of FISH50 winter and spring seasonal hindcasts with corresponding hindcasts from single tiered systems of CFSv2 and CCSM3.0 reveal the following:

The RMSE of precipitation and surface land temperature are higher in FISH50 and is least in CCSM3.0. In case of precipitation, the RMSE are larger in the tropical latitudes while for surface land temperature it appears at higher latitudes.

Skillful boreal winter (spring) season rainfall anomalies over southeastern and southwestern United States, northeastern South America and eastern Africa (northwestern United States) appears to be a forced signal as they appear in the deterministic skill analysis. It may be noted that in all these regions FISH50 displays the largest skill compared to either CFSv2 and CCSM3.0.

Similarly the correlations of the seasonal surface land temperature anomalies from the seasonal hindcasts and corresponding observations are most extensive over Africa, South and North America in FISH50 relative to the other two models in both boreal winter and spring seasons.

The probabilistic skill analysis reveals that in all three models there is more forecast skill to be gleaned for the

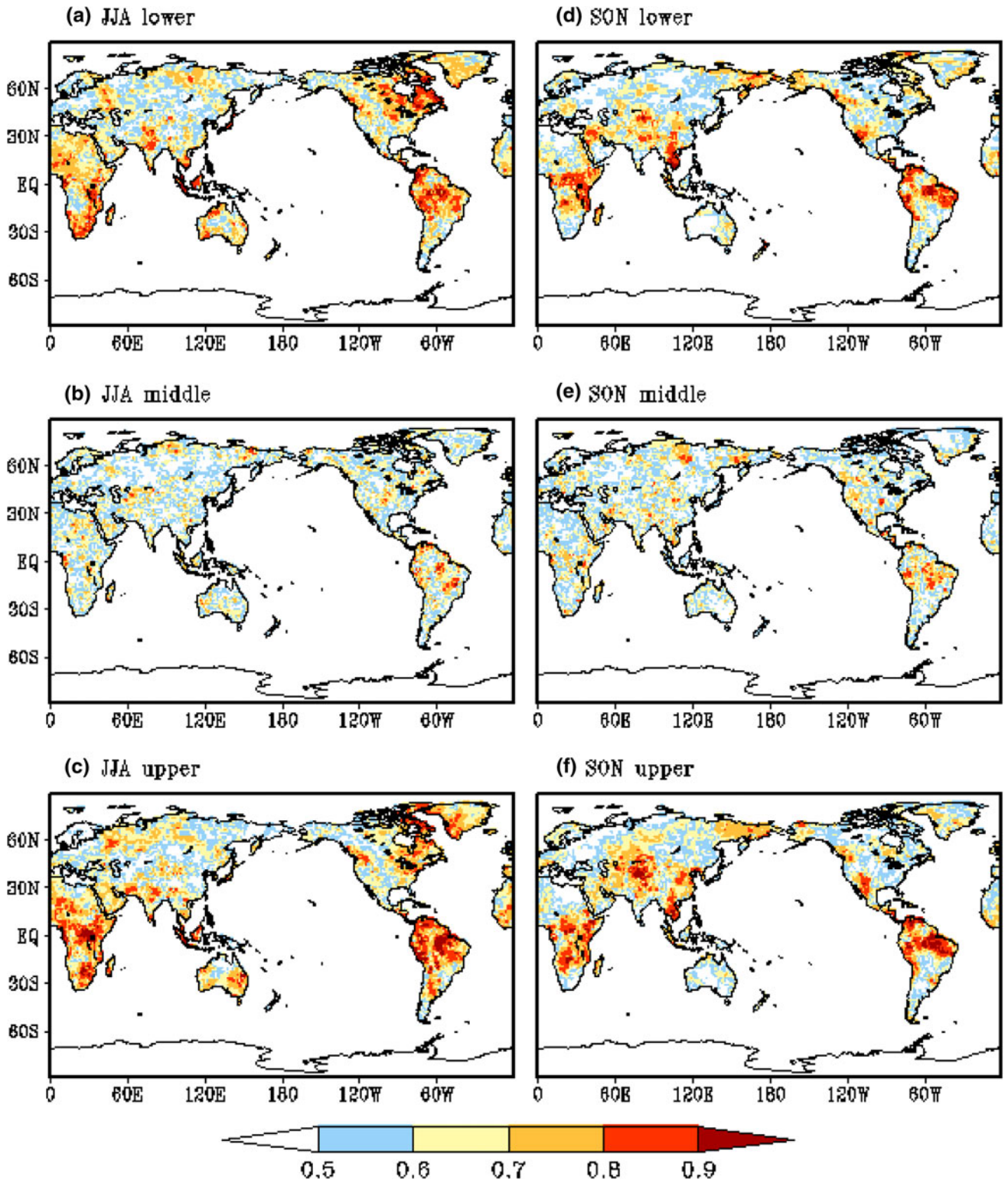

Fig. 12 The area under the relative operation characteristic curve (ROC) for **a** *lower*, **b** *middle*, and **c** *upper* tercile for DJF (zero season lead) from FISH50 T2 m, Similarly, the area under the ROC for **d** *lower*, **e** *middle*, and **f** *upper* tercile for MAM (one season lead) from FISH50 T2 m. Area over 0.5 is colored and signifies higher skill than climatology

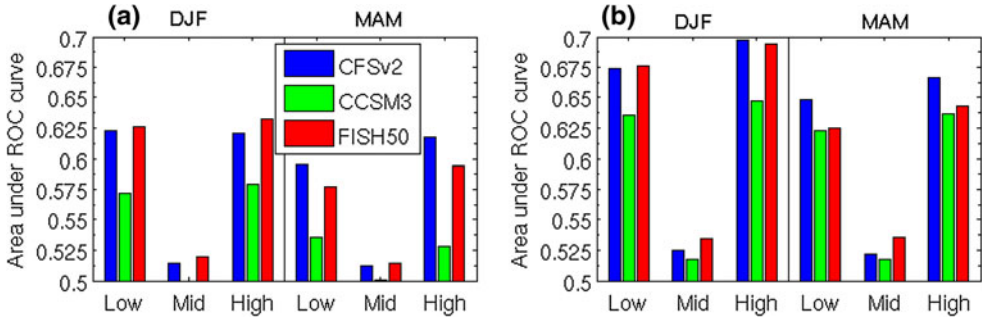

Fig. 13 Area under the ROC averaged over **a** global land and **b** tropical land for *low*, *middle*, and *upper* terciles of CFSv2, CCSM3, and FISH50 temperatures in DJF and MAM

extreme tercile events in both seasons (DJF and MAM) and for both variables (precipitation and surface land temperature), while for the middle tercile events CFSv2 and FISH50 (CCSM3.0) are marginally better (worse) than climatology.

