

Sea Surface Temperature Predictions in NCEP CFSv2 Using a Simple Ocean Initialization Scheme

JIESHUN ZHU

Climate Prediction Center, NOAA/NWS/NCEP, College Park, and Innovim, Greenbelt, Maryland

ARUN KUMAR

Climate Prediction Center, NOAA/NWS/NCEP, College Park, Maryland

HUI WANG

Climate Prediction Center, NOAA/NWS/NCEP, College Park, and Innovim, Greenbelt, Maryland

BOHUA HUANG

Department of Atmospheric, Oceanic, and Earth Sciences, College of Science, and Center for Ocean–Land–Atmosphere Studies, George Mason University, Fairfax, Virginia

(Manuscript received 10 September 2014, in final form 18 April 2015)

ABSTRACT

In contrast to operational climate predictions based on sophisticated ocean data assimilation schemes at the National Centers for Environmental Predictions (NCEP), this study applied a simple ocean initialization scheme to the NCEP latest seasonal prediction model, the Climate Forecast System, version 2 (CFSv2). In the scheme, sea surface temperature (SST) was the only observed information applied to derive ocean initial states. The physical basis for the method is that, through air–sea coupling, SST is capable of reproducing some observed features of ocean evolutions by forcing the atmospheric winds. SST predictions based on the scheme are compared against hindcasts from the National (lately North American) Multimodel Ensemble (NMME) project.

It was found that due to substantial biases in the tropical eastern Pacific in the ocean initial conditions produced by SST assimilation, ENSO SST predictions were not as good as those with sophisticated initialization schemes (e.g., hindcasts in the NMME project). However, in other basins, SST predictions based on a simple ocean initialization procedure were not worse (sometimes even better) than those with sophisticated initialization schemes. These comparisons indicate that it was helpful that subsurface ocean information be assimilated to improve the tropical Pacific SST predictions, while SST-based ocean assimilation was an effective way to enhance SST prediction capability in other ocean basins. By examining multimodel ensembles with the simple scheme-based hindcasts either included or excluded in NMME, it is also suggested that including the hindcast would generally benefit multimodel ensemble forecasts. In addition, possible ways to further improve ENSO SST predictions with the simple initialization scheme are also discussed.

1. Introduction

Sea surface temperature (SST) is one of the major factors being used to forecast the climate variations for one to a few seasons ahead. In particular, the SST variability in the tropical Pacific associated with El Niño–Southern Oscillation (ENSO) has a profound influence

on weather and climate worldwide. Since its first dynamical prediction conducted around three decades ago (Cane et al. 1986), the ability of dynamical models to predict ENSO has been improved significantly, and SSTs can be successfully predicted several seasons ahead (e.g., Ji et al. 1994; Wang et al. 2002; Zhang et al. 2003; Chen et al. 2004; Keenlyside et al. 2005; Luo et al. 2005, 2008; Jin et al. 2008; Kirtman and Min 2009; Zhu et al. 2012b, 2013a,c, 2014; Xue et al. 2013). In addition, there are also forecast studies about other ocean basins, such as the tropical Indian Ocean (e.g., Luo et al. 2007), tropical

Corresponding author address: Dr. Jieshun Zhu, Climate Prediction Center, 5830 University Research Court, College Park, MD 20740.
E-mail: jieshun.zhu@noaa.gov

Atlantic Ocean (e.g., [Hu and Huang 2007](#)), and the extratropical oceans including the North Atlantic (e.g., [Hu et al. 2013](#)), the North Pacific (e.g., [Hu et al. 2014](#)), the southern subtropical Indian and Atlantic (e.g., [Yuan et al. 2014](#)), and the southern subtropical Pacific (e.g., [Guan et al. 2014](#)).

At the National Centers for Environmental Predictions (NCEP), development of dynamical seasonal prediction systems was initiated in the early 1990s ([Ji et al. 1994](#)), targeting at predicting climate conditions (including SST) at the seasonal time scale. Since then, several prediction systems have followed, including the Seasonal Forecast System 2000 ([Kanamitsu et al. 2002](#)), the Climate Forecast System, version 1 (CFSv1; [Wang et al. 2005](#); [Saha et al. 2006](#)), and currently the Climate Forecast System, version 2 (CFSv2; [Saha et al. 2014](#)). During their development course at NCEP, coupled ocean–atmosphere model and ocean data assimilation have been two major foci. Ocean data assimilation is chosen as a focus because the memory for seasonal climate forecasting mainly resides in the ocean. At NCEP, a 3D variational technique ([Derber and Rosati 1989](#)) with gradual upgrades is applied to assimilate different surface/subsurface and satellite ocean observations.

The strategy of using sophisticated ocean data assimilation schemes to initialize seasonal prediction systems is also pursued at other operational centers, including the European Centre for Medium-Range Weather Forecasts (ECMWF) ([Stockdale et al. 1998](#)), Australian Bureau of Meteorology Research Centre (BMRC) ([Wang et al. 2002](#)), and others. The advantage of the strategy is that the derived ocean initial states (including the subsurface states) use all available ocean observations and can be closer to the reality. However, because of differences in the observed and the model-simulated states (including mean and variability), the analyzed ocean initial states can significantly deviate from the model's own mean climate, and could lead to certain disadvantages. During the prediction, starting from the initialized ocean states, such discrepancies could introduce substantial imbalances (or shocks), which may degrade the prediction performances. Another disadvantage of the use of ocean initial conditions based on ocean data assimilation systems is that seasonal hindcasts are seldom extended back prior to the 1980s (e.g., [Chen et al. 2004](#); [Deng and Tang 2009](#); [Zheng et al. 2009](#)). This limitation restricts the sample size of hindcasts that are typically over a 30-yr period.

In contrast to sophisticated ocean data assimilation schemes, there are also approaches using simple ocean initialization schemes that are proposed for seasonal predictions. For example, the ocean initial states can be derived by running a coupled model with its SST or SST anomalies strongly nudged to observations. These simple

initialization methods have been used in ENSO predictions with both intermediate complexity models (ICMs; e.g., [Zhang et al. 2003](#); [Chen et al. 2004](#)) mostly by nudging SST anomalies, and coupled global climate models (CGCMs; [Keenlyside et al. 2005](#); [Luo et al. 2005](#)) frequently by nudging full SSTs. The physical basis for the potential success of these methods is that, in addition to providing a realistic oceanic mixed layer temperature, through air–sea interaction and coupling, the observed SST information is able to reproduce at least part of realistic ocean subsurface evolutions by forcing the atmospheric winds. The simple initialization scheme used in the Scale Interaction Experiment-Frontier Research Center for Global Change (SINTEX-F) model has been documented to have skillful ENSO predictions with useful prediction skill extending up to 2 yr ([Luo et al. 2008](#)). Although, in comparison with the sophisticated assimilation schemes, the ocean initial states based on the simple scheme may have larger deviations from the reality due to deficiencies in the coupled models, they are likely to be closer to models' own mean states leading to lesser initial shock at the beginning of the forecast integration. This feature could potentially benefit climate forecasts as a result of smaller shocks.

Such simple initialization schemes have not been tried in any NCEP's seasonal prediction models. In this study we attempt to fill in this gap by applying this method to the latest seasonal prediction model at NCEP (i.e., CFSv2). However, we note that our experimental design is not optimal, that is, the model generating ocean initial conditions (i.e., CFSv1) is not the same as that used for predictions (i.e., CFSv2). Being cognizant of this discrepancy, we proceeded with the forecast experiments as a preliminary step toward (i) assessing the feasibility, and forecast skill of SSTs over different ocean basins using a simple SST based ocean data assimilation system; and more importantly (ii) exploring the possibility of extending the generation of such ocean initial states and hindcasts over a much longer period prior to 1980 because the only information required for this approach is the observed SSTs. The assessment of forecast skill, and its comparison with skill achieved in a seasonal prediction system with more advanced ocean data assimilation systems also highlights the potential importance of subsurface ocean observations and their assimilation.

By evaluating global SST predictions, we also explore the regional dependence of two factors influencing SST seasonal predictions (i.e., the accuracy of ocean initial conditions and the balance between forecast model climatology and initial states). For this purpose, hindcasts initialized by the simple scheme will be compared with hindcasts from multiple systems [i.e., hindcasts from the National (lately North American) Multimodel Ensemble

(NMME) project; [Kirtman et al. \(2014\)](#)]. All NMME hindcasts were initialized by sophisticated data assimilation systems.

The paper is organized as follows. The forecast model, the experimental design, and the datasets are described in the next section. [Section 3](#) evaluates the SST forecast results, including comparisons with NMME hindcasts. The summary and discussion are given in [section 4](#).

2. Model, hindcast experiments, and datasets

a. Model

The forecast model used in this study is NCEP CFSv2 ([Saha et al. 2014](#)). CFSv2 has been the operational forecast system for seasonal-to-interannual prediction at NCEP since March 2011, replacing its predecessor, CFSv1 ([Saha et al. 2006](#)). In CFSv2, the ocean model is the GFDL MOM version 4, which is configured for the global oceans with a horizontal grid of $0.5^\circ \times 0.5^\circ$ poleward of 30°S and 30°N and meridional resolution increasing gradually to 0.25° between 10°S and 10°N . The vertical coordinate is geopotential (z) with 40 levels (27 of them in the upper 400 m), with a maximum depth of approximately 4.5 km. The atmospheric model is the Global Forecast System (GFS), which has horizontal resolution at T126 (105-km grid spacing, a coarser resolution than that used for the GFS operational weather forecast), and 64 vertical levels in a hybrid sigma-pressure coordinate. The oceanic and atmospheric components exchange surface momentum, heat and freshwater fluxes, as well as SST, every 30 min. More details about CFSv2 can be found in [Saha et al. \(2014\)](#).

b. CFSv2 hindcasts with a simple initialization scheme

In contrast to operational forecasts with CFSv2 at NCEP where many subsurface and satellite data are assimilated in a complex way to derive its ocean initial conditions (OICs), a simple initialization scheme is applied in this study. In the scheme only SSTs are specified as the observed information in coupled model integrations (e.g., [Keenlyside et al. 2005](#); [Luo et al. 2005](#)). In particular, model SSTs in CFSv1 are nudged to the observed daily SST, with a restoring time scale of 3.3 days ([Wang et al. 2013](#); [Kumar et al. 2014](#)). The observed daily SSTs are interpolated from the weekly SSTs from the National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation SST (OISST) version 2 ([Reynolds et al. 2002](#)). [Kumar et al. \(2014\)](#) have indicated that the coupled integration generated a realistic evolution of subsurface ocean temperature, and particularly, the evolution of slow variability related to ENSO has a good resemblance to its

observational counterpart. More details about the SST nudging procedure (including a brief description of CFSv1) can be found in [Kumar et al. \(2014\)](#). As we mentioned earlier, although the experimental setup is not optimal in which an SST-based assimilation using CFSv1 is used to initialize CFSv2 forecasts, we consider this as a first step toward developing a simple procedure for initializing ocean states and extending seasonal hindcasts over a much longer period.

