Global Seasonal Climate Predictability in a Two Tiered Forecast System. Part I: Boreal Summer and Fall Seasons

Vasubandhu Misra[1, 2, 3, 4], H. Li[1, 2], Z. Wu[1, 2], and S. DiNapoli[1]

[1]Department of Earth, Ocean and Atmospheric Science, Florida State University

P.O. Box 3064520, Tallahassee, FL 32306-4520

[2]Center for Ocean-Atmospheric Prediction Studies, Florida State University

2035 E. Paul Dirac Dr., 200 RM Johnson Bldg., Tallahassee, FL 32306-2840

[3]Florida Climate Institute, Florida State University,

2035 E. Paul Dirac Dr., 200 RM Johnson Bldg., Tallahassee, FL 32306-2840

Submitted to Climate Dynamics

[4] Corresponding author; Email: vmisra@fsu.edu

Abstract

This paper shows demonstrable improvement in the global seasonal climate predictability of boreal summer (at zero lead) and fall (at one season lead) seasonal mean precipitation and surface temperature from a two-tiered seasonal hindcast forced with forecasted SST relative to two other contemporary operational coupled ocean-atmosphere climate models. The results from an extensive set of seasonal hindcasts are analyzed to come to this conclusion. This improvement is attributed to:

i) The global atmospheric model which is run at a relatively high resolution of 50km grid resolution compared to the two other coupled ocean-atmosphere models

ii) The multi-model bias corrected SST used to force the atmospheric model

The results of the seasonal hindcast are analyzed for both deterministic and probabilistic skill. The probabilistic skill analysis shows that significant forecast skill can be harvested from these seasonal hindcasts relative to the deterministic skill analysis. The paper concludes that the coupled ocean-atmosphere seasonal hindcasts have reached a reasonable fidelity to exploit their SST anomaly forecasts to force such relatively higher resolution two tier prediction experiments to glean further boreal summer and fall seasonal prediction skill.

1. Introduction

Seasonal climate prediction has been one of the primary drivers of climate science for several decades (Barnett et al. 1993; Bengtsson et al. 1993; Shukla 1998; Shukla et al. 2000; Palmer et al. 2004; Saha et al. 2006; Wang et al. 2009; Kirtman and Min 2009; Stockdale et al. 2011), although the emphasis is rapidly shifting to other time scales (e.g. intraseasonal, decadal, climate change). As a result of the rich heritage of seasonal predictability studies, there is a growing body of application studies in hydrology (Bohn et al. 2010; Clark and Hay 2004; Shukla et al. 2012), crop modeling (Cantelaube and Terres 2005; Challinor et al. 2005), human health (Morse et al. 2005; Doblas-Reyes et al. 2005) and in many other sectors which make use of these seasonal climate predictions and make them even more relevant to society.

There is a burgeoning realization in the last decade or less to move towards a "seamless" forecast system (Hurrell et al. 2007; Palmer et al. 2008; Shapiro et al. 2010; Shukla et al. 2010). A seamless system ideally entails using the same prediction tool across all time scales. This implies that under such a paradigm a fully coupled earth system model will be used for numerical weather prediction to climate change projections. Along the same lines of seamless systems there is also a debate on raising the resolution of the numerical weather and climate models significantly from their current resolution. However, with limited resources the debate is raging on the one side to move towards coupling many more climate components to develop an earth system framework (seamless in resolving interaction across climate components) and on the other side to raise the resolution to levels where parameterization of sub-grid scale processes can be totally avoided (seamless in resolution). Arguments for both strategies are persuasive and readers are referred to (Palmer et al. 2009; Shukla et al. 2009) for further discussion on this debate. A compromise solution that is now popularly used for seasonal prediction by many operational centers around the world is to use the coupled physical climate system framework that resolves the interactions of the land-atmosphere-ocean-seaice while leaving out other components (e.g., biogeophysical and biogeochemical interactions, atmospheric chemical interactions). This coupled physical climate model used for seasonal prediction is often referred as a single tiered seasonal forecast system.

In this paper we discuss the results of seasonal predictability from an Atmospheric General Circulation Model (AGCM) forced with forecasted SST from another single tiered forecast system. This set up common until recently was referred to as the two tiered seasonal forecast system. We are motivated to pursue this two-tiered approach as some recent studies have shown that with the improvement in the SST forecasts in the El Niño and Southern Oscillation (ENSO) over the equatorial Pacific region in the single tiered systems, the predictability of the Indian summer monsoon rainfall has demonstrably improved from its remote teleconnections with ENSO (DelSole and Shukla, 2012). In other words the externally (boundary) forced (or two tier paradigm) climate anomalies are still in play. So in this study we investigate the efficacy on the seasonal prediction skill from modest increase in resolution of the AGCM (with twice as high to four times as high as the single tiered systems from which the forecasted SST was borrowed) and with a new technique to bias-correct the forecasted SST.

2. Experiment Design

The two-tiered **F**lorida Climate **I**nstitute-Florida State University **S**easonal **H**indcasts at **50**km (FISH50) were conducted with the idea that no future information that corresponds to the forecast period will be used and furthermore to make a comparison with a couple of available single tiered seasonal hindcasts. At the time of conceiving this experiment two dynamical seasonal forecasts of SST that were readily available were those from the NCEP CFSv2 (Saha et al. 2010) and CCSM3.0 (Kirtman and Min 2009). Both these models are run at relatively coarser resolution of T126 spectral truncation (~100km grid spacing) for CFSv2 and at T85 spectral truncation (~150km grid spacing) for CCSM3.0 compared to T248 spectral truncation (~50km grid spacing) of FISH50. However as Kirtman and Min (2009) point out the skills of the models (CFSv2 and CCSM3.0) in forecasting the tropical SST's are somewhat orthogonal, suggesting that model average across the two models produce superior prediction skill for example over Niño3.4 region compared to either one of the models when examining the hindcasts over several years. Both these models however have systematic errors in SST (Fig. 1). When this multi-model averaged SST is used without any bias correction, it can detrimentally affect the prediction from the high resolution FISH50. For example the cold bias in the equatorial Pacific and the warm bias in the eastern subtropical oceans are persistent in both models, which when averaged across the two models will continue to persist. In order to ameliorate these errors we chose to do bias correction on the multi-model averaged SST (from the two models). Traditionally the bias correction would involve replacing the forecast climatology with the corresponding observed climatology (Drijfhout and Walsteijn 1998; Kirtman et al. 2002; Kirtman 2003). This procedure is however not operable in a true forecast environment where the observed climatology is computed over the forecast period. This issue of choosing the period of observed climatology becomes even more acute when you begin to see the significant impact of low frequency variations including secular change on the mean. For example, Fig. 2 shows the comparison of the observed climatology of SST computed over the period 1955 to 1981 (prior 27 years) and 1982 to 2008 (current 27 years that coincide with the period of the seasonal hindcasts of FISH50). Fig. 2 shows that the difference in the climatology between these two periods for the two seasons of June-July-August (JJA) and September-October-November (SON) is comparable to anomalies from El Niño and the Southern Oscillation (ENSO) in the equatorial Pacific Ocean. In order to circumvent this issue we use a time varying climatology that incorporates the linear trend along with the other low frequency variations in the following manner:

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

where SST_F is the forecast SST used to force FISH50 AGCM. SST_{OLF} is the observed low pass filtered SST and avoids any use of observations of SST during the seasonal forecast period of the year. It is however updated at the start of each seasonal forecast. SST_{OLF} is computed from the Extended Reynolds SST version 3 (ERSSTv3; Smith et al 2008) at its native 2° grid resolution. $SSTA_{CYCLE}$ is the stationary, monthly observed climatology of SST obtained from the ERSSTv3 computed between 1901-1981 at the native 2° grid resolution. $SSTA_{MME}$ is the monthly and multi-model ensemble mean SSTA made available at 1° grid resolution. The monthly mean SST_F which is then computed on the $SSTA_{MME}$ at 1° grid resolution is then linearly interpolated to the T248 gaussian grid of the FGSM. The monthly mean SST_F is then interpolated to daily time interval following Taylor et al. (2000).

SST_{OLF} is obtained using the spatial-temporal Multi-dimensional Ensemble Empirical Model Decomposition (MEEMD; Wu et al. 2009). MEEMD is based on Ensemble Empirical Mode Decomposition (EEMD; Wu and Huang 2009) which itself is based on Empirical Mode Decomposition (EMD; Huang et al. 1998). In EEMD, the SST time series at each spatial grid is adaptively decomposed into a small number of amplitude-frequency modulated components, with that number usually smaller than the base 2 logarithm of the length of the time series. An illustrative example of the EEMD decomposition of a climate time series can be found in Wu et al. (2011). After the ERSSTv3 time series at all grid points across the globe are decomposed, the multidecadal component and the secular trend are combined to generate SST_{OLF} at each grid. Although the decomposition does not use any information of spatial coherence, the obtained evolution of SST_{OLF} are both temporally and spatially coherent and exhibit large spatial scale features when the SST_{OLF} at all grids are pieced together using MEEMD.

Each season has six ensemble members in FISH50. The ensemble members in FISH50 differ only in their atmospheric initial conditions. The six initial conditions of the atmosphere are obtained from the six successive days starting from 0000UTC 28 May through 0000UTC 03 June, interpolated to the FISH50 grid from the NCEP-DOE Reanalysis (R2; Kanamitsu et al. 2002a). The land initial conditions are kept identical in all six-ensemble members and are interpolated linearly from the R2 reanalysis corresponding to 0000UTC 01 June. The $SSTA_F$ is also kept identical for all ensemble members.

3. Model Description

The AGCM used in this study has its origins from the Experimental Climate Prediction Center (ECPC) AGCM (Kanamitsu et al. 2002b; Shimpo et al. 2008). However, a few subtle but important changes were made to the AGCM to adapt to this study and henceforth referred as the **F**lorida Climate Institute-Florida State University **G**lobal **S**pectral **M**odel (FGSM). Typically the ECPC GSM was run at spectral truncation of T62 (~250km). We raised the resolution to spectral truncation T248 (~50km). This forced us to change the convection scheme, which was previously using Relaxed Arakawa Schubert (RAS; Moorthi and Suarez 1992) to Kain-Fritsch version 2 (KF2; Kain and Fritsch 1993) as the model climatology of rainfall was greatly improved with the latter scheme at T248 spectral truncation (Fig. 3). The rainfall climatology in Fig. 3 was generated from a preliminary single ensemble member 10 year seasonal hindcast for a 6-month integration period of FGSM forced with SST_F and initialized on 0000UTC 01 June conditions from R2. It is apparent from Figs. 3a, c, and e that the KF2 scheme greatly reduced the dry bias in the June-July-August (JJA) season over the equatorial oceans displayed by the RAS scheme (Fig. 3c). In the process the ITCZ and western Pacific warm pool rainfall climatology appears relatively far more realistic in the KF2 run (Fig. 3e). However the dry bias over the Indian monsoon region and the wet bias over equatorial Africa and South America and central America are accentuated further by KF2 (Fig. 3c). Similarly in the September-October-November (SON) season, KF2 continues to improve the rainfall climatology over the equatorial oceans (Fig. 3f) and greatly ameliorates the split ITCZ phenomenon displayed by the RAS scheme over the Indian Ocean (Fig. 3d). The ITCZ structure in the Pacific also appears far more realistic in the KF2 seasonal hindcasts (Fig. 3f) compared to the RAS (Fig. 3d). However the wet (dry) bias over equatorial Africa, Central America, and the Amazon region is perpetuated by the KF2 seasonal hindcasts from the earlier season (Fig. 3e). In summary, from examination of these preliminary seasonal hindcasts we were convinced that KF2 convection scheme was a better choice for this resolution of the model.

In FISH50, we also changed the concentration of CO_2 following the Mauna Loa observatory (http://www.esrl.noaa.gov/gmd/ccgg/trends/index.html#global) at the start of each seasonal hindcast. The radiation scheme for longwave follows from Chou and Suarez (1994) and that for shortwave from Chou and Lee (1996). The boundary layer parameterization is the non-local scheme (Hong and Pan 1996). The land surface scheme used is the 4-layer NOAH scheme (Chen and Dudhia 2001; Ek et al 2003).

4. Results

The results will largely focus on precipitation and surface temperature over land as they reasonably reflect on the overall fidelity of the model and are two of the most widely used variables for application studies. We will begin with examining the climatology and deterministic skill of the seasonal hindcasts followed by the probabilistic skill. It is now well recognized in the community that it is perilous and in fact inappropriate to examine only the deterministic skill of a seasonal forecast model given the imperfections of the initial conditions and the models (Mason and Graham 1998, 2002; Palmer et al. 2000). For precipitation we use the NCEP's Climate Prediction Center Merged Analysis of Precipitation (CMAP; Xie and Arkin 1997) available at monthly intervals and at 2.5° spatial resolution for validation. For validation of surface temperature we use the monthly mean Climate Research Unit version 3 (CRUv3;

Mitchell and Jones 2002) surface temperature available at 0.5° spatial resolution.

a) SST forcing

The SST_F (equation 1) biases at zero lead for JJA and at one season lead for SON are shown in Figs. 4a and b. The cold bias along the equatorial oceans and in the subtropical Pacific and Atlantic oceans displayed by both CFSv2 (Figs. 1c and d) and CCSM3.0 (Figs. 1e and f) are greatly reduced in SST_F. Similarly, the warm bias in the eastern subtropical Pacific and Atlantic Oceans displayed by both the models is also greatly reduced in SST_F. This improvement in SST_F is commendable given that the methodology to compute SST_F ensures that observations of SST during the forecast period are not used. In Fig. 5 we show the corresponding standard deviation of the seasonal mean SST from observations, SST_F, CFSv2, and CCSM3. At the outset the standard deviations seem comparable. There is however subtle but important differences in the various SST forecast products compared to the observations:

- The higher standard deviation of SST along the middle latitude storm track regions is rather strong in CCSM3 (Figs. 5d and h) while it is weak in CFSv2 (Figs. 5c and g) relative to the observations (Figs. 5a and e). SST_F in FISH50 (Figs. 5b and f) shows the SST variability in this region is somewhere in between that displayed by the two models, which is still underestimating the observed SST variability.
- Similarly in the southern hemisphere the SST variations in the polar ocean are much stronger (comparable) in CFSv2 (CCSM3) relative to observations. SST_F in FISH50 has SST variations in this region which are still higher than the observed

variations and CCSM3 but less than those in CFSv2.