The changes made to the convection scheme, the increase of resolution to T248 spectral truncation (~50 km grid resolution), and use of the unique way of bias correction of dynamically forecasted SST may all have contributed to the displayed fidelity of FISH50. The deterioration of the skill from DJF to MAM season across all three models suggests some further investigation is required to discern the role of increasing lead time and the inherent seasonal nature of the climate system (e.g. the spring predictability barrier). Nonetheless, this study shows that coupled forecast systems may have reached a stage, wherein the forced forecast systems like FISH50 could be used to exploit the superiority of the SST forecasts to glean further seasonal prediction skill. The advantage of stand alone AGCM's is that it is computationally less demanding to raise their resolution to levels that can then be more meaningful for application in other fields (e.g. hydrology, agriculture etc.). FISH50 in this study unlike in the boreal summer and fall seasons (in part I of the paper) do not seem to show as much of a benefit over the much coarser dynamical coupled ocean–atmosphere seasonal forecasts. However, it should be mentioned that all three models display hindcast skills, which is better than climatology.

Acknowledgments This paper is dedicated to the memory of Dr. Masao Kanamitsu, without whose pioneering development of the FISH50 AGCM, this work would not have been possible. We also acknowledge the help of Dr. Zhaohua Wu who provided us the methodology and the data for the bias corrected SST (SST_{OLF}). We thank Mr. Steven DiNapoli for making Figs. 14, 15, 16. This work was supported by grants from NOAA (NA12OAR4310078, NA10OAR 4310215, NA11OAR4310110), USGS (06HQGR0125), and USDA (027865). All computations for this paper were done on the computational resources provided by the Extreme Science and Engineering Discovery Environment (XSEDE) under TG-ATM120017 and TG-ATM120010.

