

Improved Seasonal Prediction of Temperature and Precipitation over Land in a High-Resolution GFDL Climate Model

LIWEI JIA,* XIAOSONG YANG,* GABRIEL A. VECCHI,⁺ RICHARD G. GUDGEL,⁺ THOMAS L. DELWORTH,⁺ ANTHONY ROSATI,[#] WILLIAM F. STERN,⁺ ANDREW T. WITTENBERG,⁺ LAKSHMI KRISHNAMURTHY,* SHAOQING ZHANG,⁺ RYM MSADEK,* SARAH KAPNICK,[@] SETH UNDERWOOD,[&] FANRONG ZENG,⁺ WHIT G. ANDERSON,⁺ VENKATRAMANI BALAJI,[@] AND KEITH DIXON⁺

* University Corporation for Atmospheric Research, Boulder, Colorado, and NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

⁺ NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

[#] University Corporation for Atmospheric Research, Boulder, Colorado

[@] Princeton University, and NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

[&] Dynamics Research Corporation, Andover, Massachusetts, and NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

(Manuscript received 5 February 2014, in final form 15 October 2014)

ABSTRACT

This study demonstrates skillful seasonal prediction of 2-m air temperature and precipitation over land in a new high-resolution climate model developed by the Geophysical Fluid Dynamics Laboratory and explores the possible sources of the skill. The authors employ a statistical optimization approach to identify the most predictable components of seasonal mean temperature and precipitation over land and demonstrate the predictive skill of these components. First, the improved skill of the high-resolution model over the previous lower-resolution model in seasonal prediction of the Niño-3.4 index and other aspects of interest is shown. Then, the skill of temperature and precipitation in the high-resolution model for boreal winter and summer is measured, and the sources of the skill are diagnosed. Last, predictions are reconstructed using a few of the most predictable components to yield more skillful predictions than the raw model predictions. Over three decades of hindcasts, the two most predictable components of temperature are characterized by a component that is likely due to changes in external radiative forcing in boreal winter and summer and an ENSO-related pattern in boreal winter. The most predictable components of precipitation in both seasons are very likely ENSO-related. These components of temperature and precipitation can be predicted with significant correlation skill at least 9 months in advance. The reconstructed predictions using only the first few predictable components from the model show considerably better skill relative to observations than raw model predictions. This study shows that the use of refined statistical analysis and a high-resolution dynamical model leads to significant skill in seasonal predictions of 2-m air temperature and precipitation over land.

1. Introduction

Motivated by a desire to represent processes at more detailed scales, and enabled by developments in supercomputing capabilities and advanced numerical techniques, high-resolution climate models have been developed at various modeling centers (Delworth et al. 2012; Jung et al. 2012; Kinter et al. 2013; Shaffrey et al. 2009). High-resolution climate models, with the ability to better represent small-scale processes, show advantages in simulating many key

aspects of climate such as the El Niño–Southern Oscillation (ENSO), the Indian monsoon (Delworth et al. 2012), tropical precipitation, atmospheric circulation, and extratropical cyclones (Jung et al. 2012). At the Geophysical Fluid Dynamics Laboratory (GFDL) a stream of model development aiming to assess the impact of resolution on simulation and prediction has led to the development of a family of coupled climate models with different resolutions. The simulated climate in the high-resolution GFDL CM2.5 (hereinafter CM2.5; Delworth et al. 2012) showed remarkable improvements, including a reduction of the double intertropical convergence zone, improved simulations of ENSO and Amazonian rainfall, and a better representation of the geographic distribution

Corresponding author address: Liwei Jia, NOAA/Geophysical Fluid Dynamics Laboratory, 201 Forrestal Rd., Princeton, NJ 08540.
E-mail: liwei.jia@noaa.gov

of snow variables (Kapnick and Delworth 2013) over the coarser-resolution GFDL CM2.1 (hereinafter CM2.1; Delworth et al. 2006).

However, high-resolution coupled climate models are computationally expensive. It is worth exploring which elements of enhanced resolution are critical for each problem of interest. Inspired by the hypothesis that increased atmosphere and land resolution is critical for many of the improvements seen in the CM2.5 simulations over its lower-resolution predecessor CM2.1, GFDL recently developed a forecast-oriented climate model based on the fully coupled high-resolution CM2.5: the Forecast-Oriented Low Ocean Resolution version of CM2.5 (CM2.5-FLOR, hereinafter FLOR). This FLOR model has a high-resolution ($\sim 0.5^\circ$) atmosphere and land as in CM2.5, but a coarser-resolution (1°) ocean and sea ice as in CM2.1. FLOR was designed to substantially reduce computing time relative to CM2.5 to enable the running of large ensembles needed for climate predictions while still maintaining high-resolution land and atmosphere to allow exploration of regional climate and extremes. FLOR is the first high-resolution climate model used for operational seasonal forecasts in the United States. The hypothesis underlying the development of FLOR is that the atmosphere and land resolution is crucial for seasonal forecasts. Testing this hypothesis is one of the objectives of this study. While the impact of initialization of land and atmosphere on seasonal temperature and precipitation predictions has been previously explored in the literature (Paolino et al. 2012; Koster and Suarez 2003), this study focuses instead on the impact of the increased resolution of land and atmosphere components on seasonal predictions.

Another objective of this study is to measure the seasonal prediction skill of 2-m air temperature and precipitation over land and investigate the sources of the skill in the new high-resolution FLOR considering the societal and scientific values of seasonal predictions over land. A statistical optimization approach, called average predictability time (APT) (DelSole and Tippett 2009a,b; DelSole et al. 2011) is utilized. The APT approach is able to extract components with maximum predictability. One can refine the raw forecasts by filtering unpredictable components determined by the APT approach. Our hypothesis is that using refined statistical techniques can improve seasonal predictions.

Unlike predictions on multiyear-to-decadal time scales that are characterized by combined signals from internal climate variations and changes in external radiative forcing (Meehl et al. 2009; Taylor et al. 2012), predictions on seasonal scales, generally up to 12 months in length, are dominated by internal variability. However, if the assessment period for seasonal predictions spans several decades, the externally forced climate signal can

be prominent. Hence, seasonal climate predictions could be a joint initial-boundary value problem (Doblas-Reyes et al. 2013), similar to the decadal predictions. Distinguishing the role of externally forced changes from internal variability in the decadal predictions has been well studied (Smith et al. 2007; Solomon et al. 2011; Yang et al. 2013) but has not been well documented in the seasonal predictions. In this study, we employ the APT method to isolate predictable patterns on different time scales in the seasonal hindcasts and investigate the roles of external forcing and internal climate variability in the seasonal predictions.

The rest of the paper is organized as follows: The model and data are introduced in section 2. The methodology is described in section 3. Our results are discussed in section 4 and summarized in section 5.

2. Model and data

a. Model description

To test the hypothesis that atmosphere and land resolution is critical for seasonal predictions, we employ the new high-resolution FLOR, which has high resolution in the land and atmosphere components. FLOR is a combination of two previously described coupled model configurations. The atmosphere [GFDL Atmospheric Model version 2.5 (AM2.5)] and land models are those used in the high-resolution CM2.5 (Delworth et al. 2012), which have an approximately 0.5° spatial resolution. However, in contrast to CM2.5, which has high resolution in both its atmosphere and ocean components, the ocean and sea ice components of FLOR are at 1° resolution. The ocean and sea ice components of FLOR are based on those of the low-resolution CM2.1 (Delworth et al. 2006), which has been used extensively for climate research, predictions, and projections for close to a decade. In FLOR, the ocean component has been slightly altered from that of CM2.1 by having a more realistic representation of the solar absorption by the ocean using a biharmonic horizontal viscosity scheme, as well as some fixes documented in Delworth et al. (2012). FLOR also incorporates a newer and higher-order advection scheme used in CM2.5 and an updated parameterization for eddies. Additional description of aspects of FLOR will be described in a series of papers, including Vecchi et al. (2014), Winton et al. (2014), Yang et al. (2015), Msadek et al. (2014), and A. Wittenberg et al. (2014, unpublished manuscript).

Two versions of FLOR (viz., A06 and B01) have joined the North American Multimodel Ensemble (NMME) seasonal forecasting system, which is currently delivering real-time seasonal-to-interannual predictions (<http://www.cpc.ncep.noaa.gov/products/NMME>). These two versions have identical atmospheric, land, and sea ice configurations but

have slightly different ocean configurations (Vecchi et al. 2014). We used the B01 version in this study. FLOR has been shown to produce skillful seasonal forecasts of tropical cyclone activity, sea ice extent, and extratropical storm tracks (Vecchi et al. 2014; Msadek et al. 2014; Yang et al. 2015). We will show in section 4 that FLOR exhibits higher seasonal skill of the Niño-3.4 index, 2-m air temperature, and precipitation than CM2.1, which is also included in the NMME and is comparable to other NMME models in seasonal forecast skill of 2-m air temperature, precipitation, sea surface temperature (SST), and the Niño-3.4 index (Becker et al. 2014).

b. Control simulations, historical simulations, and ensemble hindcasts

Control simulations of CM2.1, FLOR, and CM2.5 with constant radiative forcing agents and land surface conditions appropriate to 1990 conditions, were used to explore the temperature and precipitation regressions onto Niño-3 SST anomalies. We analyzed the data during years 11–290, 11–290, and 11–270 for CM2.1, FLOR, and CM2.5, respectively. The 20-yr running mean was removed from the control runs, which reduces climate drift and decadal variability. The variability in the control runs is merely internal variability in the climate system.

Five members of historical simulations for CM2.1, FLOR, and CM2.5 were created individually using initial conditions obtained from widely separated (40 yr) points in their corresponding multicentury control simulations. The multicentury control simulations were forced by fixed 1860 radiative forcing agents and land surface conditions. The time-varying historical forcing agents for the historical runs are atmospheric CO₂, CH₄, N₂O, halons, tropospheric and stratospheric O₃, anthropogenic tropospheric sulfates, black and organic carbon, solar irradiance, land use, and volcanic aerosols (Knutson et al. 2006). We analyzed five ensemble members over 1981–2010 for each model.

