AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE

10.1002/2016JD026391

Key Points:

- Prediction skill (predictability) is not significantly influenced (weakly influenced) by ocean-atmosphere coupling
- Errors in SSTs lead to more significant differences in prediction skill and predictability versus ocean-atmosphere coupling differences

Correspondence to:
J. M. Infanti,
jinfanti@rsmas.miami.edu

Citation:
Infanti, J. M., and B. P. Kirtman (2017), CGCM and AGCM seasonal climate predictions: A study in CCSM4, *J. Geophys. Res. Atmos.*, *122*, 7416–7432, doi:10.1002/2016JD026391.

Received 16 DEC 2016
Accepted 6 JUL 2017
Accepted article online 8 JUL 2017
Published online 26 JUL 2017

CGCM and AGCM seasonal climate predictions: A study in CCSM4

Johnna M. Infanti[1] **and Ben P. Kirtman**[1]

[1]Department of Atmospheric Sciences, University of Miami Rosenstiel School of Marine and Atmospheric Science, Miami, Florida, USA

Abstract Seasonal climate predictions are formulated from known present conditions and simulate the near-term climate for approximately a year in the future. Recent efforts in seasonal climate prediction include coupled general circulation model (CGCM) ensemble predictions, but other efforts have included atmospheric general circulation model (AGCM) ensemble predictions that are forced by time-varying sea surface temperatures (SSTs). CGCMs and AGCMs have differences in the way surface energy fluxes are simulated, which may lead to differences in skill and predictability. Concerning model biases, forecasted SSTs have errors compared to observed SSTs, which may also affect skill and predictability. This manuscript focuses on the role of the ocean in climate predictions and includes the influences of ocean-atmosphere coupling and SST errors on skill and predictability. We perform a series of prediction experiments comparing coupled and uncoupled Community Climate System Model version 4.0 (CCSM4) predictions and forecasted versus observed SSTs to determine which is the leading cause for differences in skill and predictability. Overall, prediction skill and predictability are only weakly influenced by ocean-atmosphere coupling, with the exception of the western Pacific, while errors in forecasted SSTs significantly impact skill and predictability. Comparatively, SST errors lead to more significant and robust differences in prediction skill and predictability versus inconsistencies in ocean-atmosphere coupling.

1. Introduction

Seasonal climate predictions are composed of model simulations of the near-term climate for approximately a year in the future. These predictions are based on known present conditions and simulate anticipated near-term climate information for the upcoming months and seasons. On the seasonal timescale, sea surface temperatures (SSTs) are a very important element, as predictions are predominantly influenced by slowly evolving surface boundary conditions including, but not limited to, SSTs [*Shukla*, 1998; *Koster and Suarez*, 2003; *Paolino et al.*, 2011; *Kirtman et al.*, 2014; among others].

Recent efforts in seasonal climate prediction utilize coupled general circulation models (CGCMs) initialized from observed states in the ocean, land, and atmosphere domains, such as those included in the North American Multi-Model Ensemble (NMME) [see, for example, *Kirtman and Min*, 2009; *Paolino et al.*, 2011; *Kirtman et al.*, 2014]. NMME predictions have been discussed in numerous studies. For example, *Becker et al.* [2014] considered overall prediction and predictability, *Infanti and Kirtman* [2013, 2015] assessed prediction skill and model response to tropical Pacific forcing over North America, and similar assessments have been preformed over Africa [*Shukla et al.*, 2016] and China [*Ma et al.*, 2016], among others. In other efforts, atmospheric general circulation models (AGCMs) with prescribed, time-varying, forecasted SST taken from coupled simulations or persisted SSTs are used to formulate climate predictions [*Bengtsson et al.*, 1993; *Goddard and Mason*, 2002; *Kumar et al.*, 2008] or for predictability or potential prediction studies using observed SSTs [e.g., *Graham et al.*, 2000]. Atmospheric teleconnections or oceanic heat fluxes may be misrepresented in CGCM and AGCM forecasts due to errors in SSTs, thus leading to errors in associated land-based precipitation or temperature forecasts. Differences in ocean-atmosphere coupling in AGCM and CGCM predictions can lead to differences in predictions of midlatitude oceanic surface energy fluxes, also potentially influencing predictions over land.

Biases in the oceanic and atmospheric components of climate models are common and are principally due to misrepresentation of or failure to resolve physical processes, especially in the tropics due to feedbacks in the region [*Wang et al.*, 2014]. For example, the current generation of coupled climate models suffers from an excessive cold tongue in the equatorial Pacific and a double intertropical convergence zone [*Mechoso* *et al.*, 1995; *de Szoeke and Xie*, 2008; *Li and Xie*, 2013]. The cold tongue errors can cause biases in precipitation and winds, and these biases are amplified by ocean-atmosphere interaction in coupled models in comparison to atmosphere-only simulations with observed prescribed SSTs. Biases due to the double intertropical convergence zone can be associated with excessive downward solar radiation in atmospheric models [*de Szoeke and Xie*, 2008; *Li and Xie*, 2013]. El Niño–Southern Oscillation (ENSO) performance is also biased in terms of amplitude, location, and timing of ENSO events, which are tied to atmospheric feedbacks, specifically the Bjerknes and heat flux feedbacks (particularly short wave) [*Bellenger et al.*, 2014]. The tropical Atlantic also suffers from biases in SSTs related to the meridional mode and in the equatorial cold tongue region, and the interannual variability of SSTs in the region can impact remote precipitation [*Richter et al.*, 2014, and references therein]. These biases can impact ENSO variability due to both coupled process errors and SST biases [*Richter et al.*, 2014; *Sasaki et al.*, 2014]. As SST biases can influence the atmospheric model component (including teleconnections) and, in turn, coupled processes can add to SST biases, we focus our efforts on two areas. The first is the influence of SST biases on prediction skill and predictability, and the second is the differences in ocean-atmosphere coupling in CGCM and AGCM predictions.

Due to the importance of SSTs, one of our main considerations is the influence of SST errors on hindcast skill and predictability in Community Climate System Model version 4.0 (CCSM4). SST biases or errors can impact both the skill of predictions and remote trends. Foreknowledge or skillful predictions of SST anomalies can add to predictability or prediction skill in regions with strong associations [*Shukla*, 1998; *Livezey and Timofeyeva*, 2008]. For example, SST anomalies, such as those associated with ENSO, can stimulate atmospheric teleconnections that impact North American climate variability [*Ropelewski and Halpert* 1986; *Trenberth et al.*, 1998]. Using historical CMIP5 models, *Shin and Sardeshmukh* [2011] found that coupled models have biases in tropical SSTs, which impact remote trends. In a climate prediction setting, *Shukla* [1998] showed that during the JFM1998 El Niño event, extratropical circulation was predicted with some skill even at a 6 month lead but would have been more accurate if the forecasted SSTs were closer to observed, indicating that SST biases impacted the prediction. In some regions, such as the southeastern U.S., prediction skill is highly sensitive to changes in SST, and skill can suffer due to errors in predicted SSTs [*Infanti and Kirtman*, 2015]. In contrast, and when using persisted versus observed prescribed SSTs, *Goddard and Mason* [2002] found that in regions where skill is highly tied to El Niño, the skill was similar for both simulations but that SST errors led to losses in prediction skill in other regions.

A second motivation for this manuscript is the differences in ocean-atmosphere coupling in CGCM versus AGCM predictions. We focus on large-scale fields affecting North American climate variability. Ocean-atmosphere coupling, or lack of ocean-atmosphere coupling in the case of an AGCM prediction, may have some bearing on model performance and therefore prediction skill due to atmospheric feedbacks on SSTs and any resulting changes in global circulation. Wintertime North American climate variability is largely and principally influenced by the tropical Pacific through ENSO variations [*Ropelewski and Halpert*, 1986, 1987; *Trenberth et al.*, 1998; *Mo and Schemm*, 2008a, 2008b; among many others]. However, other studies have suggested that midlatitude SSTs may play a role, particularly on seasonal timescales, such as the linkage of the North Pacific Ocean and geopotential height over North America, the Pacific North American (PNA) pattern, or impacts from the tropical Atlantic [*Palmer and Zhaobo*, 1985; *Frankignoul*, 1985; *Barsugli and Battisti*, 1998; *Saravanan*, 1998; *Straus and Shukla*, 2002; *Wu and Kirtman*, 2007; *Sasaki et al.*, 2014; among others]. The use of a AGCM can lead to inconsistencies in surface energy fluxes and thus inaccuracies in simulated climate variability [*Barsugli and Battisti*, 1998]. However, CGCMs are not without their deficiencies. They can also produce unrealistic air-sea fluxes due to biases in SSTs or winds [*Yu and Mechoso*, 1999]. It is therefore important to characterize influence of ocean-atmosphere coupling on skill and predictability in a prediction setting.

In addition to characterizing ocean-atmosphere coupling influences on skill and predictability in a prediction setting, comparison of CGCM and AGCM predictions is also useful to determine if there are any expected skill or predictability differences. Despite the potential shortcomings of an AGCM prediction system, it has been a practical choice for prediction studies for quite some time [see, for example, *Bengtsson et al.*, 1993]. Further studies such as high-resolution predictions [e.g., *Li and Misra*, 2014; *Misra et al.*, 2014], or bias correcting predicted SSTs [e.g., *Kumar et al.*, 2008] are very possible in this type of prediction system, and thus any differences between the two simulations are important to consider. Previous studies have investigated CGCM and AGCM simulations in both free-running and prediction simulations. In a comparison of the Community

Climate System Model version 3 (CCSM3) CGCM and the Community Atmosphere Model version 3 AGCM, the atmospheric responses to SST forcing as well as weather noise statistics were similar [*Chen et al.*, 2012]. Conversely, other studies found that AGCM performance strongly depended on whether or not SST forcing was dominant, for example, in the tropical Pacific versus midlatitudes, and speculate that this could affect prediction skill [*Wu and Kirtman*, 2004; *Wu et al.*, 2006; *Wu and Kirtman*, 2007]. In a prediction framework, using coupled and uncoupled experiments in the Climate Forecast System (CFS), *Kumar et al.* [2008] found that strong atmospheric responses to SSTs were sufficient enough to show similar responses in winter predictions. However, *Kumar et al.* [2008] and *Zhou et al.* [2012] also state that in other seasons/timescales, ocean-atmosphere coupling may be more important. CCSM4 predictions, in the framework of CGCM and AGCM predictions, have not yet been studied.

We focus on the role of the ocean in seasonal climate predictions of North American land-based precipitation and temperature. We aim to study the relative importance of errors in SSTs (observed versus forecasted) and differences in ocean-atmosphere coupling in predictions made with a CGCM compared to an AGCM. This study considers CGCM and AGCM prediction methods using CCSM4, of which fully coupled hindcasts are included in NMME [*Kirtman et al.*, 2014]. Given the potential impacts to skill due SSTs and/or ocean-atmosphere coupling differences in CGCM compared to AGCM predictions, a comparison of deterministic skill, probabilistic skill, and predictability is shown in a variety of CCSM4 simulations. These simulations include CCSM4 fully coupled hindcasts, hindcasts with SSTs prescribed from predictions, and "potential predictability" hindcasts with SSTs prescribed from observations (AGCM hindcasts). Regionally, our focus is on North America and the Pacific/Atlantic Oceans. This manuscript is organized as follows: Section 2 details the relevant modeling experiments and analysis methods, section 3 includes results, section 4 includes the discussion, and section 5 the concluding remarks.