Based on the SST-derived OICs, hindcasts are conducted starting from each April during 1982–2009 and last for 12 months. April is chosen because it is an important month to predict whether an El Niño event will occur in a particular year. We also note that in comparison with predictions starting from other seasons, hindcasts initialized in April generally have lower predictive skill, possibly due to the effect of the spring predictability barrier ([Jin et al. 2008](#)).

For each hindcast year, four ensemble members are used based on a combination of two different OICs and two different atmosphere–land surface conditions. The two OICs are randomly chosen from nine CFSv1 integrations constrained by observed SSTs ([Kumar et al. 2014](#)). The atmosphere–land surface conditions are the instantaneous fields from 0000 UTC of the first two days of April in the NCEP Climate Forecast System Reanalysis (CFSR; [Saha et al. 2010](#)). We acknowledge that such ensemble generation scheme might generate less balanced initial conditions between ocean and atmosphere than that in the SINTEX-F practice ([Luo et al. 2005, 2008](#)). The hindcast experiment is referred to as the SSTnudging hindcast.

[Figure 1](#) examines the consistency of the upper-ocean heat content (HC; defined as the depth-average temperature of the upper 300 m) evolutions between the SST-derived OICs and the NCEP Global Ocean Data Assimilation System (GODAS; [Behringer and Xue 2004](#)). It is surprising to see that, in addition to good consistency in the tropical Pacific, there is also high consistency in the extratropical Pacific and Atlantic. However, in the eastern tropical Pacific, there is a correlation minimum along the equator, which is likely to influence ENSO prediction. In addition, it is also noticed that the correlation pattern, particularly north of 30°S , resembles well the consistency map of HC anomalies in different ocean analyses [see [Fig. 1](#) in [Zhu et al. \(2012a\)](#) for the map of signal-to-noise ratios], which implies that the uncertainties in the SST-derived ocean states might also be similar to those in present generation of ocean analyses.

c. NMME hindcast data and data for validations

To evaluate the prediction skill in CFSv2 with the simple initialization scheme, SSTnudging is compared with hindcasts in the NMME project ([Kirtman et al.](#)

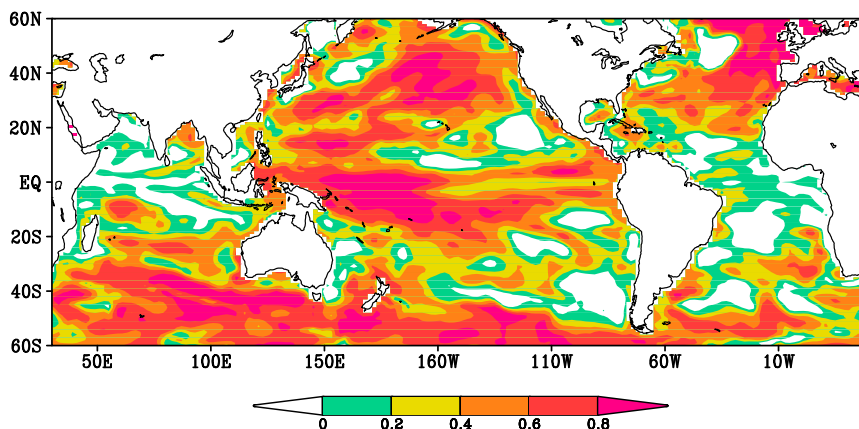


FIG. 1. Distribution of anomaly correlations between the GODAS-analyzed upper-ocean 300-m heat-content anomalies and those derived by CFSv1 with SSTs nudged to observations.

2014). In the Phase-1 NMME, seven U.S. models participated, with several decades of hindcast data available to the public. These models include COLA-RSMAS-CCSM3 (Kirtman and Min, 2009), GFDL-CM2.2 (Zhang et al. 2007), IRI-ECHAM4.5a (DeWitt, 2005), IRI-ECHAM4.5f (DeWitt, 2005), NASA-GMAO-GEOS-5 (G. Vernieres et al. 2011, unpublished manuscript), NCEP-CFSv1 (Saha et al. 2006), and NCEP-CFSv2 (Saha et al. 2010). In this paper, hindcasts from the seven systems are referred to as CCSM3, CM2.2, ECHAM4a, ECHAM4f, GEOS-5, CFSv1, and CFSv2, respectively. Specifically, there are 29-yr (1982–2010) hindcast data, with ensemble members of each model ranging from 6 to 24, and forecasts up to 8–12 months. All the NMME hindcasts are initialized by sophisticated ocean data assimilation systems. More details about NMME and their hindcast descriptions can be found in Kirtman et al. (2014).

It is worthwhile to clarify that the CFSv2 hindcasts conducted by NCEP as part of the NCEP CFS Reanalysis and Reforecast project (CFSRRR; Saha et al. 2014) is also a member of NMME. The CFSv2 hindcasts in NMME (referred to as CFSv2) were initialized by a sophisticated data assimilation system (i.e., CFSR; Saha et al. 2010). Thus, the major difference between hindcasts CFSv2 and SSTnudging lies in their respective ocean initializations.

Considering that seasonal predictability of SSTs, for example ENSO SSTs (Kumar and Hu 2014), mostly resides in the mean shift of the probability density function (PDF) with its spreads as a secondary source, the following analyses will focus on the ensemble mean SST as the forecast. In addition, the multimodel ensemble mean forecast based on the seven systems in the NMME is referred to as hindcast NMME. We also constructed another multimodel ensemble forecast (referred to as NMMEpSSTndg) by additionally including SSTnudging (i.e., totaling eight system predictions). By comparing

hindcast NMME with hindcast NMMEpSSTndg, the purpose is to examine whether there is a gain or loss of skill when SSTnudging is included as part of NMME. In constructing the two multimodel ensemble forecasts, equal weights are given to each system instead of each forecast member. This choice prevents the multimodel ensemble from strongly leaning toward forecasts with large ensemble members, like CFSv2 with 24 members.

The above strategy of skill validation reflects a practical view; that is, making the most skillful forecast with largest ensemble size available for the seasonal prediction system. Such validations, however, are influenced by differences in ensemble size. To rule out skill differences due to ensemble size, we also constructed another group of hindcasts with four ensemble members for each NMME model (referred to as CCSM3_4E, CM2.2_4E, ECHAM4a_4E, ECHAM4f_4E, GEOS-5_4E, CFSv1_4E, and CFSv2_4E, respectively). Accordingly, these hindcasts have the same ensemble size as SSTnudging. The multimodel ensemble mean forecast based on the new group of hindcasts is referred to as hindcast NMME4E. Also, to rule out skill differences in multimodel ensembles due to different model numbers, we also constructed multimodel ensemble mean forecast (referred to as NMME4EwSSTndg), which is the same as NMME4E but with its CFSv2_4E component replaced by SSTnudging.

All the analyses are based on forecasts over the common period (i.e., 1982–2009). The predicted SST anomalies (SSTAs) are derived by subtracting lead-time-dependent climatologies from the total SSTs, but no additional time smoothing is applied. The climatologies are based on the whole hindcast period of 1982–2009. The observation-based monthly SST analysis used for validation is from OISST V2 on a $1.0^\circ \times 1.0^\circ$ (latitude \times longitude) grid. The subsurface ocean temperatures from the NCEP GODAS (Behringer and Xue 2004) are used for HC validations.

SST Predictive Skill (April ICs, 1982–2009): Correlation (3-month lead)