A series of ensemble seasonal hindcasts were generated for FLOR and CM2.1 separately. The initial conditions for CM2.1 are based on the GFDL's Ensemble Coupled Data Assimilation (ECDA) system (Zhang et al. 2007; Zhang and Rosati 2010; Chang et al. 2013). The ECDA covers the period from 1960 to the present and is being updated monthly for GFDL's seasonal-to-decadal experimental forecasts (Yang et al. 2013; Vecchi et al. 2013). A comprehensive assessment of oceanic variability from the latest version of ECDA analyzed from 1960 to 2010 can be found in Chang et al. (2013). The initial conditions for the ocean and ice components in FLOR are taken from the ECDA developed specifically for CM2.1. The initial conditions for the atmosphere and land components in FLOR are taken from AM2.5 simulations forced by observed SST. Before 2005, FLOR and CM2.1

are forced with observationally based estimates of changing concentrations of greenhouse gases, aerosols, land-use changes, solar irradiance variations, and volcanic aerosols. After 2005, FLOR and CM2.1 are forced with estimates of changing greenhouse gases and aerosols based on the representative concentration pathway 4.5 (RCP4.5) scenario (Meinshausen et al. 2011). Such time-varying external forcings are usually considered in operational seasonal forecast (Saha et al. 2014). For FLOR (CM2.1), 12 (10) ensemble members are initialized on the first day of each month from the 1980s to the present and are integrated for 12 months. Ensemble hindcasts during the period of 1980–2012 (1982–2011) for FLOR (CM2.1) are analyzed to estimate seasonal skill.

c. Observational data

The observations used in this study are precipitation at 0.5° resolution from the National Oceanic and Atmospheric Administration (NOAA)'s precipitation reconstruction over land (Chen et al. 2002); the CPC Merged Analysis of Precipitation (CMAP) at 2.5° resolution; the Global Historical Climatology Network (GHCN) gridded v2 dataset 2-m air temperature over land at 0.5° resolution (Fan and van den Dool 2008); the Hadley Centre Sea Ice and Sea Surface Temperature dataset (Rayner et al. 2003); and the NCEP–NCAR Reanalysis-1 surface temperature and precipitation at 2.5° resolution. Velocity and sea level pressure data are from the Modern-Era Retrospective Analysis for Research and Applications (MERRA) (Rienecker et al. 2011). The observed Niño-3.4 index was downloaded from NOAA's website (<http://www.cpc.ncep.noaa.gov/data/indices/ersst3b.nino.mth.81-10.ascii>).

3. Review of methodology

To identify predictable components, we employ the technique of average predictability time, proposed by DelSole and Tippett (2009a,b). This approach determines components that maximize predictability. To find such components, one must first define a measure of predictability. Consider a set of ensemble forecasts, initialized at distinct start times. Ensemble members at each start time are generated with slightly different initial conditions. In this context, a standard measure of predictability is

$$P(\tau) = 1 - \frac{\sigma_{\tau}^2}{\sigma_{\text{clim}}^2}, \quad (1)$$

where σ_{τ}^2 is the forecast variance at lead time τ , averaged over all start times. As lead time increases, the forecast variance tends to increase and approaches the climatological variance σ_{clim}^2 as lead time τ approaches

infinity. The climatological variance is practically the variance of all ensemble forecasts from different start times. The measure $P(\tau)$ typically is close to one initially and decreases with lead time until it vanishes when the forecast variance equals the climatological variance. It can be shown that $P(\tau)$ is related to signal-to-noise ratio (Jia 2011), which is often used to measure predictability. APT is then defined as twice the sum of $P(\tau)$ over all lead times:

$$\text{APT} = 2 \sum_{\tau=1}^{\infty} \left(1 - \frac{\sigma_{\tau}^2}{\sigma_{\text{clim}}^2} \right). \quad (2)$$

It yields a time scale that agrees with the usual e -folding time for an exponentially decaying forecast signal. In reality, the sum in (2) is generally from lead time 1 to finite lead times. In this study, the sum is from lead 1 to 10 months. Let $\mathbf{x}(\tau, t, e)$ be the state vector specifying the amplitudes of ensemble forecasts at fixed lead time τ , start time t , and ensemble member e . We seek the projection \mathbf{q} such that $\mathbf{q}^T \mathbf{x}(\tau, t, e)$ maximizes APT, where the superscript T denotes the transpose operation. The forecast variance of $\mathbf{q}^T \mathbf{x}(\tau, t, e)$ at lead time τ is

$$\begin{aligned} \sigma_{\tau}^2 &= \mathbf{q}^T [\overline{\mathbf{x}(\tau, t, e) - \langle \mathbf{x}(\tau, t, e) \rangle}][\overline{\mathbf{x}(\tau, t, e) - \langle \mathbf{x}(\tau, t, e) \rangle}]^T \mathbf{q} \\ &= \mathbf{q}^T \hat{\Sigma}_{\tau} \mathbf{q}, \end{aligned} \quad (3)$$

where the angle brackets denote the average over ensemble members. The overline denotes the average over start times, and $\hat{\Sigma}_{\tau}$ denotes an estimate of the forecast covariance matrix at lead time τ . The climatological variance is just the forecast variance as the lead time approaches infinity. It is denoted as

$$\sigma_{\text{clim}}^2 = \mathbf{q}^T \hat{\Sigma}_{\text{clim}} \mathbf{q}, \quad (4)$$

where $\hat{\Sigma}_{\text{clim}}$ is an estimate of the climatological covariance matrix. Substituting (3) and (4) into (2) gives

$$\text{APT} = 2 \sum_{\tau=1}^{\infty} \left[\frac{\mathbf{q}^T (\hat{\Sigma}_{\text{clim}} - \hat{\Sigma}_{\tau}) \mathbf{q}}{\mathbf{q}^T \hat{\Sigma}_{\text{clim}} \mathbf{q}} \right]. \quad (5)$$

It can be shown that maximizing (5) leads to a generalized eigenvalue problem:

$$2 \sum_{\tau=1}^{\infty} (\hat{\Sigma}_{\text{clim}} - \hat{\Sigma}_{\tau}) \mathbf{q} = \lambda \hat{\Sigma}_{\text{clim}} \mathbf{q}. \quad (6)$$

The eigenvalue λ gives APT value, and each eigenvector \mathbf{q} corresponds to a component. Equation (6) above has more than a single solution; thus, it yields multiple components. The time series of each component is uncorrelated with

one another because of the symmetry characteristic of $\hat{\Sigma}_{\tau}$ and $\hat{\Sigma}_{\text{clim}}$. We order the eigenvalues and their associated eigenvectors by decreasing order such that the first component maximizes APT, the second maximizes APT subject to being uncorrelated with the first, and so on. This decomposition based on APT is analogous to empirical orthogonal function (EOF) analysis, but instead of decomposing variance, we decompose predictability. APT is essentially the sum of signal-to-total ratio over lead times that can be used to measure predictability. In general, components that are persistent (e.g., trend) or oscillate over narrow frequencies have large APT. Components with different APT values often vary on different time scales. The spatial pattern \mathbf{p} associated with each component is obtained by regressing the time series $\mathbf{q}^T \mathbf{x}(\tau, t, e)$ with $\mathbf{x}(\tau, t, e)$, which gives $\mathbf{p} = \hat{\Sigma}_{\text{clim}} \mathbf{q}$. The spatial patterns are not orthogonal with one another. Note that the APT method is discussed here in the context of ensemble forecasts, as we will apply it to the ensemble hindcast data in this study. The APT approach is also applicable to a single simulation: for instance, a control simulation. More details of this APT technique are found in DelSole and Tippett (2009a,b), and the application of this technique can be found in Jia and DelSole (2011, 2012) and Yang et al. (2013).

For typical global forecast data, the number of grid points exceeds the number of samples (i.e., time steps), so the above covariance matrices are singular, and the eigenvalue problem cannot be solved. A standard approach is to project the forecast data onto a few leading principal components (PCs) and then to maximize APT only in the subspace spanned by the leading PCs. In this paper, we chose 30 PCs for 2-m air temperature and precipitation. The sensitivity of APT values to the number of PCs has been tested. The APT values are not sensitive when using more than 20 PCs. Since precipitation variance can vary substantially even between neighboring grid points, the PCs, which maximize variance, might over-emphasize grid points with large variance and capture small-scale structures. The APT analysis, which includes a linear combination of PCs, hence, is influenced. We thus derive the PCs from precipitation hindcasts in which each grid point is normalized by its own standard deviation. The normalized precipitation is able to capture large-scale precipitation structures.

To determine whether a component is predictable or not, we tested the significance of the APT value of the component using Monte Carlo methods under the null hypothesis that the component is unpredictable. A component is then considered to be predictable if its APT value is statistically significant at a certain significance level. The application of Monte Carlo methods can be found in DelSole et al. (2011), Jia and DelSole (2011),

and Yang et al. (2015). Basically, for M spatial dimensions, N time steps, E ensemble members, I initial conditions, and L lead times, we generated a $M \times N \times E \times I \times L$ data matrix by drawing independent random numbers from a normal distribution with zero mean and unit variance. APT analysis was applied to this matrix to produce an ordered sequence of APT values. The procedure was repeated 1000 times. The 5% significance level for each APT associated with each component was determined by selecting the 95th percentile of the APT values derived from the random data.

The squared error skill score (hereinafter SESS) is used to measure skill (Goddard et al. 2013), which is defined as

$$\text{SESS} = 1 - \frac{\sum_n (O_n - P_n)^2}{\sum_n (O_n - \bar{O})^2}, \quad (7)$$

where O_n is the centered observation at time n , P_n is the centered prediction of O_n , and \bar{O} is the time mean of O_n for all years (i.e., climatological mean). The value of SESS is one for a perfect prediction and is negative if the error of a prediction is larger than the error of a prediction based on the climatological mean.

4. Results

a. Climate mean state, variability, and prediction skill

Figure 1 shows annual mean precipitation and 2-m air temperature in observations and historical simulations of FLOR and CM2.1 during 1981–2010, as well as the bias in FLOR and CM2.1. Compared to observed mean precipitation, FLOR and CM2.1 show dry biases over most of South America, although FLOR shows less bias than CM2.1. FLOR also simulates mean precipitation better than CM2.1 over tropical Africa, eastern China, and the southeastern United States. As for annual mean 2-m air temperature, the observed pattern is well simulated in FLOR and CM2.1, but both models show cold biases in most areas. FLOR shows less bias along the Andes and in equatorial Africa than CM2.1.

Besides the annual mean temperature and precipitation over land, we also examined seasonal mean climate and variability of several important variables over the globe. Figure 2 shows the scatterplot of pattern correlation of seasonal mean climate and standard deviation between observation and historical simulations for FLOR versus CM2.1 and FLOR versus CM2.5. A correlation coefficient above the diagonal line indicates that FLOR has higher correlation skill than CM2.1 or CM2.5. As seen from Figs. 2a and 2c, the correlation skill is higher in FLOR than in CM2.1 for nearly all variables

and seasons, both in mean climate and standard deviation. In contrast, the correlations in FLOR are comparable to those in CM2.5 for seasonal mean climate (Fig. 2b) and are slightly higher than CM2.5 for standard deviation (Fig. 2d), although the ocean resolution is increased in CM2.5. The fact that mean climate and climate variability are improved considerably in FLOR over CM2.1, but moderately in CM2.5 over FLOR, supports the hypothesis that atmosphere and land resolution, rather than ocean resolution, is critical to the improvements in these quantities.