- Over the equatorial Pacific, the SST_F in FISH50 displays higher SST variations off the coast of Peru-Ecuador than either of the two models in both seasons.
- SST_F in FISH50 also tends to reduce the SST variability between 150°W-120°W relative to either of the other two models along the equatorial Pacific and is more akin to observations.

b) FISH50 Climatology

The summer (at zero lead) and fall (one season lead) climatology of precipitation from observations and root mean square errors (RMSE) from FISH50, CFSv2 and CCSM3.0 are shown in Fig. 6. It is apparent from this figure that the RMSE in the tropical regions is much higher in FISH50 compared to either of the two models. FISH50 seems to ubiquitously rain more in the global tropics. This is also illustrated in the zonal mean climatology of rainfall shown in Fig. 7. FISH50 in both seasons rains the most in the deep tropics. In fact FISH50 rains more than the observations in nearly all latitude bands but is comparable to CFSv2 in the higher latitudes (Fig. 7). This excessive wet bias may be a result of the absence of the damping effect of an interactive SST in FISH50, which is otherwise present in the coupled ocean-atmosphere models.

Similarly the climatology of the surface temperature from observations and the corresponding RMSE from the three model seasonal hindcasts are shown in Fig. 8. The RMSE errors of FISH50 are comparable to the other two models in the deep tropics. But in the subtropical and middle latitudes FISH50 exhibits a higher RMSE error especially over Eurasia, the central United States and over North Africa. In fact in the two seasons

FISH50 displays a weak warm bias in the deep tropical land areas, a relatively colder bias in the subtropical land area and a comparatively large warm bias in the mid-latitude regions (not shown).

c) Deterministic Predictability

In this section we will examine the variability of the anomalies from the ensemble mean of the seasonal hindcasts, which is a deterministic approach to skill assessment of a forecast model. It is sometimes useful to examine this deterministic skill of the model as it can quickly point to the existence of any region of high seasonal predictability from external forcing like SST anomalies. Although it undermines the potential to harvest skill from the ensemble spread of the model, which will be demonstrated in the subsequent section.

Fig. 9 shows the correlation of the seasonal mean rainfall from the three models with the corresponding observations for both seasons. It is sobering to note that statistically significant skill (at 90% confidence interval according to t-test) is most prominent in the equatorial Atlantic and Pacific Oceans while they are nearly absent in the land areas. There is however some display of significant positive correlations over some of the islands in the maritime continent in the western Pacific Ocean. Furthermore by the subsequent SON season (at one season lead forecast) there is significant diminishment of the spatial extent of these correlations in all three models. In comparing the three models it is observed that they display comparable skill in this metric in the first season of hindcast (JJA). In SON as well it could be argued that the skills are comparable although FISH50 hindcasts seem to deteriorate slightly more than the other the models.

The signal to noise ratio (see appendix I) of the rainfall (Fig. 10) shows that all models exhibit significantly higher predictability over the tropical oceans compared to that over land with the exception of the Amazon region. This signal to noise ratio in all three models shows significant reduction in the second season (SON) of the forecast over these regions. There are, however, several interesting points that one can observe from comparing these figures:

- CCSM3 exhibits higher predictability in the tropical Indian and Atlantic Ocean than either of the two models in both seasons.
- CCSM3 persists with this higher signal/noise ratio in the SON season from the previous JJA season far more than either of the two models.
- Over Amazon, FISH50 exhibits the highest signal/noise ratio compared to the other two models in both seasons.

Similarly the correlations of the seasonal mean surface temperature anomalies over land with corresponding observations from the three models are shown in Fig. 11. We note here that FISH50 exhibits higher skill than the other two models over South America, Africa, Northern Australia and the southwestern United States in the first season of the hindcast (JJA). CCSM3 surprisingly displays barely any skill in either of the two seasons. Over Asia, CFSv2 shows a higher skill than FISH50 in JJA. However, similarly to rainfall (Fig. 9), we observe that there is significant deterioration of the positive correlations of the seasonal surface temperature anomalies in the subsequent season of SON. The signal to noise ratio of the surface temperature from the three models is shown in Fig. 12. As in the case of rainfall anomalies, we observe higher ratios for surface temperature anomalies over the tropical oceans, Amazon and over equatorial

Africa. But compared to rainfall anomalies (Fig. 10), the surface temperature anomalies exhibit higher ratios over land. In comparing the three models in Fig. 12 we observe:

- All three models exhibit a reduction in the ratio from the first season (JJA) to the second season (SON). This reduction is much higher over land than the oceanic regions.
- In both seasons the signal/noise ratio of the surface temperature anomalies in the tropics is much higher in FISH50 than in either of the two models.
- In the higher latitudes CCSM3 exhibits higher ratio than either of the two models. Although in SON the ratio in FISH50 in the southern hemisphere mid-latitudes is comparable to that in CCSM3.

d) Probabilistic Prediction Skill

We compute the Area under the Relative Operating characteristic Curve (AROC; see appendix II) to examine the probabilistic skill of these seasonal hindcasts[&]. We examined the AROC for the predictability of the lower, middle and upper terciles of the summer and fall seasonal anomalies of precipitation (Figs.13-16) and surface temperature (Figs. 17-20). The tercile thresholds for the models and observations were based on the respective model hindcast and observed seasonal anomalies for the period 1982-2008. The benefit of examining probabilistic skill is immediately apparent from Figs. 13-20 by noticing that useful skill could be harvested from these models over the land areas, which was otherwise noted to be bereft of any skill in terms of the ensemble mean seasonal anomalies. In both the seasonal precipitation and surface temperature anomalies, FISH50

[&] We also compare the FISH50 AROC with the rest of the National Multi-Model Ensemble (NMME) models in Appendix III displays generally a higher skill than either of the two coupled models. More specifically we see from these figures that

- The AROC for the seasonal precipitation and temperature anomalies reduces in most parts from JJA to SON for all three models as was observed earlier with the deterministic skill analysis.
- All three models show superior skill to predict the lower and upper tercile events than the middle tercile for both the variables.
- In all three models, the AROC for precipitation and surface temperature anomalies are largest over the tropical oceans and tropical land respectively.
- The improvement of FISH50 over the other two models for both variables is most pronounced in the first season of the hindcast (JJA) and most apparent for the extreme terciles than the middle terciles.
- The improvement of FISH50 over the other two models is more apparent in precipitation than in surface temperature anomalies.
- The improvement of FISH50 over the other two models for both variables is best seen over the land areas especially in the JJA season.

5. Summary and conclusions

In this paper we introduce a relatively new AGCM, the FGSM, which was used to conduct an extensive set of seasonal hindcasts for boreal summer and fall season. The FGSM follows from the extensive development of the model conducted at previously Experimental Climate Prediction Center at Scripps Institute of Oceanography (Kanamitsu et al. 2002; Shimpo et al. 2008; Kanamitsu et al. 2010). The FGSM was integrated at a spectral truncation of T248 (~50km grid resolution). Another important modification made to FGSM was that we used the Kain-Fritsch version 2 (KF2; Kain and Fritsch 1993) convection scheme. The previous versions of the AGCM typically used the Relaxed Arakawa Schubert (RAS; Moorthi and Suarez 1992) convection scheme. In addition we introduced time varying CO_2 concentration in the FGSM forecasts. In this study we showed that for the same SST forcing, we get dramatic improvement in the seasonal rainfall climatology of the FGSM using the KF2 scheme.

Unlike previous two-tiered seasonal hindcasts or flux corrected coupled integrations we took extensive care to ensure that no future information (that is coincident with the forecast period) is used in making the bias correction for SST. We showed that low frequency variations and climate change has a significant influence on the period chosen for computing the observed climatology. We adopted the technique of MEEMD (Wu et al. 2009) to develop a low pass filtered SST from the period prior to the hindcast period (SST_{OLF}) to which the seasonally hindcasted coupled ocean-atmosphere multi-model (CFSv2 and CCSM3.0) averaged monthly SST anomalies (http://www.cpc.ncep.noaa.gov/products/NMME/) and the observed stationary seasonal cycle were added to obtain a true two-tiered SST forecast field (SST_F). The SST_F was used to force the **F**lorida Climate **I**nstitute-Florida State University **S**easonal **H**indcasts at 50km resolution (FISH50). The SST_F bias was found to be systematically lower than either of the two coupled model seasonal hindcast SST climatologies, with significant improvements in the cold bias along the equatorial oceans and the warm bias over the subtropical eastern Pacific and Atlantic Oceans.

We made a systematic comparison of the global and tropical seasonal predictability of surface temperature and precipitation between the three models and made the following salient observations:

1. The RMSE of the seasonal mean rainfall (and surface temperature) in both seasons is higher over the tropics (global land) in FISH50 compared to either CFSv2 or CCSM3.
2. The positive correlations of the ensemble mean seasonal precipitation anomalies with corresponding observations are significantly large over the equatorial Pacific and Atlantic Oceans in JJA, which get significantly diminished in SON in all three models. These correlations are comparable in magnitude and spatial extent in all three models.
3. FISH50 exhibits higher positive correlations of the ensemble mean seasonal surface temperature anomalies with corresponding observations compared to either of the two models over South America, Africa, northern Australia and western United states in the first season of the hindcast (JJA).
4. All three models exhibit higher probabilistic skill for the lower and upper terciles for both seasons than the middle tercile.
5. The probabilistic skill analysis of seasonal rainfall anomalies shows that the benefit of FISH50 over the other two models is best realized in the first season (JJA) and is realized the most compared to the other two models over the land regions. FISH50 displays relatively improved seasonal prediction skill of precipitation in FISH50 over certain land areas (e.g. United States, tropical South America, maritime continent).

6. The probabilistic skill in seasonal surface temperature anomalies reveal that FISH50 improves significantly over the seasonal hindcasts from the other two models over the tropical land areas in both seasons.

In conclusion we have shown in this study, a two-tiered forecast with the proposed bias correction for SST and using an AGCM with a modest increase in resolution and with reasonable climatology can yield useful seasonal forecast skill that exceeds in some respects to that attained from dynamical coupled ocean-atmosphere models from which the forecasted SST was borrowed. In other words we suggest that we may be at a stage where single tiered dynamical SST forecasts are of reasonable fidelity that they can be used with appropriate bias corrections to pursue with higher resolution two tiered AGCMs to yield superior seasonal forecasts. We contend that efforts like FISH50 could prove to harvest more seasonal predictability. This study also clearly reveals that there is useful model predictability to harvest when we adopt a probabilistic skill analysis as an alternative to deterministic skill analysis. There is however room for further development of FISH50 in areas of land surface initialization, higher spatial resolution (both in the vertical especially in the stratosphere and in the horizontal), and incorporation of direct/indirect effects of aerosols that could benefit in realizing more useful seasonal predictability.

Acknowledgements

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. This work was supported by grants from NOAA (NA12OAR4310078,

NA10OAR4310215, 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 I: Signal to noise ratio

The signal to noise ratio is an objective measure of an AGCM's predictability (Straus and Shukla 2000). It is basically a measure of the variance displayed by the ensemble mean relative to the ensemble spread of the seasonal hindcast. So higher values of this ratio correspond to higher predictability of the phenomenon by the AGCM. This measure of predictability however does not reflect on the verification of the hindcast or forecast. The ensemble mean for a given climate variable (say Y), for a given climate model, and for a given year j is:

$$\overline{Y}_j = \frac{1}{K}\sum_{i=1}^{K} Y_{ji}$$

where i is the index for number of ensemble members and K is the total number of ensemble members (in our study it is 6).

The variance of the ensemble spread for a given year j is given by:

$$\sigma_j^2 = \frac{1}{K}\sum_{i=1}^{K} (Y_{ji} - \overline{Y}_j)^2$$

The variance of the ensemble spread is a measure of the noise in the forecast system, which is averaged over all years of the hindcast to obtain:

$$\sigma_{noise}^2 = \frac{1}{L}\sum_{j=1}^{L}\sigma_j^2$$

The variance of the signal component is given by:

$$\sigma_{signal}^2 = \frac{1}{L}\sum_{j=1}^{L}(\overline{Y}_j - \overline{Y})^2$$

where

$$\overline{Y} = \frac{1}{LK}\sum_{j=1}^{L}\sum_{i=1}^{K}Y_{ji}$$

Then, predictability (Π) or signal to noise ratio is defined as:

$$\Pi = \frac{\sigma_{signal}^2}{\sigma_{signal}^2 + \sigma_{noise}^2}$$

Appendix II: Area under the relative operating characteristic curve

The area under the relative operating characteristic curve (AROC) is obtained by plotting the relative operating characteristic curve (ROC) for every grid point of the three models. A contingency table along the lines shown in Table A below is prepared first:

Table A: Contingency Table

Is the event observed?	Does the ensemble probability for the event exceed the threshold?	
	Yes	No
Yes	Hit (H)	Miss (M)

No	False Alarm (FA)	Correct Rejection (CR)

The ROC is then plotted as the Hit Rate (HR) against the False Alarm Rate (FAR) for every grid point. HR and FAR are defined as follows:

$$HR = \frac{H}{(H + M)}$$

and

$$FAR = \frac{FA}{(FA + CR)}$$

To define an event in Table A we first rank the (seasonal mean surface temperature and precipitation) observations and hindcasts independent of each other to develop separate thresholds to define the lower, middle, and upper terciles for each of the variables. Then for each tercile we construct the contingency table with several points plotted on the ROC by defining the event for discrete number of ensemble members. For example, to develop the ROC for lower tercile of JJA seasonal precipitation anomaly, we fill the contingency table by seeking an answer for a given grid point to the question: Do 20% of the ensemble members of the seasonal hindcast show a lower tercile event? Likewise we will fill the contingency table for the same grid point by seeking the answer to the question: Do 40% of the ensemble members of the seasonal hindcast show a lower tercile event? This is done similarly for 60% of ensemble members and so on. By following this series of questions, we are able to construct ROC for each grid point. Then using the trapezoidal rule we compute AROC. By definition, any value lower than or equal to 0.5 for AROC would suggest that the seasonal hindcast is no better than climatology.

Therefore for skillful probabilistic forecast, the seasonal hindcasts should have a value greater than 0.5 for AROC.

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

A multi-institutional NMME (http://www.cpc.ncep.noaa.gov/products/ctb/nmme/) project to conduct seasonal retrospective forecasts and also maintained in real time is hosted on the International Research Institute for Climate and Society, Columbia University. Here we compare the AROC across these models for the two seasons JJA and SON at zero and one season lead for precipitation and surface temperature with FISH50.