2. Methods

2.1. Modeling Experiments

The goal of this manuscript is to compare predictions with differing SSTs (observed versus forecasted) and with differing ocean-atmosphere coupling (CGCM versus AGCM predictions). The CGCM predictions are performed in the fully coupled CCSM4 [*Gent et al.*, 2011] and are those included in NMME [*Kirtman et al.*, 2014]. AGCM predictions are performed in the standalone atmospheric model of CCSM4, the Community Atmosphere Model version 4 (CAM4), which is coupled to the Community Land Model version 4 (CLM4). Both CGCM and AGCM predictions use initial states taken from the Climate Forecast System Reanalysis (CFSR) [*Saha et al.*, 2010], approximating observations. Prescribed SST data used in the AGCM predictions are detailed below. Both CCSM4 AGCM and CGCM predictions are identical in hindcast initialization and model setup, but SSTs are prescribed in the CCSM4 AGCM prediction experiments (no bias correction is performed on prescribed SSTs). Experiments and relevant notation follow.

1. FC: Fully coupled CCSM4 hindcasts. Ten ensemble members generated from observed atmosphere, land, and ocean initial states for a period of 1982–2009. More information about these fully coupled hindcasts can be found in *Paolino et al.* [2011] and *Kirtman et al.* [2014].
2. CAM4_OBS: CAM4 "hindcasts" with observed, prescribed SST [*Rayner et al.*, 2003]. Ten ensemble members generated from observed atmosphere and land initial states for a period of 1982–2009. Hindcasts are initialized every December Lead times up to 6 months are considered.
3. CAM4_FC: CAM4 hindcasts with SSTs prescribed from the first to tenth ensemble member FC predictions, where CAM4_FC ensemble member 1 has prescribed SST consistent with FC ensemble member 1, CAM4_FC ensemble member 2 has prescribed SST consistent with FC ensemble member 2, and so on to ensemble member 10. Ensemble generation, initialization, and lead times as above.

This experimental design is such that comparing CAM4_OBS and CAM4_FC shows any differences expected if SSTs could be perfectly predicted; however, the experiments have energetic inconsistencies as SSTs are prescribed. As CAM4_OBS uses observed, prescribed SSTs, it is more aptly referred to as a "potential prediction" experiment; however, we refer to it as a prediction in the interest of brevity. Systematic error between FC predicted SSTs and observations is shown in Figure 1, and we see that SST error increases as lead time increases (see Figures 1a–1d) and that there is a cold bias over much of the tropical ocean(s). In comparing CAM4_FC and FC experiments, we consider the impact of ocean-atmosphere coupling on predictions, which have the

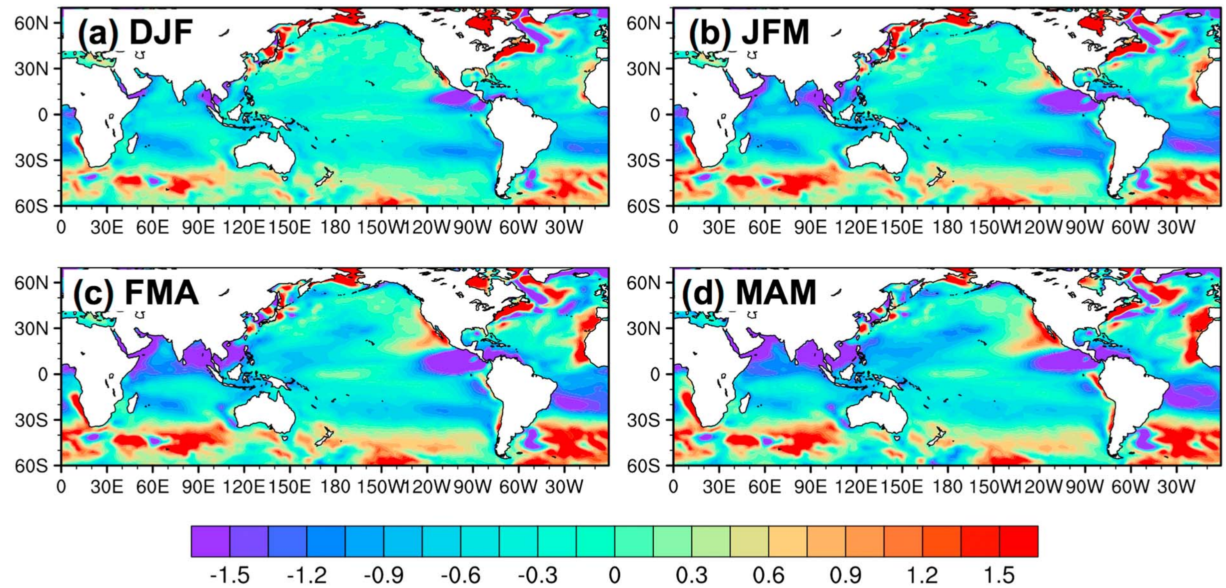

Figure 1. Systematic error of FC ensemble mean predicted SSTs minus observed (NCDC) SSTs.

same SSTs but differing ocean-atmosphere coupling. As the SSTs are prescribed in the CAM4_FC experiment, this experiment isolates the SST-forced component of atmospheric anomalies. However, we note that the prescribed SSTs include some coupling information, as they are output from the coupled model. This experiment is very similar to *Wu et al.* [2006] but utilizes forecasted SSTs.

Initialization of CCSM4 is similar to CCSM3 [*Paolino et al.*, 2011], and we briefly discuss the procedure for initialization of CAM4 and CLM4. All initial data are taken from the CFSR [*Saha et al.*, 2010]. CAM4 is initialized from multilevel fields of temperature, zonal and meridional winds, specific humidity, cloud liquid water content, cloud ice water content, and cloud fraction and from single level fields of surface pressure, surface geopotential, surface temperature, and planetary boundary layer height. The data are regridded to the 0.9×1.25 degree grid and 26-hybrid sigma pressure levels used by CAM4.

CLM4 is initialized from daily fields of soil moisture, soil temperature, snow depth, snow temperature, vegetation temperature, and canopy moisture. These fields are normalized by their standard deviations and combined with the mean and standard deviation of soil climatology from 30 years of CLM4 output data, sampled after 100 year spin-up. Initial data south of 60S are set to model climatology. As observed data do not exist, vegetation temperature and canopy moisture initial fields are produced from a 7 day CCSM4 spin-up forecast where CAM is initialized as above to produce fields influenced by the initial atmospheric state.

Daily data from the end of the prior month are used to initialize ensemble members. A hindcast initialized on 1 December uses CFSR data from 26 November 00Z to initialize ensemble member 1, data from 26 November 12Z to initialize ensemble member 2, and so on. We refer to the first season after initialization as seasonal lead 1 (December initialized hindcast predicting DJF), second season as seasonal lead 2 (December initialized hindcast predicting JFM), etc. In CGCM predictions, SSTs are freely evolving and initialized from observed ocean states. In AGCM predictions, SSTs are prescribed. Ocean-atmosphere coupling frequency in FC hindcasts is once per day; thus, CAM4_FC experiments are forced by daily SST output as in *Chen et al.* [2012]. CAM4_OBS experiments are forced by monthly SSTs as is typical in observed SST experiments. Although we use daily SST in CAM4_FC experiments for consistency with the coupling frequency of FC hindcasts, we saw only small differences when monthly SSTs were used, and the errors in FC SSTs are large compared to the error when using monthly as opposed to daily observed SSTs.

2.2. Methods of Analysis

Observational estimates considered are the Climate Prediction Center Merged Analysis of Precipitation (CMAP) [*Xie and Arkin*, 1997] and the Global Historical Climatology Network/Climate Anomaly Monitoring System 2 m temperature (T2m) [*Fan and van den Dool*, 2008]. Model anomalies are calculated with respect

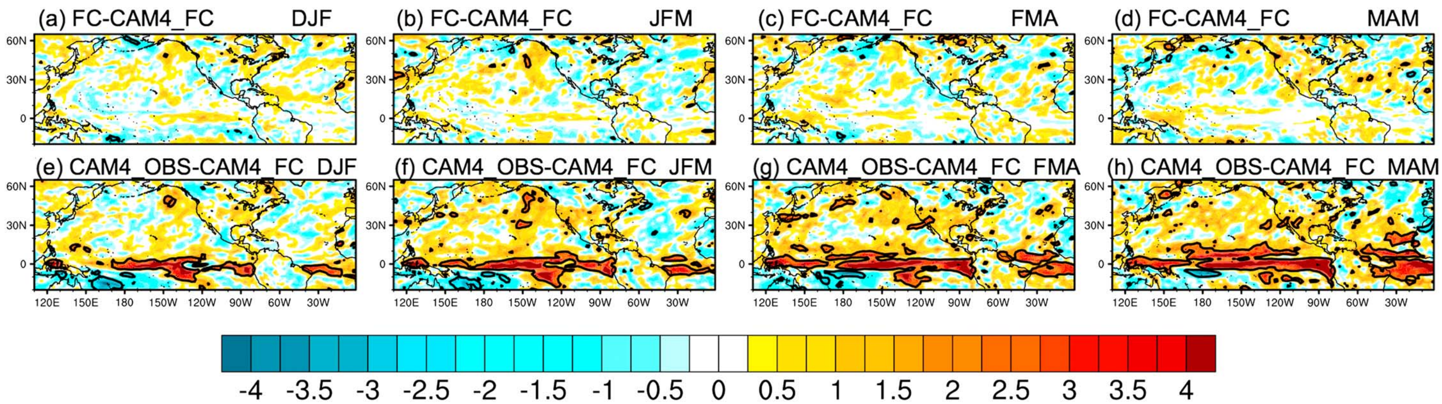

Figure 2. Transformed (using Fisher's R-to-Z transformation) difference in precipitation anomaly correlation (AC). (a–d) FC − CAM4_FC December initialized hindcasts predicting DJF − MAM. (e–h) CAM4_OBS − CAM4_FC December initialized hindcasts predicting DJF − MAM. Red (blue) shading indicates regions where FC or CAM4_OBS has stronger (weaker) skill than CAM4_FC. Contours indicate significance of difference at 95% confidence level. Anomaly correlation is calculated with respect to observed (CMAP) precipitation.

to model climatology and observed anomalies with respect to observed climatology. The hindcast period considered is 1982–2009. The ensemble mean is calculated by averaging all 10 ensemble members without weighting [e.g., *Infanti and Kirtman*, 2013].