Another phenomenon worth examining is ENSO and its teleconnections to remote regions. Figure 3 shows the patterns of correlation between Niño-3 sea surface temperature anomalies (SSTA) and the global anomalies of surface temperature and precipitation for observations, a 280-yr 1990-control simulation of FLOR and CM2.1, and a 260-yr 1990-control simulation of CM2.5. In the observations, Niño-3 SSTA is strongly correlated with both surface temperature and precipitation over the equatorial central and eastern Pacific. These strong correlations are well simulated in all three models. However, this zone of positive correlation extends farther west than observed in all three simulations, though the westward extent is somewhat reduced in FLOR and CM2.5 relative to CM2.1. Compared to CM2.1, FLOR and CM2.5 also show less extreme temperature correlations over Australia, the Amazon region, southern Africa, the tropical Atlantic and Indian Oceans, and the Southern Ocean, which are more in line with observed values. In all three simulations, the negative temperature correlations over the contiguous United States, Argentina, China, the off-equatorial west Pacific, and the southern Atlantic appear to be too strong, as are the positive temperature correlations over equatorial land areas. For precipitation, regions of both positive and negative correlations are generally stronger than estimated from observations. In particular, the negative precipitation correlations over the Maritime Continent and tropical Atlantic are too strong in all three simulations. The negative precipitation correlations over Australia, the Amazon, and southern Africa and positive correlations over the equatorial Pacific and Indian Oceans, contiguous United States, Argentina, and western Asia are also too strong in all three simulations, but these are somewhat reduced in FLOR and CM2.5. Consistent with our earlier results, the simulated temperature and precipitation correlation patterns appear to be influenced more by the increase in atmosphere and land resolution (going from CM2.1 to FLOR) than by the increase in ocean resolution (going from FLOR to CM2.5). In addition, the spectrum of the Niño-3.4 SSTA in FLOR (not shown) exhibits a prominent, broad-spectrum spectral peak near 2.8 years, which is a somewhat shorter time scale than the observed peak of 3.2

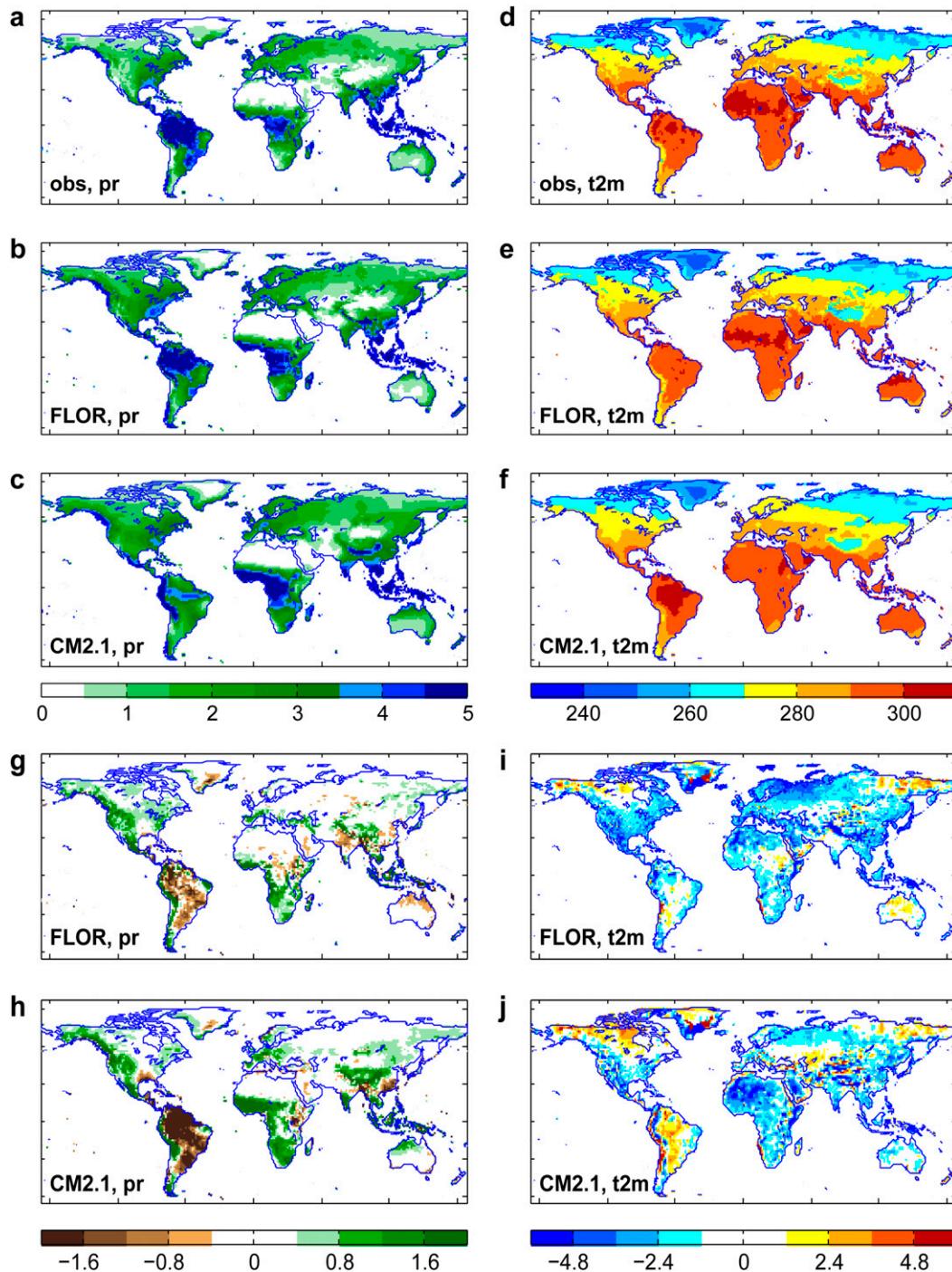


FIG. 1. Annual mean (a)–(c) precipitation (in mm day^{-1}) and (d)–(f) 2-m air temperature (in K) in observations, FLOR, and CM2.1 historical simulations during 1981–2010 and the bias of annual mean (g),(h) precipitation (in mm day^{-1}) and (i),(j) 2-m air temperature (in K) in FLOR and CM2.1.

years and the CM2.1 peak of 3.3 years. The ENSO SST anomaly variance along the equator in FLOR is stronger than observed, especially in the central Pacific, although both biases are reduced compared to CM2.1 (not shown).

To examine ENSO prediction skill, we show in Fig. 4 the SESS values and anomaly correlations of the Niño-3.4 index as a function of initial month and target month in FLOR and CM2.1. Both FLOR and CM2.1

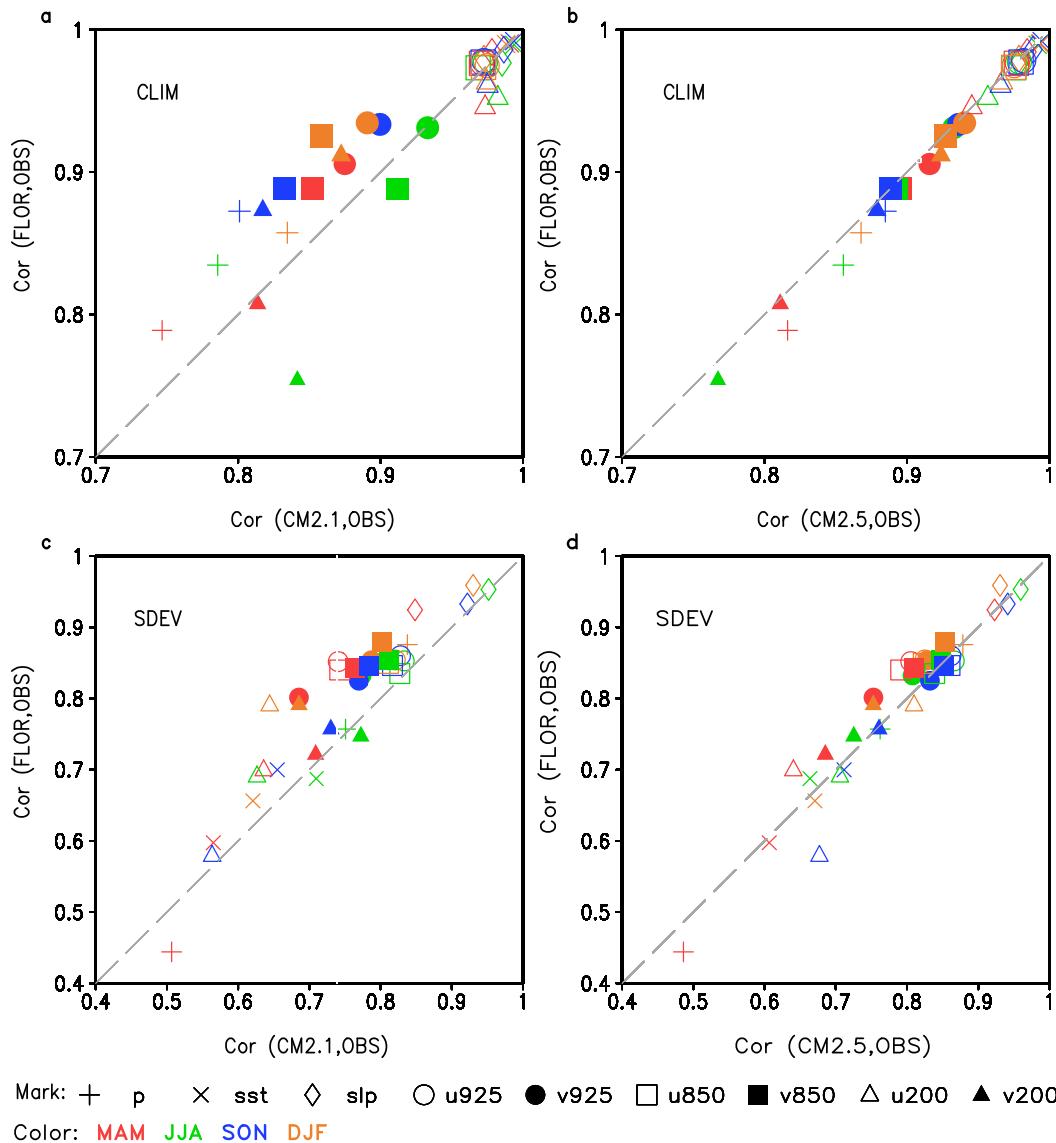


FIG. 2. Scatterplot of pattern correlation between CM2.1 historical runs (1981–2010) and observation (x axis) vs FLOR historical runs (1981–2010) and observation (y axis) for (a) seasonal mean climate and (c) std dev; and between CM2.5 historical runs (1981–2010) and observation (x axis) vs FLOR historical runs and observation (y axis) for (b) seasonal mean climate and (d) std dev of the following variables: global precipitation, sea surface temperature, sea level pressure, and zonal and meridional velocities at 925, 850, and 200 hPa. Different colors indicate different seasons. Each symbol represents a particular variable.

show very high correlation skill. FLOR shows slightly higher correlations than CM2.1 at short leads for initial months from August to December. However, the SESS values are much larger in FLOR than in CM2.1 in boreal winter and spring. It follows that FLOR has higher skill in predicting the Niño-3.4 index than CM2.1 in those two seasons. Such skill improvements in SESS are not seen in anomaly correlation, indicating that conditional bias is reduced in FLOR. In other words, CM2.1 and FLOR

have comparable skill in capturing the phasing of ENSO, but the amplitude of ENSO is captured better by FLOR than by CM2.1. We also found that the SESS values at long leads initialized in November–January are lower in FLOR than those in CM2.1. We suspect this low skill might be because the initial conditions for FLOR are taken from the ECDA developed for CM2.1, rather than a coupled assimilation with FLOR itself. It will be shown shortly that the reduced SESS at long leads in FLOR