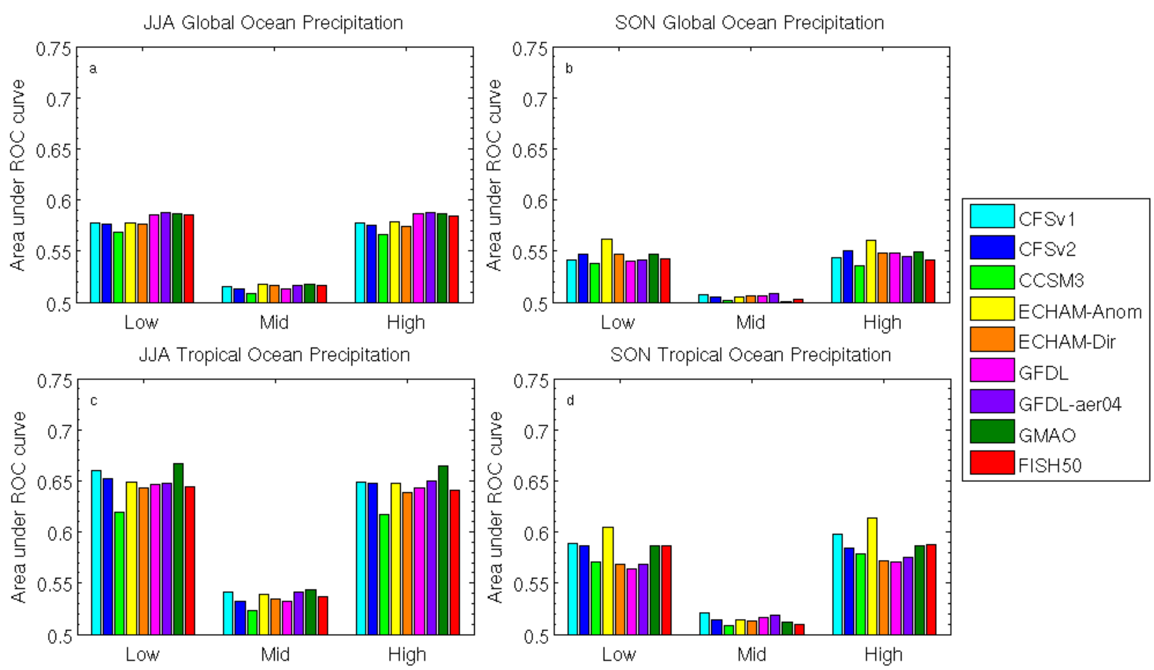

Figure AIII.1: Area under the ROC averaged over global oceans for (a) JJA, (b) SON, over tropical oceans for (c) JJA, and (d) SON for low, middle, and upper terciles of NMME and FISH50 precipitation.

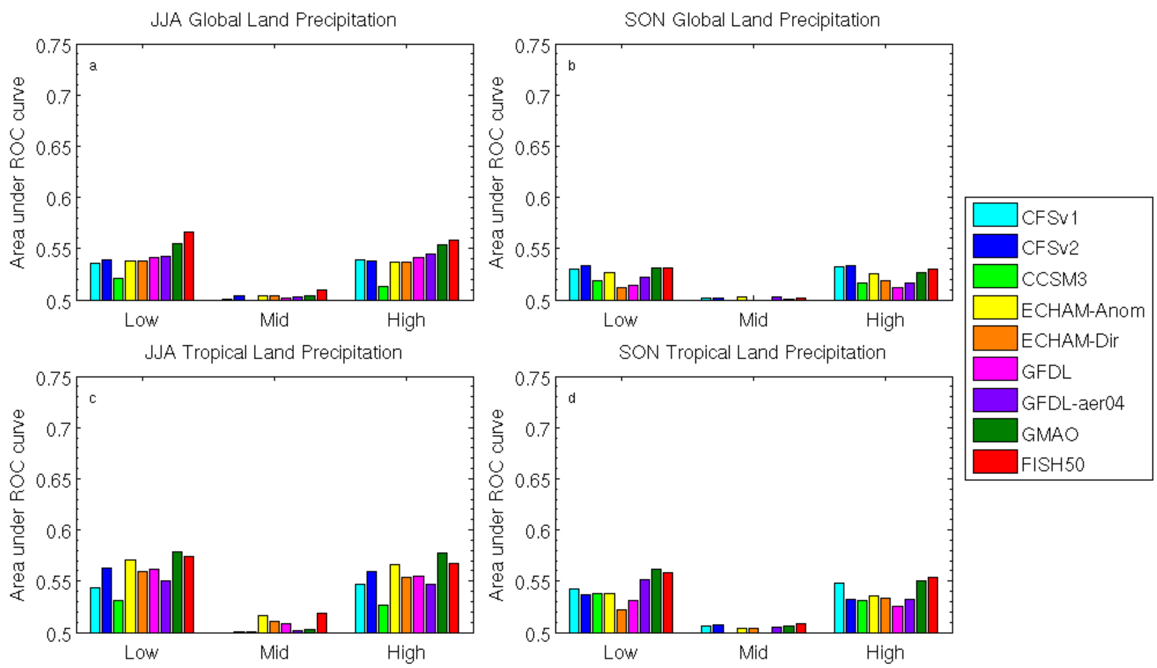

Figure AIII.2: Area under the ROC averaged over global land for (a) JJA, (b) SON, over tropical land for (c) JJA, and (d) SON for low, middle, and upper terciles of NMME and FISH50 precipitation.

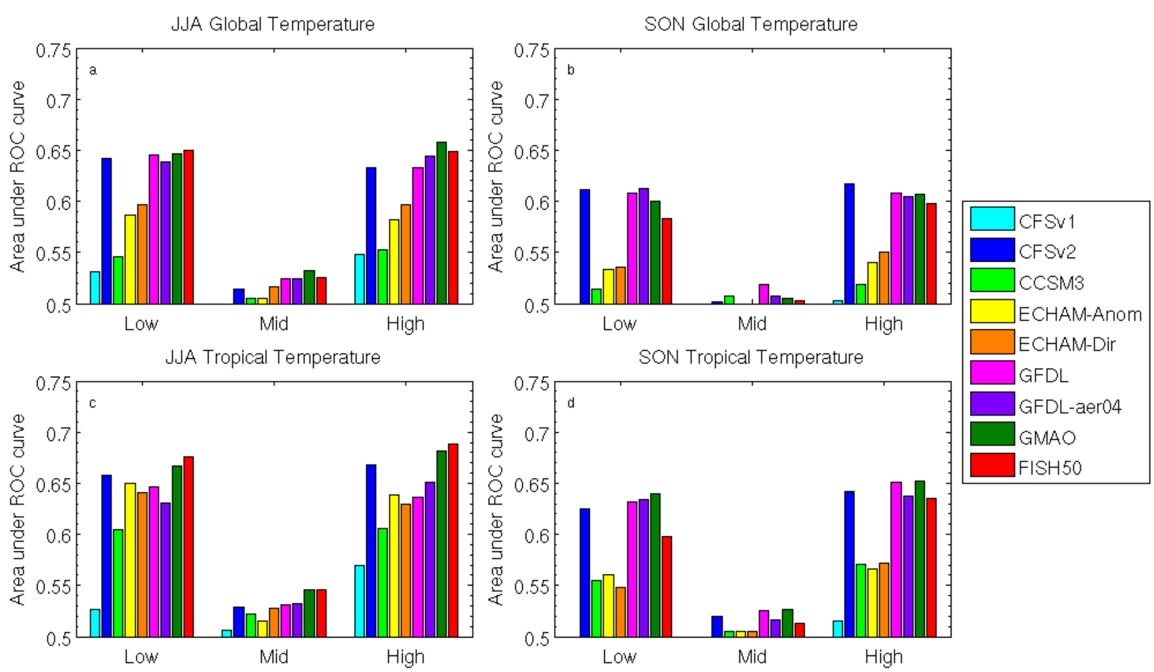

Figure AIII.3: Same as Fig. AIII.3 but for surface temperature.

References:

Bengtsson L, Schlese U, Roeckner E, Latif M, Barnett T, Graham N (1993) A two-tiered approach to long-range climate forecasting. Science 261: 1026-1029.

Bohn TJ, Sonessa MY, Lettenmaier DP (2010) Do Multi-model ensemble average always yield improvements in forecast skill? J. Hydromet., doi:10.1175/2010JHM1267.1

Barnett, TP, Bengtsson L, Arpe K, Flugel M, Graham N, Latif M, Ritchie J, Roeckner E, Schlese U, Schulzweida U, Tyree M (1993) Forecasting global ENSO-related climate anomalies. Tellus 46A: 381-397.

Cantelaube P and Terres J-M (2005) Seasonal weather forecast for crop yield modeling in Europe. Tellus A, 57, 476-487.

Challinor AJ, Slingo J M. Wheeler TJ, and Doblas-Reyes FJ (2005) Probabilistic simulations of crop yield over western India using the DEMETER seasonal hindcast ensembles. Tellus A, 57, 498-512.

Chen F, Dudhia J (2001) Coupling and advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. Mon Wea Rev 129: 569-585

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

Chou M-D, Suarez MJ (1994) An efficient thermal infrared radiation parameterization for use in general circulation models. Technical report series on global modeling and data assimilation, NASA/TM-1994-104606, 3, 85pp.

Clark, MP and Hay LE (2004) Use of medium-range numerical weather prediction model output to produce forecasts of streamflow. J. Hydrometeorol., 5, 15-32.

DelSole T, Shukla J (2012) Climate models produce skillful predictions of Indian summer monsoon rainfall. Geophys Res Letts 39: L09703, doi: 10.1029/2012GL051279.

Doblas-Reyes, J. and co-authors, 2005: A forecast quality assessment and end-to-end probabilistic multi-model seasonal forecast system using a malaria model. Tellus A, 57, 464-475.

Drijfhout SS, Walsteijn FH (1998) Eddy induced heat transport in a coupled ocean-atmosphere anomaly model. J. Phys. Oceanography, 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

Hong S-Y, Pan H-L (1996) Nonlinear boundary layer vertical diffusion in a medium-range forecast model. Mon Weather Rev 124: 2322-2339

Hurrell JW, Bader D, Delworth T, Kirtman B, Meehl J, Pan HL, Wielicki B (2007) White Paper on Seamless Prediction. Available at http://www.cgd.ucar.edu/cas/jhurrell/Docs/SeamlessModellingDraft03302007.pdf

Kain JS, Fritsch JM (1993) Convective parameterization for mesoscale models: The Kain-Fritsch scheme. The Representation of Cumulus convection in Numerical Models, Meteorol. Monogr., No. 46, Amer Meteorol Soc, 165-170.

Kanamitsu M, Ebusuzaki W, Woollen J, Yang S-K, Hnilo J, Fiorino M, Potter GL (2002a) NCEP-DOE AMIP-II reanalysis (R-2). Bull Amer Meteorol Soc 83: 1631-1643

Kanamitsu M, and Coauthors (2002b) NCEP Dynamical Seasonal Forecast System 2000. Bull Amer Meteorol Soc 83: 1019-1037

Kanamitsu M, Yoshimura K, Yhang Y-B, Hong S-Y (2010) Errors of interannual variability and multi-decadal trend in dynamical regional climate downscaling and its corrections. J Geophys Res 115: D17115, doi: 10.1029/2009JD013511.

Kirtman BP, (2003) The COLA anomaly coupled model: Ensemble ENSO prediction. Mon. Wea. Rev., 10, 2324-2341.

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

Kirtman, B. P., and D. Min, (2009) Multimodel Ensemble ENSO Prediction with CCSM and CFS. *Monthly Weather Review*, **137,** 2908-2930.

Mason SJ, Graham NE (1999) Conditional probabilities, relative operating characteristics, and relative operating levels. Wea Forecasting 14: 713-725

Mason SJ, Graham NE (2002) Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: significance and interpretation. Quart J Roy Meteorol Soc 128: 2145-2166

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

Moorthi S, Suarez MJ (1992) Relaxed Arakawa-Schubert. A Parameterization of Moist Convection for General Circulation Models. Mon Wea Rev 120: 978-1002

Morse AP, Doblas-Reyes FJ, Hoshen MB, Hagedorn R, Palmer TN (2005) A forecast quality assessment of an end-to-end probabilistic multi-model seasonal forecast system using a malaria model. Tellus A, 57, 464-475.

Palmer TN, Brankovic C, and Richardson DS, (2000) A probability and decision model analysis of PROVOST seasonal multimodel ensemble integrations. Quart. J. Roy. Meteor. Soc., 126, 2013–2034

Palmer TN, and Coauthors (2004) Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER). Bull Amer Meteorol Soc 85: 853-872.

Palmer TN, Doblas-Reyes FJ , Weisheimer A, Rodwell MJ (2009) *Bulletin of the American Meteorological Society*, 89, 459-470.

Palmer TN, Doblas-Reyes FJ , Weisheimer A, Rodwell MJ (2009) Reply. *Bulletin of the American Meteorological Society* Volume 90, 1551-1554.

Saha, S., et al. (2006) The climate forecast system at NCEP, J. Clim., 19, 3483 – 3517, doi:10.1175/JCLI3812.1.

Saha S et al (2010) The NCEP climate forecast system reanalysis. Bull Amer Meteorol Soc 91: 1015-1057

Shapiro M, et al. (2010) An Earth-System Prediction Initiative for the Twenty-First Century. Bull. Amer. Soc., 91, 1377-1388.

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., J. Anderson, D. Baumhefner, C. Brankovic, Y. Chang, E. Kalnay, L. Marx, T. Palmer, D. A. Paolino, J. Ploshay, S. Schubert, D. M. Straus, M. Suarez, J. Tribbia, (2000) Dynamical Seasonal Prediction, *Bulletin of the American Meteorological Society,* 81, 2593-2606.

Shukla, J. and co-authors, 2009: Revolution in climate prediction is both necessary and possible. A declaration at the world modeling summit for climate prediction. Bull. Amer. Soc., 2, 175-178.

Shukla J, Palmer TN, Hagedorn R, Hoskins B, Kinter J, Marotzke J, Miller M, Slingo J (2010) Toward a new generation of world climate research and computing facilities. *Bulletin of the American Meteorological Society*, 91, 1407-1412.

Shukla S, Voisin N and Lettenmaier DP (2012) Value of medium range weather forecasts in the improvement of seasonal hydrologic prediction skill. Hydrol. Earth Sys. Sci. Discuss., 9, 1827-1857.

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

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. Climate Dyn 37: 455-471.

Taylor, K. E., William, D., and F. Zwiers, 2000: The SST and seaice boundary conditions for AMIPII simulation. PCMDI report 60. 24 pp. Available from http://www-pcmdi.llnl.gov/projects/amip/AMIP2EXPDSN/BCS/amip2bcs.php

Wang B, and Coauthors (2009) Advance and prospectus of seasonal prediction: Assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980–2004). Climate Dyn 33: 93-117.