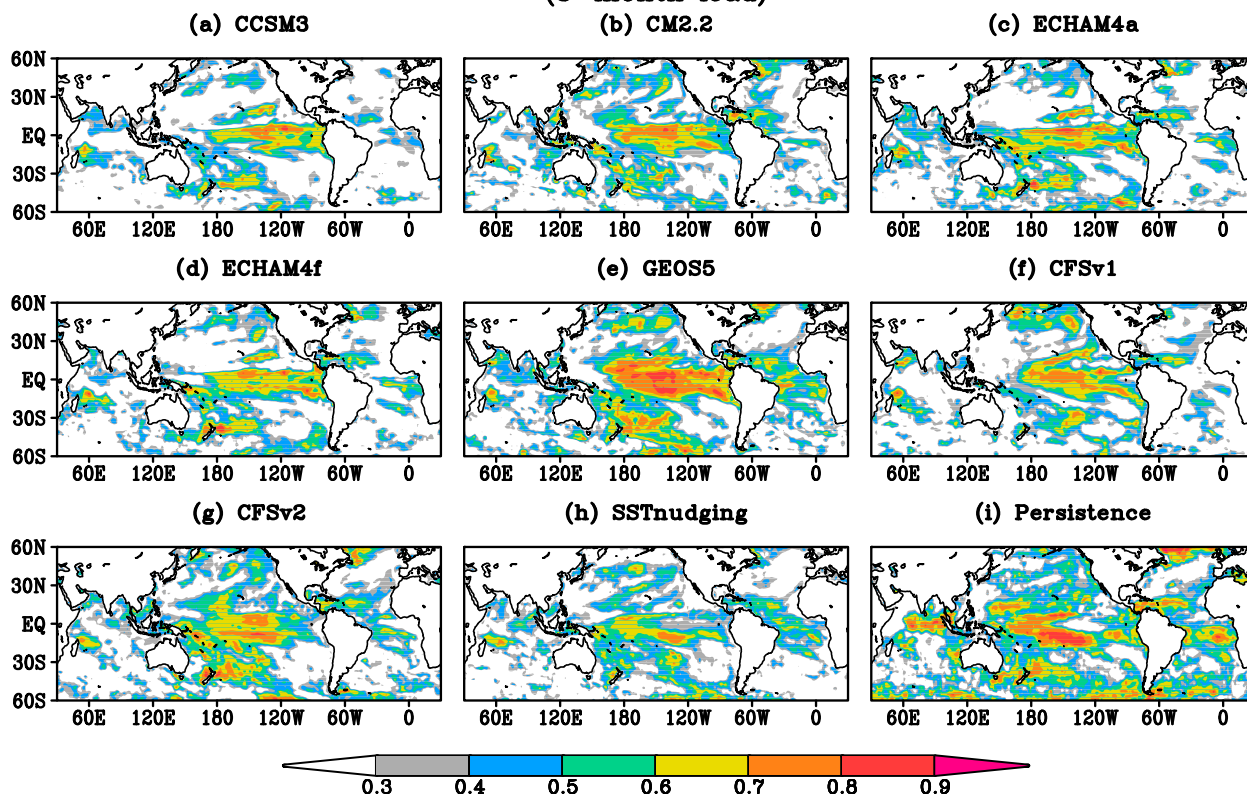


FIG. 2. Distribution of anomaly correlations between observed and predicted SST anomalies at the 3-month lead time in (a) CCSM3, (b) CM2.2, (c) ECHAM4a, (d) ECHAM4f, (e) GEOS-5, (f) CFSv1, (g) CFSv2, (h) SSTnudging, and (i) persistence. The hindcasts start from April initial conditions.

3. Results

Figures 2 and 3 show the horizontal distributions of quasi-global SST prediction skill at the lead times of 3 and 6 months, respectively, with each system using their respective ensemble sizes. Validations with each NMME model also using four ensemble members generally show little difference in the tropical Pacific and some difference in the other two tropical basins (figures not shown), but relatively large differences over the tropical Pacific are found for some models (e.g., CFSv2). The reason for small sensitivity to ensemble size in the tropical Pacific is due to high predictability and signal-to-noise ratio, a scenario where small ensemble size is adequate to realize expected value of skill (Kumar and Hoerling 2000).

For both lead times, regions with the highest correlation are located in the tropical Pacific in all systems. Among hindcasts from different NMME component models (Figs. 2a–g and 3a–g), however, it is interesting to notice substantial spreads in their SST prediction skill, even in the tropical Pacific. For example, at the 3-month lead time, GEOS-5 seems better than others; its correlation skill is

greater than 0.7 over a large area of the central and eastern tropical Pacific with few patches even greater than 0.8, while in other systems regions of correlations greater than 0.7 are clearly smaller. With the increase in lead times, the correlation first drops gradually in the tropical Pacific during the ENSO developing phase (figures not shown), then increases modestly when approaching the boreal winter when ENSO events generally peak. As a result, the ENSO SSTA prediction skill does not drop much in all systems at the 6-month lead time (Fig. 3) in comparison with at the 3-month lead time (Fig. 2). Compared with other systems, CFSv1 and GEOS-5 seem to have better skill at the 6-month lead time, with correlations larger than 0.7 over a sizable region of the central and eastern Pacific. For CFSv1, there is even a discernible region with correlation greater than 0.8.

As for the CFSv2 skill, it should be noted that a better SST prediction skill could be achieved in the tropical Pacific if two climatologies are applied as suggested by Xue et al. (2013). The resulting skill can be comparable to (or even slightly better than) CFSv1 (Xue et al. 2013). The procedure takes into account the factor that there

**SST Predictive Skill (April ICs, 1982–2009): Correlation
(6-month lead)**

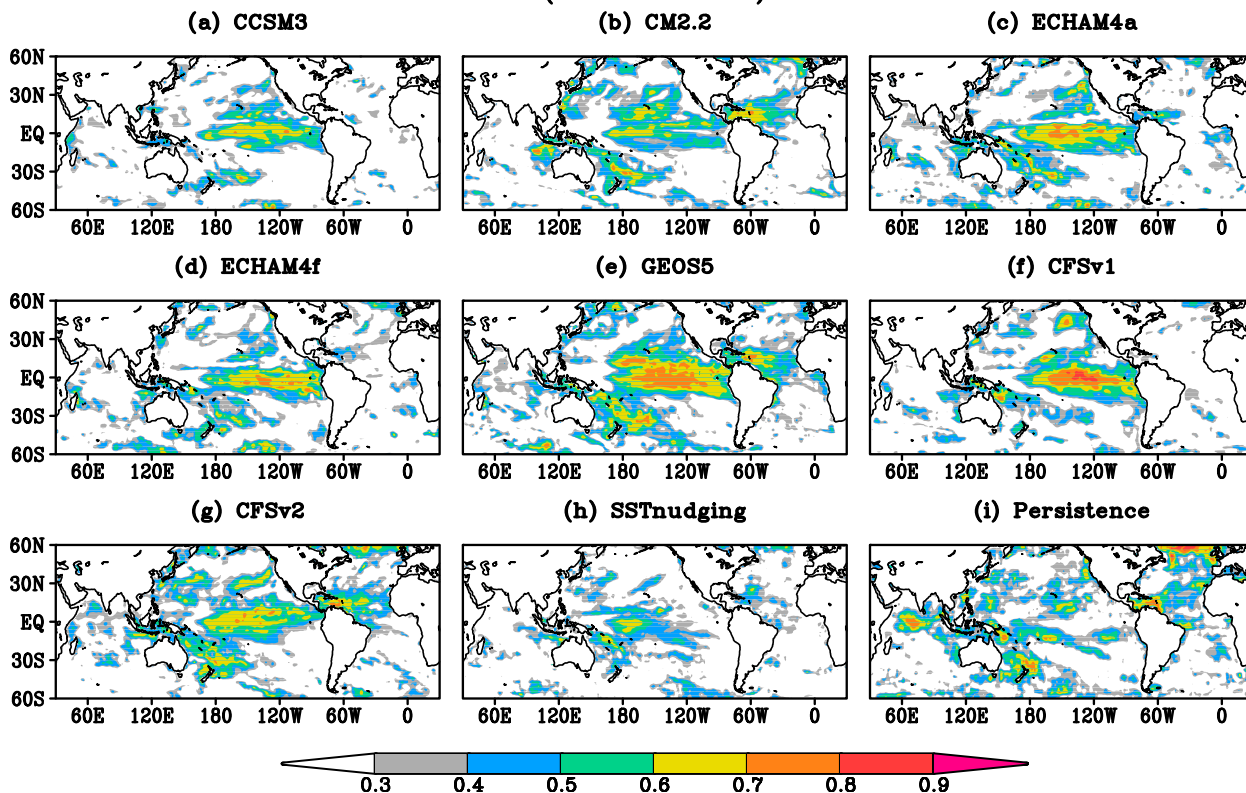


FIG. 3. As in Fig. 2, but at the 6-month lead time.

is a discontinuity in the CFSv2 reforecast and reanalysis data with significantly warmer mean SST appearing in the tropical Pacific after 1999 than before. The discontinuity likely results from the assimilation of new satellite observations in the CFSR initial conditions starting from 1999 (Kumar et al. 2012), which should be taken into account in defining the predicted SST anomalies. However, it was also found that applying two climatologies clearly degrades SST prediction skill in other ocean basins, particularly in the extratropical oceans. Because the global SST predictions are evaluated in this study, we chose to use one climatology derived from the whole forecast period. This choice is also more consistent with evaluations of other hindcasts.

Compared with the NMME component systems (Figs. 2a–g and 3a–g), SST prediction in the SSTnudging hindcast (Figs. 2h and 3h) generally shows equivalent skill in the western and central Pacific at both 3- and 6-month lead times. On the other hand, the SSTnudging skill is clearly lower than those from the others in the eastern equatorial Pacific. At the 3-month lead time, there is a region with correlation below 0.4 in SSTnudging (Fig. 2h), extending from the central to eastern equatorial Pacific

and farther southward along the coastal South America, which is lower than all NMME component systems (Figs. 2a–g). The correlation skill also drops faster from the 3- to 6-month lead time in SSTnudging over these regions. For 6-month lead time, SSTnudging correlation skill is mostly below 0.4, clearly lower than the NMME component systems. However, SSTnudging skill is still generally better than persistence (i.e., Figs. 2i and 3i) in the eastern equatorial Pacific, especially at the 6-month lead time.

The relatively poor skill of SSTnudging in the eastern Pacific could be related to the biases in its ocean initial conditions. Analysis in Kumar et al. (2014, see their Figs. 8, 10 and 11; also the current Fig. 1) shows that nudging SST does not replicate observed variability in the eastern equatorial Pacific well, with the simulated variability significantly weaker than the reality along the thermocline. This is also consistent with the assessment of HC in Fig. 1. The bias is likely due to a deficiency commonly appearing in current ocean general models, that is, the vertical processes are poorly represented, which is particularly important in regions with shallow thermoclines like the eastern equatorial Pacific. Because

of the deficiency, when simulations are only constrained by observed SSTs, the observed subsurface variability is likely weaker in the eastern equatorial Pacific, not to mention the errors in the SST-forced wind stresses influencing coupled interactions. These initial errors in SSTnudging can persist and even grow during the forecast period, while signals propagating eastward from the areas near date line (where ocean variability in the SSTnudged simulations replicates observations well) may also be weakened by the diffused thermocline. Both may contribute to the low prediction skill in the eastern equatorial Pacific in hindcast SSTnudging. In contrast, hindcasts initialized from sophisticated ocean data assimilation systems (e.g., the NMME hindcasts) with better representation of oceanic variability in the eastern Pacific reflect this advantage as better prediction skill. This skill comparison between the SSTnudging hindcast and the NMME component hindcasts suggests that assimilation of subsurface information contributes to improving the SST predictions in the tropical Pacific, while the balance between forecast model and initial conditions may be a secondary factor. The latter argument is also supported by a recent experiment showing little skill difference in ENSO SST predictions when either full or anomaly ocean initializations are applied (Zhu et al. 2012b).