Prediction skill is determined using anomaly correlation (AC; deterministic) and rank probability skill score (RPSS; probabilistic) as using both methods gives a more comprehensive assessment of prediction skill [*Kirtman*, 2003]. AC is a deterministic measure of skill comparing the predicted ensemble mean to observations [e.g., *Kirtman and Min*, 2009; *Wilks*, 2011; *Infanti and Kirtman*, 2013, 2015]. RPSS is a probabilistic measure of skill based on comparison of the cumulative squared probability error of the prediction versus that of a reference forecast for three tercile categories [e.g., *Mason*, 2004; *Weigel et al.*, 2007; *Wilks*, 2011]. We also consider estimates of "perfect model" predictability, or the skill of the simulation in predicting itself; i.e., there is no systematic error or biases with respect to observations. This can show regions or seasons in which we might expect a given variable to be predictable [*Boer*, 2004; *Cheng et al.*, 2011]. We retain one ensemble member as "truth" and estimate predictability based on the remaining nine ensemble members [e.g., *Infanti and Kirtman*, 2016].

Our interests lie in comparing FC to CAM4_FC and in comparing CAM4_OBS to CAM4_FC to determine differences in skill and predictability due to ocean-atmosphere coupling and SST errors. For AC, we use Fisher's R-to-Z transformation to transform the correlations to approximately normal [*Wilks*, 2011; *Infanti and Kirtman*, 2013; *DelSole and Tippett*, 2014; *Infanti and Kirtman*, 2016] and then difference the normalized ACs. RPSS comparison is calculated using CAM4_FC RPSS as the reference forecast versus FC and CAM4_OBS.

3. Results

3.1. Comparison of Skill

We consider two representations of skill, AC (deterministic) and RPSS (probabilistic). AC and RPSS are calculated with respect to observations for precipitation and temperature; however, we are interested in a comparison of skill, or regions where FC or CAM4_OBS skill differs from CAM4_FC. Comparison of FC to CAM4_FC indicates regions where skill increase or decrease is due to ocean-atmosphere coupling differences. Comparison of CAM4_OBS to CAM4_FC indicates regions where skill increase or decrease is due to SST errors.

Figure 2 shows the difference in AC for FC minus CAM4_FC (Figures 2a–2d) and for CAM4_OBS minus CAM4_FC (Figures 2e–2h) for precipitation. All ACs have been converted to approximately normal using Fisher's R-to-Z transformation in order to easily subtract and assign significance levels (significance of difference at 95% confidence level shown); thus, this figure indicates regions in which FC and CAM4_OBS significantly differ from CAM4_FC. While there are slight differences in skill between FC and CAM4_FC (Figures 2a–2d), these differences are rarely significant, even over the midlatitude oceans where we may expect differences in ocean-atmosphere coupling in AGCM versus CGCM experiments to matter greatly

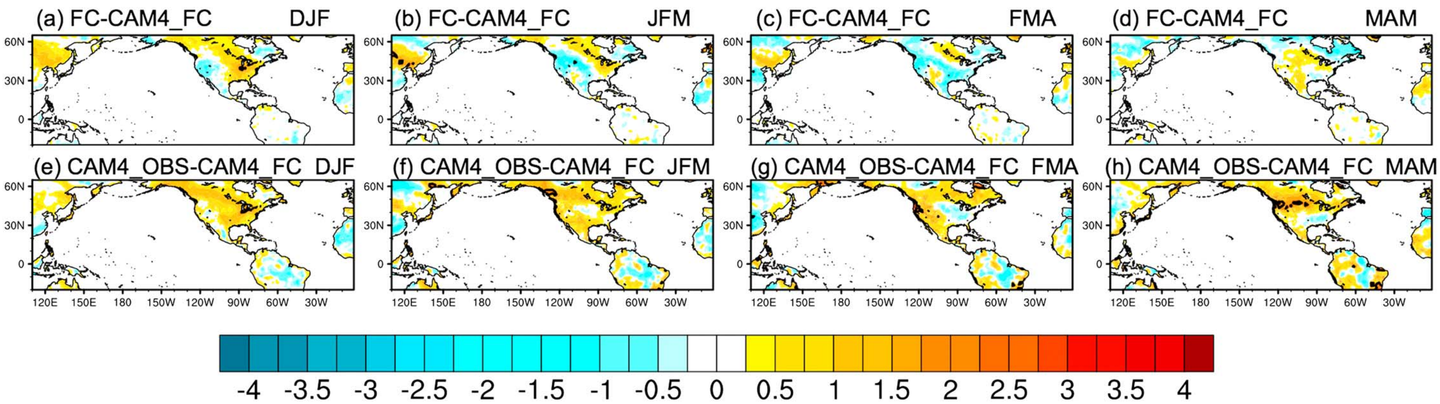

Figure 3. As in Figure 2 but for 2 m temperature.

[*Wu et al.*, 2006; *Wu and Kirtman*, 2007]. However, differences are mainly positive, and we do not find any evidence of ocean-atmosphere coupling causing *significant* skill decrease. Land based 2 m temperature results show similarly insignificant differences, but the differences are larger over land than what is seen for precipitation (Figure 3). These findings suggest that in a realistic prediction setting, ocean-atmosphere coupling inconsistencies between CGCM and AGCM experiments do not have a significant effect on deterministic skill, given the current initialization strategy and the specific model in question. Nevertheless, the influence of ocean-atmosphere coupling is mainly positive.

On the other hand, there is a significant increase in precipitation skill when comparing CAM4_OBS and CAM4_FC that increases with lead time, particularly in the tropical Pacific, tropical Atlantic, and off the coast of the northwestern U.S. in the region of the Aleutian Low (Figures 2e–2h). Land-based 2 m temperature differences are shown in Figures 3e–3h, and we also find increased skill, although it is not significant. Thus, errors in SSTs have a significant negative effect on precipitation skill in the tropical Pacific and negatively impact precipitation and 2 m temperature skill over the midlatitudes/land-based regions, more so at long leads. Large SST biases exist in some regions, for example, near Central America or the eastern coast of South America, see Figure 1, although much of the tropical ocean region shows some bias. Although the bias in the tropical oceans is weaker, biases could cause modulation of associated tropical-extratropical teleconnections, thus impacting skill of remote regions. This relationship between biases and skill is discussed further in section 4.2.

We also consider probabilistic skill comparison measured with RPSS. Typically, RPSS is calculated for three terciles (lower, middle, and upper) and compared to climatology, or equally probable outcomes in each category. To facilitate comparison between experiments, we use CAM4_FC probabilities as a reference in place of climatology. RPSS comparison figures are very similar to the AC difference figures in interpretation. Probabilistic results are consistent with the deterministic assessment; thus, we only show RPSS comparison for precipitation in Figure 4. We again find that there is very little difference between FC and CAM4_FC relative to CAM4_OBS and CAM4_FC, which shows larger differences. This indicates that any energetic inconsistencies in CGCM versus AGCM predictions do not lead to significant difference in deterministic or probabilistic skill, but errors in SSTs contribute significantly to skill decrease, particularly in the tropical Pacific.

3.2. Comparison of Predictability

One might assume that the lack of significant difference in skill between FC and CAM4_FC is due to overall model errors. Predictability, as considered here, is the "ability to predict" a given variable in a perfect model prediction setting [e.g., *Boer*, 2004] and estimates the prediction skill if hindcasts were without systematic error or bias [e.g., *Cheng et al.*, 2011]. Model predictability is an estimate, as models can overestimate predictability relative to nature [*DelSole*, 2004, 2005] but can be used as guidance for expected skill given low or no forecast errors. *Wu and Kirtman* [2007] showed that in an idealized setting, energetic inconsistencies between CGCM and AGCM simulations led to significant differences over midlatitude oceans, suggesting that systematic error may be playing a role in the skill results presented above. As predictability minimizes error, we can determine if model errors are the cause of the similarities between the experiments. CAM4_OBS can be seen

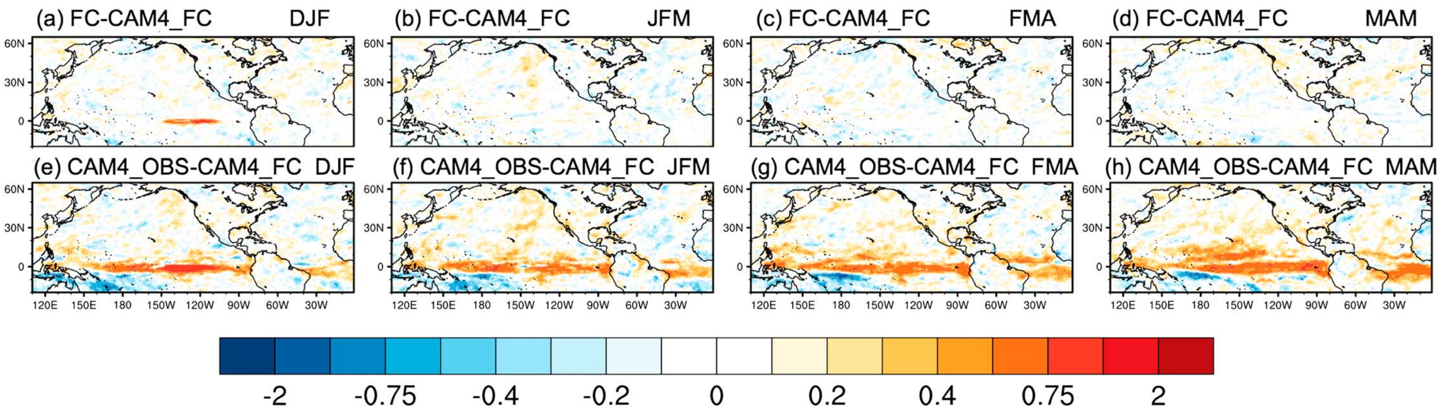

Figure 4. Rank probability skill score (RPSS). (a–d) FC − CAM4_FC December initialized hindcasts predicting DJF − MAM. (e–h) CAM4_OBS − CAM4_FC December initialized hindcasts predicting DJF − MAM. Red (blue) shading indicates regions where FC or CAM4_OBS has stronger (weaker) skill than CAM4_FC. RPSS is calculated with respect to observed (CMAP) precipitation.

as a "potential predictability" experiment given perfectly predicted SSTs, but there may still be model errors related to oceanic fluxes, etc.; thus, comparison is still warranted. Predictability is defined specifically with respect to CCSM4 and estimates may differ for other models.

Assessment of deterministic predictability is similar to Figure 2, but we retain one ensemble member as "truth" in place of observations and the remaining nine ensemble members form the ensemble mean. To avoid sampling errors, we repeat this calculation for each possible combination of one ensemble member and the mean of the remaining nine. The average of all possible combinations of ensemble members is shown, and stippling indicates robustness of predictability increase of the FC or CAM4_OBS simulation versus CAM4_FC (7/10 ensemble combinations in agreement that the difference is positive). Deterministic predictability differences for precipitation are shown in Figure 5 (2 m temperature results not shown).