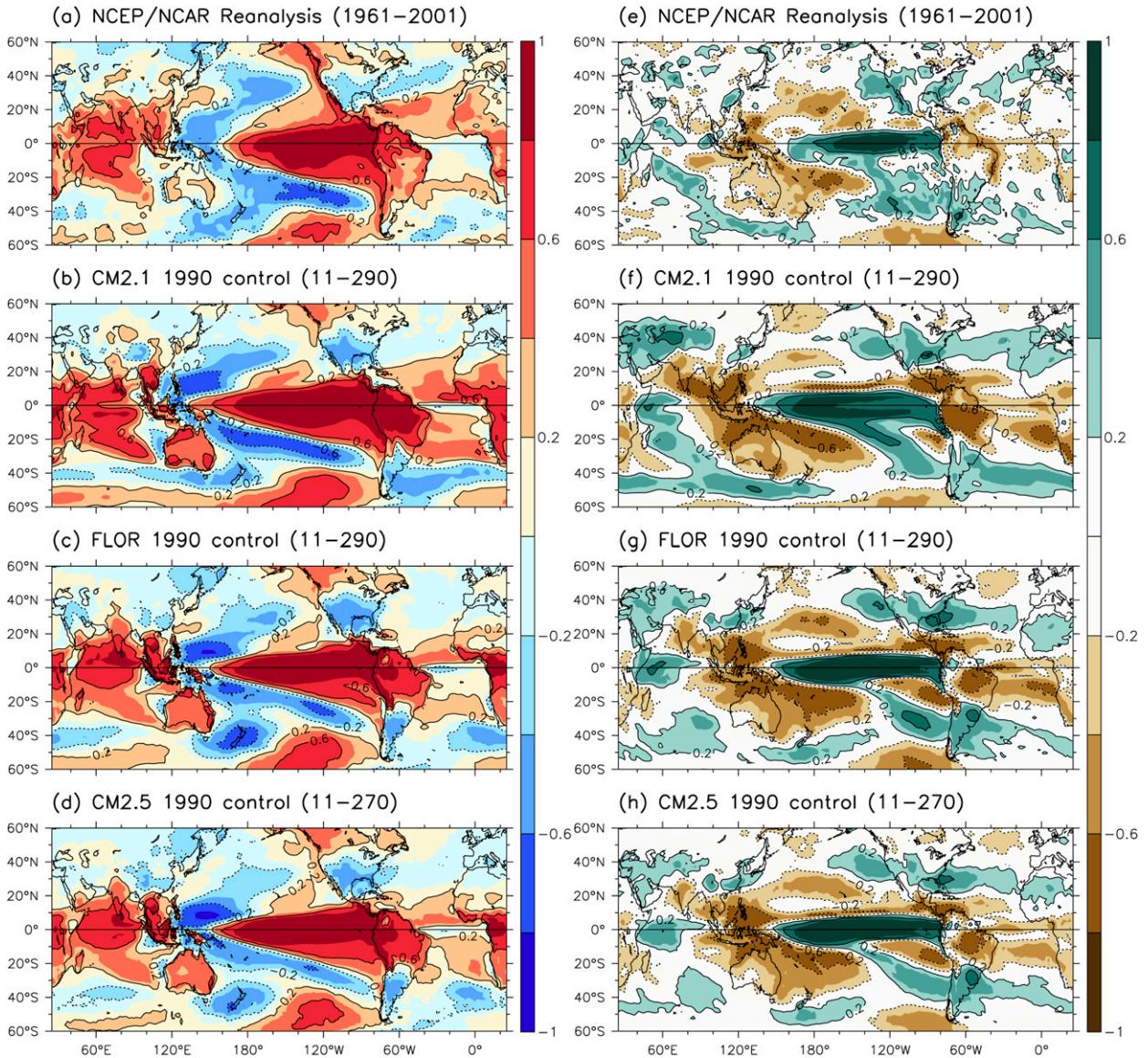


FIG. 3. Correlation between (left) local surface temperature and (right) precipitation anomalies and Niño-3 (150° – 90° W, 5° S– 5° N) SST anomalies, for annual (June–May) means from (a),(e) NCEP–NCAR Reanalysis-1 (1961–2001), and 1990-control runs from (b),(f) CM2.1 (years 11–290); (c),(g) FLOR (years 11–290); and (d),(h) CM2.5 (years 11–270). Anomalies are computed by subtracting a 20-yr running mean from the original June–May annual mean temperature time series, which, in addition to removing 25%, 50%, and 75% of the amplitude at periods of 25, 33, and 49 yr, also truncates the initial and final decades from the anomaly time series.

might account for the low skill of the most predictable precipitation pattern at long leads in FLOR, since the most predictable precipitation pattern is significantly correlated with the Niño-3.4 index.

We now identify predictable components of precipitation over land in FLOR and CM2.1 hindcasts using APT analysis. A 3-month running mean was applied to the hindcasts to remove subseasonal variability. Hence, for each initial month, there are 10-months leads. The hindcasts for all 12 initial months and 10 lead months

were used in the APT analysis, which gives a dataset of 3960 (3600) months in length for FLOR (CM2.1). The most predictable patterns of precipitation for FLOR and CM2.1 are shown in Figs. 5b and 5c, respectively. The APT value of the most predictable pattern for FLOR (CM2.1) is 16.1 (15.8) months, which is statistically significant at a 5% significance level based on the Monte Carlo methods. The most predictable patterns are strikingly similar to the precipitation regression patterns on the Niño-3.4 index in both models (not shown). Moreover, the time series of the

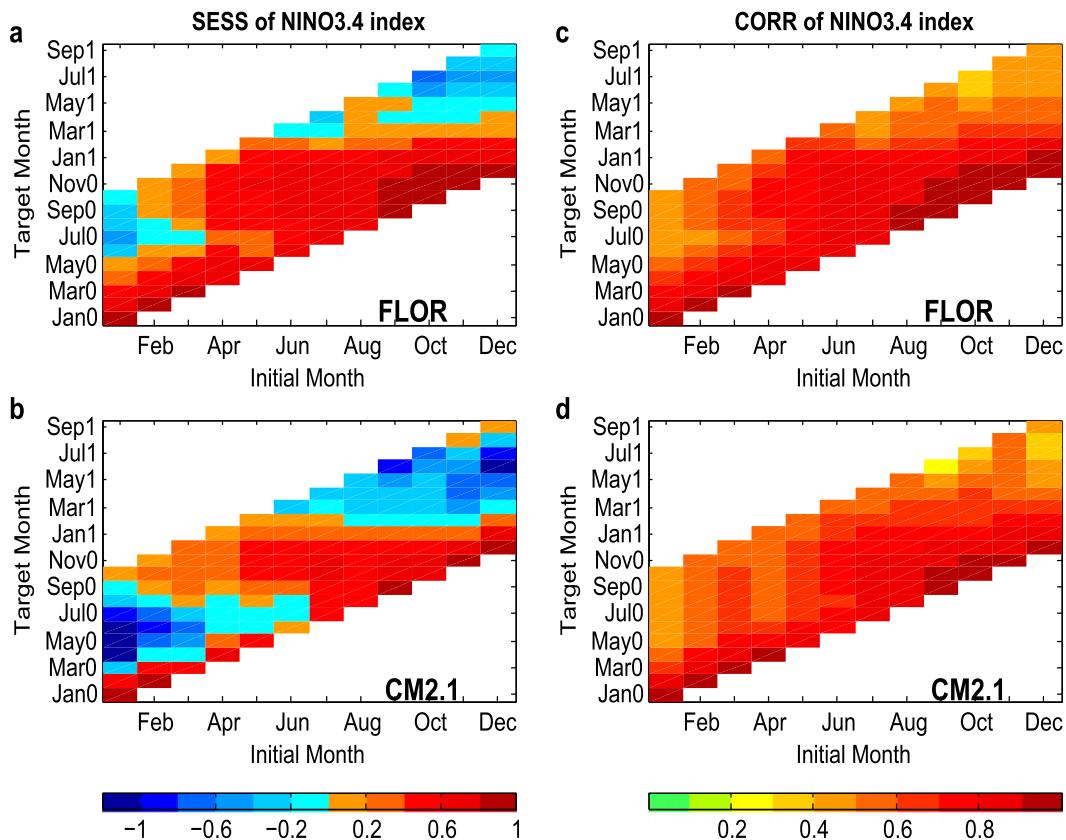


FIG. 4. (a),(b) Squared error skill score and (c),(d) anomaly correlation of the Niño-3.4 index in FLOR and CM2.1 for each initial month and target month during 1981–2010 (1983–2010) in FLOR (CM2.1). Each target month indicates a 3-month mean (e.g., target month of January denotes January–March mean).

most predictable patterns in observations are significantly correlated with the observed Niño-3.4 index (correlation coefficient is 0.7 for FLOR and 0.61 for CM2.1) in both models, where the time series in observations for each model is derived by projecting the projection vector of the most predictable component onto the observed precipitation. Thus, to evaluate the predictable patterns emerging from the model, we compare the patterns to the observed precipitation regression pattern with the Niño-3.4 index (Fig. 5a). The most predictable pattern in FLOR is much closer to the observed precipitation regression pattern than CM2.1, particularly in South America, the southern United States, eastern China, Australia, and eastern Africa. A close comparison of the patterns in low latitudes of the Americas and Asia reveals that FLOR is able to capture small-scale structures near the Andes and in tropical Asian islands. The arc-shaped pattern in northern Australia is also better captured in FLOR than CM2.1. The wet areas in eastern China are weaker in CM2.1 than those in FLOR and observations, and CM2.1 predicts the east coast of equatorial Africa with the wrong sign.