Wu, Z., and N. E Huang, 2009: Ensemble Empirical Mode Decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis*. **1**, 1-41.

Wu, Z., N. E. Huang, and X. Chen, 2009: The multi-dimensional Ensemble Empirical Mode Decomposition method. *Advances in Adaptive Data Analysis*, **1**, 339-372.

Wu Z., N. E. Huang, J. M. Wallace, B. Smoliak, X. Chen, 2011: On the time-varying trend in global-mean surface temperature. *Clim. Dyn.* **37**, 759-773, DOI: 10.1007/s00382-011-1128-8.

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

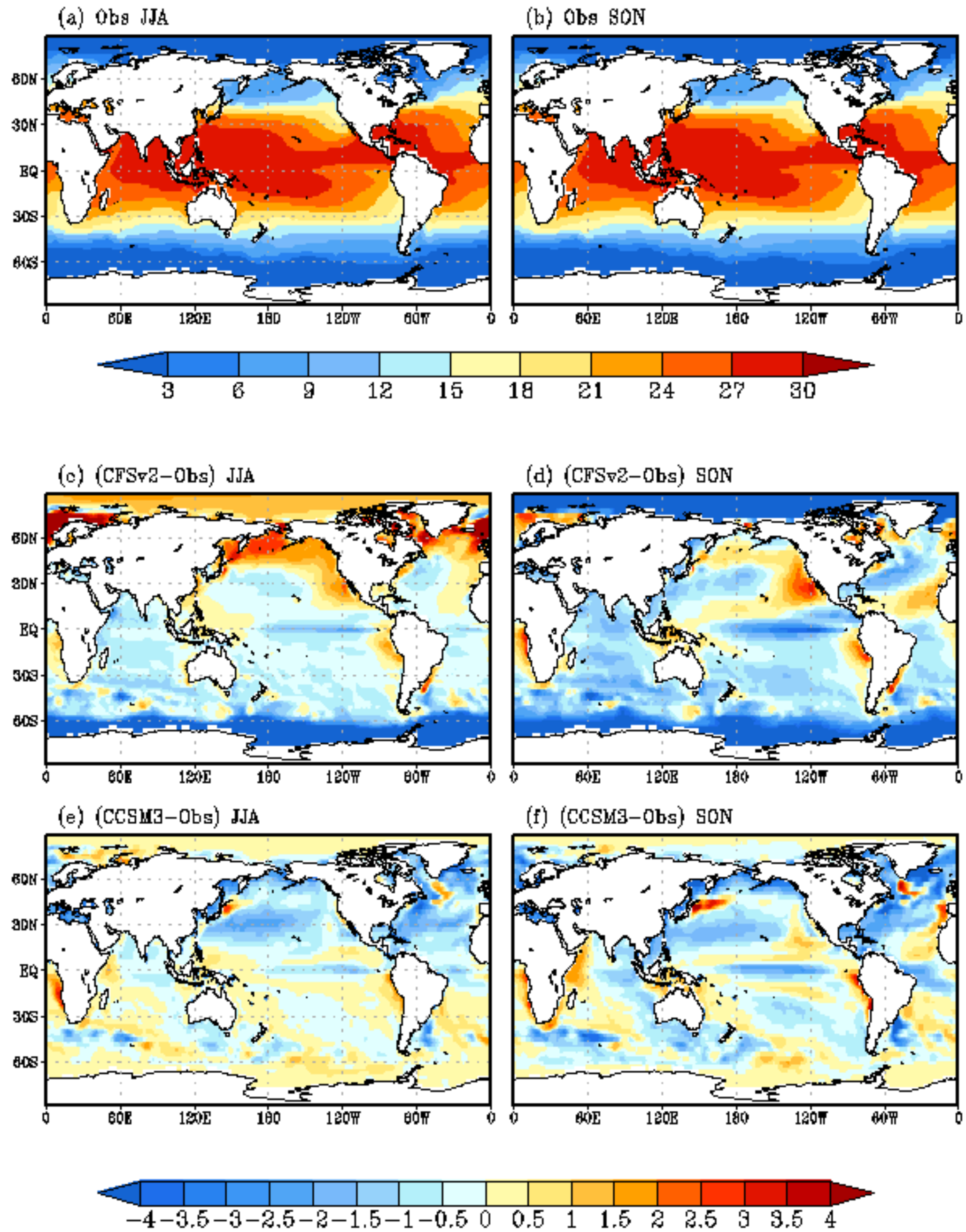

Figure 1. The observed climatological SST for boreal (a) summer (JJA) season and (b) fall (SON) season. The bias of hindcasted SST at zero lead for boreal summer season from (c) CFSv2, (e) CCSM3. Similarly, the bias of hindcasted SST at one season lead for boreal fall season from (d) CFSv2 and (f) CCSM3.0. The units are in °C.

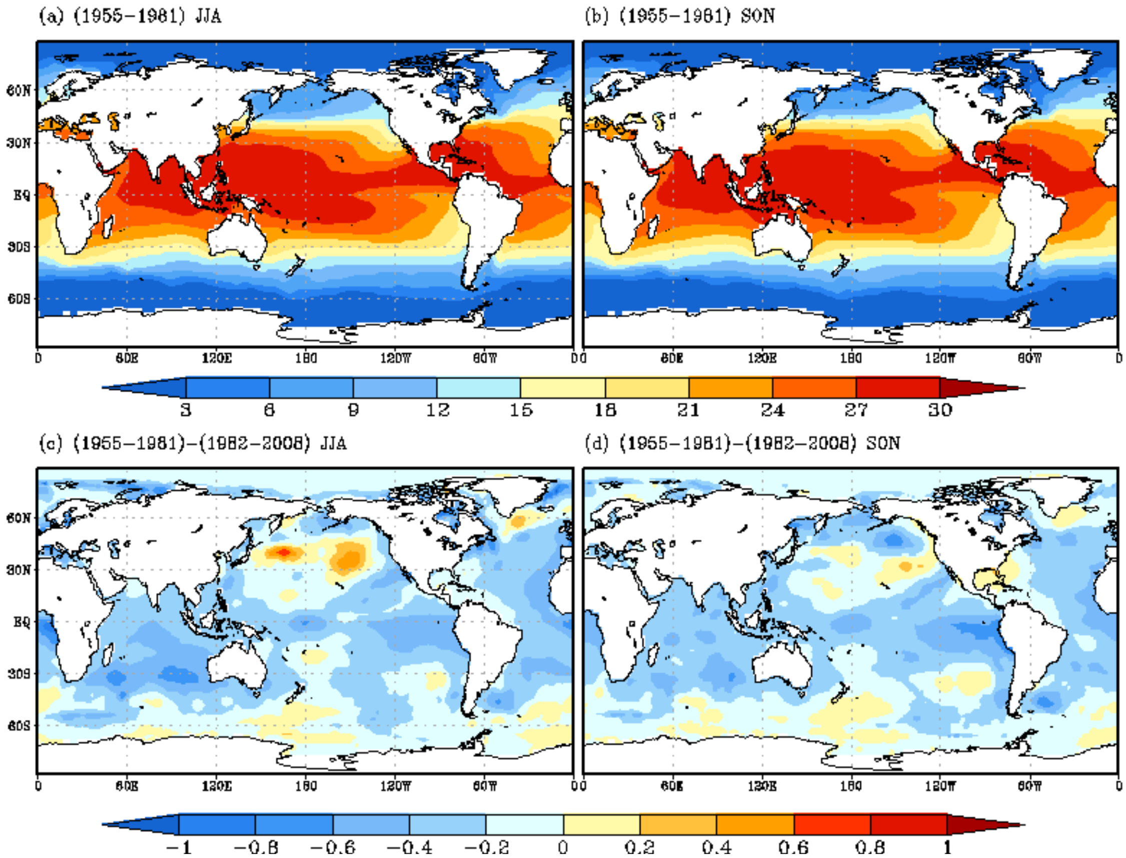

Figure 2. The observed climatology SST computed over a period of 1955-1981 for (a) JJA season and (b) SON season, and their corresponding differences with climatology computed over the period 1982-2008 for (c) JJA and (d) SON. The units are in °C.

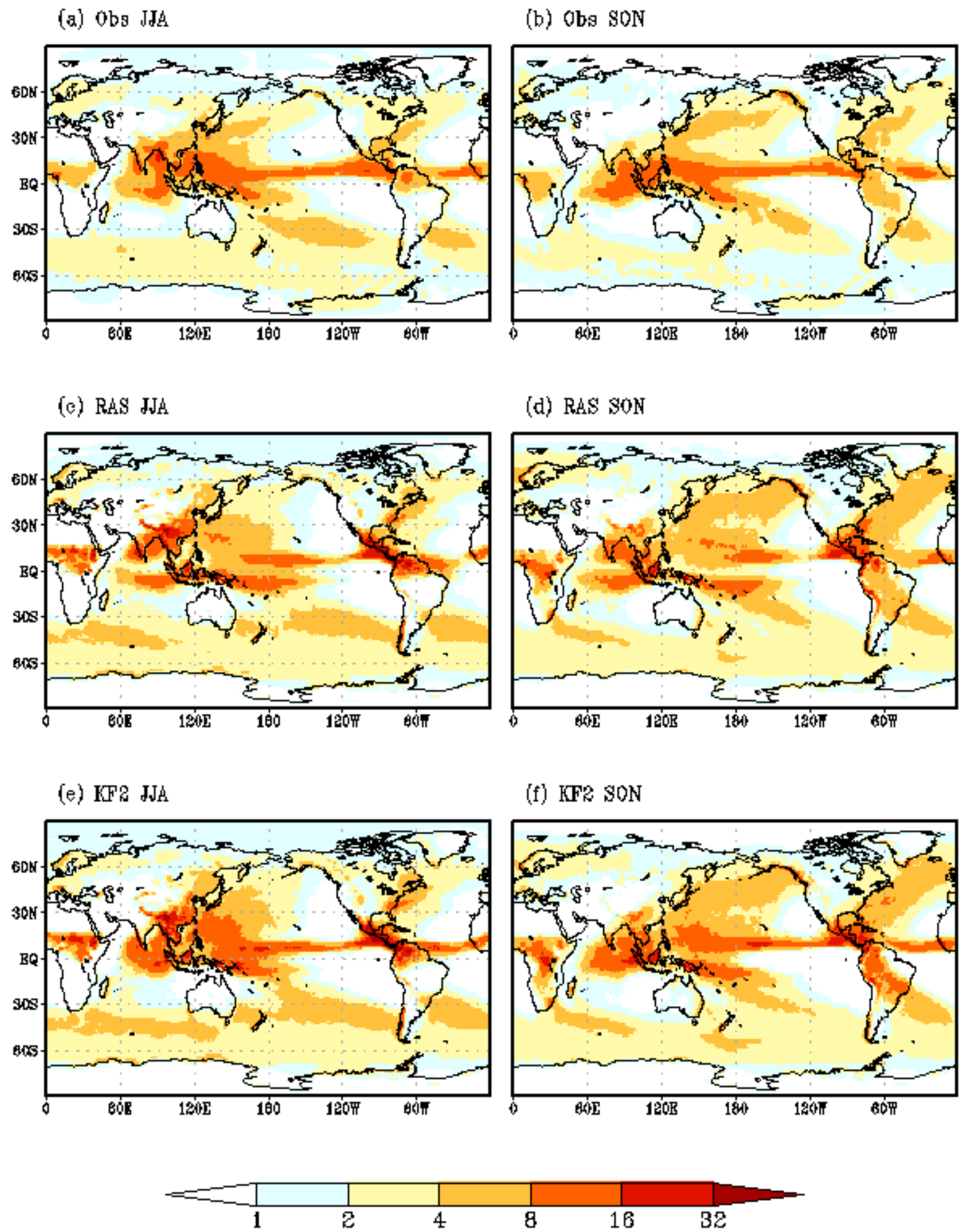

Figure 3. The observed climatology of precipitation computed over a period of 1982-1993 for (a) JJA and (b) SON seasons. The corresponding climatology of precipitation from a single member seasonal hindcast for the period of 1982-1993 using the RAS convection scheme for (c) JJA (at zero lead) and (d) SON (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) JJA (at zero lead) and (f) SON (one season lead). The units are in mm/day.

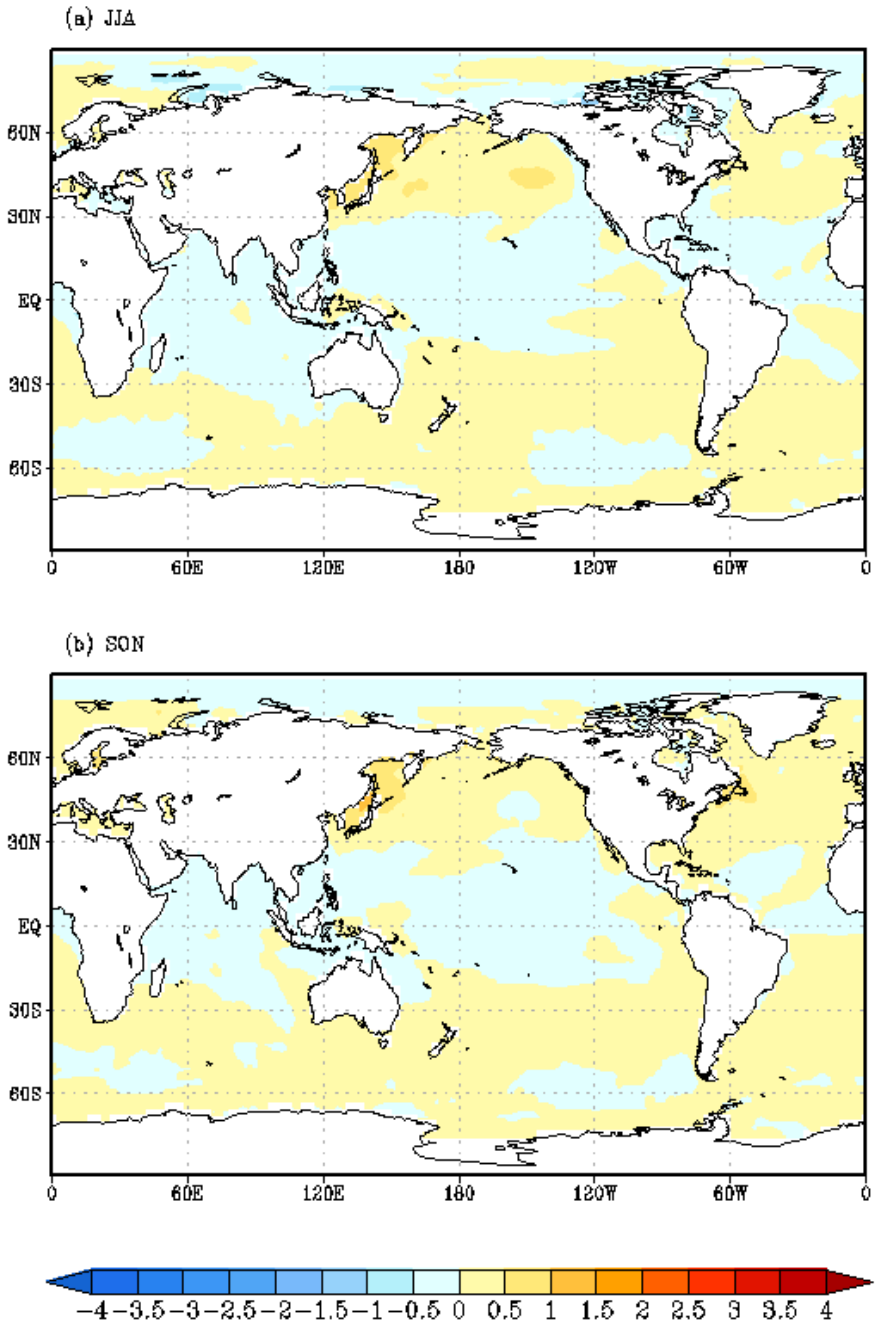

Figure 4. The climatological SST bias computed for (a) JJA season (at zero lead) and (b) SON season (at one season lead) from FISH50. The observed SST climatology was computed over the period 1982-2008 as shown in Figs. 1a and 1b. The units are in °C.