In addition to the relatively high skill in the tropical Pacific, there are also other regions with reasonable skill for most forecast systems. At the 3-month lead time (Fig. 2), all hindcasts including SSTnudging show equivalent skill in the tropical southwestern Indian Ocean where the Seychelles Dome (Yokoi et al. 2008) is located. In the hindcasts of GEOS-5, CFSv2, and SSTnudging, skill is also present in the tropical Atlantic Ocean, better than other hindcasts. In addition, almost all hindcasts exhibit some skill in the extratropical ocean basins, like the southern subtropical Pacific, the North Atlantic, and the North Pacific. At the 6-month lead time (Fig. 3), relatively high correlation skill appears in the northern tropical Atlantic in GEOS-5, CM2.2, and CFSv2, and more or less in SSTnudging, which has been suggested to be associated with the remote effect from ENSO (Hu and Huang 2007). Overall, it is encouraging that SSTnudging can forecast the SSTAs as well as the NMME component systems do in ocean basins other than the tropical Pacific at these two lead times.

The above finding suggests that sophisticated ocean data assimilation do not provide much improvement in the SST predictions in basins other than the tropical Pacific. There are many factors that might potentially contribute to this. First, it is possible that because of biases the current generation of models is unable to benefit from sophisticated ocean data assimilation

procedures and the availability of subsurface ocean observations. In fact, current climate models indeed present substantial biases in these basins, such as the tropical Atlantic (Huang et al. 2007). Further, in comparison with the tropical Pacific, current ocean analysis systems also show much higher uncertainties in estimating ocean variability in other ocean basins [e.g., Fig. 1 in Zhu et al. (2012a)]. These uncertainties will undoubtedly bring uncertainties in the predicted SSTs and thus degrade SST predictions, which will be discussed more when evaluating HC predictions below (Fig. 7). It is also possible that, outside major upwelling zones, the oceanic mixed layer plays a more dominant role than the thermocline fluctuations on seasonal time scales. No matter which factor contributes more, our skill comparison indicates that at present SST nudging is an effective way to enhance SST prediction capability in ocean basins other than the tropical Pacific. The method, by ensuring a balance between forecast models and initial conditions, might be more effective than use of subsurface ocean observations in advanced ocean data assimilation schemes to improve the accuracy of ocean initial conditions for these basins. It is only with reduction in model biases and improvement in ocean data assimilations over these basins that the advantages from increasing ocean observations can be recognized in our prediction practice.

To further evaluate the SST predictions of SSTnudging, Figs. 4–6 present the anomaly correlation and root-mean-square error (RMSE) between the observed and predicted SSTA time series as a function of lead time based on the hindcasts from SSTnudging and NMME (including both full ensemble members and four members). The chosen SSTA time series represent major modes in three tropical oceans, and in these modes ocean dynamics plays a critical role for their evolutions. In particular, Fig. 4 is for ENSO in the tropical Pacific, represented by the averaged SSTA over the Niño-3.4 region (5°S – 5°N , 170° – 120°W); Fig. 5 is for the ATL3 index (i.e., averaged SSTA in 3°S – 3°N , 20°W – 0°), representing the zonal equatorial mode in the tropical Atlantic; Fig. 6 examines the eastern Indian Ocean (EIO) (10°S – 0° , 90° – 110°E) and western Indian Ocean (WIO) (10°S – 10°N , 50° – 70°E) indices, which respectively represent the eastern and western poles of the Indian Ocean dipole in the tropical Indian Ocean. In addition to the skill comparison of one model versus another, comparisons are also done in the framework of multimodel ensemble forecast (i.e., NMME vs NMMEpSSTndg and NMME4E vs NMME4EwSSTndg). The difference between NMME and NMMEpSSTndg is whether SSTnudging is included as an extra member or not. Again, this comparison reflects a practical view, and is conducted to examine

Predictive Skill of Niño3.4 index

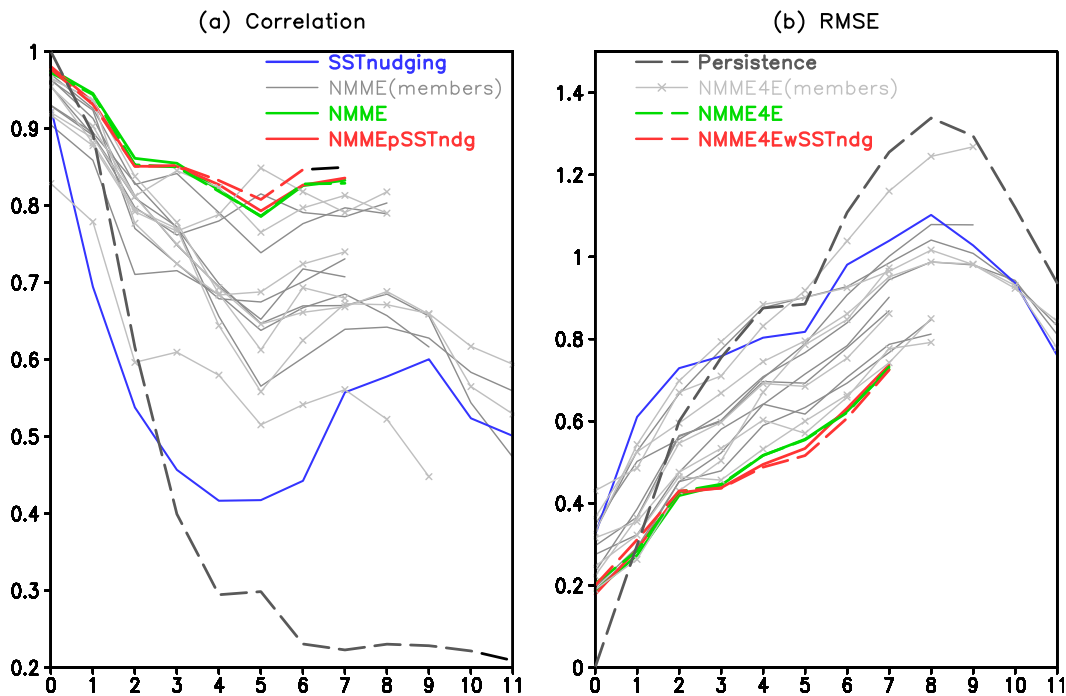


FIG. 4. (a) Anomaly correlation coefficients and (b) RMSEs ($^{\circ}\text{C}$) of Niño-3.4 index as a function of forecast lead months (x axis) after removing the mean bias. Solid curves in blue (gray) are for forecasts of SSTnudging (each NMME member model with full ensemble size). Solid curves in green (red) are for multisystem ensemble forecasts of NMME (NMMEpSSTndg). Dashed curves in dark gray are persistence forecast. Solid gray curves with \times marks are for each NMME member model with four ensemble members. Dashed curves in green (red) are for multisystem ensemble forecasts of NMME4E (NMME4EwSSTndg).

whether there are any gains or losses if SSTnudging is included as part of the overall NMME multimodel ensemble system. On the other hand, the comparison between NMME4E and NMME4EwSSTndg does not include the positive influence from different ensemble size, and is a stricter validation about the effect because of inclusion of forecasts with the simple ocean initialization scheme. However, because only one out of seven (or eight) members is different between NMME and NMMEpSSTndg (between NMME4E and NMME4EwSSTndg), the differences are expected to be marginal. In fact, significance tests based on the two-tailed z test indicate that differences in correlation skill are not significant at the 90% confidence level. Thus, only a qualitative conclusion can be drawn.

For the Niño-3.4 index hindcasts for each NMME model (Fig. 4), there is little skill difference between constructions using their full ensemble members (extending from 6 to 24) and those using 4 members, which is consistent with studies by DelSole et al. (2014) and Zhu et al. (2015), and is due to the high signal-to-noise ratio in ENSO (Kumar and Hoerling 2000). However, there are substantial differences evident among the

NMME component systems. In particular, in terms of anomaly correlation for the NMME models using their full ensemble members, there are systems with correlations of around 0.8 at all lead times of 8 months, but there are also systems for which correlation skill drops quickly to around 0.6 at the lead times of longer than 4 months. The difference in RMSE can also reach above 0.2°C . For SSTnudging, as expected from the large biases in its initial conditions in the eastern Pacific (Kumar et al. 2014; Fig. 1), the prediction skill of Niño-3.4 index is generally below all the NMME component systems, but clearly better than persistence at lead times longer than 3 months. The multimodel ensemble hindcasts (i.e., NMME, NMMEpSSTndg, NMME4E, and NMME4EwSSTndg) generally improve on all component hindcasts in terms of both anomaly correlation and RMSE, which proves the utility of the multimodel ensemble method (Kirtman and Min 2009; Kirtman et al. 2014). The difference between both NMME and NMMEpSSTndg and that between NMME4E and NMME4EwSSTndg are small as expected. Comparing them more closely, NMME (NMMEpSSTndg) looks slightly better at short (long) lead times. What is more interesting is that NMME4EwSSTndg shows