To measure the prediction skill of the most predictable pattern, we project the projection vector \mathbf{q} of the

most predictable pattern onto the observed precipitation. The resulting time series was used to calculate the SESS of the most predictable pattern. The SESS values of the ensemble mean prediction of the most predictable pattern for FLOR and CM2.1 are shown in Figs. 5e and 5f as a function of initial month and target month. Recall that a 3-month running mean was applied to the data; each target month in the figure indicates the start month of a 3-month mean. The SESS values are much larger in FLOR than in CM2.1 in nearly all initial and target months, especially for target months from October to the following March. To gain insight into the striking improvements of retrospective skill in FLOR over CM2.1, we compute the standard deviation of the time series of the most predictable pattern and normalize it relative to the time series of the pattern in observations for FLOR and CM2.1, respectively. A resulting value close to one implies that the predicted variability is close to the observed variability, although the predictions and observations can be out of phase. But the out-of-phase case is penalized in the measure of SESS (i.e., leads to small SESS values). Figure 5d shows the normalized standard deviation as a function of target

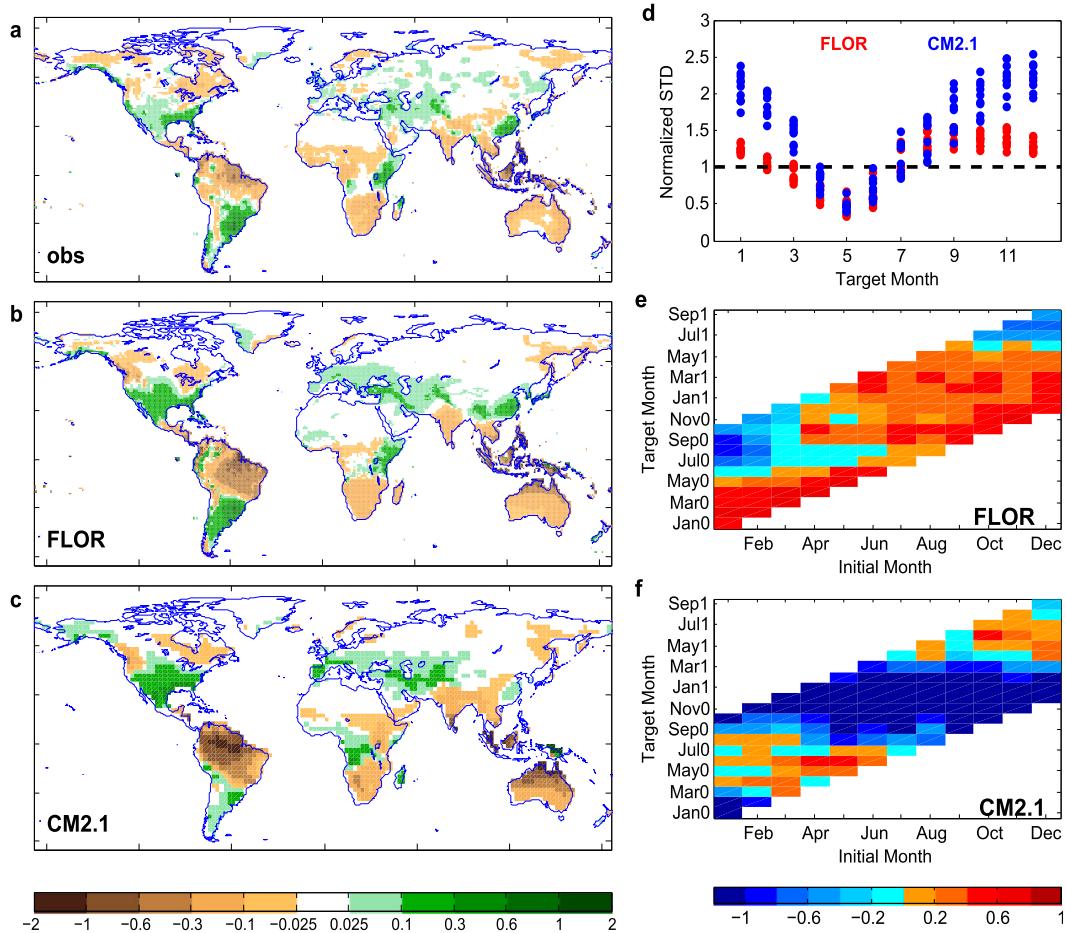


FIG. 5. (a) Observed precipitation regression pattern with the Niño-3.4 index (in mm day^{-1} per unit std dev); spatial structure of the most predictable component of precipitation over land (in mm day^{-1} per unit std dev) from (b) FLOR and (c) CM2.1; (d) std dev of time series of the most predictable pattern for different target months and initial months, normalized relative to the observations; squared error skill score of the most predictable pattern for each initial month and target month in (e) FLOR and (f) CM2.1. Each dot in (d) represents a particular initial month.

month at different lead months in FLOR and CM2.1. At a specific target month, each dot in the figure denotes a particular lead month from 0 to 9. The normalized standard deviations in FLOR are closer to 1 than those of CM2.1 at target months from October to the following March, and the corresponding SESS values are much larger in FLOR compared to CM2.1, implying that the variability of the predictable pattern in those months is better predicted by FLOR than by CM2.1. In other words, the conditional biases are considerably reduced in FLOR relative to CM2.1, which leads to higher SESS values. Skill improvements of FLOR over CM2.1 are also found for predictable components of 2-m air temperature over land (not shown). As described in section 2, the data assimilation used in FLOR was taken from the ECDA in CM2.1. We expect that skill can be further improved once the

ECDA based on FLOR is available. Note that, although the prediction skill improves in most cases in FLOR, we do find a few cases where the SESS values are lower than CM2.1, such as the long leads initialized in October, November, December, and January. Such low skill in FLOR might be associated with the low skill in ENSO prediction, as shown in Fig. 4.

b. Predictable components of precipitation over land in FLOR on seasonal scales

The improved predictions in FLOR motivate us to further explore the prediction skill of precipitation and temperature over land in this high-resolution model on seasonal scales, considering that patterns of precipitation and temperature vary with seasons. Results for seasonal mean predictions in December–February (DJF) and June–August (JJA) are discussed in this section.

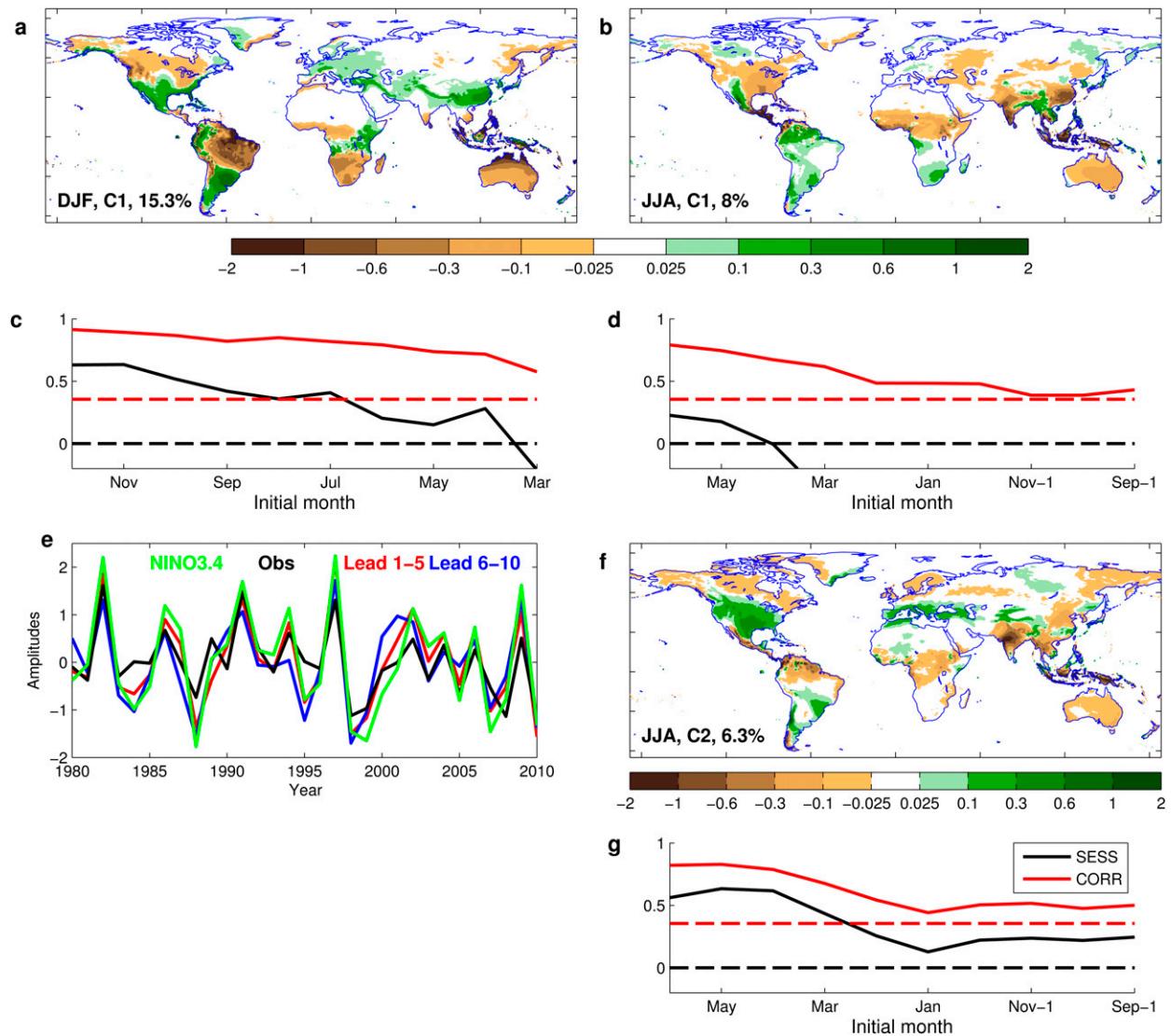


FIG. 6. Spatial structure of (a) the first predictable component of precipitation over land (in mm day^{-1} per unit std dev) in DJF and (b),(f) the first two predictable components in JJA; (c),(d),(g) the corresponding SESS (black solid) and anomaly correlation (red solid) skill as a function of initial month; (e) the time series of the most predictable component in DJF for observation (black), leads 1–5 (red), leads 6–10 (blue), and the observed Niño-3.4 index. The red dash lines indicate the 5% significance level for anomaly correlation. The black dash lines indicate zero SESS. The percentage denotes the variance explained by the component.

The spatial patterns of the first predictable component in DJF and the first two predictable components in JJA derived from hindcasts during 1980–2012 are shown in Figs. 6a, 6b, and 6f, as these components have significant skill. Here, only the hindcasts at a specific target season (DJF or JJA) were utilized to estimate the pattern, which gives a dataset of 330 months in length. The first predictable pattern in DJF shows wet anomalies in eastern China, southern North America into South America, southeast Africa, the Andes, and dry anomalies in northern Australia, southern Africa, and northeastern South America. In JJA, the first predictable component shows

dry anomalies over India, eastern China, eastern Australia, the Sahel, and Central America. The second component in JJA shows dry conditions over India and northern South America but wet conditions over large areas of the United States. The predictable patterns diagnosed here are in large agreement with the land precipitation teleconnection pattern to ENSO, as shown in previous studies (Ropelewski and Halpert 1996; Yang and DelSole 2012). In fact, the SST regression pattern on these components displays a classic ENSO structure with the largest loadings over the tropical Pacific (not shown). To confirm the relationship between these

predictable components and ENSO, we show in Fig. 6e the observed Niño-3.4 index and the time series of the most predictable pattern in DJF for observation and model hindcasts (averaged over the first 5 leads and the second 5 leads). Both the time series of the pattern in model and observations are highly correlated with the observed Niño-3.4 index. The correlation coefficient between the time series of the pattern in observations and the observed Niño-3.4 index is 0.88. In JJA, the correlations between the time series of the first two components in observations and the observed Niño-3.4 index are also statistically significant with correlation coefficients of 0.43 and 0.69 for the first and the second component, respectively. Therefore, these predictable components of precipitation over land are likely ENSO related.