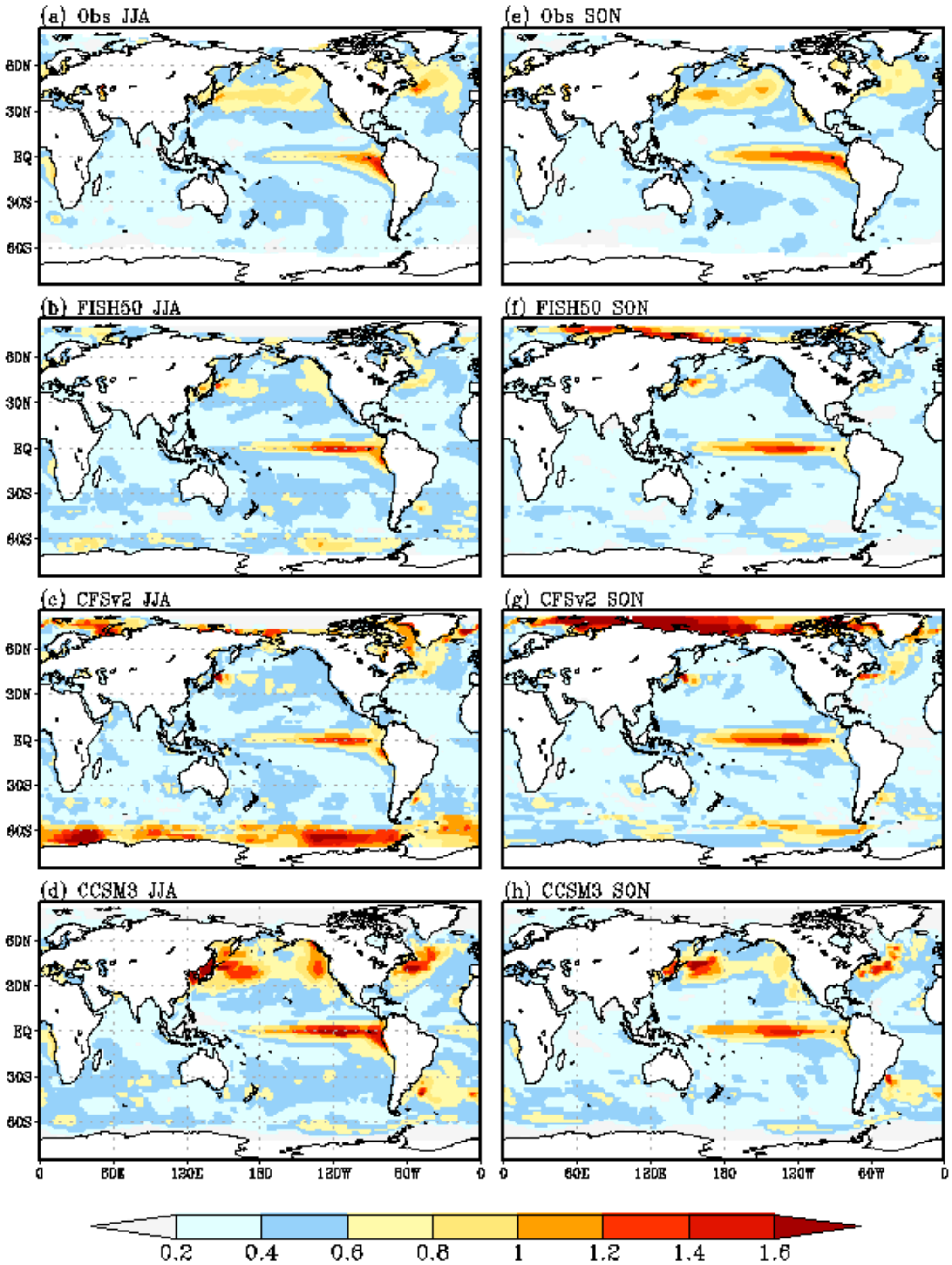

Figure 5. The standard deviation of JJA seasonal mean SST from (a) observations, and seasonal hindcasts at zero lead from (b) FISH50, (c) CFSv2, and (d) CCSM3. Similarly the standard deviation of SON seasonal mean SST from (e) observations, and seasonal hindcasts at one season lead from (f) FISH50, (g)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

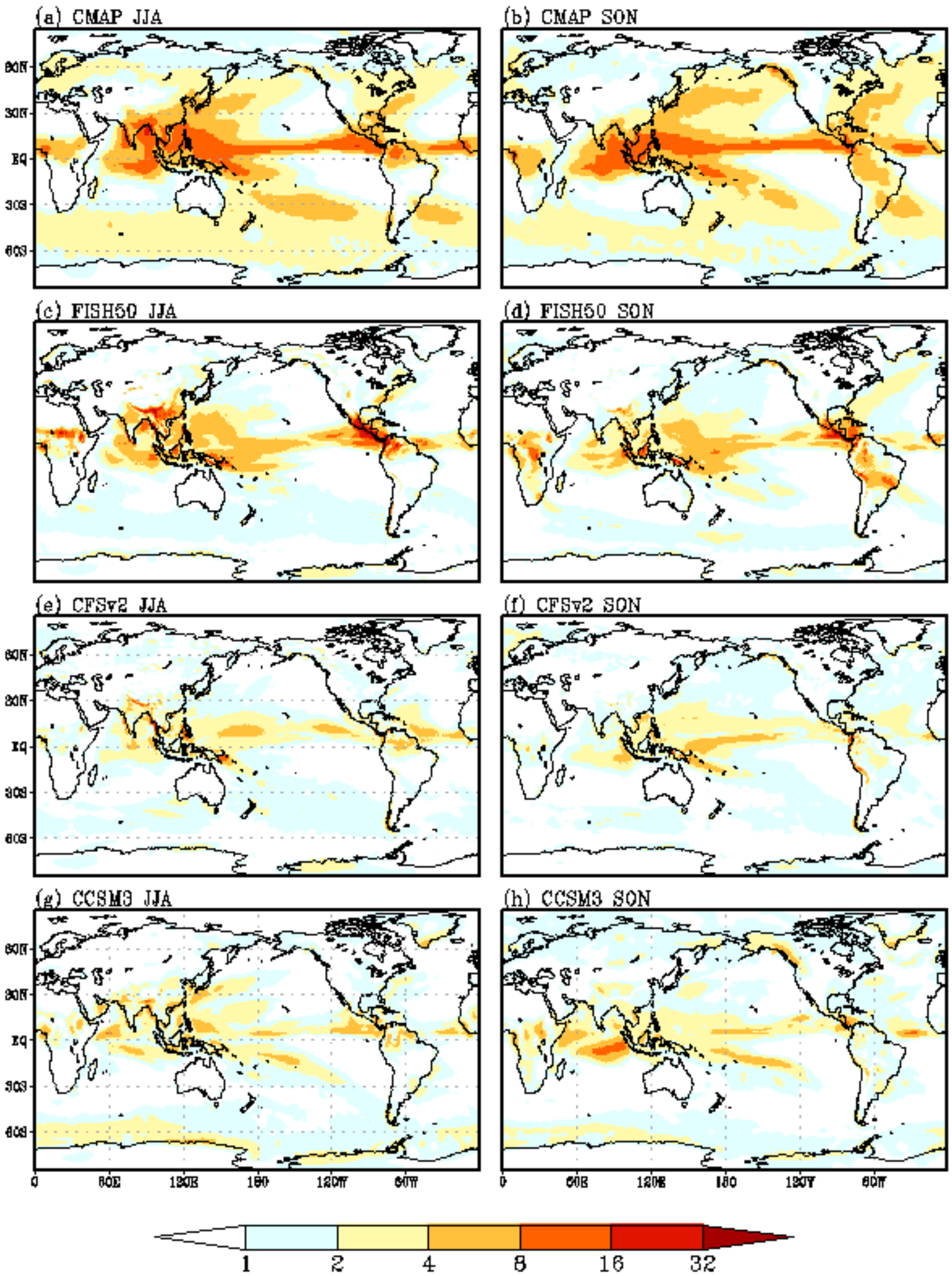

Figure 6. The observed climatology of precipitation (1982-2008) in (a) JJA, and (b) SON. The root mean square error of the ensemble mean precipitation for JJA (zero lead) for seasonal hindcasts from (c) FISH50, (e) CFSv2, and (g) CCSM3. Likewise, the root mean square error of the ensemble mean precipitation for SON (one season lead) for seasonal hindcasts from (d) FISH50, (f) CFSv2, and (h) CCSM3. The units are in mm day^{-1}.

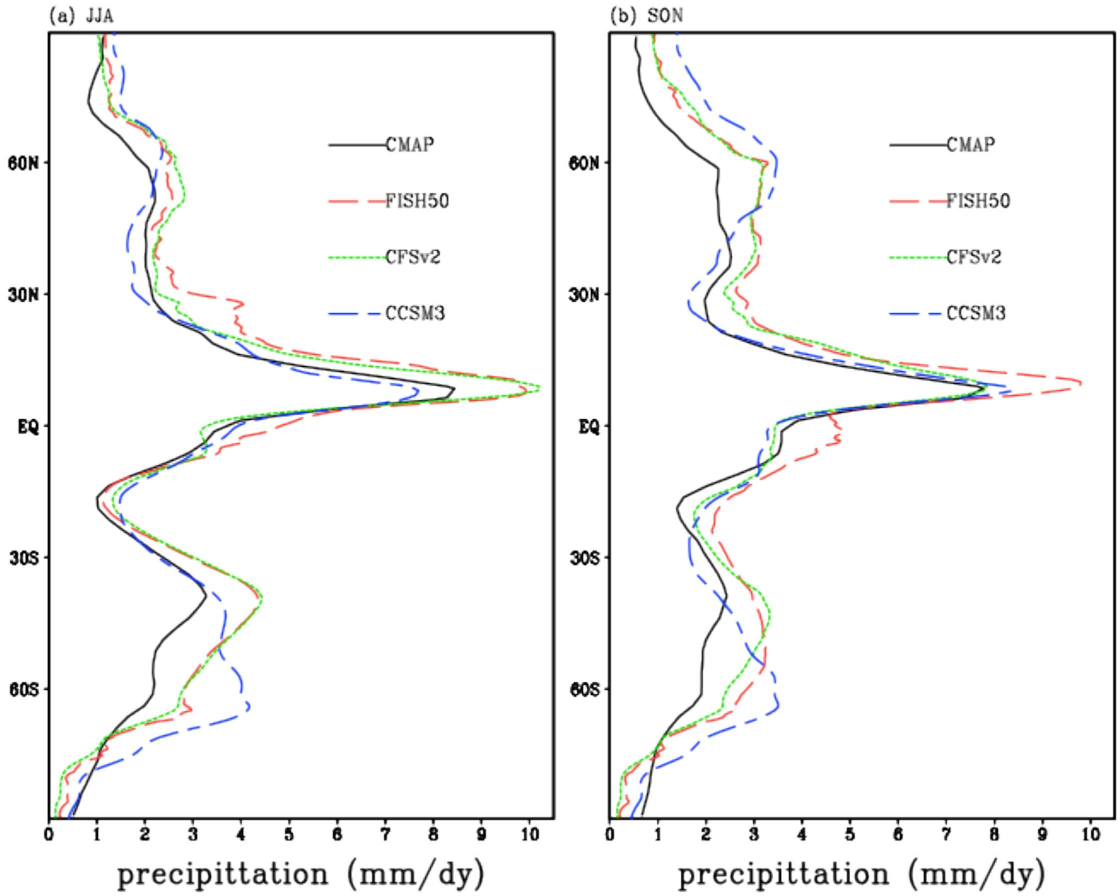

Figure 7. The zonal mean climtological (a) summer (JJA) and (b) fall (SON) precipitation from observations and the three seasonal hindcasts.