Predictive Skill of ATL3 index

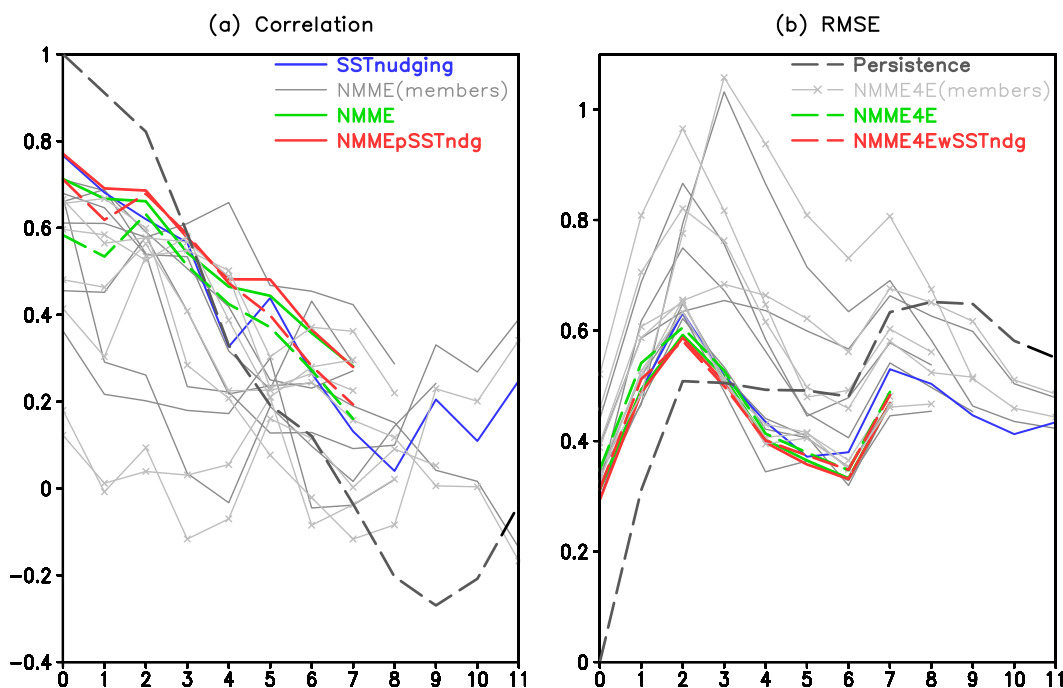


FIG. 5. As in Fig. 4, but for the ATL3 index.

certain improvements that are different than NMME4E, which seems to suggest that the CFSv2 model can better cancel errors in other NMME models when using the simple ocean initialization scheme than using the CFSR (Saha et al. 2010). In addition, there is a skill rebound at the lead times of 7–9 months in all hindcasts, which reflects that prediction systems are generally more skillful in predicting the SSTs during the mature ENSO state (when slowly evolving subsurface heat-content anomalies emerge at the surface as SSTs) than the transient ones during developing stage.

For the ATL3 index (Fig. 5), prediction skill is clearly lower than that of the Niño-3.4 index for all systems, and further, could not even beat persistence for short lead forecasts. In contrast to little difference in Niño-3.4 (Fig. 4), substantial differences appear in some NMME component systems when using four ensembles instead of their full ensemble, and are likely due to the low signal-to-noise ratio in ATL3 (Kumar and Hoerling 2000). Further, the skill difference among the NMME component systems is also larger than that for the Niño-3.4 index, which can reach 0.6 (0.5°C) in anomaly correlation (RMSE) at some lead times. It is encouraging that the prediction skill (in terms of both anomaly correlation and RMSE) of SSTnudging lies within the skill range of the NMME component systems. Furthermore, the SSTnudging skill is better than the median skill of all

hindcasts at most lead times. Compared with persistence, skill for SSTnudging is clearly better at lead times longer than 4 months, and also is not below it much at short lead times. In addition, the multimodel ensemble hindcasts (i.e., NMME, NMMEpSSTndg, NMME4E, and NMME4EwSSTndg) generally have better skills than individual system. However, NMME is clearly better than NMME4E, which confirms that the low signal-to-noise ratio in ATL3 requires larger ensemble sizes that in tropical Pacific to realize skill. More encouragingly, NMMEpSSTndg (NMME4EwSSTndg) is better than NMME (NMME4E) in terms of anomaly correlation metric for almost all lead times of 0–7 months, which is indicative of the benefit from SSTnudging.

Predictions of EIO and WIO indices (Fig. 6) are similar to ATL3, but the beneficial effect from SSTnudging is larger for EIO than WIO. In addition, we analyzed the prediction skill (figures not shown) for the averaged SSTA in the tropical southwestern Indian Ocean (SWIO; 20°–10°S, 50°–70°E) and the Atlantic's main development region for hurricanes (MDR; 10°–20°N, 80°–20°W). It was found that their predictive skill in SSTnudging was comparable to the NMME component systems.

To better understand SST prediction skill, HC predictions were also evaluated by comparing CFSv2 with SSTnudging (note that HC fields or three-dimensional ocean temperatures from NMME are unavailable). The

Predictive Skill of EIO (upper) and WIO (low) indices

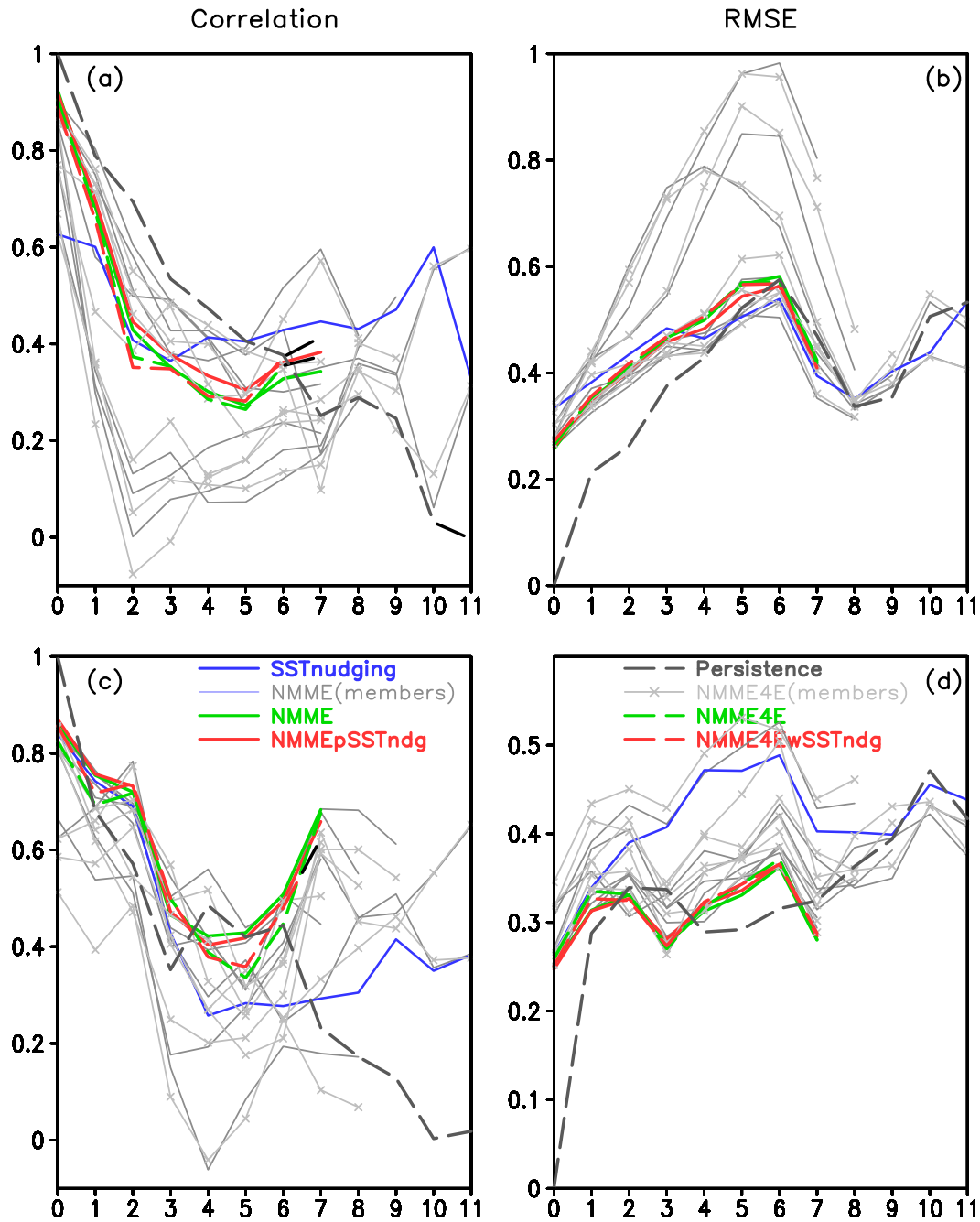


FIG. 6. As in Fig. 4, but for (a),(b) EIO and (c),(d) WIO indices.

HC forecasts at the 0-month lead time reflect the quality of ocean initial conditions in the upper oceans, which is evidenced by high spatial resemblance between Figs. 1 and 7b. By comparing them between CFSv2 and SSTnudging, it is clear that HC states are better initialized in CFSv2 due to assimilations of subsurface observations. As prediction continues, however, the HC

prediction skill drops faster in CFSv2 than SSTnudging, which might suggest a smaller initial shock in SSTnudging. For the 6-month lead time the HC prediction skills become comparable in a global view, even though some regional differences can be noted. Specifically, in the eastern tropical Pacific, HC evolutions are always predicted worse in SSTnudging than CFSv2 mainly due to