Appendix: Comparison of FISH50 with the other National Multi-Model Ensemble (NMME) models

The NMME project (Kirtman et al. 2013; http://www.cpc.ncep.noaa.gov/products/ctb/nmme/) hosted by International Research Institute for Climate and Society, Columbia University and maintained in real time at the NCEP Climate Prediction Center (http://www.cpc.ncep.noaa.gov/products/NMME/) are eight single tiered coupled ocean–atmosphere models, which have conducted extensive seasonal hindcasts over the same time period as FISH50 and more. In fact NMME models have completed seasonal hindcasts for several lead times throughout the year and here we compare the AROC for tercile events of seasonal mean surface land temperature and precipitation from FISH50 at zero (one season) lead time for JJA (SON) with the corresponding hindcasts of the NMME. The horizontal and vertical resolutions of the NMME models along with their references are shown in the Table below.

National Multi-Model Ensemble (NMME) models

Model	Horizontal resolution	Vertical resolution	References
CFSv1	T62 (~200 km)	64 sigma	Saha et al. (2006)
CFSv2	T126 (~100 km)	64 sigma-pressure	Saha et al. (2010)
CCSM3	T85 (~140 km)	26 sigma-pressure	Kirtman and Min (2009)
ECHAM-Anom	T42 (~250 km)	19 sigma-pressure	DeWitt (2005)
ECHAM-Dir	T42 (~250 km)	19 sigma-pressure	DeWitt (2005)
GFDL	2 × 2.5 degrees	24 Layers	Zhang et al. (2007)
GFDL-aer04	2 × 2.5 degrees	24 Layers	Zhang et al. (2007)
GMAO	2 × 2.5 degrees	34 Layers	Bacmeister et al. (2000)

See Figs. 14, 15, 16.

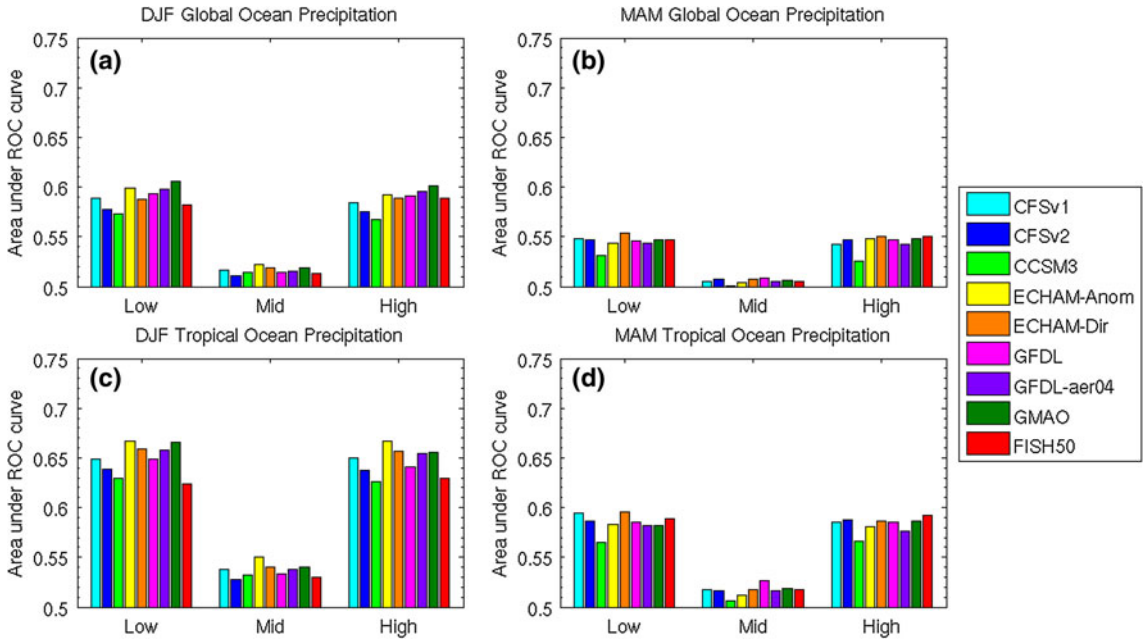

Fig. 14 AROC averaged over global oceans for **a** DJF, **b** MAM, over tropical oceans for **c** DJF, and **d** MAM for *low*, *middle*, and *upper* terciles of NMME and FISH50 precipitation