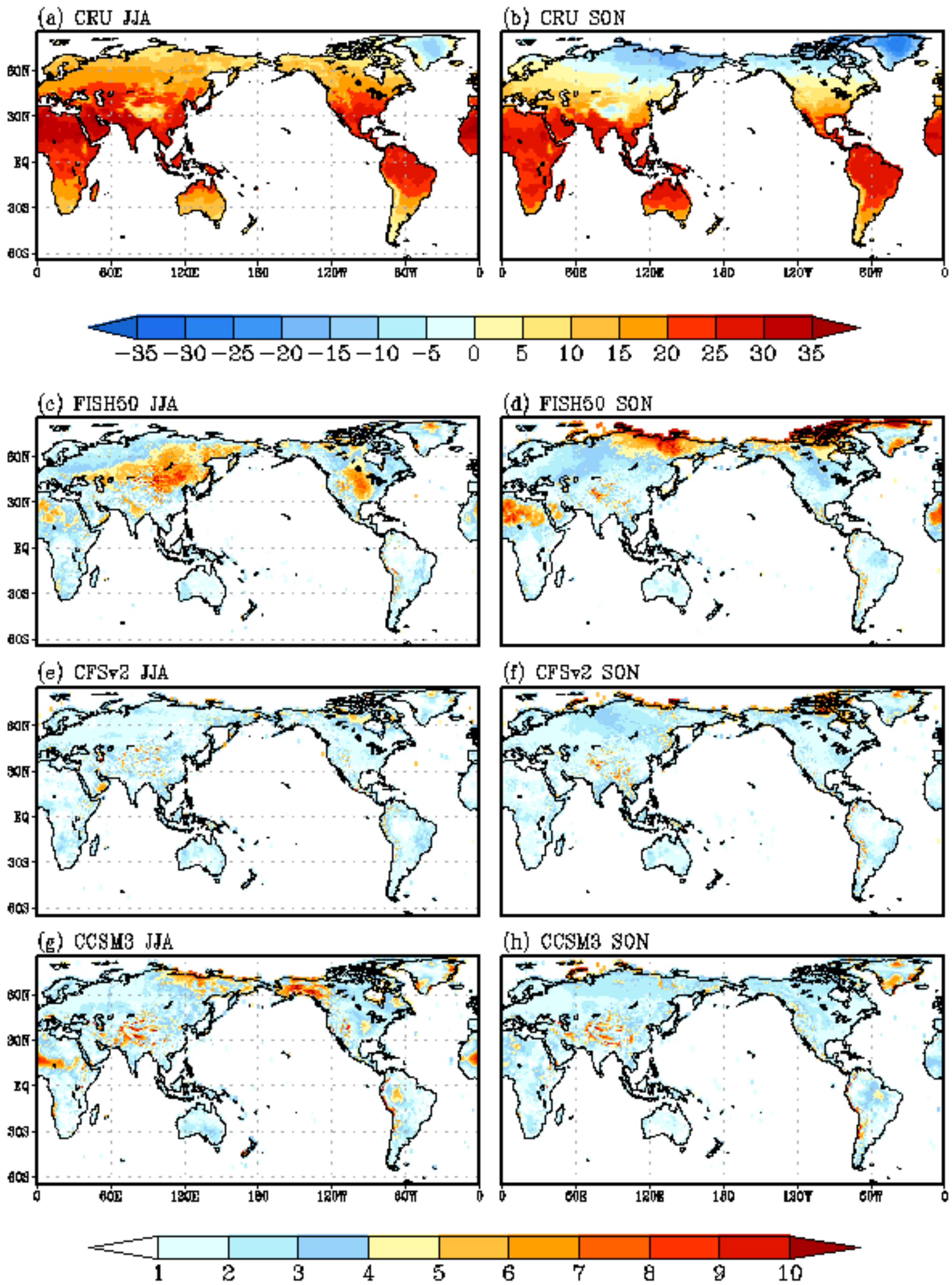

Figure 8. The observed climatology of surface temperature (1982-2008) in (a) JJA, and (b) SON. The root mean square error of the ensemble mean precipitation for JJA (zero lead) for seasonal hindcasts from (c) FISH50, (e) CFSv2, and (g) CCSM3. Likewise, the root mean square error of the ensemble mean precipitation for SON (one season lead) for seasonal hindcasts from (d) FISH50, (f) CFSv2, and (h) CCSM3. The units are in °C.

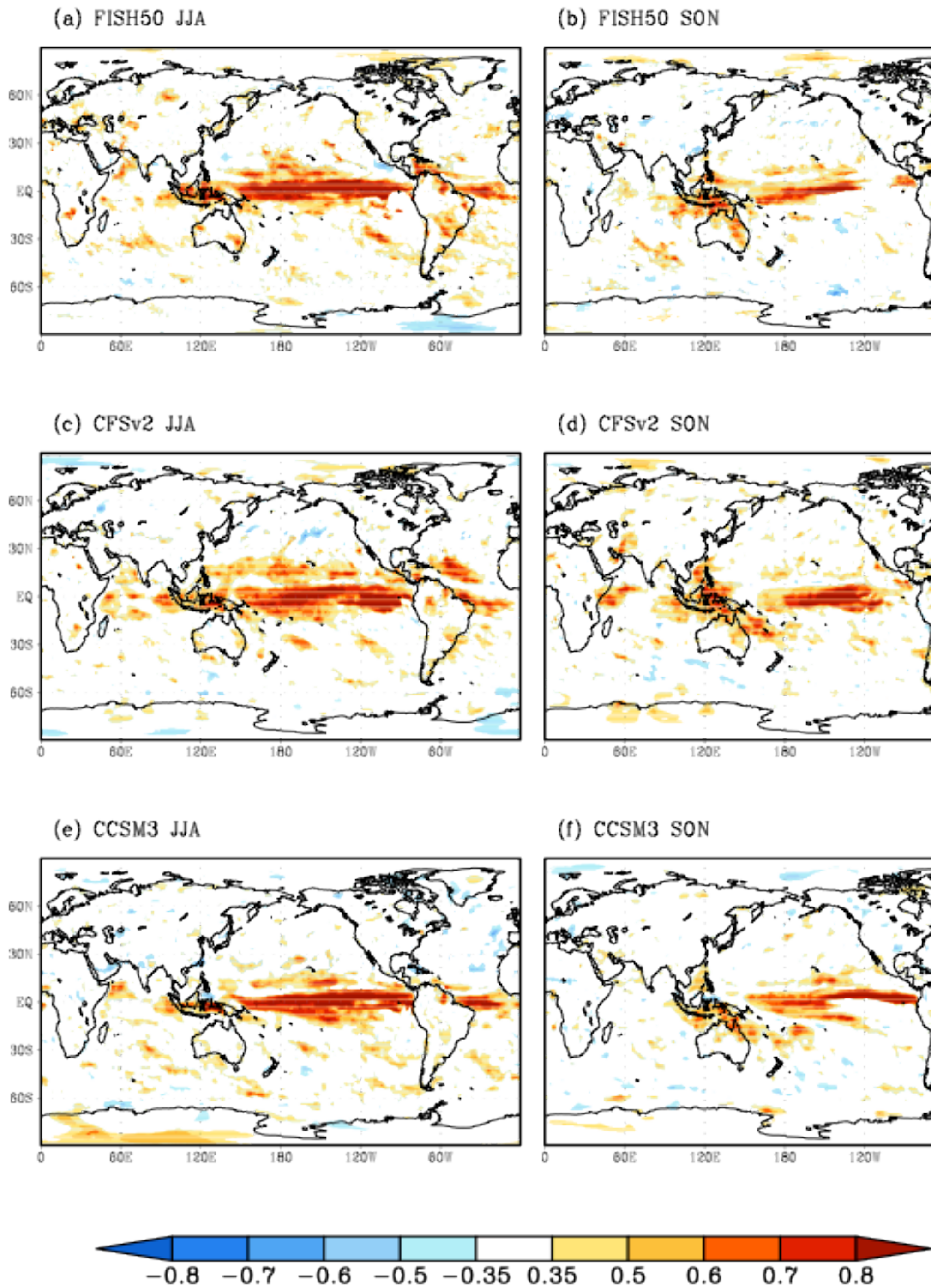

Figure 9. The correlation of the ensemble mean precipitation for JJA (zero lead) from (a) FISH50, (c) CFSv2, and (e) CCSM3. Similarly, the correlation of the ensemble mean precipitation for SON (one season lead) from (b) FISH50, (d) CFSv2, and (f) CCSM3.

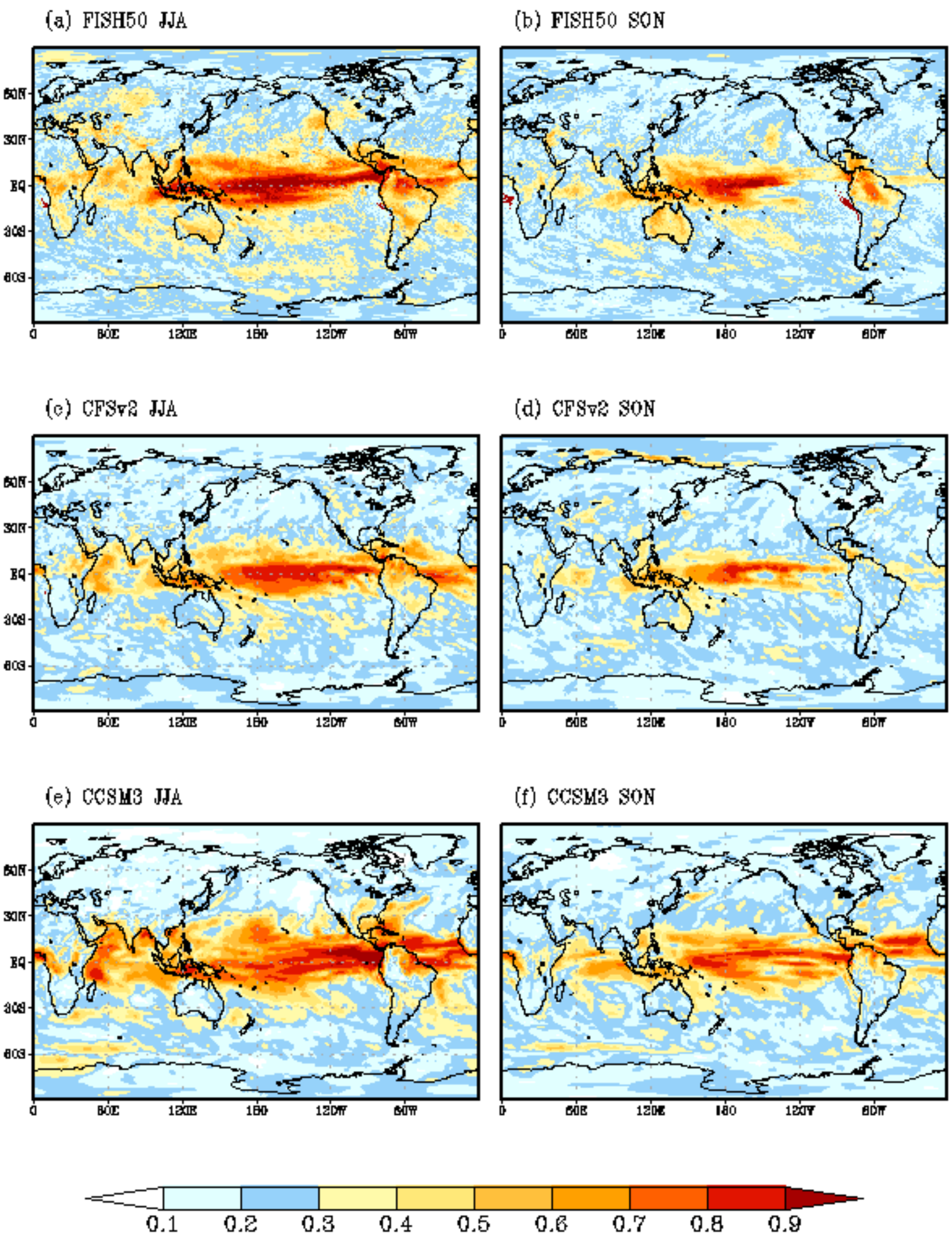

Figure 10. The signal to noise ratio of precipitation for JJA season (zero lead) for (a) FISH50, (c) CFSv2, and (e) CCSM3. The signal to noise ratio of precipitation for SON (one season lead) from (b) FISH50, (d) CFSv2, and (f) CCSM3.

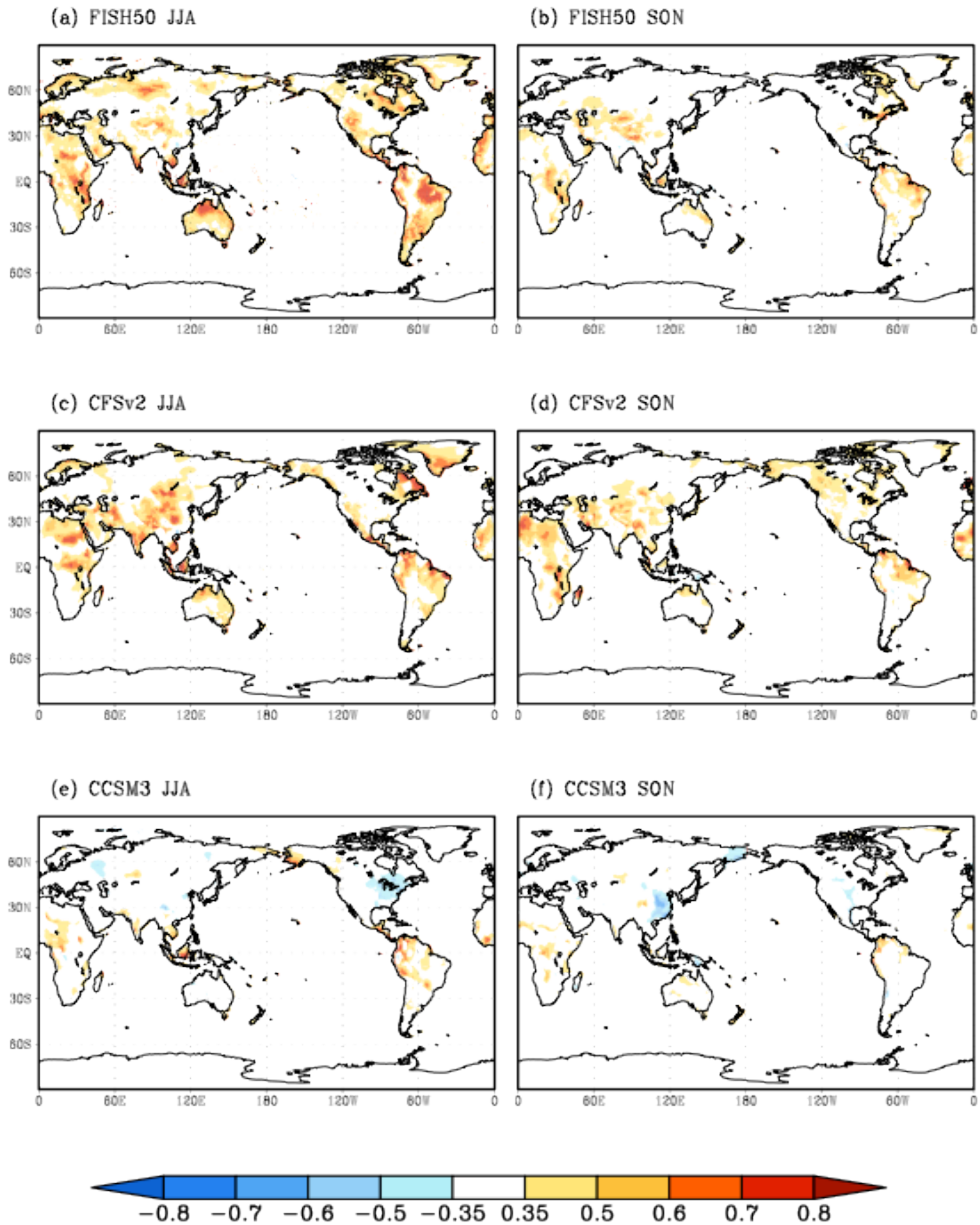

Figure 11. The correlation of the ensemble mean T2m for JJA (zero lead) from (a) FISH50, (c) CFSv2, and (e) CCSM3. Similarly, the correlation of the ensemble mean precipitation for SON (one season lead) from (b) FISH50, (d) CFSv2, and (f) CCSM3.