There are few land-based regions in which FC has increased predictability compared to CAM4_FC (Figures 5a–5d). However, the difference in predictability is generally positive over the oceans, and there is a sizable and robust predictability increase in the western Pacific in the FC predictions compared to CAM4_FC predictions. This region was also highlighted by *Wu and Kirtman* [2007]. The western Pacific is an important source for the PNA pattern reaching North America [e.g., *Lau and Nath*, 1996] which could lead to teleconnection impacts. The equatorial Atlantic also shows some robust predictability increase in the FC simulation, indicating that ocean-atmosphere coupling adds to predictability in the region, likely due to better representation of ocean-atmosphere feedbacks [e.g., *Richter et al.*, 2014]. In turn, due to links to ENSO [e.g.,

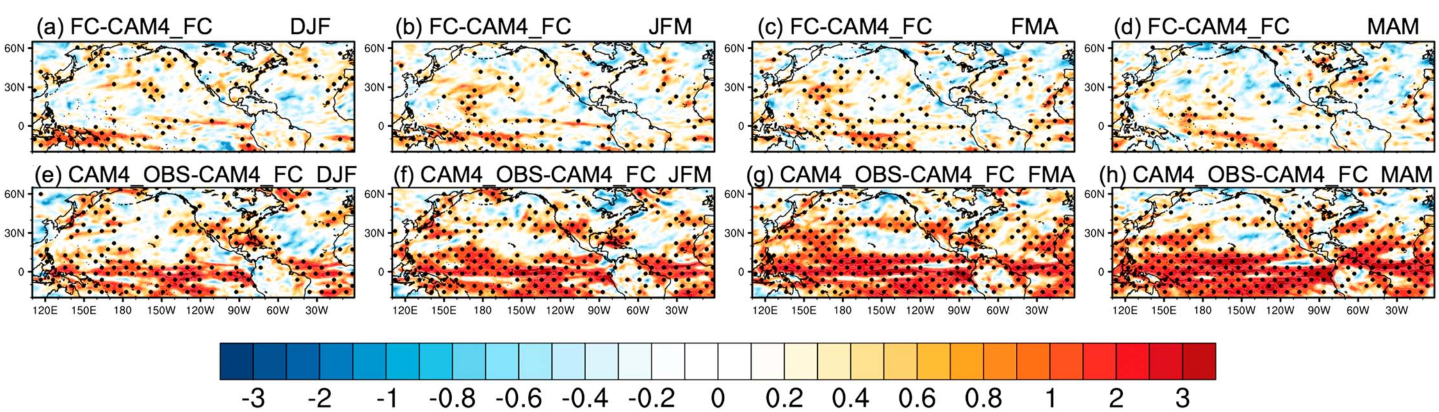

Figure 5. Deterministic precipitation predictability comparison using anomaly correlation of all possible combinations of nine ensemble members predicting the remaining. (a–d) FC − CAM4_FC December initialized hindcasts predicting DJF − MAM. (e–h) CAM4_OBS − CAM4_FC December initialized hindcasts predicting DJF − MAM. Red (blue) shading indicates regions where FC or CAM4_OBS has stronger (weaker) predictability than CAM4_FC. Stippling indicates "robustness" of deterministic predictability, in which 7/10 ensemble combinations are in agreement that the difference is positive.

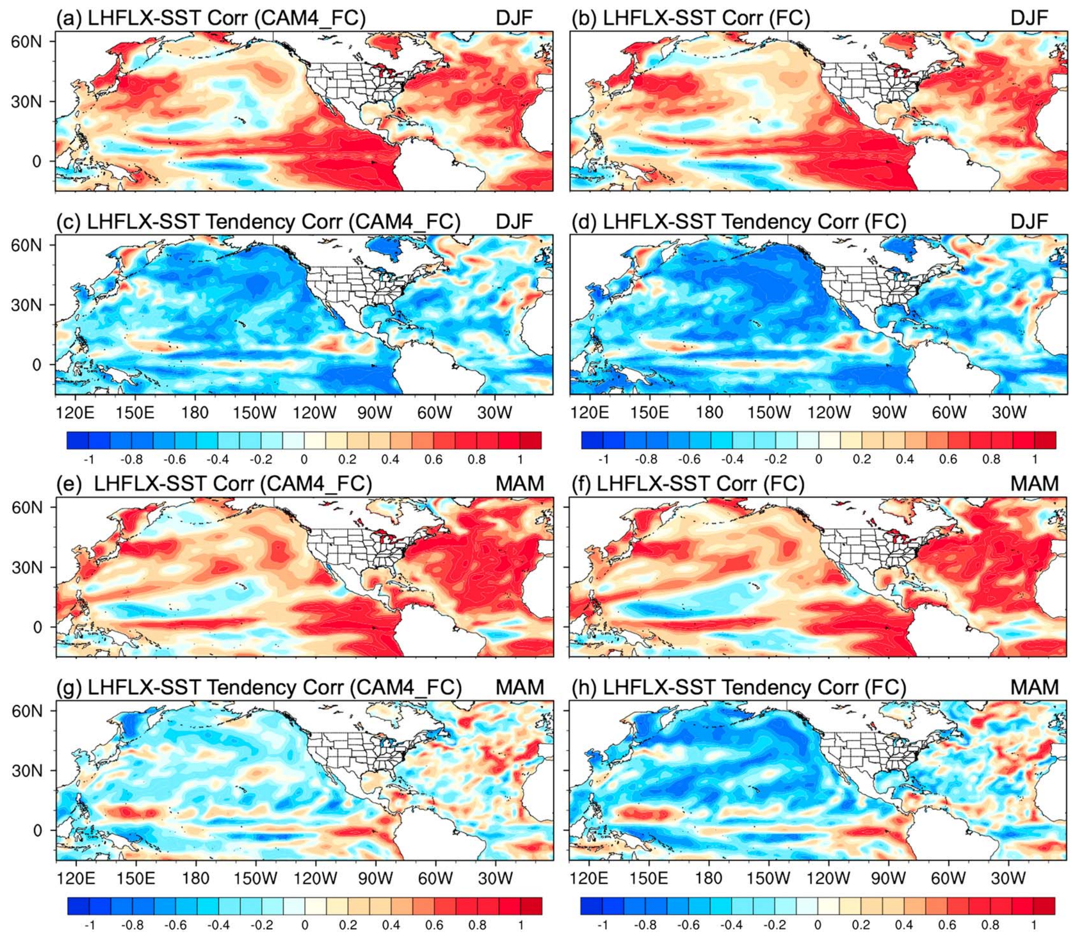

Figure 6. Latent heat flux-SST and latent heat flux-SST tendency correlation for (a–d) DJF and (e–h) MAM. Calculation is based on the ensemble mean. Simultaneous correlation for DJF (MAM) is shown in Figures 6a and 6b (Figures 6e and 6f). Tendency correlation for DJF (MAM) is shown in Figures 6c and 6d (Figures 6g and 6h). CAM4_FC is shown on the left and FC shown on the right.

Handoh et al., 2006], this enhanced Atlantic predictability may have some bearing on teleconnection patterns. Two-meter temperature (not shown) again shows few regions of robust predictability increase when comparing FC to CAM4_FC.

In comparison, there are large and robust differences in predictability when observed SST are used in place of forecasted SSTs (Figures 5e–5h), which are seen over the tropical Pacific, some midlatitude oceanic regions, and in the southern tier of the U.S. This is more pronounced at longer leads as SST errors increase. Two-meter temperature predictability results are similar to prediction skill, though there are some regions of robust predictability increase at longer leads (not shown). We also consider probabilistic assessment using RPSS, computed similarly to AC predictability and retaining one ensemble member as truth. As results are very similar to deterministic predictability, we do not show them here; however, the similarity of deterministic and probabilistic predictability results further supports the above conclusions.

Overall, CGCM versus AGCM prediction strategies do not lead to (comparatively) robust differences in predictability excepting select regions, specifically the western Pacific, for precipitation. However, FC predictability is generally increased over CAM4_FC predictability. As the predictability results minimize model errors, the similarity in skill between FC and CAM4_FC is not due to the presence of model errors. The most pronounced difference in predictability is when observed SSTs are included in the forecast, similar to the skill assessment in section 3.1. The remaining section discusses potential reasons why there are only a small difference in coupled versus uncoupled predictions and a comparatively larger difference when observed SSTs are used.

4. Discussion

In section 3, we found that the largest and most significant skill differences existed for tropical Pacific precipitation when observed versus forecasted SST was prescribed. In the extratropics, skill differences were not significant, but prescribing observed SSTs still led to more skill. Conversely, energetic inconsistencies over the oceans in AGCM versus CGCM hindcasts did not lead to any significant difference in skill. Predictability results are similar, though with some predictability difference in the western Pacific. This section discusses the potential reasons for why there is little difference in skill and predictability for CGCM versus AGCM predictions, and larger differences when accurate SSTs are used.

4.1. CGCM Versus AGCM Predictions

Energetics, more specifically the transfer (or flux) of latent heat into or out of the atmosphere, is an important component of climate simulations as it is the driving force of anomalies. For example, specific to CCSM4, excessive trade winds in the atmospheric model lead to an erroneously large loss in latent heat flux, weak equatorial zonal winds lead to a warm SST bias in the southeast Atlantic, and insufficient low-level clouds lead to erroneously high SSTs, among other issues [*Grodsky et al.*, 2012]. By comparing the rainfall-SST (or evaporation-SST) simultaneous correlation to the rainfall-SST (or evaporation-SST) tendency correlation from CGCM and AGCM simulations, one can determine if ocean forcing of the atmosphere is dominant or if atmosphere forcing of the ocean is dominant [*Wu et al*., 2006; *Wu and Kirtman*, 2007]. The atmosphere responds very quickly to oceanic forcing; thus, if the simultaneous rainfall-SST correlation is large and positive compared to a weak rainfall-SST tendency correlation, ocean forcing of the atmosphere is dominant. Conversely, if the rainfall-SST tendency correlation is strong and negative compared to weak simultaneous correlation, this indicates that atmosphere forcing of ocean anomalies is dominant. The broad conclusions from *Wu et al*. [2006], *Wu and Kirtman* [2007] using idealized CCSM3 simulations are that ocean forcing is dominant in the tropics and atmosphere forcing is dominant in the midlatitudes. We investigate this in CCSM4 predictions.

While rainfall can be used as a proxy for heat transfer in the tropics, this does not specifically hold true for the extratropics. We thus show the latent heat flux-SST simultaneous and tendency correlations for CAM4_FC and FC prediction experiments in Figure 6. Correlations are calculated using the ensemble mean, and as taking the ensemble mean minimizes noise, this represents the *signal* component of the coupling. The mean of all possible simultaneous and tendency correlations for each ensemble member is also computed, shown in Figure 7. This represents the *noise component* of the coupling. Finally, we include results from a fully coupled control run (no initialization) to determine the expected character of coupling outside of a prediction setting (Figure 8). Figure 8 is thus very similar to [*Wu et al*., 2006; *Wu and Kirtman*, 2007], but for CCSM4.

The signal and noise definitions used in this manuscript follow roughly from *Straus and Shukla* [2002]. External variability is the variability of the ensemble mean, corresponding to the predictable component, or signal. Internal variability is the variability about the ensemble mean, which can be due to nonlinear dynamics and is the unpredictable component, or noise. Here we simply define the signal component of the coupling based on the ensemble mean, and the noise component based on calculation for each ensemble member and subsequent averaging. The FC control simulation is a free-running, multidecade simulation completed in CCSM4 that utilizes year 2000 forcing but does not simulate any particular observed period. The simulation was run for approximately 300 years, and we use a sample size of 50 years.