An important question is whether the predictable components diagnosed in model system exist in the real world, since components that emerge in a dynamical model need not exist in nature (i.e., observations). To address this question, we show in Figs. 6c, 6d, and 6g the anomaly correlations between the time series of the predictable components from model hindcasts and those from observations, as well as the SESS of the components. The correlations are statistically significant at the 5% significance level at all initial months in both seasons based on Student's *t* test. However, the SESS values that take into account conditional bias, drop considerably for March initial condition in all cases. The skill drop of predictions initialized in March is presumably due to the spring barrier of ENSO prediction (Barnston et al. 2012). Based on the strong correlations between the time series of the predictable components in model and observations, we argue that the predictable components diagnosed in the dynamical model do exist in the real world and can be predicted with significant skill at least 9 months in advance.

c. Predictable components of air temperature over land in FLOR on seasonal scales

The spatial patterns of the most predictable component of 2-m air temperature in DJF and JJA derived from hindcasts during 1980–2012 are shown in Figs. 7a and 7b. The most predictable components in both seasons show positive amplitudes nearly everywhere, except for a few limited areas in DJF. Areas with maximum amplitudes vary with seasons. In DJF, large loadings are located in the high latitudes of North America, central South America, South Africa, and Australia. In JJA, maximum loadings are located in central North America, Greenland, northern Africa, and central Eurasia. The time series of the most predictable component averaged over lead 1–5 and lead 6–10, as well as the time series in

observations show an increasing trend in both seasons (Figs. 7e,f). Therefore, the most predictable component of 2-m air temperature indicates a multidecadal warming signal in both seasons.

To explore the mechanism of the most predictable component, we estimated externally forced patterns of 2-m air temperature over land in DJF and JJA from FLOR historical runs, using a signal-to-total maximizing EOF method (Ting et al. 2009). As shown in Fig. 8, the externally forced patterns bear great similarity to the most predictable patterns in both seasons, with a pattern correlation of 0.68 for DJF and 0.76 for JJA, suggesting that the most predictable components are likely the response to the changes in external radiative forcing. However, this does not exclude the possibility that other mechanisms might be responsible for the low-frequency variability of the most predictable components. For instance, in JJA, in spite of the similarities between the most predictable pattern and the externally forced pattern, differences are seen in North America and Europe. The differences might be attributable to the Atlantic multidecadal oscillation (AMO), as found by Sutton and Hodson (2005) and R. Zhang (2014, personal communication).

The anomaly correlations and SESS values of the first predictable component in DJF and JJA are shown in Figs. 7c and 7d as a function of initial month. Different initial months also indicate different lead times from 0 to 9 months. For instance, for DJF predictions, the initial month of December indicates a 0-month lead; the initial month of November indicates a 1-month lead, and so on. Very high correlations are seen for all initial months in both seasons and are nearly independent of initial month. The SESS values are higher than 0.8 for most initial conditions. Thus, the most predictable component of 2-m air temperature in DJF and JJA can be predicted with significant skill at least 9 months in advance.

The spatial pattern of the second predictable component in DJF (Fig. 9a) shows dipole structures in North America, South America, and Africa; a positive sign in Australia and southern Asia; and a negative sign in the mid to high latitudes of Eurasia. The time series of the second component in observations is significantly correlated with the observed Niño-3.4 index (correlation coefficient is 0.7), and the regressed SST pattern on this component reveals a classic ENSO pattern (not shown), implying that this component is associated with ENSO. The spatial structure of this component is consistent with the findings of the temperature teleconnection pattern with ENSO (Yang and DelSole 2012; Zhang et al. 2011). In JJA, relatively weak amplitudes in the second predictable pattern (Fig. 9b) are found compared to those in

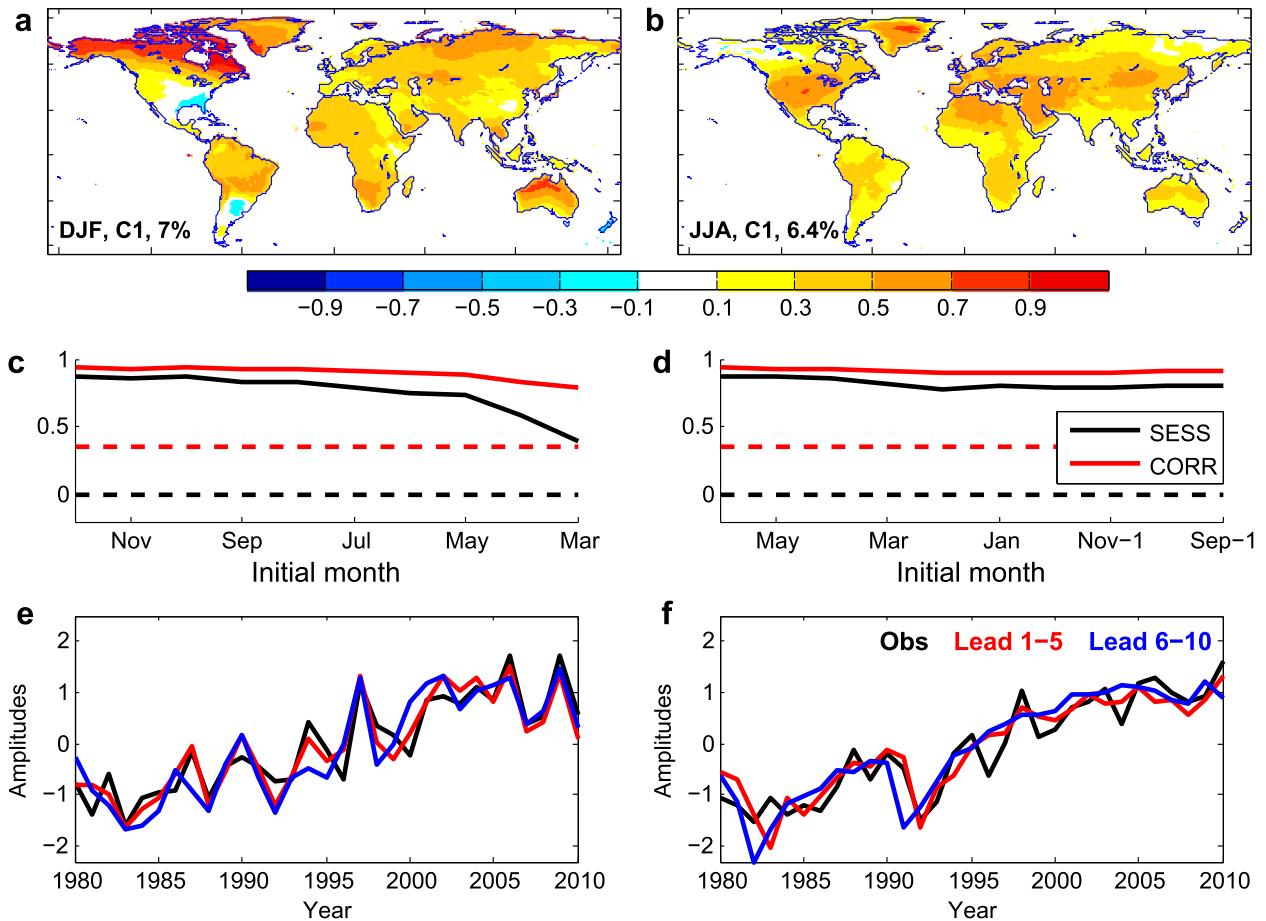


FIG. 7. Spatial structure of the first predictable component of 2-m air temperature (in K per unit std dev) in (a) DJF and (b) JJA; (c),(d) the corresponding SESS (black solid) and anomaly correlation (red solid) skill as a function of initial month; (e),(f) the time series of the most predictable component in DJF and JJA for observation (black), leads 1–5 (red), and leads 6–10 (blue). The red dash lines indicate the 5% significance level for anomaly correlation. The black dash lines indicate zero SESS. The percentage denotes the variance explained by the component.

DJF. Unlike in DJF, the correlation between the time series of the second predictable component in observations and the observed Niño-3.4 index is not statistically significant in JJA, suggesting that processes other than ENSO contribute to the predictability of JJA

temperature. The mechanism of this component remains to be studied.

The anomaly correlations of the second predictable components, shown in Figs. 9c and 9d, are significant for all initial months in both seasons but are smaller than

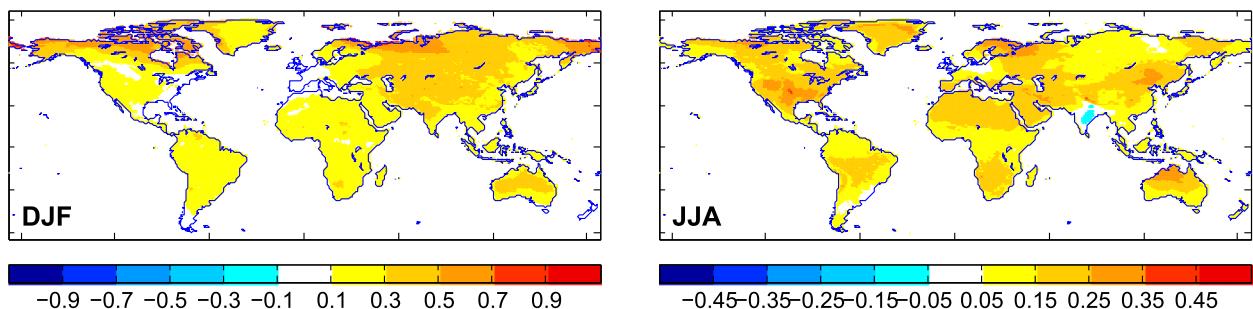


FIG. 8. Externally forced pattern of 2-m air temperature (in K per unit std dev) over land in (left) DJF and (right) JJA derived from the 5-member historical runs of FLOR from 1981 to 2010.