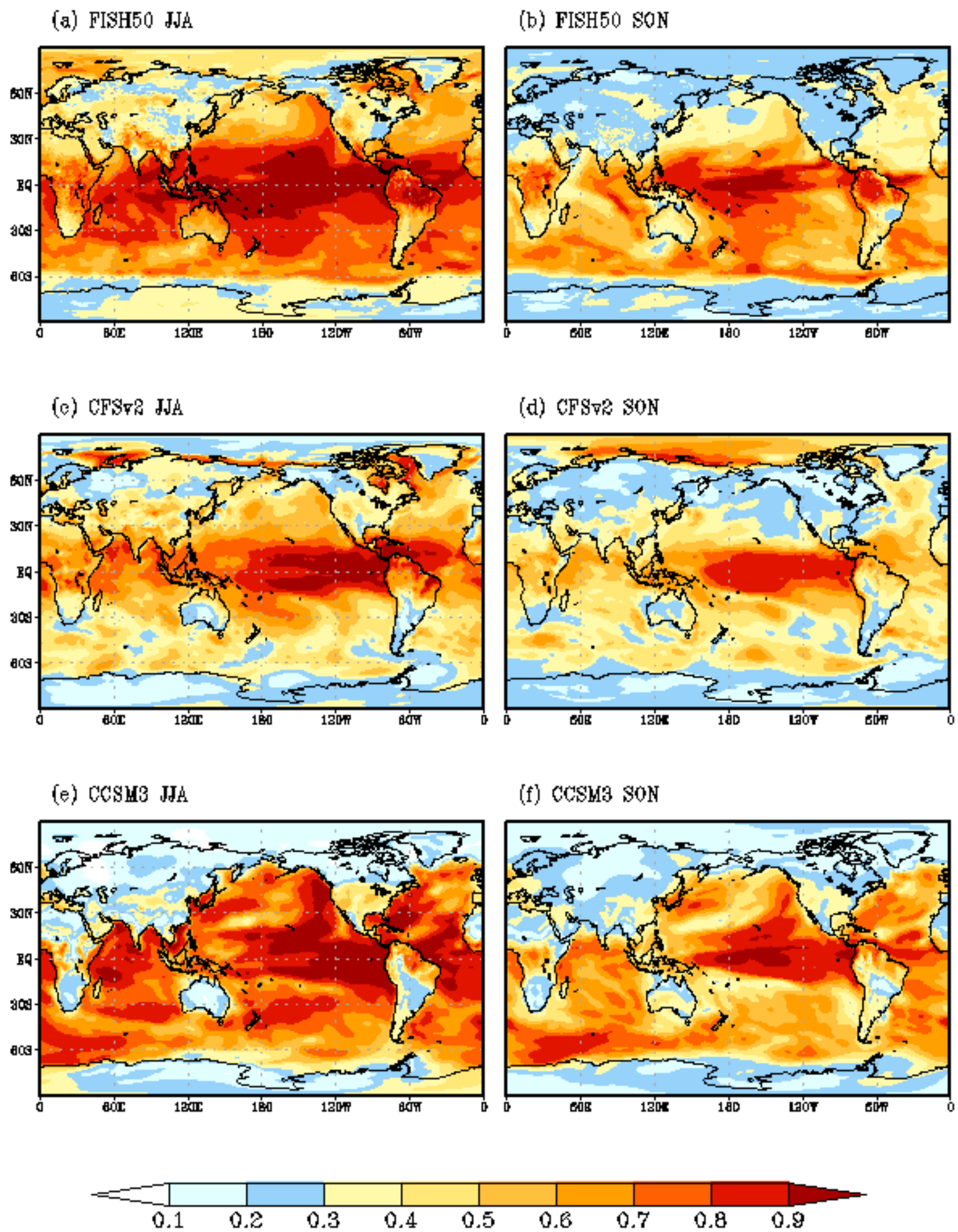

Figure 12. The signal to noise ratio of T2m for JJA season (zero lead) for (a) FISH50, (c) CFSv2, and (e) CCSM3. The signal to noise ratio of precipitation for SON (one season lead) from (b) FISH50, (d) CFSv2, and (f) CCSM3.

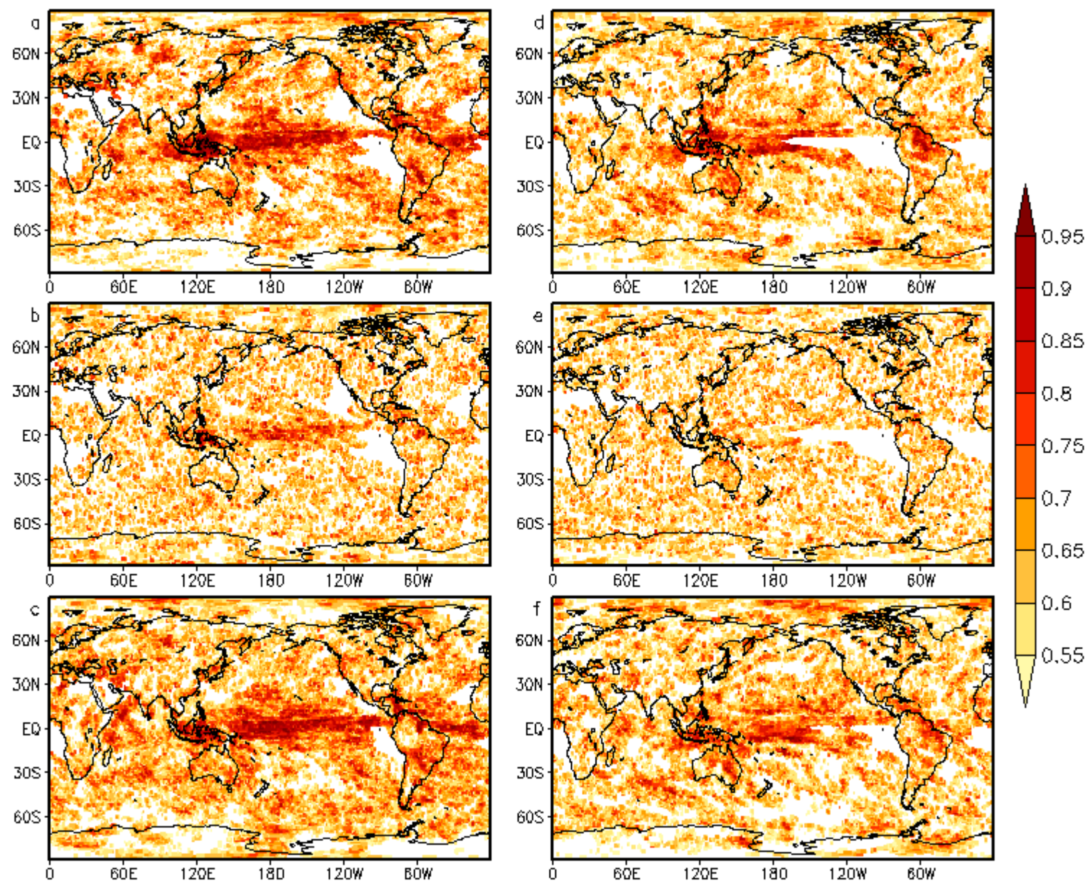

Figure 13. The area under the relative operation characteristic curve (ROC) for (a) lower, (b) middle, and (c) upper tercile for JJA (zero season lead) from FISH50 precipitation, Similarly, the area under the ROC for (d) lower, (e) middle, and (f) upper tercile for SON (one season lead) from FISH50 precipitation.

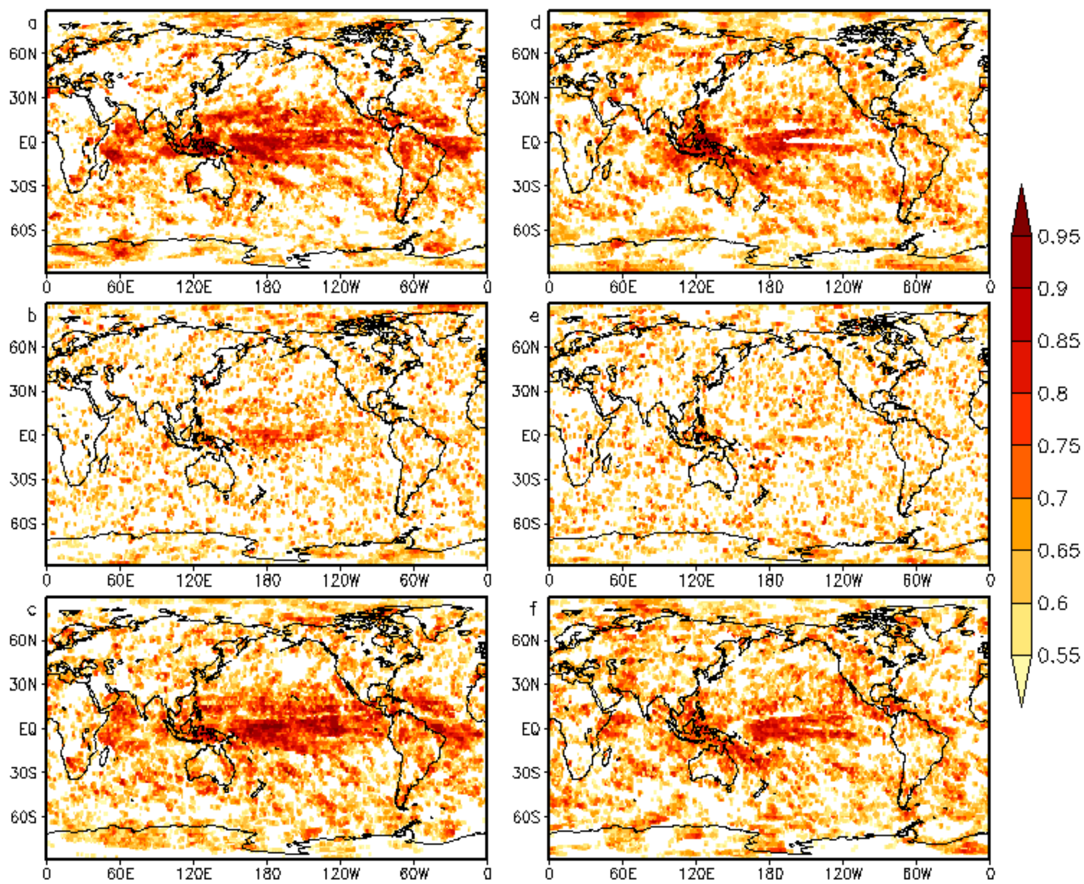

Figure 14. The area under the relative operation characteristic curve (ROC) for (a) lower, (b) middle, and (c) upper tercile for JJA (zero season lead) from CFSv2 precipitation, Similarly, the area under the ROC for (d) lower, (e) middle, and (f) upper tercile for SON (one season lead) from CFSv2 precipitation.

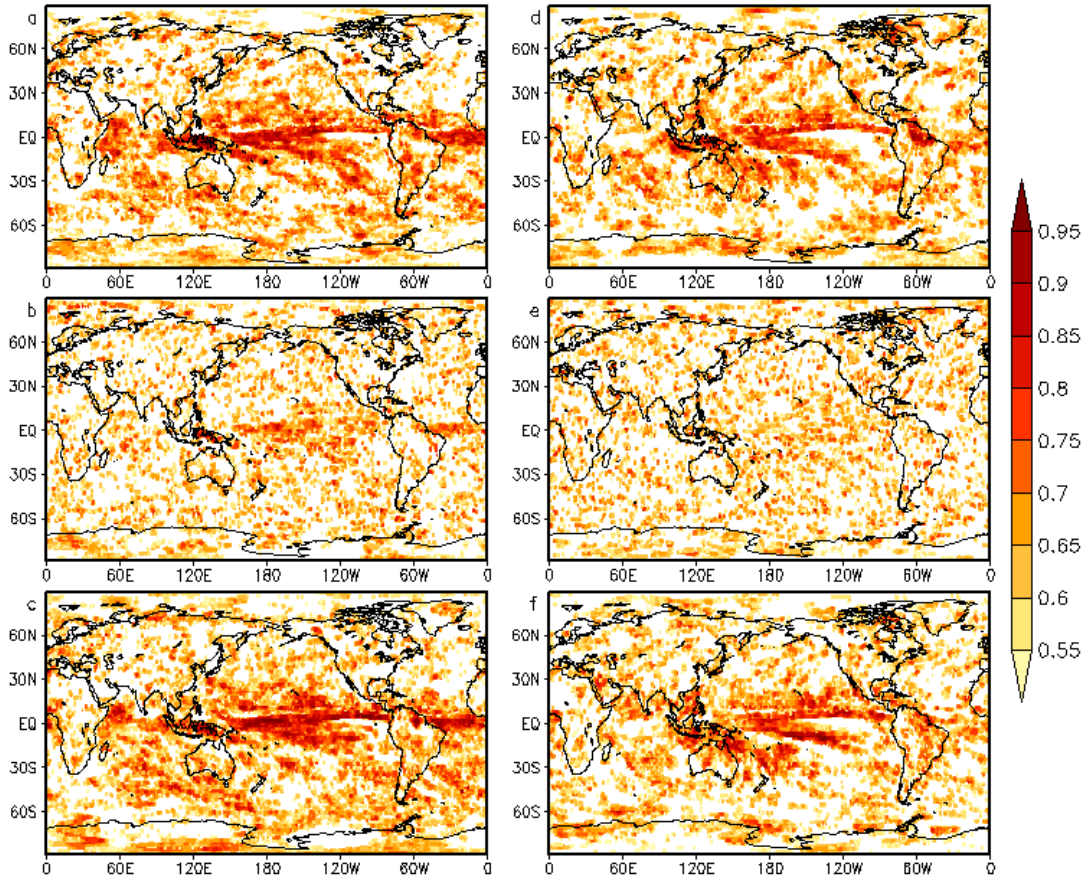

Figure 15. The area under the relative operation characteristic curve (ROC) for (a) lower, (b) middle, and (c) upper tercile for JJA (zero season lead) from CCSM3 precipitation, Similarly, the area under the ROC for (d) lower, (e) middle, and (f) upper tercile for SON (one season lead) from CCSM3 precipitation.