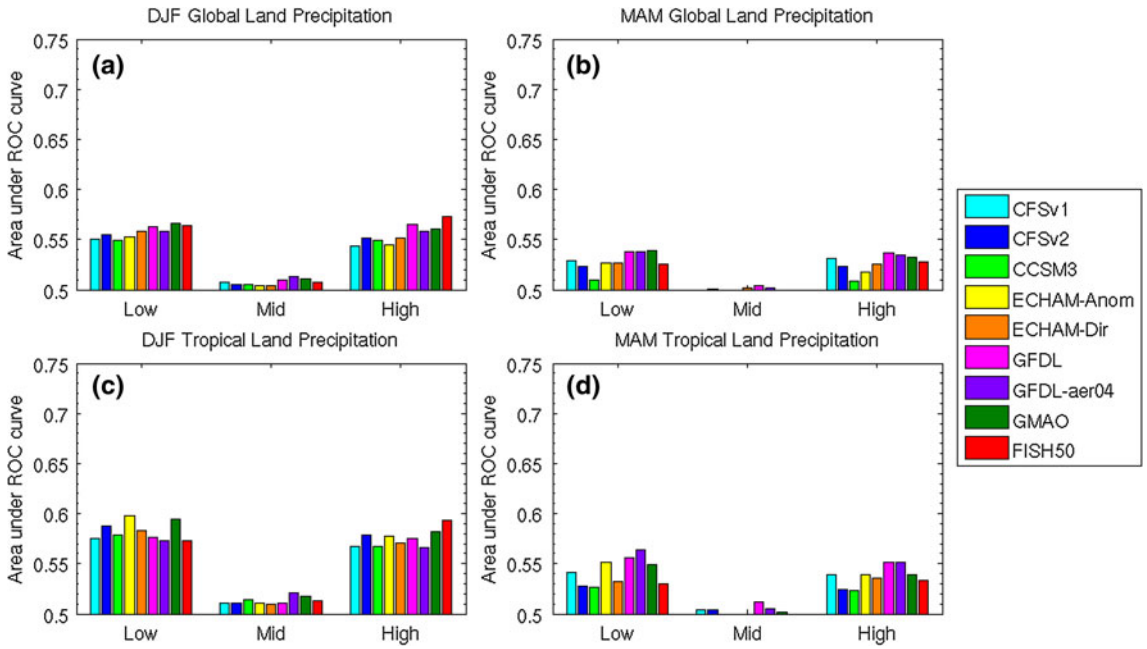

Fig. 15 AROC averaged over global land for **a** DJF, **b** MAM, over tropical land for **c** DJF, and **d** MAM for *low*, *middle*, and *upper* terciles of NMME and FISH50 precipitation