Figure 6 indicates that for the signal component of coupling (i.e., when the coupling is assessed based on the ensemble mean and noise is reduced), the response is very similar in both the FC and the CAM4_FC predictions. For example, when comparing the DJF latent heat flux-SST simultaneous and tendency correlations for CAM4_FC and FC simulations (Figures 6a–6d), the relative sizes of the simultaneous and tendency correlations are similar between the two simulations. This result differs from the control simulation in Figure 8, which clearly shows stronger latent heat flux-SST tendency correlation (Figures 8c and 8d) in midlatitudes compared to the simultaneous correlation (Figures 8a and 8b). There are larger differences when we consider the noise component of the coupling (i.e., when the coupling is assessed for each ensemble member individually and subsequently averaged; thus, the noise is not reduced as in the ensemble mean; Figure 7) and results more closely match the control simulation. For example, consider Figures 7b and 7d, which show the noise component of the coupling for the FC simulation, in contrast to the signal component, the noise component shows the expected larger ocean forcing of atmosphere anomalies in the midlatitudes.

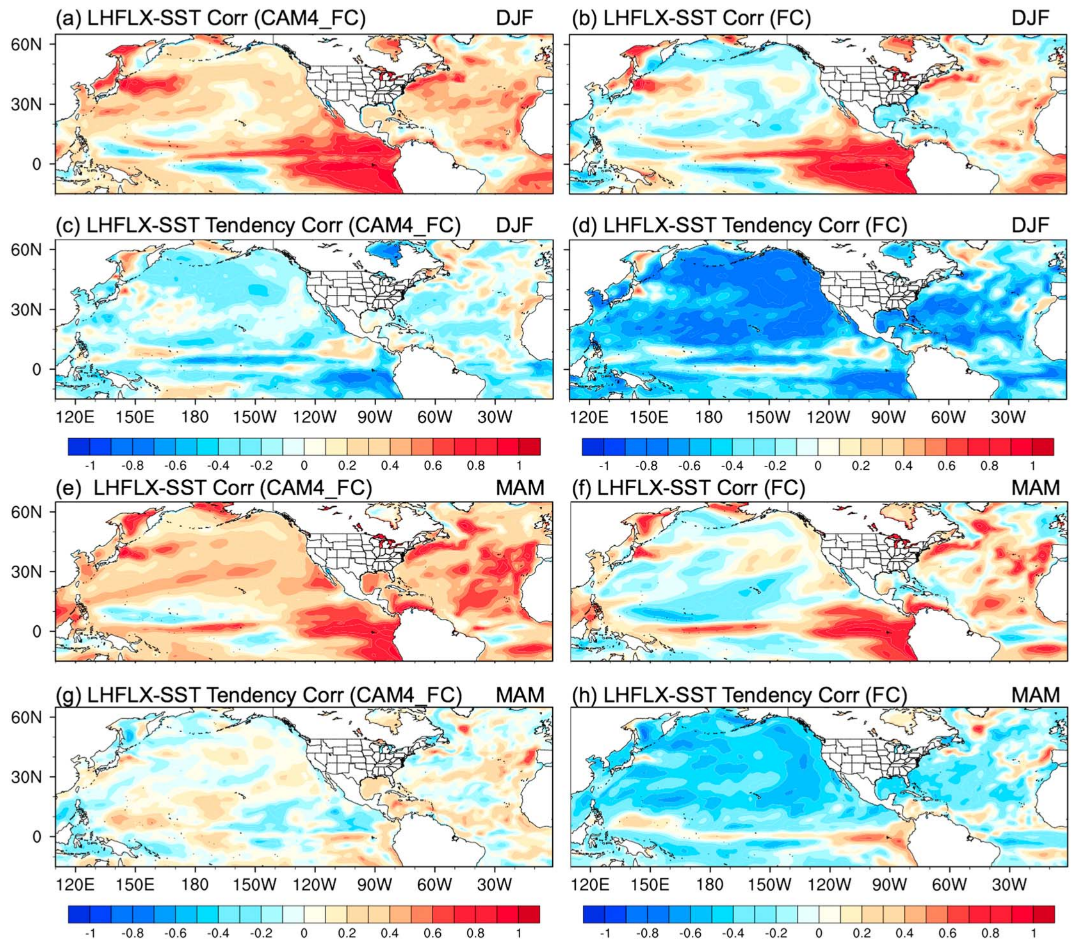

Figure 7. As in Figure 6, but based on computation of these values for each ensemble member individually and subsequent averaging.

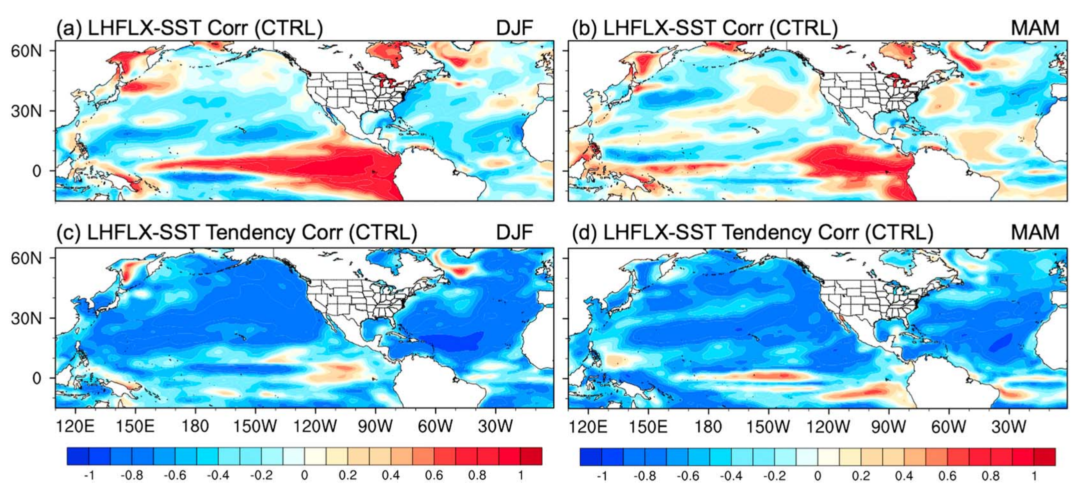

Figure 8. (a and b) Latent heat flux-SST and (c and d) latent heat flux-SST tendency correlation for DJF and MAM for 50 years of FC_CTRL.

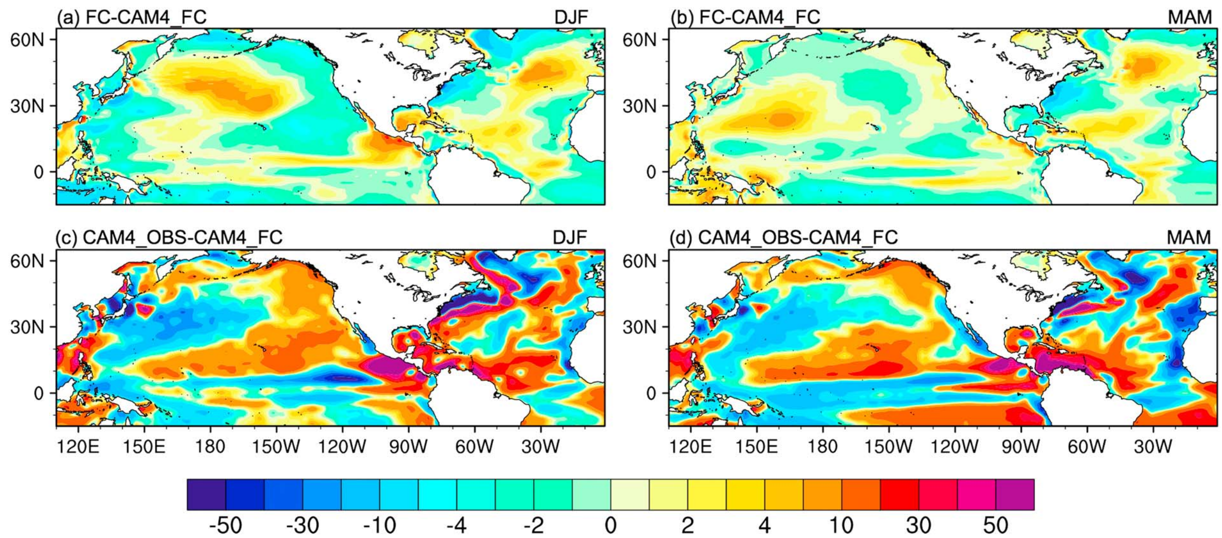

Figure 9. Ensemble mean latent heat flux climatology differences for (a and b) FC minus CAM4_FC (panels a and b) and (c and d) CAM4_OBS minus FC. DJF shown on the left and MAM on the right.

Conversely, the CAM4_FC simulation shows weak ocean forcing of atmosphere anomalies in the midlatitudes (for example, Figures 7a and 7c). However, as expected, all simulations are similar in the tropical Pacific, where there is strong ocean forcing of atmosphere anomalies.

Wu and Kirtman [2005] showed that coupled simulations may have lower atmospheric variability, leading to a reduction in both noise and total variance, compared to uncoupled simulations. These authors explain that in FC experiments, ocean-atmosphere coupling introduces a "damping" affect where SST forced changes also act to reduce SST forcing, in turn weakening atmospheric anomalies and persistence. By design, this damping feedback is suppressed in any prescribed SST experiment, allowing SST-forced atmospheric anomalies to persist for a longer time, enhancing atmospheric variability. Although we find differences in the noise component of the coupling in the CCSM4 predictions, this does not lead to significant differences in skill or predictability (excepting select regions of precipitation predictability, such as the western Pacific).

The tropical Pacific is typically cited as the main contributor to wintertime North American climate variability; however, as noted in the introduction, many references have pointed to some influence of midlatitude SST coupling on winter seasonal timescales. While we find some differences in the noise component of the coupling in the two models in the midlatitudes, this has minimal bearing on the skill or predictability of precipitation and 2 m temperature. As the coupling associated with the climate signal is very similar between the two simulations, we expect that the resulting climate response will be similar as well. Prediction skill and predictability is dominated by coupling associated with the signal, and the differences in noise in the midlatitudes do not significantly contribute to differences in skill or predictability. While not specifically tested here, it is possible that strong tropical Pacific forcing is overwhelming the midlatitude response, although we found similar results for both DJF and MAM. We hope to test this with the addition of more initial months and seasons.