HCA Predictive Skill (April ICs, 1982–2009): Correlation

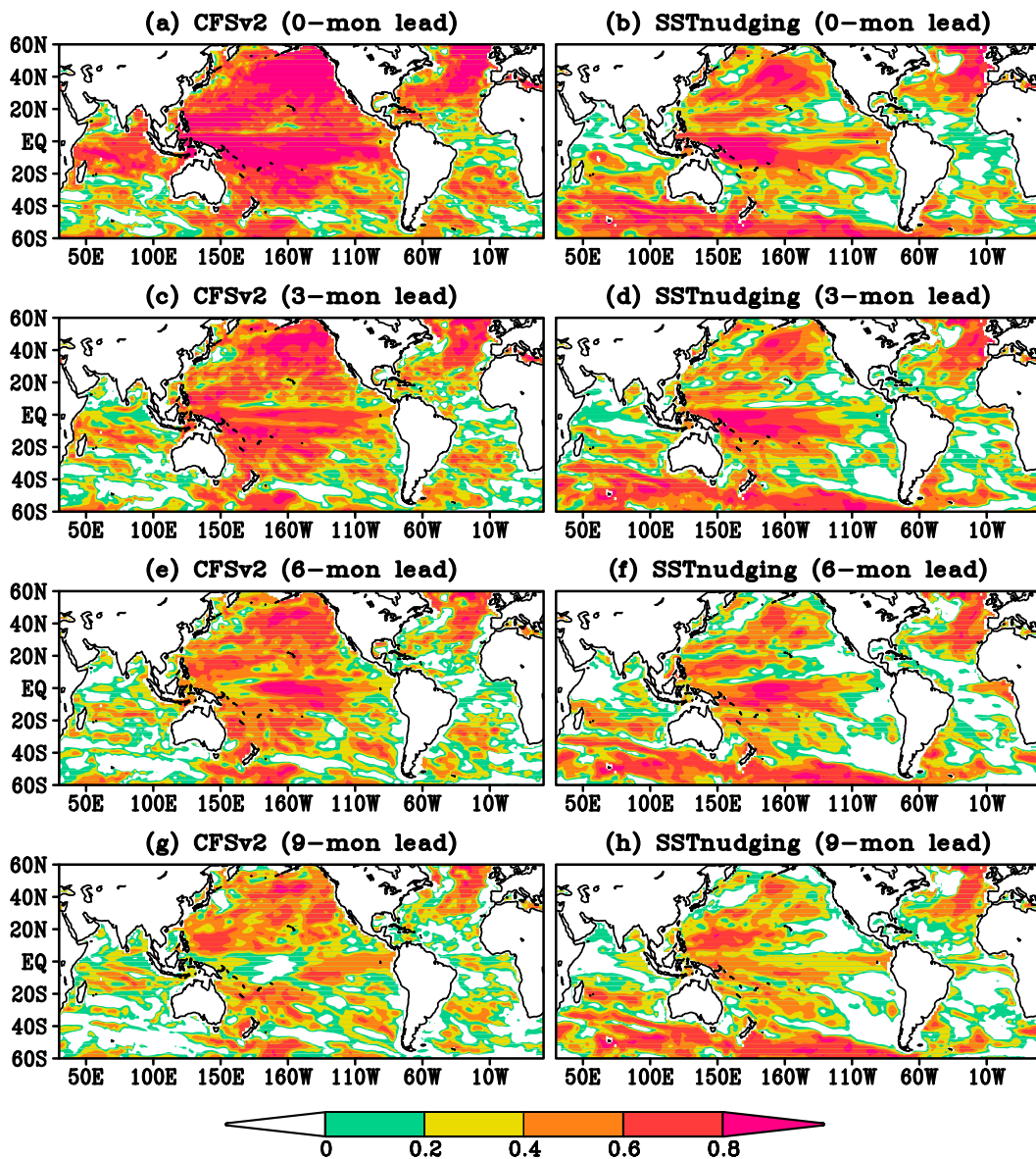


FIG. 7. Distribution of anomaly correlations between GODAS-analyzed and predicted heat-content anomalies in the upper 300 m at a (a), (b) 0-, (c), (d) 3-, (e), (f) 6-, and (g), (h) 9-month lead time; (left) CFSv2 and (right) SSTnudging.

the larger biases in SSTnudging's initial states (Kumar et al. 2014; Fig. 1), which explain the lower SST prediction skills in SSTnudging shown above (i.e., Figs. 2, 3, and 4). On the other hand, it should be noted that, because more propagating signals appear in HCs than in SST, a rather large portion of HC prediction skill might not rely on the regional HC initial states. The skill over the equatorial Atlantic in SSTnudging is a good example of this (Figs. 7b,d,f,h). The skill seems to propagate from the northern tropical Atlantic into the equatorial

waveguide, and propagate farther along the equator into the eastern Atlantic, which also reflects westward after reaching the eastern boundary. All these processes look consistent with the physics identified previously (Huang and Shukla 1997; Zhu et al. 2012c), which, however, are less evident in hindcast CFSv2 (Figs. 7a,c,e,g).

The differences between hindcasts for SSTnudging and CFSv2 might also be due to the influence of assimilation of subsurface observations in generating the initial states of hindcast CFSv2 (Saha et al. 2010).

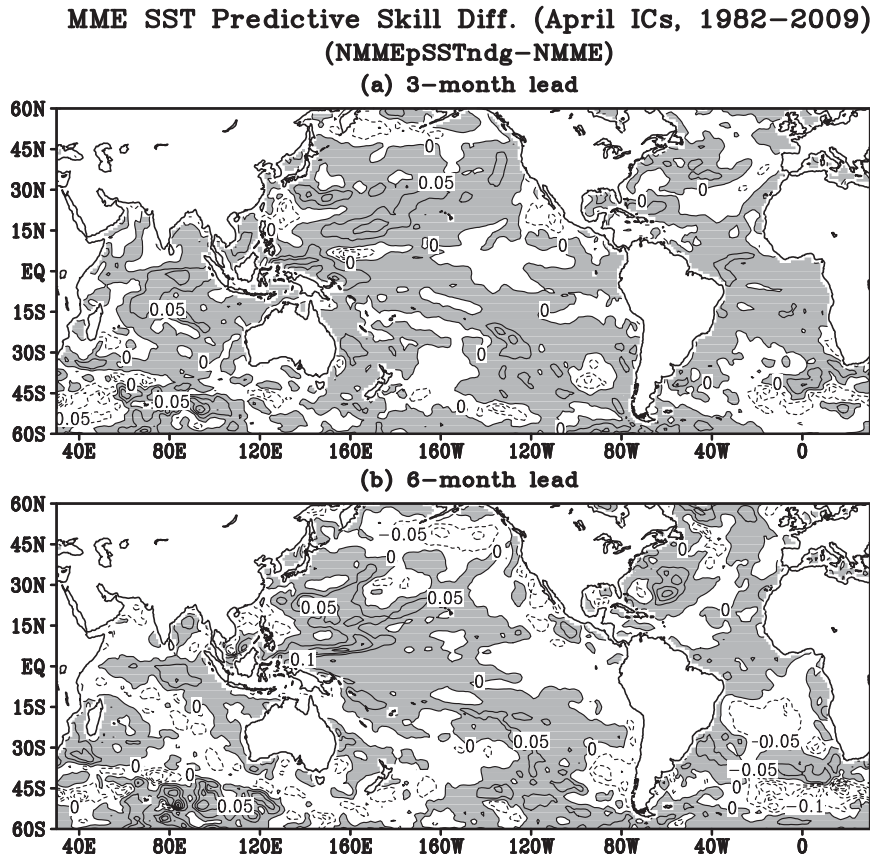


FIG. 8. The correlation skill difference of two multimodel at (a) 3- and (b) 6-month lead time. The contour interval is 0.05 with negative contours dashed. The shaded regions represent positive values.

Assimilating subsurface observations could introduce unphysical (or numerical) noises. For basins like the tropical Atlantic where the signal-to-noise ratios are relatively low, the numerical noises could explain a large amount of variance (Zhu et al. 2012a). Accordingly, the derived ocean state might deviate substantially from the physics of prediction models, which could bring inconsistencies (or shocks) during the prediction, and consequently contaminate prediction skill. In contrast, in the simple initialization scheme, the subsurface conditions are purely derived from the ocean model forced by SST-driven fluxes. This may lead to larger consistency between the initial subsurface ocean and the physics of prediction model, potentially benefitting both HC and SST predictions. Therefore, we argue that it is an effective way to enhance SST prediction capability in most basins other than tropical Pacific by pursuing the balance between forecast models and initial conditions.

Finally, we evaluate the value of SSTnudging in the frame of multimodel ensemble forecast (Kirtman and Min 2009; Kirtman et al. 2014). Figure 8 presents the global correlation skill differences between NMME and

NMMEpSSTndg for the 3- and 6-month lead times. Again, significance tests based on the z test indicate that differences in correlation skill are significant at the 90% confidence level over few regions, and thus only a qualitative conclusion can be drawn. At the 3-month lead time, except for a few locations the skill improvement by including SSTnudging is almost globally distributed. The largest improvement appears in the tropical western Pacific–Indian Oceans and the extratropical Pacific. At the 6-month lead time, while improvements over the above regions are still evident, there are more regions with decreased skill, including the eastern Pacific. This suggests that, in a global view, SSTnudging benefits the multimodel ensemble more at the short than long lead times, probably because it takes time for the initial large subsurface biases of SSTnudging to be evident in the surface temperature. It is also probable that smaller initial shock in SSTnudging benefits the multimodel ensemble at the short lead times. In general, the global skill difference maps suggest that SSTnudging would benefit the multimodel ensemble forecast. Figure 9 presents the skill difference between NMME4E and NMME4EwSSTndg, in

MME SST Predictive Skill Diff. (April ICs, 1982–2009)
(NMME4EwSSTndg–NMME4E)

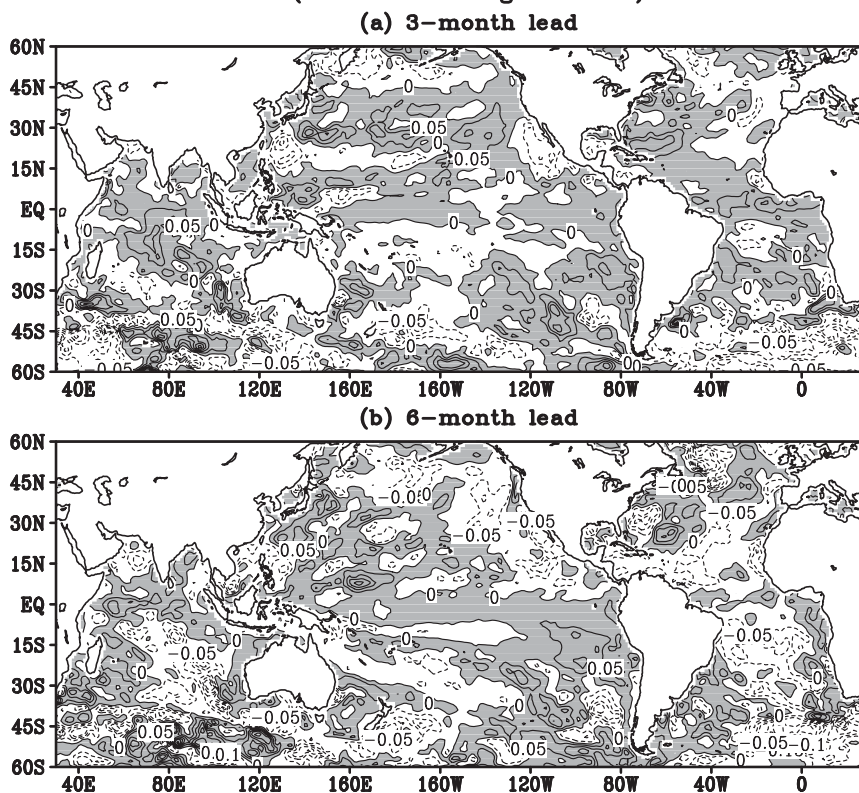


FIG. 9. As in Fig. 8, but for the difference between NMME4EwSSTndg and NMME4E.

which skill differences due to ensemble size were ruled out. Comparing Fig. 9 with Fig. 8, it can be found that the major findings remain unchanged. From Fig. 9, however, it is surprising to see that ENSO prediction would also slightly benefit from the replacement of hindcast CFSv2 with SSTnudging.