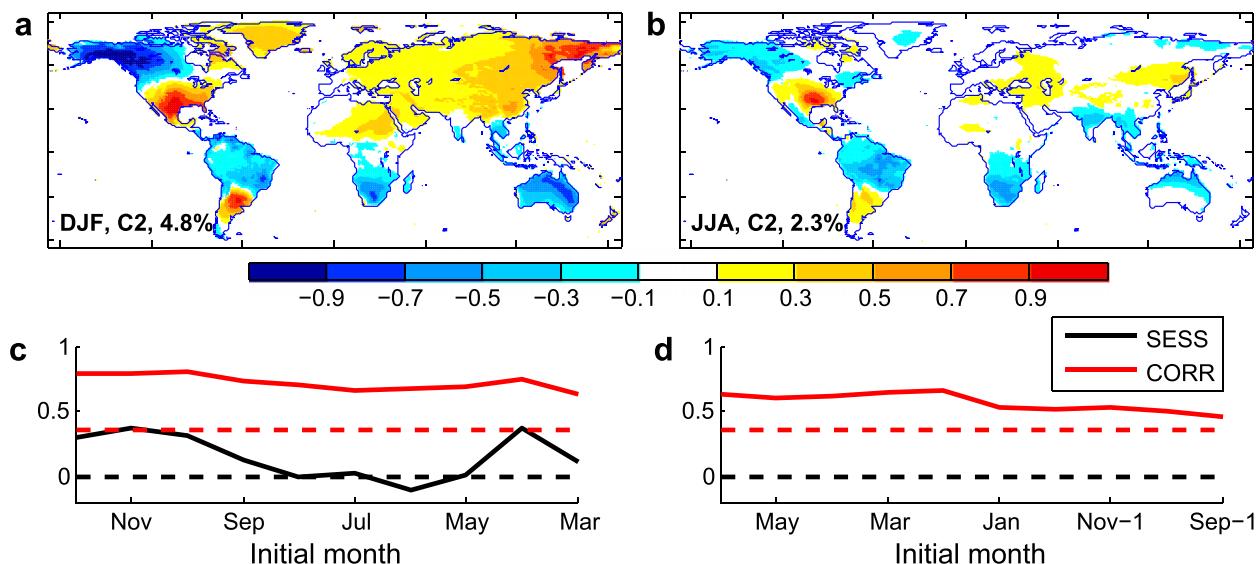


FIG. 9. Spatial structure of the second predictable component of 2-m air temperature (in K per unit std dev) in (a) DJF and (b) JJA; (c),(d) the corresponding SESS (black solid) and anomaly correlation (red solid) skill as a function of initial month. The red dash lines indicate the 5% significance level for anomaly correlation. The black dash lines indicate zero SESS. The percentage denotes the variance explained by the component.

those of the first component. The SESS values for all initial months are negative in JJA. A close scrutiny of the associated time series of the second component in JJA reveals that the amplitudes in the model are larger than those in observations (not shown). Therefore, the negative SESS values are due to the overestimation of the predictable pattern in the model compared to the observations.

The above analysis based upon APT-isolated predictable components with different time scales and mechanisms (i.e., the externally forced trend component on multidecadal scales and the ENSO-related component on interannual scales). The ability of the APT method to isolate these components cleanly is impressive. The trend component explains 7% (6.4%) of total variance in DJF (JJA), the ENSO-related component in DJF explains 4.8% of total variance, and the second component in JJA explains 2.3% of the total variance. It is noteworthy that both the forced component (i.e., trend) and the unforced (i.e., ENSO) internal variability contribute to the predictions of temperature on seasonal time scales, consistent with the findings in Jia et al. (2014).

d. Reconstructing predictions from predictable components

Having identified predictable components of seasonal temperature and precipitation over land and demonstrated prediction skill of these components, it is compelling to reconstruct predictions based upon the first few predictable components. We hypothesize that since

the reconstructed predictions filter out unpredictable components in the model, they will be more likely to have higher skill when compared with observations than the raw predictions directly from the model. Thus, by ignoring unpredictable elements of the model predictions, we expect improvements in skill, even though we are throwing out some elements of the model predictions.

The geographic distribution of SESS averaged over lead times is computed for predictions reconstructed from the first few predictable components (as those shown in Figs. 6, 7, 9) of temperature and precipitation, respectively, and are compared to those from raw predictions of FLOR in DJF and JJA (Fig. 10). The reason for averaging SESS over lead times is that both the geographical distribution and the magnitude of SESS are very similar among different lead times. Overall, the actual SESS values of the reconstructed and the raw predictions are larger in temperature than precipitation. The map of SESS difference (Fig. 10, right column), defined as the SESS of reconstructed predictions minus SESS of raw predictions, shows positive values nearly everywhere over the globe in both air temperature and precipitation and for both seasons, indicating skill improvements in reconstructed predictions for temperature and precipitation in both seasons. These results are impressive in that reconstructed predictions using only 1–2 predictable components outperform raw predictions. The improvements in precipitation predictions are generally higher than those in temperature, as indicated by the darker color in the skill

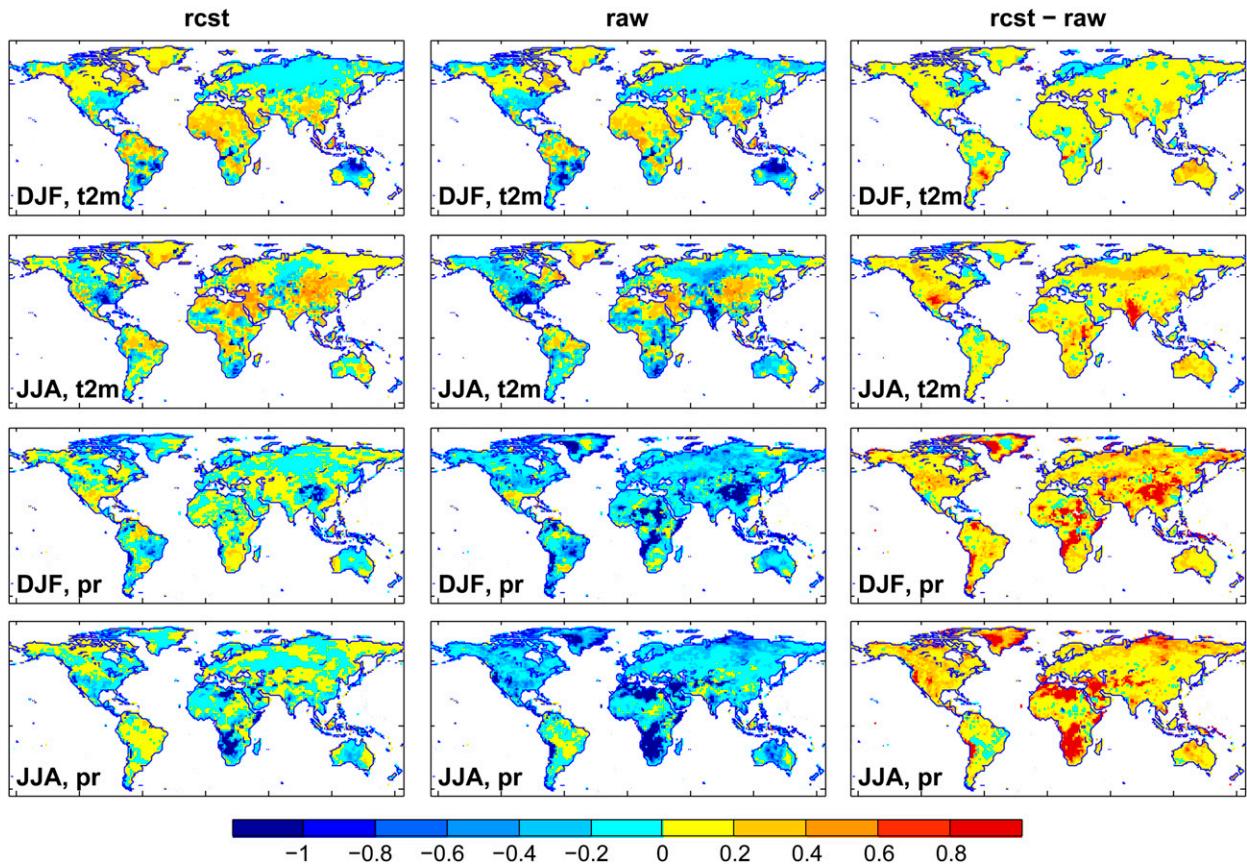


FIG. 10. SESS of reconstructed predictions of 2-m air temperature and precipitation from (left) the first 1–2 predictable components, (middle) raw predictions directly from FLOR, and (right) SESS of reconstructed predictions minus SESS of raw predictions. The SESS is averaged over lead times from 0 to 9 months.

difference map of precipitation. Note that the skill improvements in precipitation are mostly over areas with negative SESS values in the raw predictions. And a large number of those areas with negative SESS show positive SESS in reconstructed predictions. Similar geographic distribution of anomaly correlation skill shows that the correlation difference between the reconstructed and the raw predictions is small (Fig. 11). Only moderate improvements are found, and even decreases in correlation skill are seen in certain areas. The fact that there are moderate to no improvements in correlation skill but considerable improvements in SESS over nearly the whole globe reveals that reconstructing predictions based on predictable components substantially reduces conditional biases, although the significance of the improvements in SESS is difficult to test.

A question might be raised as to whether the higher skill in reconstructed predictions versus raw predictions is a result of optimal filtering of unpredictable components or merely a result of filtering out PCs with small variances (recall that the leading 30 PCs were used in the APT analysis). To address this question, we examined the SESS

of the reconstructed predictions based on the leading 30 PCs. This is equivalent to using all 30 predictable components without any optimal filtering. The resulting SESS values (averaged over the globe) in DJF and JJA were lower than the reconstructed predictions from the first few predictable components for both temperature and precipitation. Therefore, optimal filtering of unpredictable components does contribute to the skill improvements of temperature and precipitation over land. However, we note that such results can be model dependent. A similar study by Jia et al. (2014) reconstructed predictions based on skillful components derived from the ENSEMBLES multimodel dataset and showed that, except for South America and Africa, most of the skill improvements of seasonal mean 2-m air temperature come not from filtering out unskillful components, but from filtering out PCs.

To further compare the reconstructed and the raw predictions, we show in Fig. 12 the percentage of grid points in each bin (interval of 0.04) for SESS and the anomaly correlation of reconstructed versus raw predictions. For example, a value of 0.5 indicates 0.5% of

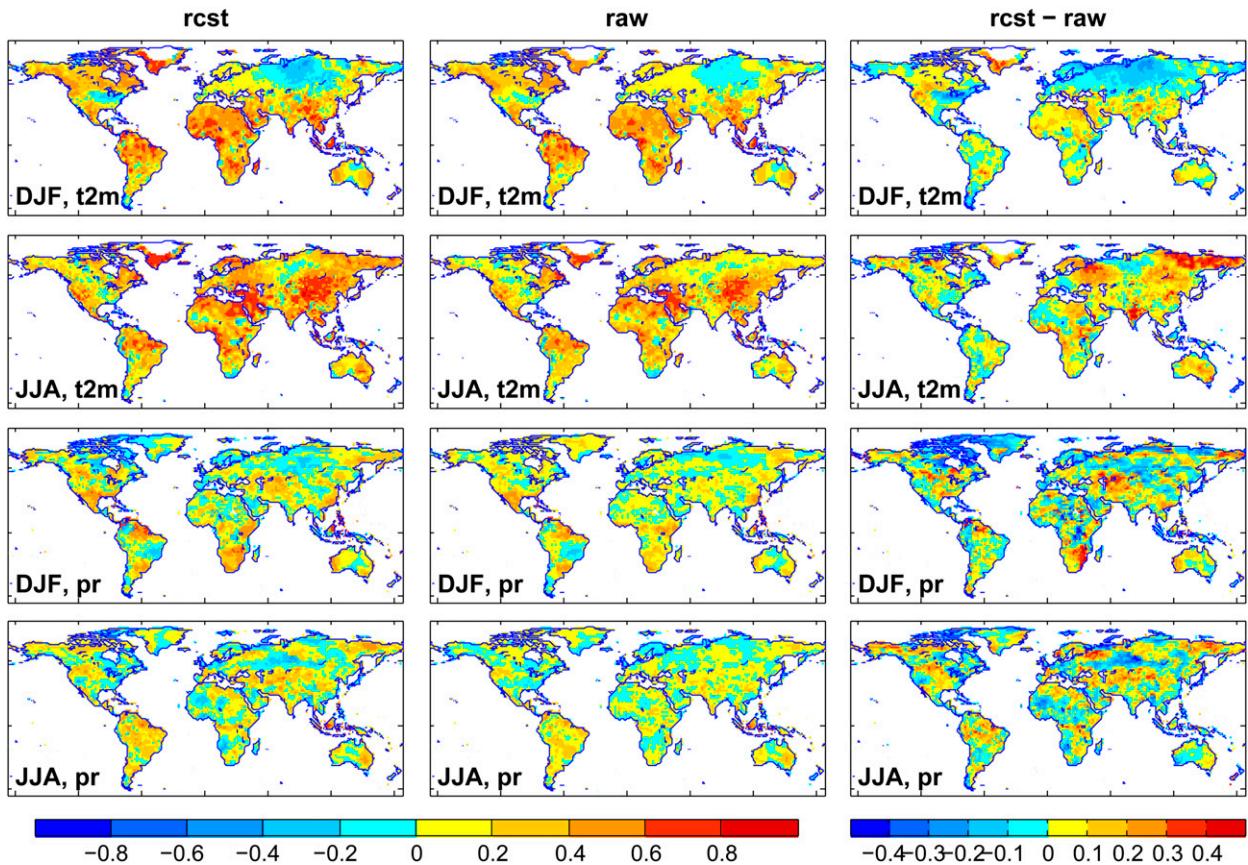


FIG. 11. Correlation of reconstructed predictions of 2-m air temperature and precipitation from (left) the first 1–2 predictable components, (middle) raw predictions directly from FLOR, and (right) correlation of reconstructed predictions minus correlation of raw predictions. The correlation is averaged over lead times from 0 to 9 months.