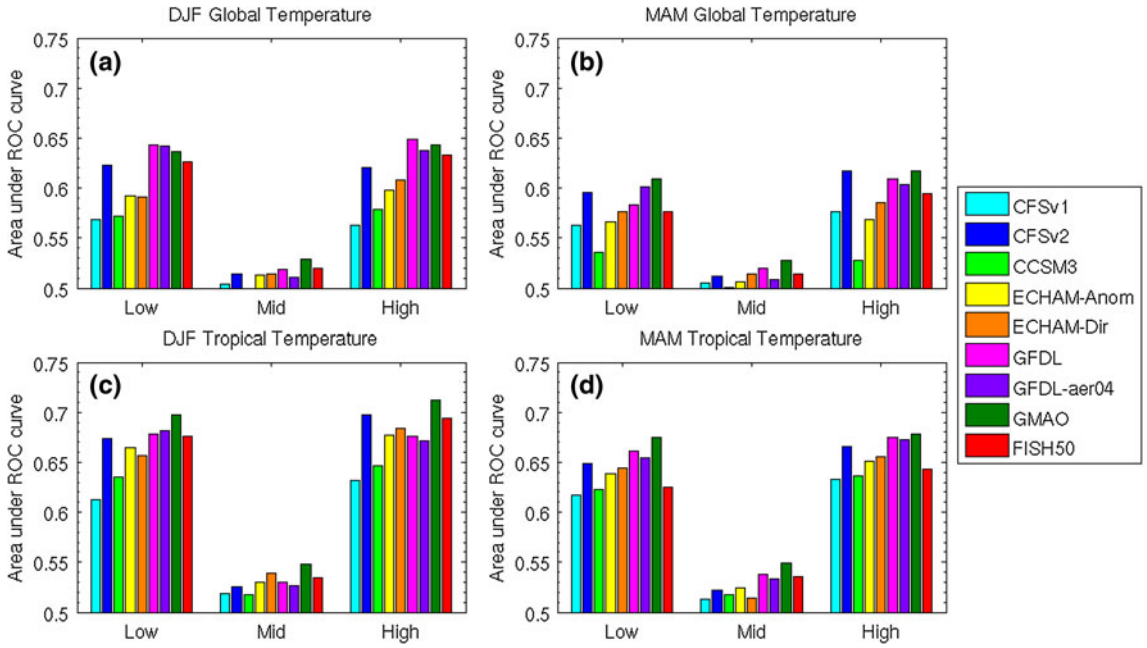

Fig. 16 AROC averaged over global land for **a** DJF, **b** MAM, over tropical land for **c** DJF, and **d** MAM for *low*, *middle*, and *upper* terciles of NMME and FISH50 surface land temperature

References

AchutaRao KM, Sperber KR (2006) ENSO simulation in coupled ocean-atmosphere models: are the current models better? Clim Dyn. doi:10.1007/s00382-006-0119-7

Oleson KW et al (2004) Technical description of the Community Land Model (CLM). Tech. Rep. NCAR/TN-461 + STR, NCAR, Boulder, CO, 174 p

Kirtman BP et al (2013) The North American multi-model ensemble (NMME) for intra-seasonal to interannual prediction. Bull Am Met Soc (Submitted)

Saha S et al (2013) The NCEP climate forecast system version 2. J Climate (Submitted) Available from http://cfs.ncep.noaa.gov/cfsv2.info/CFSv2_paper.pdf

Alpert JC, Kanamitsu M, Caplan PM, Sela JG, White GH, Kalnay E (1988) Mountain induced gravity wave drag parameterization in the NMC medium-range model. Preprints, Eighth conference on numerical weather prediction, Baltimore, MD. Am Meteor Soc 726–733

Bacmeister J, Pegion PJ, Schubert SD, Suarez MJ (2000) An atlas of seasonal means simulated by the NSIPP 1 atmospheric GCM. Vol. 17. NASA Tech. Memo. 104606, Goddard Space Flight Center, Greenbelt, MD, 194 p

Barnston AG et al (1994) Long lead seasonal forecasts-where do we stand? Bull Am Meteorol Soc 75:2097–2114

Bengtsson L, Schlese U, Roeckner E (1993) A two-tiered approach to long-range climate forecasting. Science 261:1026–1029

Berger AL (1978) Long-term variations of daily insolation and quaternary climate changes. J Atmos Sci 25:2362–2367

Bohn TJ, Sonessa MY, Lettenmaier DP (2010) Seasonal hydrologic forecasting: do multi-model ensemble averages always yield improvements in forecast skill? J Hydrometeorol 11:1358–1372

Chou M-D, Suarez MJ (1994) An efficient thermal infrared radiation parameterization for use in general circulation models. In: NASA Technical report series on global modeling and data assimilation, NASA/TM-1994-104606, vol 3. Goddard Space Flight Center, Greenbelt, USA

Chou M-D, Lee K-T (1996) Parameterizations for the absorption of solar radiation by water vapor and ozone. J Atmos Sci 53:1203–1208