4. Conclusions and discussion

In this study, we applied a simple ocean initialization scheme to the NCEP latest seasonal prediction model, CFSv2. The scheme is much simpler than the sophisticated ocean data assimilation schemes as applied in operational climate predictions at NCEP. In the scheme, SST is the only observed information applied to derive ocean initial states. The physical basis for the method is that, through air–sea interaction, SST is able to reproduce some realistic ocean evolutions by forcing the atmospheric winds.

We examined the SST predictions based on the scheme and also compared with hindcasts from the NMME project. It was found that, due to substantial biases in the eastern Pacific in the ocean initial conditions, its SST predictions for this region are not as good

as those with sophisticated initialization schemes (e.g., hindcasts in the NMME). However, in other basins, SST predictions seem comparable to those with sophisticated initialization schemes. These comparisons indicate that, at the present, assimilating subsurface information greatly improves the prediction skill in the tropical Pacific, while nudging observed SST information is an effective way to enhance SST prediction capability in other ocean basins. In addition, by skill evaluations in the frame of multimodel ensemble forecast, we found that the hindcast using the simple initialization scheme would also generally benefit the NMME forecast.

As for ENSO predictions, we note that, using the similar initialization scheme, SINTEX-F (Luo et al. 2005, 2008) reported skills clearly higher than ours. There are many factors that can potentially contribute to the difference. First, our predictions start from April only, which is the most challenging month for ENSO prediction possibly due to the existence of the spring predictability barrier (e.g., Luo et al. 2005; Jin et al. 2008). We also found our SST-derived subsurface conditions in the eastern equatorial Pacific were the least accurate in April (Fig. 10), which will definitely degrade the SST

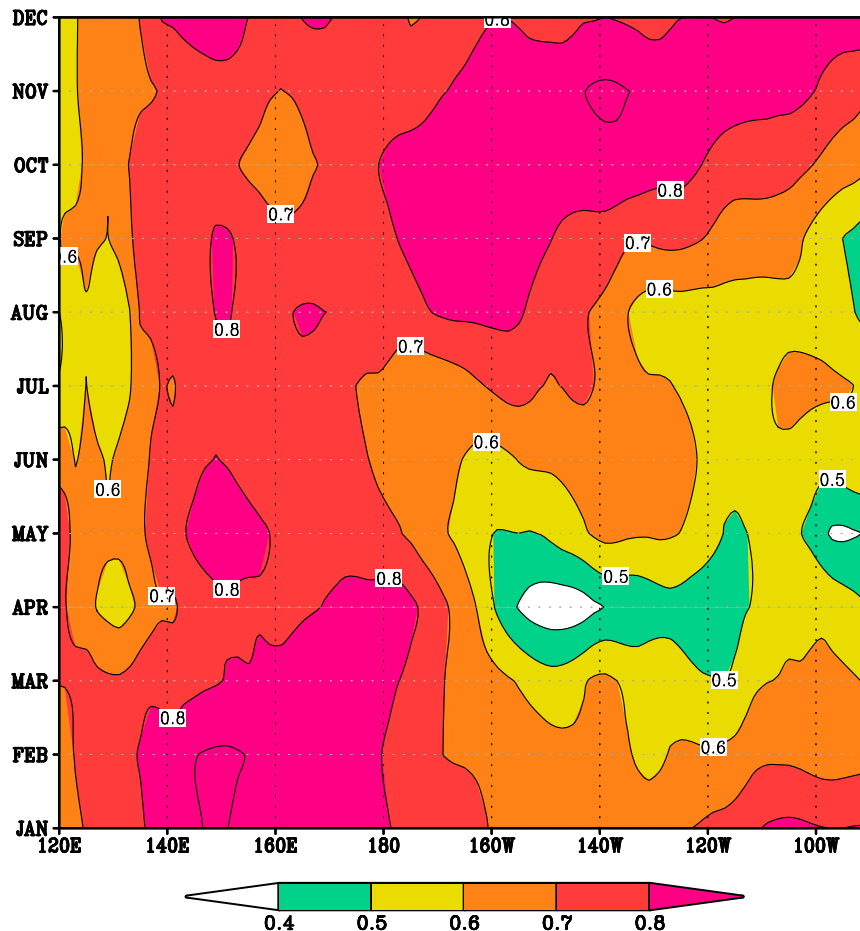


FIG. 10. Seasonality (monthly variation) of anomaly correlations between GODAS-analyzed heat-content anomalies and those derived by CFSv1 with SSTs nudged to observations along the equator (averaged over 5°S–5°N) in the Pacific.

predictions starting from April. Therefore, it is reasonable to expect that higher skills could be achieved for forecasts starting from all seasons as in Luo et al. (2005, 2008). Second, as mentioned, our application of the simple initialization scheme is not optimal in its design [i.e., the model generating OICs (i.e., CFSv1) is not the same as that used for predictions (i.e., CFSv2)]. Even though the two models show some common biases, like the warm SST biases in the southeast Pacific and Atlantic (figures not shown), each model has its own bias distribution. There could still be an initial shock because of the differences between the climatologies of CFSv1 and CFSv2. Third, our atmospheric initial conditions are directly taken from CFSR (Saha et al. 2010), different from the practice in Luo et al. (2005, 2008) that adopted both atmospheric and oceanic initial conditions from a continuous run with SST nudged to observations. Thus, our scheme generates less balance initial conditions between the ocean and atmosphere than that in SINTEX-F

(Luo et al. 2005, 2008), even though the effect from the imbalance might be smaller in CFSv2 than other models (Zhu et al. 2012b). However, it is also noted that, even for the same start month (i.e., April), SINTEX-F still seems to have better skills than found in our analysis [Fig. 5a in Luo et al. (2008)]. Thus, the last two factors may be the major reasons for the difference between SINTEX-F and our results. In addition, there is another issue, which could potentially further reduce the initial shocks, that is, nudging the model to the observed SST anomalies instead of the full SST fields. This procedure could reduce the shocks due to the differences in climatology between model and the reality.

In summary, the skill presented here is a baseline estimate for this approach, as we implied in introduction. The present analysis is a first step toward testing the feasibility of obtaining ocean initial states based on SST nudging, and exploring the possibility of extending the hindcasts over a longer period. We expect a better

balance between the forecast model and initial conditions, if the same model (e.g., CFSv2) is used for both generating OICs (and/or atmospheric initial conditions) and predictions. The better balance is likely to bring improvements over the present SST prediction skill in the tropical Pacific. The corresponding experiments with CFSv2 are under consideration. These experiments will also help quantify usefulness of the subsurface oceans in predicting ENSO. Further, recent studies have identified the importance of perturbing OICs in climate predictions (Zhu et al. 2012b, 2013a,b; Bellucci et al. 2013), and we argue that such a simple ocean initialization scheme can be another way to generate different set of OICs, which would be a good supplement to most current operational practices with sophisticated schemes.

Acknowledgments. We thank NOAA's Climate Program Office, Climate Observation Division for their support. B. Huang is supported by grants from NSF (Grant ATM-0830068), NOAA (Grant NA09OAR4310058), and NASA (Grant NNX09AN50G).