An important distinction in the FC versus CAM4_FC comparison is that we are comparing two predictions with the *same SSTs*, thus the influence of coupling irrespective of any bias between the two SST fields. However, the forecasted SSTs have errors compared to observations. We further assess ocean-atmosphere coupling with the inclusion of SST bias in Figure 9, which shows the ensemble mean systematic latent heat flux difference between FC and CAM4_FC simulations (Figures 9a and 9b) and between CAM4_OBS and CAM4_FC (Figures 9c and 9d) for DJF and MAM. There are weak differences in latent heat flux between the FC and CAM4_FC simulations in the midlatitude Pacific and Atlantic oceans, reflecting some small difference in the character of ocean-atmosphere coupling. However, Figure 9 indicates that should SSTs be more accurate (possibly through better simulation of relevant energetics), this could strongly influence latent heating and the resulting atmospheric fields. Thus, while using the same SSTs but differing ocean-atmosphere

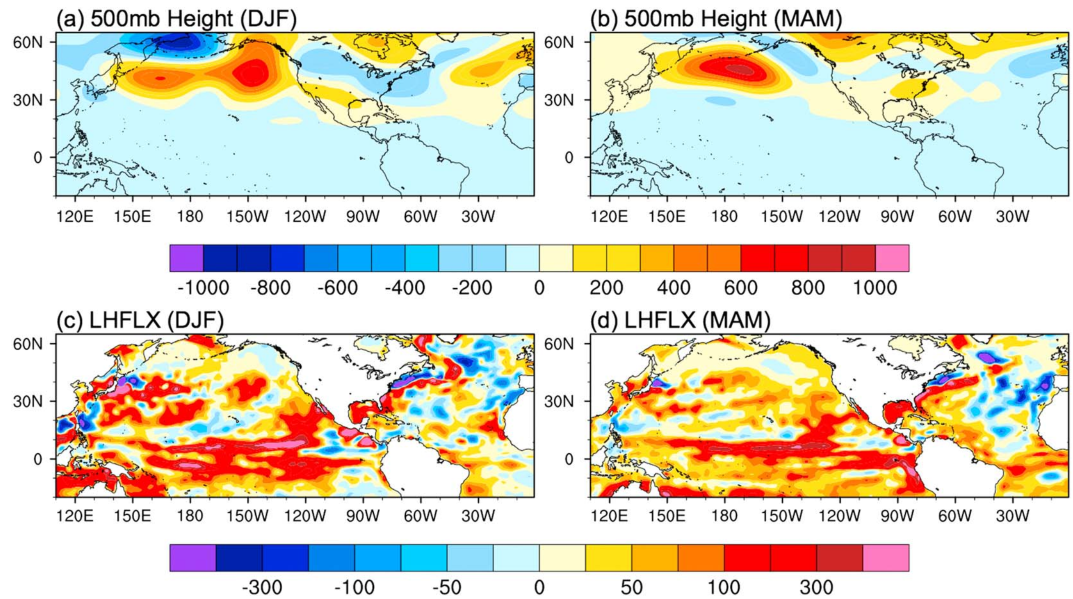

Figure 10. Difference in signal variance for DJF (left) and MAM (right). Difference depicted is CAM4_OBS minus CAM4_FC. Red (blue) shading indicates regions where CAM4_OBS has larger signal variance than CAM4_FC for the given variable. (a and b) The 500 mb height signal variance is shown and (c and d) latent heat flux.

coupling does not lead to a (comparatively) large impact on skill or predictability of precipitation or 2 m temperature, it cannot be discounted as a source of potential predictability *for the SSTs themselves*. This subtlety is not directly tested in this manuscript; however, the remaining section discusses the reasoning why accurate SSTs add to prediction skill and predictability, albeit in an AGCM prediction setting.

4.2. Forecasted Versus Observed SST Predictions

In contrast to CGCM versus AGCM predictions, forecasted versus observed SSTs in predictions show larger skill and predictability differences in some regions, where observed SSTs positively and robustly increase skill and predictability. The use of observed SSTs effectively leads to the assumption that the accurate SSTs cause more accurate atmospheric response(s) and fluxes (though with possible energetic inconsistencies due to lack of coupling) and thus more accurate teleconnections. Also, in a prediction setting, because the SSTs and initial states are in agreement, persistence of initial anomalies is extended which can lead to enhanced predictability and prediction skill [e.g., *Schubert et al.*, 2007]. To further diagnose why differences in skill and predictability are large between CAM4_FC and CAM4_OBS, we focus on differences in latent heat flux climatology and signal variance of atmospheric anomalies.

Large differences in latent heat flux exist between the CAM4_OBS and CAM4_FC simulations, including in the tropical Pacific and tropical Atlantic where we expect forecasted SSTs to be most skillful and accurate (Figures 9c and 9d). As atmospheric anomalies are tied to latent heat fluxes, we expect that dissimilarities in latent heat flux between the two simulations are one of the leading causes in the robust and significant differences between CAM4_OBS and CAM4_FC skill and predictability. For example, we might expect that predicted SSTs that are colder (warmer) than observations will lead to less (more) flux of latent heat than what would be found should observed SSTs be prescribed (or forecasted SSTs be more accurate), leading to differences in skill. Figures 9c and 9d show that CAM4_OBS has more latent heating than CAM4_FC in much of the extra-tropical Pacific and tropical Atlantic and less in the equatorial Pacific. These are similar to the regions in Figure 1 that show biases between observed SSTs and FC SSTs, particularly in the tropical Atlantic.

The tropical Atlantic may influence the tropical Pacific wherein anomalous equatorial Atlantic warming impacts the circulatory pattern stretching from the Atlantic to the Pacific [*Chiang et al.*, 2000; *Rodríguez-Fonseca et al.*, 2009; *Ding et al.*, 2012; *Martín-Rey et al.*, 2014; *Polo et al.*, 2015]. This has important implications for ENSO prediction and predictability, as the Atlantic can influence the amplitude and frequency of ENSO during certain decades, such as when the AMO is negative [*Frauen and Dommenget*, 2012; *Martín-Rey* *et al.*, 2014]. Moreover, correcting systematic biases in Atlantic SSTs can improve ENSO forecasts due to remote forcing [*Keenlyside et al.*, 2013]. As climate models have biases in reproducing tropical Atlantic climate, the large biases seen here are likely impacting the tropical Pacific as well [*Sasaki et al.*, 2014; *Martín-Rey et al.*, 2015], which in turn could impact teleconnections. Our assumption is that using observed SSTs will allow for the best possible simulation of fluxes; however, energetic inconsistencies between CGCM and AGCM simulations, while very weak and related only to noise, still exist within the predictions.

Finally, Figure 10 shows the difference in DJF and MAM 500 mb geopotential height and latent heat flux unbiased signal variance for CAM4_OBS minus CAM4_FC. Unbiased signal variance is calculated following methodology from *Rowell* [1998] and *Schubert et al.* [2002]. The figure depicts regions in which CAM4_OBS signal variance is larger than (red shading) or smaller than (blue shading) CAM4_FC signal variance. Here signal variance represents the "slowly varying" response to sea surface temperatures, etc. Simply stated, if signal variance is larger in CAM4_OBS, there is more potential predictability of the given variable. Signal variance is larger in CAM4_OBS for midlatitude 500 mb heights and overall for latent heat flux. Signal variance is mildly weaker for 500 mb heights outside of the midlatitudes. We also find enhanced signal variance of latent heat flux over much of the depicted region, indicating that latent heat flux is more predictable in CAM4_OBS overall. The enhanced signal variance is the likely contributor to increased skill and predictability, as the dynamical quantities impact precipitation and temperature patterns.

We have also calculated estimates of the total variance for SST (not shown) to determine if the increase in signal variance is due to an overall increase in CAM4_OBS variance (i.e., that both signal and noise variance are increased). For SST in DJF, the total variance is very similar for both CAM4_OBS and CAM4_FC, with very few regions where the variance is significantly different. The total variance of CAM4_OBS and CAM4_FC differs more significantly as lead-time increases, though this is regional, and CAM4_OBS variance is significantly larger mainly in the tropical Pacific. Thus, the CAM4_OBS total SST variance is not (significantly) modified in the forecasts at short leads, but there is some regional modification at longer leads. By design, use of the same SST field in each ensemble member could cause larger (weaker) signal (noise) variance in CAM4_OBS. However, we believe that the increase in signal variance in regions outside the tropics seen in Figure 10 is due to more consistent, and possibly better, representation of surface heat fluxes in CAM4_OBS, which adds to predictability of these variables. As the signal variance is (mainly) larger in CAM4_OBS versus CAM4_FC, this supports our above conclusions using traditional skill and predictability assessments that use of observed SSTs causes increased skill and predictability, as the SST errors act to cause errors in dynamical quantities such as latent heat flux and 500 mb heights.

5. Conclusions

We provide a comparison of fully coupled predictions versus prescribed SST predictions and of forecasted versus observed SSTs. Our intent is to determine the relative importance of ocean-atmosphere coupling and SST errors for prediction skill and predictability of precipitation and 2 m temperature. Prediction skill and predictability are examined through deterministic (anomaly correlation) and probabilistic (RPSS) methods for precipitation and 2 m temperature. We identify three main conclusions from this work related to prediction skill and predictability of 2 m temperature and precipitation in CCSM4:

1. Prediction skill (predictability) is not significantly influenced (weakly influenced) by ocean-atmosphere coupling when the same SSTs are used, except for the western Pacific.
2. Prediction skill and predictability are significantly and robustly influenced by errors in SSTs when comparing simulations with forecasted versus observed SSTs.
3. Comparatively, errors in SSTs lead to more significant and robust differences in prediction skill and predictability versus inconsistencies in ocean-atmosphere coupling.

These results for CCSM4 largely agree with work in other prediction and idealized settings using various models [for example, *Kumar et al.*, 2008]. However, it was noted in *Wu and Kirtman* [2007] that the performance of forced (AGCM) model simulations depends on the local atmosphere-ocean interaction and that poor performance is seen when atmospheric forcing of SSTs is dominant, such as in the midlatitudes or in the western Indo-Pacific region. In a prediction setting, as shown here, the character of ocean-atmosphere coupling is very similar in forced and coupled predictions for the part of the coupling associated with the signal, which is the dominant factor in the predictions. The character differs for the part of the coupling associated with noise, which is a comparatively weak contributor to skill and predictability. In either case, differences in coupling do not impact skill and weakly impact predictability in select regions, such as the western Pacific and tropical Atlantic. This indicates that a two-tiered forecast system, in which SSTs are predicted by the fully coupled model and subsequently prescribed in an AGCM setting, is a possibility for CCSM4, at least for DJF − MAM. The benefits of using prescribed SSTs rather than FC predictions include potential for additional ensemble members, experiments using a higher resolution atmosphere model with less computational demand than a fully coupled hindcast, bias correction of prescribed SSTs, use of enhanced versions of the atmosphere model such as CAM5, or a combination of these.

Conversely, we do find large and significant differences in skill when using observed (thus perfectly predicted) SSTs rather than forecasted SSTs for both prediction and predictability, particularly for precipitation. There are biases in SSTs in various regions, including the tropical Pacific, and while we see a significant increase in skill and a general increase overall when observed SSTs are used, this does not necessarily match spatially with the regions of largest bias. However, the reduction of bias when considering observed SSTs may lead to more realistic teleconnection patterns and thus enhanced prediction skill. The assumption here is that using observed SSTs would also allow for the best approximation of appropriate surface fluxes. *Saravanan* [1998] notes that this assumption can be problematic as benefits are mainly related to specifying tropical SSTs; however, we find an overall positive increase when observed SSTs are used. While we did not find a strong difference in *ensemble mean* ocean-atmosphere coupling in midlatitudes and found only subtle differences in latent heating when the same SST field is used, we presume that coupling in midlatitudes could lead to more accurate predicted SSTs, thus more accurate midlatitude forcing, etc. Our methodology to test any skill or predictability due to differences in SSTs is to prescribe them, and thus the simulations do not have fully coupled energetics. Nevertheless, we find that there are large differences in signal variance of oceanic latent heat flux and of large-scale height fields when considering predicted versus observed SSTs, where many regions show larger signal variance when observed SSTs are used. We also find differences in latent heat flux climatologies, which likely contribute to errors and lack of skillfulness in the forecasts, as SST errors can cause errors in latent heat fluxes.