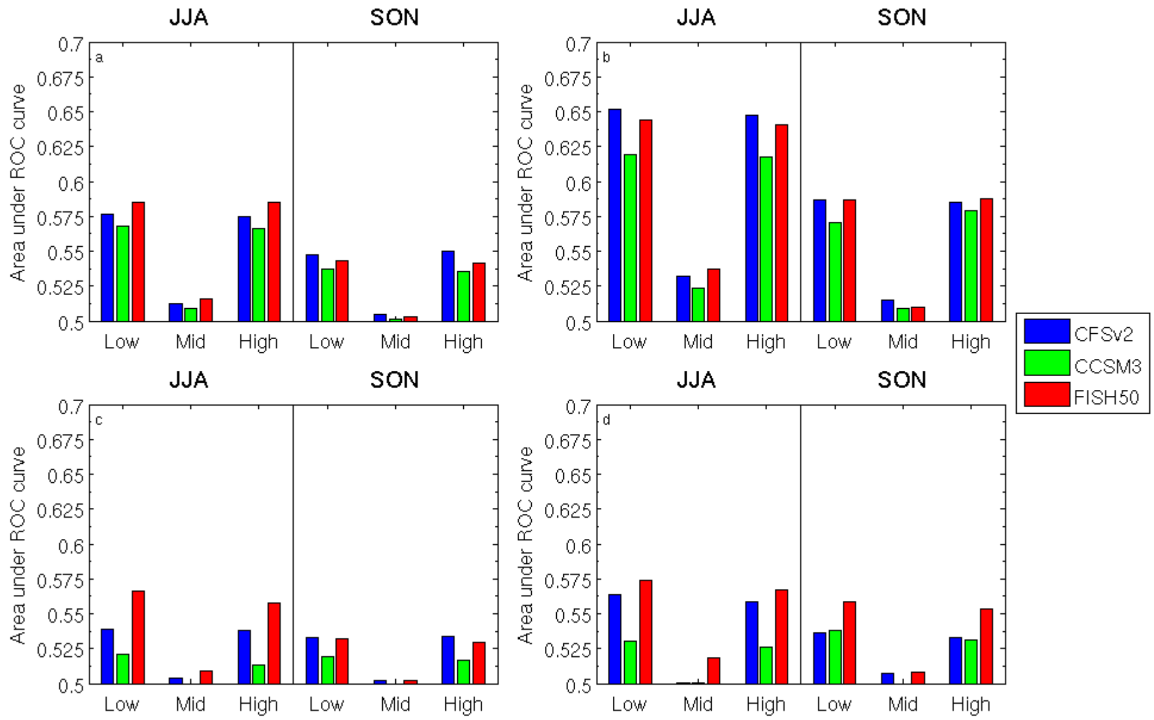

Figure 16. Area under the ROC averaged over (a) global oceans, (b) tropical oceans, (c) global land, and (d) tropical land for low, middle, and upper terciles of CFSv2, CCSM3, and FISH50 precipitation in JJA and SON.

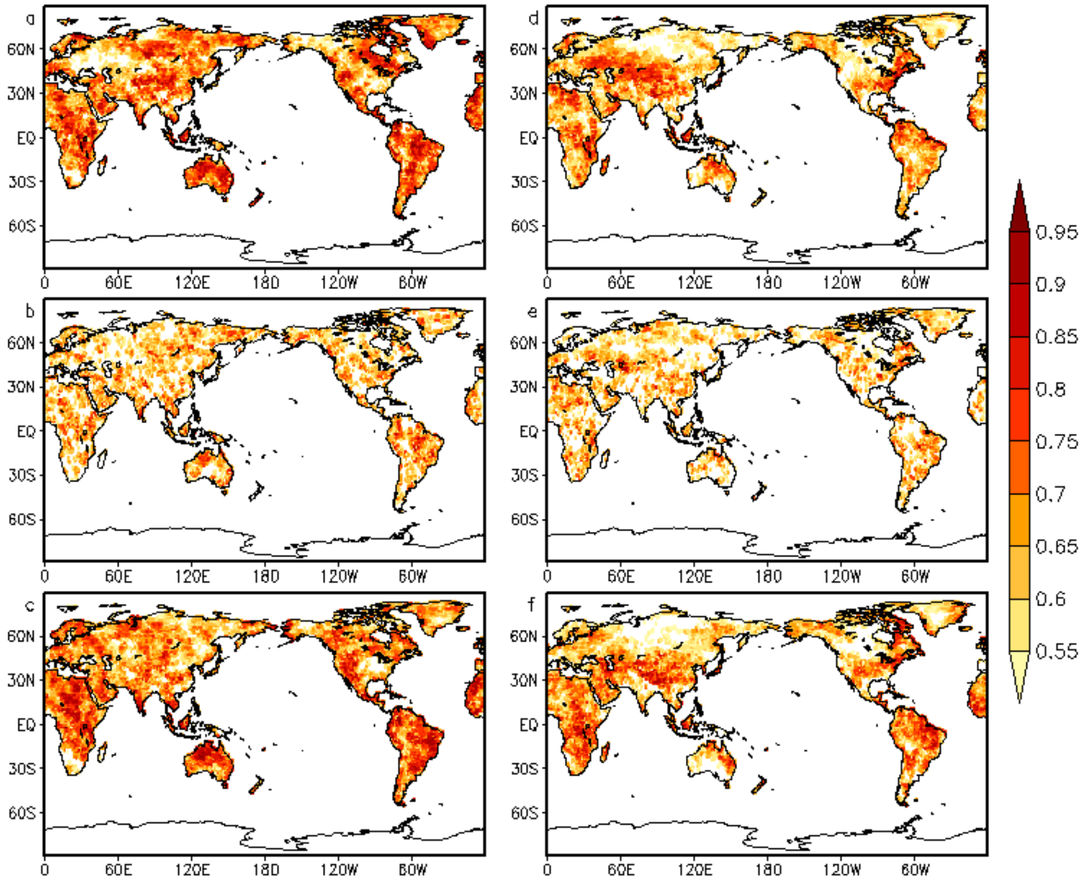

Figure 17. The area under the relative operation characteristic curve (ROC) for (a) lower, (b) middle, and (c) upper tercile for JJA (zero season lead) from FISH50 T2m, Similarly, the area under the ROC for (d) lower, (e) middle, and (f) upper tercile for SON (one season lead) from FISH50 T2m.

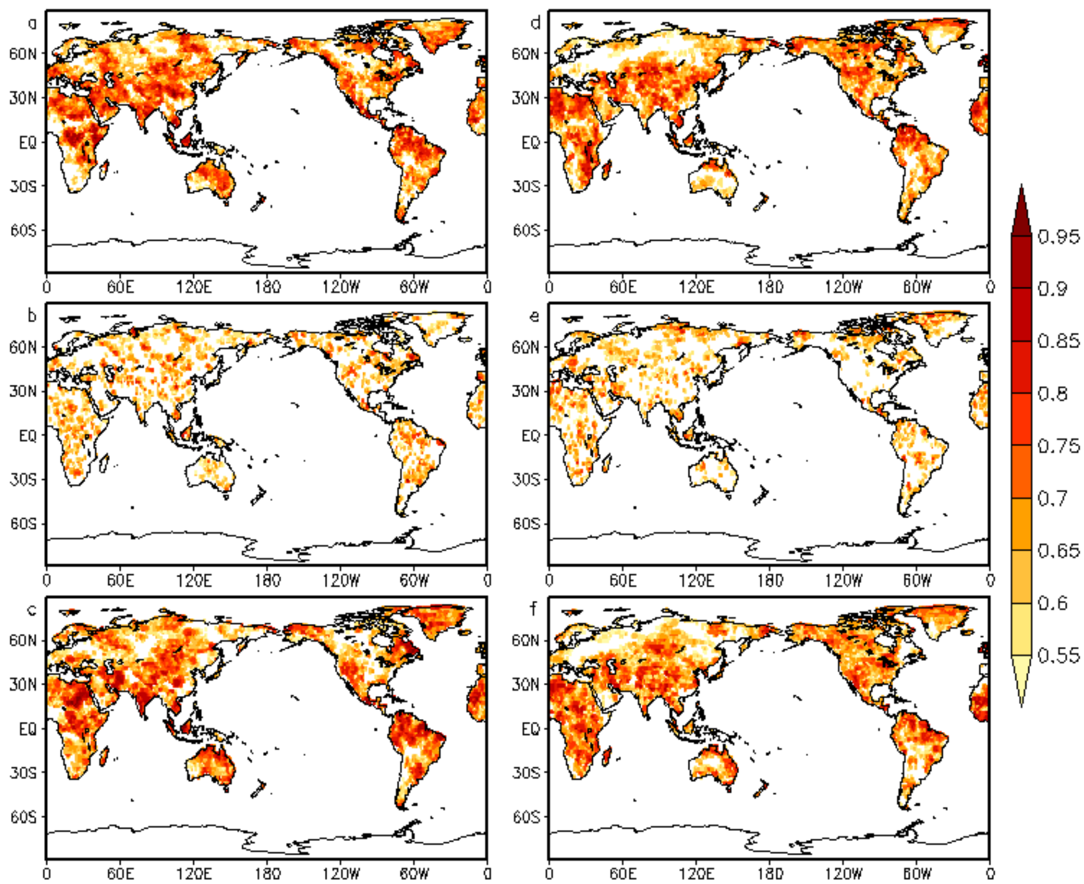

Figure 18. The area under the relative operation characteristic curve (ROC) for (a) lower, (b) middle, and (c) upper tercile for JJA (zero season lead) from CFSv2 T2m, Similarly, the area under the ROC for (d) lower, (e) middle, and (f) upper tercile for SON (one season lead) from CFSv2 T2m.

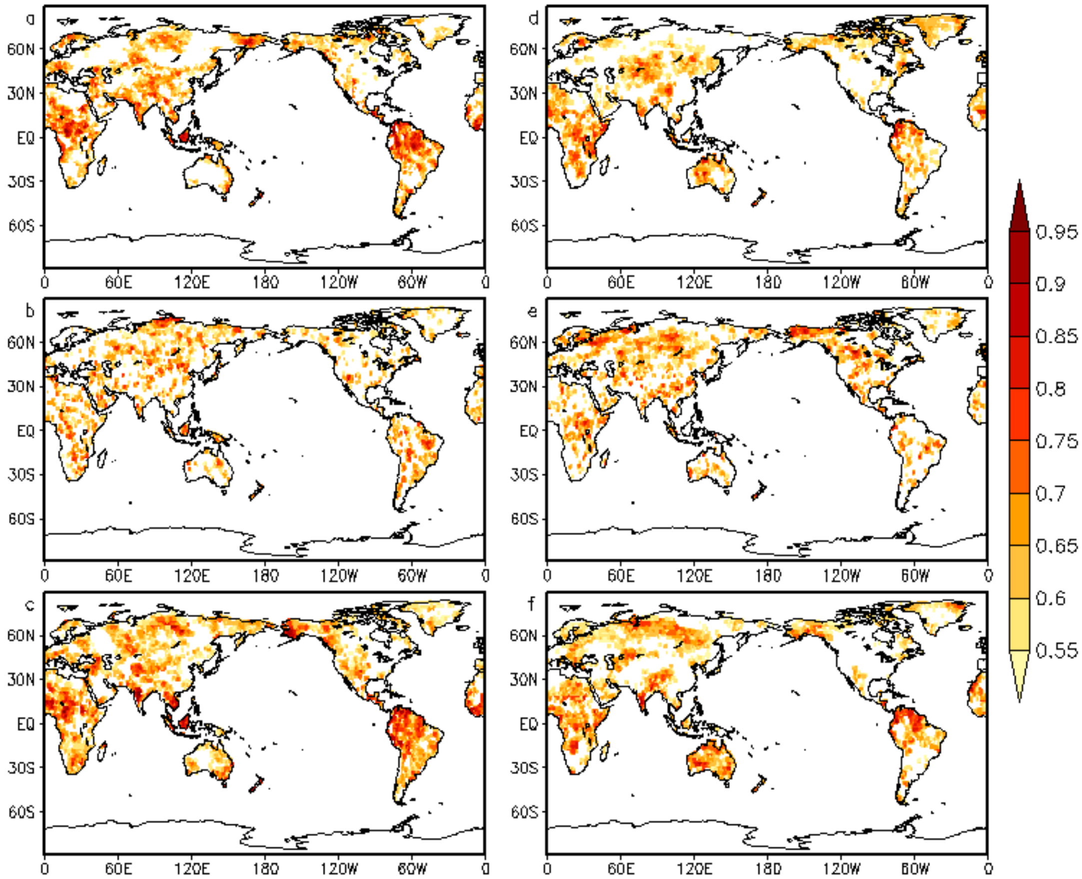

Figure 19. The area under the relative operation characteristic curve (ROC) for (a) lower, (b) middle, and (c) upper tercile for JJA (zero season lead) from CCSM3 T2m, Similarly, the area under the ROC for (d) lower, (e) middle, and (f) upper tercile for SON (one season lead) from CCSM3 T2m.

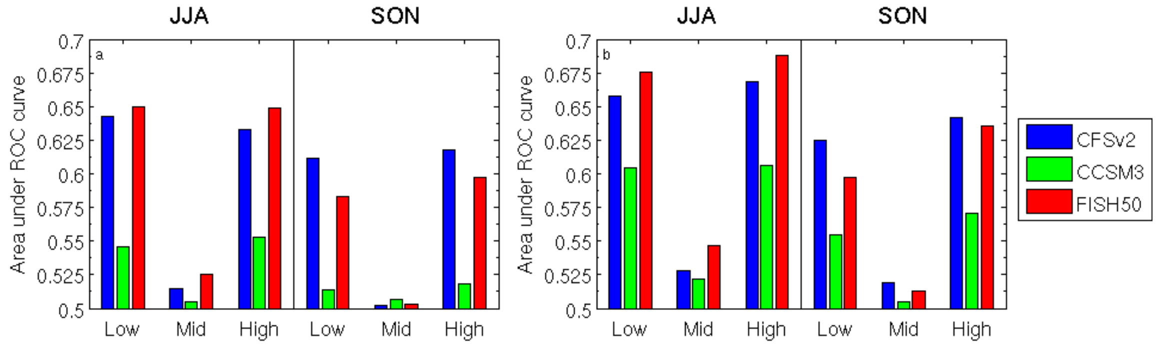

Figure 20. Area under the ROC averaged over (a) global land and (b) tropical land for low, middle, and upper terciles of CFSv2, CCSM3, and FISH50 temperatures in JJA and SON.