Chun H, Baik J (1998) Momentum flux by thermally induced interval gravity wave and its approximation for large-scale models. J Atmos Sci 55:3299–3310

Clark MP, Hay LE (2004) Use of medium-range numerical weather prediction model output to produce forecasts of stream-flow. J Hydrometeor 5:15–32

Clough SA, Shephard MW, Mlawer EJ, Delamere JS, Iacono MJ, Cady-Pereira K, Boaukabara S, Brown PD (2005) Atmospheric radiative transfter modeling: a summary of the AER codes. J Quant Spectrosc Radiat Transfer 91:233–244

Collins WD et al (2006a) The formulation and atmospheric simulation of the community atmosphere model version 3 (CAM3). J Climate 19:2144–2161

Collins WD et al (2006b) The community climate system model version 3 (CCSM3). J Climate 19:2122–2143

DelSole T, Shukla J (2012) Climate models produce skillful predictions of Indian summer monsoon rainfall. Geophys Res Lett 39. doi:10.1029/2012GL051279

DeWitt DG (2005) Retrospective forecasts of interannual sea surface temperature anomalies from 1982 to present using a directly coupled atmosphere–ocean general circulation model. Mon Wea Rev 133:2972–2995

Drijfhout SS, Walsteijn FH (1998) Eddy-induced heat transport in a coupled ocean–atmospheric anomaly model. J Phys Oceanogr 28:250–265

Ek MB, Mitchell KE, Lin Y, Rogers E, Grunmann P, Koren V, Gayno G, Tarpley JD (2003) Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale 437 Eta model. J Geophys Res 108:8851. doi:10.1029/2002JD003296

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

Hack JJ (1994) Parameterization of moist convection in the National Center for Atmospheric Research Community Climate Model (CCM2). J Geophys Res 99:5551–5568

Hoerling MP, Hurrell JW, Xu T (2001) Tropical origins for recent north Atlantic climate change. Science 292:90–92

Holtsalag AAM, Boville BA (1993) Local versus nonlocal boundary-layer diffusion in a global climate model. J Climate 6:1825–1842

Hong S-Y, Pan H-L (1996) Nonlocal boundary layer vertical diffusion in a medium-range forecast model. Mon Wea Rev 122:3–26

Hong S-Y, Pan H-L (1998) Convective trigger function for a mass-flux cumulus parameterization scheme. Mon Wea Rev 126:2599–2620

Kain JS (2004) The Kain-Fritsch convective parameterization: an update. J Appl Meteor 43:170–181

Kain JS, Fritsch JM (1993) Convective parameterization for mesoscale models: the Kain-Fritsch scheme. The representation of cumulus convection in numerical models. Meteor. Monogr. No. 46. Am Meteor Soc 165–170

Kanamitsu M et al (2002a) NCEP dynamical seasonal forecast system 2000. Bull Am Meteor Soc 83:1019–1037

Kanamitsu M, Ebisuzaki W, Wollen J, Yang S-K, Hnilo JJ, Fiorino M, Potter GL (2002b) NCEP-DOE AMIP-II reanalysis. Bull Am Meteor Soc 83:1631–1643

Kirtman BP (2003) The COLA anomaly coupled model: ensemble ENSO prediction. Mon Wea Rev 131:2324–2341

Kirtman BP, Min D (2009) Multimodel ensemble ENSO prediction with CCSM and CFS. Mon Wea Rev 137:2908–2930

Kirtman BP, Fan Y, Schneider EK (2002) The COLA global coupled and anomaly coupled ocean-atmosphere GCM. J Climate 15:2301–2320

Koster RD, Suarez MJ, Heiser M (2000) Variance and predictability of precipitation at seasonal-to-interannual time scales. J Hydrometeorol 1:26–64

Kumar A, Hoerling MP (1995) Prospects and limitations of seasonal atmospheric GCM predictions. Bull Am Meteorol Soc 76:335–345

LaRow TE (2013) The impact of SST bias correction in north Atlantic hurricane retrospective forecasts. Mon Wea Rev 141:490–498