There are some caveats to these results. We focus on the initialization and season(s) that traditionally have large predictability and skill, not necessarily the warmest ocean, which could affect the results. Midlatitude forcing has been cited as being a possible contributor to land-based precipitation and 2 m temperature, but the tropical Pacific forcing may be overwhelming any weak positive or negative contribution to skill or predictability from the midlatitudes. Because we have focused on the initialization and season(s) with larger predictability and skill, the nonlinear component of the coupling is small compared to if we were to consider the spring season, and any shortwave nonlinearities may impact the results in spring [e.g., *Bellenger et al.*, 2014]. Ocean model resolution can effect air-sea feedbacks due to differences in coupling between SSTs and latent heat flux in lower and higher resolution CGCM experiments [*Kirtman et al.*, 2012]. In addition to ocean model resolution, a high-resolution atmosphere model (CAM5) was able to capture a robust local atmospheric response that was dominated by changes in eddy heat and moisture transports over the Oyashio Extension SST front [*Smirnov et al.*, 2014]. In the tropical Atlantic, increased ocean model resolution that is able to resolve oceanic mesoscale variability can lead to smaller biases in the equatorial cold tongue [*Seo et al.*, 2006]. Atmospheric biases can be transmitted to the ocean and increase biases there, mainly due to incorrect representation of winds [*Voldoire et al.*, 2014], and increased horizontal and vertical resolution of an atmospheric model can also lead to substantial improvements in SST biases [*Harlaß et al.*, 2015]. Finally, we note that SST errors influence skill by comparing forecasted to observed SSTs. An alternative to this approach is to use the patterns of SST bias as the prescribed SST field in the AGCM simulations, which may aid in determining the impacts to heat flux, etc. This could be an interesting direction for future research.

This analysis shows that prescribed SST predictions in CCSM4 are a practical choice for targeted studies for winter initialized predictions for DJF − MAM. However, we still caution the use of this framework as the results may not necessarily be generalized to other seasons, initializations, or models. We also find that SSTs that are more accurate lead to a comparatively large increase in skill and predictability for the predictions, mainly for precipitation. Though not performed here, bias correction of SSTs may be a potential future direction leading to more skillful forecasts, keeping in mind the caveat that a coupled system with more accurate SSTs may still be the most skillful representation of climate variability.

Acknowledgments

The authors would like to thank Dughong Min and the Center for Ocean-Land-Atmosphere studies for creating and maintaining the initial states for CCSM4 climate prediction studies. The free software NCAR Command Language was used to create the plots and analyze data. The authors would also like to thank the NMME program partners. NMME data, including CCSM4 fully coupled hindcasts, can be found at http://www.cpc.ncep.noaa.gov/products/NMME/. CCSM4 model code is available from NCAR at http://www.cesm.ucar.edu/models/ccsm4.0/. The prescribed SST CAM4 simulation data are archived at the Center for Computational Science at the University of Miami and are available upon request. We would like to thank two anonymous reviewers whose comments greatly improved this manuscript. The authors acknowledge support from NOAA NA10OAR4320143, NA15OAR4320064, NA16OAR4310149, and NA16OAR4310141.

References

Barsugli, J. J., and D. S. Battisti (1998), The basic effects of atmosphere–ocean thermal coupling on midlatitude variability, *J. Atmos. Sci.*, *55*, 477–493, doi:10.1175/1520-0469(1998)055<0477:TBEOAO>2.0.CO;2.

Becker, E., H. van den Dool, and Q. Zhang (2014), Predictability and forecast skill in NMME, *J. Clim.*, *27*, 5891–5906, doi:10.1175/JCLI-D-13-00597.1.

Bellenger, H., E. Guilyardi, J. Leloup, and J. Vialard (2014), ENSO representation in climate models: From CMIP3 to CMIP5, *Clim. Dyn.*, *42*, 1999–2018, doi:10.1007/s00382-013-1783-z.

Bengtsson, L., U. Schlese, E. Roeckner, M. Latif, T. P. Barnett, and N. Graham (1993), A two-tiered approach to long-range climate forecasting, *Science*, *261*, 1026–1029, doi:10.1126/science.261.5124.1026.

Boer, G. J. (2004), Long time-scale potential predictability in an ensemble of coupled climate models, *Clim. Dyn.*, *23*, 29–44, doi:10.1007/s00382-004-0419-8.

Chen, H., E. K. Schneider, B. P. Kirtman, and I. Colfescu (2012), Evaluation of weather noise and its role in climate model simulations, *J. Clim.*, *26*, 3766–3784, doi:10.1175/JCLI-D-12-00292.1.

Cheng, Y., Y. Tang, and D. Chen (2011), Relationship between predictability and forecast skill of ENSO on various time scales, *J. Geophys. Res.*, *116*, C12006, doi:10.1029/2011JC007249.

Chiang, J. C. H., Y. Kushnir, and S. E. Zebiak (2000), Interdecadal changes in eastern Pacific ITCZ variability and its influence on the Atlantic ITCZ, *Geophys. Res. Lett.*, *27*, 3687–3690, doi:10.1029/1999GL011268.

de Szoeke, S. P., and S.-P. Xie (2008), The tropical eastern Pacific seasonal cycle: Assessment of errors and mechanisms in IPCC AR4 coupled ocean–atmosphere general circulation models, *J. Clim.*, *21*, 2573–2590, doi:10.1175/2007JCLI1975.1.

DelSole, T. (2004), Predictability and information theory. Part I: Measures of predictability, *J. Atmos. Sci.*, *61*, 2425–2440, doi:10.1175/1520-0469(2004)061<2425:PAITPI>2.0.CO;2.

DelSole, T. (2005), Predictability and information theory. Part II: Imperfect forecasts, *J. Atmos. Sci.*, *62*, 3368–3381, doi:10.1175/JAS3522.1.

DelSole, T., and M. K. Tippett (2014), Comparing forecast skill, *Mon. Weather Rev.*, *142*, 4658–4678, doi:10.1175/MWR-D-14-00045.1.

Ding, H., N. S. Keenlyside, and M. Latif (2012), Impact of the equatorial Atlantic on the El Niño Southern Oscillation, *Clim. Dyn.*, *38*, 1965–1972, doi:10.1007/s00382-011-1097-y.

Fan, Y., and H. van den Dool (2008), A global monthly land surface air temperature analysis for 1948–present, *J. Geophys. Res.*, *113*, D01103, doi:10.1029/2007JD008470.

Frankignoul, C. (1985), Sea surface temperature anomalies, planetary waves, and air-sea feedback in the middle latitudes, *Rev. Geophys.*, *23*, 357–390, doi:10.1029/RG023i004p00357.

Frauen, C., and D. Dommenget (2012), Influences of the tropical Indian and Atlantic Oceans on the predictability of ENSO, *Geophys. Res. Lett.*, *39*, L02706, doi:10.1029/2011GL050520.

Gent, P. R., et al. (2011), The Community Climate System Model version 4, *J. Clim.*, *24*, 4973–4991, doi:10.1175/2011JCLI4083.1.

Goddard, L., and S. Mason (2002), Sensitivity of seasonal climate forecasts to persisted SST anomalies, *Clim. Dyn.*, *19*, 619–632, doi:10.1007/s00382-002-0251-y.

Graham, R. J., A. D. L. Evans, K. R. Mylne, M. S. J. Harrison, and K. B. Robertson (2000), An assessment of seasonal predictability using atmospheric general circulation models, *Q. J. R. Meteorol. Soc.*, *126*, 2211–2240, doi:10.1002/qj.49712656712.

Grodsky, S. A., J. A. Carton, S. Nigam, and Y. M. Okumura (2012), Tropical Atlantic biases in CCSM4, *J. Clim.*, *25*, 3684–3701, doi:10.1175/JCLI-D-11-00315.1.

Handoh, I. C., A. J. Matthews, G. R. Bigg, and D. P. Stevens (2006), Interannual variability of the tropical Atlantic independent of and associated with ENSO: Part I. The north tropical Atlantic, *Int. J. Climatol.*, *26*, 1937–1956, doi:10.1002/joc.1343.

Harlaß, J., M. Latif, and W. Park (2015), Improving climate model simulation of tropical Atlantic sea surface temperature: The importance of enhanced vertical atmosphere model resolution, *Geophys. Res. Lett.*, *42*(7), 2401–2408, doi:10.1002/2015GL063310.

Infanti, J. M., and B. P. Kirtman (2013), Southeastern U.S. rainfall prediction in the North American Multi-Model Ensemble, *J. Hydrometeorol.*, *15*, 529–550, doi:10.1175/JHM-D-13-072.1.

Infanti, J. M., and B. P. Kirtman (2015), North American rainfall and temperature prediction response to the diversity of ENSO, *Clim. Dyn.*, 1–17, doi:10.1007/s00382-015-2749-0.

Infanti J. M., and B. P. Kirtman (2016), Prediction and predictability of land and atmosphere initialized CCSM4 climate forecasts over North America, *J. Geophys. Res. Atmos.*, *121*(21), 12,690–12,701, doi:10.1002/2016JD024932.

Keenlyside, N. S., H. Ding, and M. Latif (2013), Potential of equatorial Atlantic variability to enhance El Niño prediction, *Geophys. Res. Lett.*, *40*, 2278–2283, doi:10.1002/grl.50362.

Kirtman, B., et al. (2012), Impact of ocean model resolution on CCSM climate simulations, *Clim. Dyn.*, *39*, 1303–1328, doi:10.1007/s00382-012-1500-3.

Kirtman, B. P. (2003), The COLA anomaly coupled model: Ensemble ENSO prediction, *Mon. Weather Rev.*, *131*, 2324–2341, doi:10.1175/1520-0493(2003)131<2324:TCACME>2.0.CO;2.

Kirtman, B. P., and D. Min (2009), Multimodel ensemble ENSO prediction with CCSM and CFS, *Mon. Weather Rev.*, *137*, 2908–2930, doi:10.1175/2009MWR2672.1.

Kirtman, B. P., et al. (2014), The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction: phase-2 toward developing intraseasonal prediction, *Bull. Am. Meteorol. Soc.*, *95*, 585–601, doi:10.1175/BAMS-D-12-00050.1.

Koster, R. D., and M. J. Suarez (2003), Impact of land surface initialization on seasonal precipitation and temperature prediction, *J. Hydrometeorol.*, *4*, 408–423, doi:10.1175/1525-7541(2003)4<408:IOLSIO>2.0.CO;2.

Kumar, A., Q. Zhang, J.-K. E. Schemm, M. L'Heureux, and K.-H. Seo (2008), An assessment of errors in the simulation of atmospheric interannual variability in uncoupled AGCM simulations, *J. Clim.*, *21*, 2204–2217, doi:10.1175/2007JCLI1743.1.