REFERENCES

- Behringer, D. W., and Y. Xue, 2004: Evaluation of the global ocean data assimilation system at NCEP: The Pacific Ocean. *Eighth Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface*, Seattle, WA, Amer. Meteor. Soc., 2.3. [Available online at http://ams.confex.com/ams/84Annual/techprogram/paper_70720.htm.]
- Bellucci, A., and Coauthors, 2013: Decadal climate predictions with a coupled OAGCM initialized with oceanic reanalyses. *Climate Dyn.*, **40**, 1483–1497, doi:10.1007/s00382-012-1468-z.
- Cane, M. A., S. E. Zebiak, and S. C. Dolan, 1986: Experimental forecast of El Niño. *Nature*, **321**, 827–832, doi:10.1038/321827a0.
- Chen, D., M. A. Cane, A. Kaplan, S. E. Zebiak, and D. J. Huang, 2004: Predictability of El Niño over the past 148 years. *Nature*, **428**, 733–736, doi:10.1038/nature02439.
- DelSole, T., J. Nattala, and M. K. Tippett, 2014: Skill improvement from increased ensemble size and model diversity. *Geophys. Res. Lett.*, **41**, 7331–7342, doi:10.1002/2014GL060133.
- Deng, Z., and Y. Tang, 2009: The retrospective prediction of ENSO from 1881 to 2000 by a hybrid coupled model: (II) Interdecadal and decadal variations in predictability. *Climate Dyn.*, **32**, 415–428, doi:10.1007/s00382-008-0398-2.
- Derber, J. D., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**, 1333–1347, doi:10.1175/1520-0485(1989)019<1333:AGODAS>2.0.CO;2.
- DeWitt, D. G., 2005: Retrospective forecasts of interannual sea surface temperature anomalies from 1982 to present using a directly coupled atmosphere–ocean general circulation model. *Mon. Wea. Rev.*, **133**, 2972–2995, doi:10.1175/MWR3016.1.
- Guan, Y., J. Zhu, B. Huang, Z.-Z. Hu, and J. L. Kinter III, 2014: South Pacific Ocean dipole: A predictable mode on multi-seasonal time scales. *J. Climate*, **27**, 1648–1658, doi:10.1175/JCLI-D-13-00293.1.
- Hu, Z.-Z., and B. Huang, 2007: The predictive skill and the most predictable pattern in the tropical Atlantic: The effect of ENSO. *Mon. Wea. Rev.*, **135**, 1786–1806, doi:10.1175/MWR3393.1.
- , A. Kumar, B. Huang, W. Wang, J. Zhu, and C. Wen, 2013: Prediction skill of monthly SST in the North Atlantic Ocean in NCEP Climate Forecast System version 2. *Climate Dyn.*, **40**, 2745–2756, doi:10.1007/s00382-012-1431-z.
- , —, —, J. Zhu, and Y. Guan, 2014: Prediction skill of North Pacific variability in NCEP Climate Forecast System version 2: Impact of ENSO and beyond. *J. Climate*, **27**, 4263–4272, doi:10.1175/JCLI-D-13-00633.1.
- Huang, B., and J. Shukla, 1997: Characteristics of the interannual and decadal variability in a general circulation model of the tropical Atlantic Ocean. *J. Phys. Oceanogr.*, **27**, 1693–1712, doi:10.1175/1520-0485(1997)027<1693:COTIAD>2.0.CO;2.
- , Z.-Z. Hu, and B. Jha, 2007: Evolution of model systematic errors in the tropical Atlantic Basin from coupled climate hindcasts. *Climate Dyn.*, **28**, 661–682, doi:10.1007/s00382-006-0223-8.
- Ji, M., A. Kumar, and A. Leetmaa, 1994: A multiseason climate forecast system at the National Meteorological Center. *Bull. Amer. Meteor. Soc.*, **75**, 569–577, doi:10.1175/1520-0477(1994)075<0569:AMCFSA>2.0.CO;2.
- Jin, E. K., and Coauthors, 2008: Current status of ENSO prediction skill in coupled ocean–atmosphere models. *Climate Dyn.*, **31**, 647–664, doi:10.1007/s00382-008-0397-3.
- Kanamitsu, M., and Coauthors, 2002: NCEP Dynamical Seasonal Forecast System 2000. *Bull. Amer. Meteor. Soc.*, **83**, 1019–1037, doi:10.1175/1520-0477(2002)083<1019:NDSFS>2.3.CO;2.
- Keenlyside, N. S., M. Latif, M. Botzet, J. Jungclaus, and U. Schulzweida, 2005: A coupled method for initializing El Niño–Southern Oscillation forecasts using sea surface temperature. *Tellus*, **57A**, 340–356, doi:10.1111/j.1600-0870.2005.00107.x.
- Kirtman, B. P., and D. Min, 2009: Multimodel ensemble ENSO prediction with CCSM and CFS. *Mon. Wea. Rev.*, **137**, 2908–2930, doi:10.1175/2009MWR2672.1.
- , and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, **95**, 585–601, doi:10.1175/BAMS-D-12-00050.1.
- Kumar, A., and M. P. Hoerling, 2000: Analysis of a conceptual model of seasonal climate variability and implications for seasonal prediction. *Bull. Amer. Meteor. Soc.*, **81**, 255–264, doi:10.1175/1520-0477(2000)081<0255:AOACMO>2.3.CO;2.
- , and Z.-Z. Hu, 2014: How variable is the uncertainty in ENSO sea surface temperature prediction? *J. Climate*, **27**, 2779–2788, doi:10.1175/JCLI-D-13-00576.1.
- , M. Chen, L. Zhang, W. Wang, Y. Xue, C. Wen, L. Marx, and B. Huang, 2012: An analysis of the nonstationarity in the bias of sea surface temperature forecasts for the NCEP Climate Forecast System (CFS) version 2. *Mon. Wea. Rev.*, **140**, 3003–3016, doi:10.1175/MWR-D-11-00335.1.
- , H. Wang, Y. Xue, and W. Wang, 2014: How much of monthly subsurface temperature variability in the equatorial Pacific can be recovered by the specification of sea surface temperatures? *J. Climate*, **27**, 1559–1577, doi:10.1175/JCLI-D-13-00258.1.
- Luo, J.-J., S. Masson, S. Behera, S. Shingu, and T. Yamagata, 2005: Seasonal climate predictability in a coupled OAGCM using a different approach for ensemble forecasts. *J. Climate*, **18**, 4474–4497, doi:10.1175/JCLI3526.1.
- , —, S. K. Behera, and T. Yamagata, 2007: Experimental forecasts of the Indian Ocean dipole using a coupled OAGCM. *J. Climate*, **20**, 2178–2190, doi:10.1175/JCLI4132.1.

- , —, —, and —, 2008: Extended ENSO predictions using a fully coupled ocean–atmosphere model. *J. Climate*, **21**, 84–93, doi:10.1175/2007JCLI1412.1.
- Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate*, **15**, 1609–1625, doi:10.1175/1520-0442(2002)015<1609:AHSAS>2.0.CO;2.
- Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483–3517, doi:10.1175/JCLI3812.1.
- , and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, doi:10.1175/2010BAMS3001.1.
- , and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, doi:10.1175/JCLI-D-12-00823.1.
- Stockdale, T. N., D. L. T. Anderson, J. O. S. Alves, and M. A. Balmaseda, 1998: Global seasonal rainfall forecasts using a coupled ocean–atmosphere model. *Nature*, **392**, 370–373, doi:10.1038/32861.
- Wang, G., R. Kleeman, N. Smith, and F. Tseitkin, 2002: The BMRC coupled general circulation model ENSO forecast system. *Mon. Wea. Rev.*, **130**, 975–991, doi:10.1175/1520-0493(2002)130<0975:TBCGCM>2.0.CO;2.
- Wang, H., A. Kumar, and W. Wang, 2013: Characteristics of subsurface ocean response to ENSO assessed from simulations with the NCEP Climate Forecast System. *J. Climate*, **26**, 8065–8083, doi:10.1175/JCLI-D-12-00795.1.
- Wang, W., S. Saha, H.-L. Pan, S. Nadiga, and G. White, 2005: Simulation of ENSO in the new NCEP Coupled Forecast System Model (CFS03). *Mon. Wea. Rev.*, **133**, 1574–1593, doi:10.1175/MWR2936.1.
- Xue, Y., M. Chen, A. Kumar, Z.-Z. Hu, and W. Wang, 2013: Prediction skill and bias of tropical Pacific sea surface temperatures in the NCEP Climate Forecast System version 2. *J. Climate*, **26**, 5358–5378, doi:10.1175/JCLI-D-12-00600.1.
- Yokoi, T., T. Tozuka, and T. Yamagata, 2008: Seasonal variation of the Seychelles Dome. *J. Climate*, **21**, 3740–3754, doi:10.1175/2008JCLI1957.1.
- Yuan, C., T. Tozuka, J.-J. Luo, and T. Yamagata, 2014: Predictability of the subtropical dipole modes in a coupled ocean–atmosphere model. *Climate Dyn.*, **42**, 1291–1308, doi:10.1007/s00382-013-1704-1.
- Zhang, R.-H., S. E. Zebiak, R. Kleeman, and N. Keenlyside, 2003: A new intermediate coupled model for El Niño simulation and prediction. *Geophys. Res. Lett.*, **30**, 2012, doi:10.1029/2003GL018010.
- Zhang, S., M. J. Harrison, A. Rosati, and A. Wittenberg, 2007: System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Mon. Wea. Rev.*, **135**, 3541–3564, doi:10.1175/MWR3466.1.
- Zheng, F., J. Zhu, H. Wang, and R.-H. Zhang, 2009: Ensemble hindcasts of ENSO events over the past 120 years using a large number of ensembles. *Adv. Atmos. Sci.*, **26**, 359–372, doi:10.1007/s00376-009-0359-7.
- Zhu, J., B. Huang, and M. A. Balmaseda, 2012a: An ensemble estimation of the variability of upper-ocean heat content over the tropical Atlantic Ocean with multi-ocean reanalysis products. *Climate Dyn.*, **39**, 1001–1020, doi:10.1007/s00382-011-1189-8.
- , —, L. Marx, J. L. Kinter III, M. A. Balmaseda, R.-H. Zhang, and Z.-Z. Hu, 2012b: Ensemble ENSO hindcasts initialized from multiple ocean analyses. *Geophys. Res. Lett.*, **39**, L09602, doi:10.1029/2012GL051503.
- , —, and Z. Wu, 2012c: The role of ocean dynamics in the interaction between the Atlantic meridional and equatorial modes. *J. Climate*, **25**, 3583–3598, doi:10.1175/JCLI-D-11-00364.1.
- , —, M. A. Balmaseda, J. L. Kinter III, P. Peng, Z.-Z. Hu, and L. Marx, 2013a: Improved reliability of ENSO hindcasts with multi-ocean analyses ensemble initialization. *Climate Dyn.*, **41**, 2785–2795, doi:10.1007/s00382-013-1965-8.
- , —, Z.-Z. Hu, J. L. Kinter III, and L. Marx, 2013b: Predicting U.S. summer precipitation using NCEP Climate Forecast System version 2 initialized by multiple ocean analyses. *Climate Dyn.*, **41**, 1941–1954, doi:10.1007/s00382-013-1785-x.
- , G.-Q. Zhou, R.-H. Zhang, and Z. Sun, 2013c: Improving ENSO prediction in a hybrid coupled model with an embedded entrainment temperature parameterization. *Int. J. Climatol.*, **33**, 343–355, doi:10.1002/joc.3426.
- , B. Huang, R.-H. Zhang, Z.-Z. Hu, A. Kumar, M. A. Balmaseda, L. Marx, and J. L. Kinter III, 2014: Salinity anomaly as a trigger for ENSO events. *Sci. Rep.*, **4**, 6821, doi:10.1038/srep06821.
- , and Coauthors, 2015: ENSO prediction in Project Minerva: Sensitivity to atmospheric horizontal resolution and ensemble size. *J. Climate*, **28**, 2080–2095, doi:10.1175/JCLI-D-14-00302.1.