Mason SJ, Graham NE (1999) Conditional probabilities, relative operating characteristics, and relative operating levels. Weather Forecast 14:713–725

Mason SJ, Graham NE (2002) Areas beneath the relative operating characteristics (ROC) and levels (RROL) curves: statistical significance and interpretations. Quart J R Meteorol Soc 128:2145–2166

McFarlane NA (1987) The effects of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere. J Atmos Sci 44:1775–1800

Misra V, Li H, Wu Z, Dinapoli S (2013) Global seasonal climate predictability in a two tiered forecast system: part I : boreal summer and fall seasons. Clim Dyn. doi:10.1007/s00382-013-1812-y

Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. Int J Climatol 25:693–712

Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough SA (1997) Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the longwave. J Geophys Res 102(D14):16663–16682

Moorthi S, Suarez MJ (1992) Relaxed Arakawa-Schubert. A parameterization of moist convection for general circulation models. Mon Wea Rev 120:978–1002

Moura AD, Hastenrath S (2004) Climate prediction for Brazil's Nordeste: performance of empirical and numerical modeling methods. J Climate 17:2667–2672

Palmer TN, Brankovi CC, Richardson DS (2000) A probability and decision-model analysis of PROVOST seasonal multi-model ensemble integrations. Quart J R Meteorol Soc 126:2013–2034

Ramanathan V, Downey P (1986) A nonisothermal emissivity and absorptivity formulation for water vapor. J Geophys Res 91:8649–8666

Ropelewski CF, Halpert MS (1986) North American precipitation and temperature patterns associated with the El Nino/Southern Oscillation (ENSO). Mon Wea Rev 114:2352–2362

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

Saha S et al (2006) The NCEP climate forecast system. J Climate 19:3483–3517

Saha S et al (2010) The NCEP climate forecast system reanalysis. Bull Am Meteor Soc 91:1015–1057

Shimpo A, Kanamitsu M, Iacobellis SF, Hong S-Y (2008) Comparison of four cloud schemes in simulating the seasonal mean field forced by the observed sea surface temperature. Mon Wea Rev 136:2557–2575

Shukla J (1998) Predictability in the midst of chaos: a scientific basis for climate forecasting. Science 282:728–731

Shukla J, Anderson J, Baumhefner D, Brankovic C, Chang Y, Kalnay E, Marx L, Palmer T, Paolino DA, Ploshay J, Schubert S, Straus DM, Suarez M, Tribbia J (2000) Dynamical seasonal prediction. Bull Am Meteor Soc 81:2593–2606

Shukla J, Palmer TN, Hagedorn R, Hoskins B, Kinter J, Marotzke J, Miller M, Slingo J (2010) Towards a new generation of world climate research and computing facilities. Bull Am Meteor Soc 91:1407–1412

Smith TM, Reynolds RW, Peterson TC, Lawrimore J (2008) Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880–2006). J Climate 21:2283–2296

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

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

Taylor KE, William D, Zwiers F (2000) The SST and seaice boundary conditions for AMIPII simulation. PCMDI report 60. http://www-pcmdi.llnl.gov/publications/ab60.html

Tiedtke M (1983) The sensitivity of the time-mean large-scale flow o cumulus convection in the ECMWF model. In: Proceedings ECMWF Workshop on Convection in Large-Scale Models. European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, United Kingdom, pp 297–316

Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 1:1–41

Wu Z, Huang NE, Chen X (2009) The multi-dimensional ensemble empirical model decomposition method. Adv Adapt Data Anal 1:272–339

Xie and Arkin (1997) Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. Bull Am Meteorol Soc 78:2539–2558

Zhang GJ, McFarlane NA (1995) Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Centre general circulation model. Atmos Ocean 33:407–446

Zhang S, Harrison MJ. Rosati A, Wittenberg AT (2007) System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. Mon Wea Rev 135(10). doi: 10.1175/MWR3466.1

Zhu J, Huang B, Marx L, Kinter JL III, Balmaseda MA, Zhang R-H, Hu Z-Z (2012) Ensemble ENSO hindcasts initialized from multiple ocean analyses. Geophys Res Lett 39:L09602. doi: 10.1029/2012GL051503