Lau, N.-C., and M. J. Nath (1996), The role of the "atmospheric bridge" in linking tropical Pacific ENSO events to extratropical SST anomalies, *J. Clim.*, *9*, 2036–2057, doi:10.1175/1520-0442(1996)009<2036:TROTBI>2.0.CO;2.

Li, H., and V. Misra (2014), Global seasonal climate predictability in a two tiered forecast system. Part II: Boreal winter and spring seasons, *Clim. Dyn.*, *42*, 1449–1468, doi:10.1007/s00382-013-1813-x.

Li, G., and S.-P. Xie (2013), Tropical biases in CMIP5 multimodel ensemble: The excessive equatorial Pacific cold tongue and double ITCZ problems, *J. Clim.*, *27*, 1765–1780, doi:10.1175/JCLI-D-13-00337.1.

Livezey, R. E., and M. M. Timofeyeva (2008), The first decade of long-lead U.S. seasonal forecasts, *Bull. Am. Meteorol. Soc.*, *89*, 843–854, doi:10.1175/2008BAMS2488.1.

Ma, F., A. Ye, X. Deng, Z. Zhou, X. Liu, Q. Duan, J. Xu, C. Miao, Z. Di, and W. Gong (2016), Evaluating the skill of NMME seasonal precipitation ensemble predictions for 17 hydroclimatic regions in continental China, *Int. J. Climatol.*, *36*, 132–144, doi:10.1002/joc.4333.

Martín-Rey, M., B. Rodríguez-Fonseca, I. Polo, and F. Kucharski (2014), On the Atlantic–Pacific Niños connection: A multidecadal modulated mode, *Clim. Dyn.*, *43*, 3163–3178.

Martín-Rey, M., B. Rodríguez-Fonseca, and I. Polo (2015), Atlantic opportunities for ENSO prediction, *Geophys. Res. Lett.*, *42*, 6802–6810, doi:10.1002/2015GL065062.

Mason, S. J. (2004), On using "climatology" as a reference strategy in the Brier and ranked probability skill scores, *Mon. Weather Rev.*, *132*, 1891–1895, doi:10.1175/1520-0493(2004)132<1891:OUCAAR>2.0.CO;2.

Mechoso, C. R., et al. (1995), The seasonal cycle over the tropical Pacific in coupled ocean–atmosphere general circulation models, *Mon. Weather Rev.*, *123*, 2825–2838, doi:10.1175/1520-0493(1995)123<2825:TSCOTT>2.0.CO;2.

Misra, V., H. Li, Z. Wu, and S. DiNapoli (2014), Global seasonal climate predictability in a two tiered forecast system: Part I: Boreal summer and fall seasons, *Clim. Dyn.*, *42*, 1425–1448, doi:10.1007/s00382-013-1812-y.

Mo, K. C., and J. E. Schemm (2008a), Droughts and persistent wet spells over the United States and Mexico, *J. Clim.*, *21*, 980–994, doi:10.1175/2007JCLI1616.1.

Mo, K. C., and J. E. Schemm (2008b), Relationships between ENSO and drought over the southeastern United States, *Geophys. Res. Lett.*, *35*, L15701, doi:10.1029/2008GL034656.

Palmer, T. N., and S. Zhaobo (1985), A modelling and observational study of the relationship between sea surface temperature in the north-west Atlantic and the atmospheric general circulation, *Q. J. R. Meteorol. Soc.*, *111*, 947–975, doi:10.1002/qj.49711147003.

Paolino, D. A., J. L. Kinter, B. P. Kirtman, D. Min, and D. M. Straus (2011), The impact of land surface and atmospheric initialization on seasonal forecasts with CCSM, *J. Clim.*, *25*, 1007–1021, doi:10.1175/2011JCLI3934.1.

Polo, I., M. Martin-Rey, B. Rodriguez-Fonseca, F. Kucharski, and C. R. Mechoso (2015), Processes in the Pacific La Niña onset triggered by the Atlantic Niño, *Clim. Dyn.*, *44*, 115–131, doi:10.1007/s00382-014-2354-7.

Rayner, N., D. Parker, E. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, *108*(D14), 4407, doi:10.1029/2002JD002670.

Richter, I., S.-P. Xie, S. K. Behera, T. Doi, and Y. Masumoto (2014), Equatorial Atlantic variability and its relation to mean state biases in CMIP5, *Clim. Dyn.*, *42*, 171–188, doi:10.1007/s00382-012-1624-5.

Rodríguez-Fonseca, B., I. Polo, J. García-Serrano, T. Losada, E. Mohino, C. R. Mechoso, and F. Kucharski (2009), Are Atlantic Niños enhancing Pacific ENSO events in recent decades?, *Geophys. Res. Lett.*, *36*, L20705, doi:10.1029/2009GL040048.

Ropelewski, C. F., and M. S. Halpert (1986), North American precipitation and temperature patterns associated with the El Niño/Southern Oscillation (ENSO), *Mon. Weather Rev.*, *114*, 2352–2362, doi:10.1175/1520-0493(1986)114<2352:NAPATP>2.0.CO;2.

Ropelewski, C. F., and M. S. Halpert (1987), Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation, *Mon. Weather Rev.*, *115*, 1606–1626, doi:10.1175/1520-0493(1987)115<1606:GARSPP>2.0.CO;2.

Rowell, D. P. (1998), Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations, *J. Clim.*, *11*, 109–120, doi:10.1175/1520-0442(1998)011<0109:APSPWA>2.0.CO;2.

Saha, S., et al. (2010), The NCEP climate forecast system reanalysis, *Bull. Am. Meteorol. Soc.*, *91*, 1015–1057, doi:10.1175/2010BAMS3001.1.

Saravanan, R. (1998), Atmospheric low-frequency variability and its relationship to midlatitude SST variability: Studies using the NCAR climate system model, *J. Clim.*, *11*, 1386–1404, doi:10.1175/1520-0442(1998)011<1386:ALFVAI>2.0.CO;2.

Sasaki, W., T. Doi, K. J. Richards, and Y. Masumoto (2014), Impact of the equatorial Atlantic sea surface temperature on the tropical Pacific in a CGCM, *Clim. Dyn.*, *43*, 2539–2552, doi:10.1007/s00382-014-2072-1.

Schubert, S. D., M. J. Suarez, P. J. Pegion, M. A. Kistler, and A. Kumar (2002), Predictability of zonal means during boreal summer, *J. Clim.*, *15*, 420–434, doi:10.1175/1520-0442(2002)015<0420:POZMDB>2.0.CO;2.

Schubert, S., R. Koster, M. Hoerling, R. Seager, D. Lettenmaier, A. Kumar, and D. Gutzler (2007), Predicting drought on seasonal-to-decadal time scales, *Bull. Am. Meteorol. Soc.*, *88*, 1625–1630, doi:10.1175/BAMS-88-10-1625.

Seo, H., M. Jochum, R. Murtugudde, and A. J. Miller (2006), Effect of ocean mesoscale variability on the mean state of tropical Atlantic climate, *Geophys. Res. Lett.*, *33*(9), L09606, doi:10.1029/2005GL025651.

Shin, S.-I., and P. D. Sardeshmukh (2011), Critical influence of the pattern of tropical ocean warming on remote climate trends, *Clim. Dyn.*, *36*, 1577–1591, doi:10.1007/s00382-009-0732-3.

Shukla, J. (1998), Predictability in the midst of chaos: A scientific basis for climate forecasting, *Science*, *282*, 728–731, doi:10.1126/science.282.5389.728.

Shukla, S., J. Roberts, A. Hoell, C. C. Funk, F. Robertson, and B. Kirtman (2016), Assessing North American Multimodel Ensemble (NMME) seasonal forecast skill to assist in the early warning of anomalous hydrometeorological events over East Africa, *Clim. Dyn.*, 1–17, doi:10.1007/s00382-016-3296-z.

Smirnov, D., M. Newman, M. A. Alexander, Y.-O. Kwon, and C. Frankignoul (2014), Investigating the local atmospheric response to a realistic shift in the Oyashio Sea surface temperature front, *J. Clim.*, *28*, 1126–1147, doi:10.1175/JCLI-D-14-00285.1.

Straus, D. M., and J. Shukla (2002), Does ENSO force the PNA?, *J. Clim.*, *15*, 2340–2358, doi:10.1175/1520-0442(2002)015<2340:DEFTP>2.0.CO;2.

Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski (1998), Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures, *J. Geophys. Res.*, *103*, 14,291–14,324, doi:10.1029/97JC01444.

Voldoire, A., M. Claudon, G. Caniaux, H. Giordani, and R. Roehrig (2014), Are atmospheric biases responsible for the tropical Atlantic SST biases in the CNRM-CM5 coupled model?, *Clim. Dyn.*, *43*, 2963–2984, doi:10.1007/s00382-013-2036-x.

Wang, C., L. Zhang, S.-K. Lee, and C. R. Mechoso (2014), A global perspective on CMIP5 climate model biases, *Nat. Clim. Change*, *4*, 201–205.

Weigel, A. P., M. A. Liniger, and C. Appenzeller (2007), The discrete Brier and ranked probability skill scores, *Mon. Weather Rev.*, *135*, 118–124, doi:10.1175/MWR3280.1.

Wilks, D. S. (2011), *Statistical Methods in the Atmospheric Sciences*, 3rd ed., pp. 301–394, Elsevier/Academic Press, Amsterdam, Boston.

Wu, R., and B. P. Kirtman (2004), Impacts of the Indian ocean on the Indian summer monsoon–ENSO relationship, *J. Clim.*, *17*, 3037–3054, doi:10.1175/1520-0442(2004)017<3037:IOTIOO>2.0.CO;2.

Wu, R., and B. Kirtman (2005), Roles of Indian and Pacific ocean air–sea coupling in tropical atmospheric variability, *Clim. Dyn.*, *25*, 155–170, doi:10.1007/s00382-005-0003-x.

Wu, R., and B. Kirtman (2007), Regimes of seasonal air–sea interaction and implications for performance of forced simulations, *Clim. Dyn.*, *29*, 393–410, doi:10.1007/s00382-007-0246-9.

Wu, R., B. P. Kirtman, and K. Pegion (2006), Local air–sea relationship in observations and model simulations, *J. Clim.*, *19*, 4914–4932, doi:10.1175/JCLI3904.1.

Xie, P., and P. A. Arkin (1997), Global precipitation: A 17-year monthly analysis based on gauge estimates, and numerical model outputs, *Bull. Am. Meteorol. Soc.*, *78*, 2539–2558, doi:10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2.

Yu, J.-Y., and C. R. Mechoso (1999), A discussion on the errors in the surface heat fluxes simulated by a coupled GCM, *J. Clim.*, *12*, 416–426, doi:10.1175/1520-0442(1999)012<0416:ADOTEI>2.0.CO;2.

Zhou, S., M. L'Heureux, S. Weaver, and A. Kumar (2012), A composite study of the MJO influence on the surface air temperature and precipitation over the continental United States, *Clim. Dyn.*, *38*, 1459–1471, doi:10.1007/s00382-011-1001-9.