total grid points in that bin. A value above the diagonal line implies that the SESS/correlation of the reconstructed predictions is higher than that from the raw predictions. Nearly all values for SESS are above the diagonal lines in temperature and precipitation. The improvements in SESS are considerable, particularly for precipitation grid points with negative SESS in raw predictions, consistent with the results from geographical distribution maps shown in Fig. 10. The improvements in correlation are smaller than those in SESS. As SESS takes into account conditional bias, the higher improvements in SESS than in correlation again imply reduced conditional bias in reconstructed predictions.

5. Summary and discussion

This study investigated the seasonal prediction skill of 2-m air temperature and precipitation over land in a new high-resolution climate model (FLOR) using a statistical optimization technique: APT. We first showed that this model, with high-resolution in the atmosphere and land, simulates seasonal mean climate and variability

(including ENSO teleconnection patterns) better than the lower-resolution CM2.1. In addition, FLOR exhibits higher skill in predicting the Niño-3.4 index and the most predictable component of temperature and precipitation than CM2.1, even with ocean initial conditions that are optimized to CM2.1 and without atmospheric data assimilation in the FLOR experiments.

The improvements in FLOR motivated us to further examine the seasonal skill of temperature and precipitation over land in FLOR for DJF and JJA separately. It is shown that the two most predictable components for 2-m air temperature over land are characterized by an externally forced multidecadal warming component in DJF and JJA and an ENSO-related pattern in DJF. We emphasize that our technique is able to isolate components on different time scales that are associated with different physical mechanisms. The most predictable components of precipitation over land are ENSO-related in both seasons. These predictable components of temperature and precipitation show significant correlation skill for all leads from 0 to 9 months. The negative SESS values of the most predictable component of precipitation in both seasons

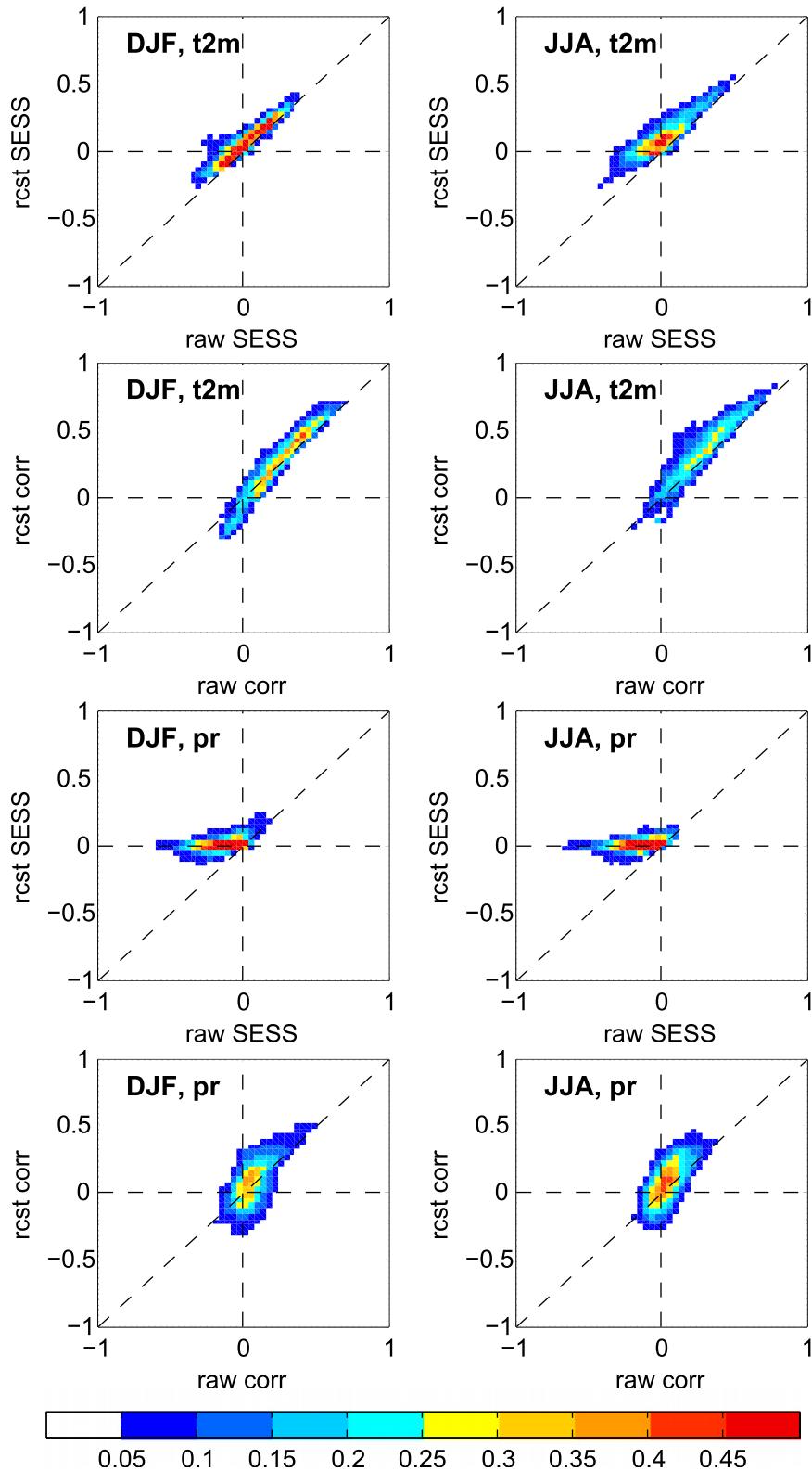


FIG. 12. Scatterplot of the percentage of grid points in each bin from -1 to 1 (interval of 0.04) for SESS and anomaly correlation of reconstructed predictions vs raw predictions for 2-m air temperature and precipitation in (left) DJF and (right) JJA.

for the March initial condition might be related to the spring barrier of ENSO prediction.

The skill of reconstructed predictions based on the first few predictable components was compared to the skill of raw predictions directly from the model in both temperature and precipitation and for both seasons. The results showed considerable improvements in SESS nearly everywhere over the globe but moderate to no improvements in correlation. This reveals that conditional bias is significantly reduced in the reconstructed predictions.

Our results suggest that increasing atmosphere and land resolution in the dynamical model and employing refined statistical optimization techniques improve seasonal predictions of 2-m air temperature and precipitation over land. The increased resolution in FLOR leads to better simulations of mean climate and variability and improved predictions of ENSO, 2-m air temperature, and precipitation over land. The statistical optimization method (APT) is able to isolate predictable components on different time scales that are associated with different physical mechanisms. It is noteworthy that both the externally forced multidecadal trend component and the internal ENSO-related component contribute to the seasonal predictions of 2-m air temperature. Reconstructing predictions based on predictable components provides a strategy to improve seasonal predictions. Further improvements in seasonal skill are expected when the data assimilation system is available for FLOR. Our results are based on the specific FLOR model, so they could be model dependent.

Acknowledgments. We thank Thomas Knutson and Charles Stock for helpful reviews of an earlier draft. We also thank Timothy DelSole and three anonymous reviewers for insightful comments that helped to improve this manuscript. This research was supported by the Visiting Scientist Program at the National Oceanic and Atmospheric Administration's Geophysical Fluid Dynamics Laboratory administered by the University Corporation for Atmospheric Research and the National Oceanic and Atmospheric Administration's Climate Program Office.

REFERENCES

- Barnston, A. G., M. K. Tippett, M. L. L'Heureux, S. Li, and D. G. DeWitt, 2012: Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing? *Bull. Amer. Meteor. Soc.*, **93**, 631–651, doi:10.1175/BAMS-D-11-00111.1.
- Becker, E., H. van den Dool, and Q. Zhang, 2014: Predictability and forecast skill in NMME. *J. Climate*, **27**, 5891–5906, doi:10.1175/JCLI-D-13-00597.1.
- Chang, Y.-S., S. Zhang, A. Rosati, T. L. Delworth, and W. F. Stern, 2013: An assessment of oceanic variability for 1960–2010 from the GFDL ensemble coupled data assimilation. *Climate Dyn.*, **40**, 775–803, doi:10.1007/s00382-012-1412-2.
- Chen, M., P. Xie, J. E. Janowiak, and P. A. Arkin, 2002: Global land precipitation: A 50-yr monthly analysis based on gauge observations. *J. Hydrometeorol.*, **3**, 249–266, doi:10.1175/1525-7541(2002)003<0249:GLPAYM>2.0.CO;2.
- DelSole, T., and M. K. Tippett, 2009a: Average predictability time. Part I: Theory. *J. Atmos. Sci.*, **66**, 1172–1187, doi:10.1175/2008JAS2868.1.
- , and —, 2009b: Average predictability time. Part II: Seamless diagnoses of predictability on multiple time scales. *J. Atmos. Sci.*, **66**, 1188–1204, doi:10.1175/2008JAS2869.1.
- , —, and J. Shukla, 2011: A significant component of unforced multidecadal variability in the recent acceleration of global warming. *J. Climate*, **24**, 909–926, doi:10.1175/2010JCLI3659.1.
- Delworth, T. L., and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *J. Climate*, **19**, 643–674, doi:10.1175/JCLI3629.1.
- , and Coauthors, 2012: Simulated climate and climate change in the GFDL CM2.5 high-resolution coupled climate model. *J. Climate*, **25**, 2755–2781, doi:10.1175/JCLI-D-11-00316.1.
- Doblas-Reyes, F. J., J. García-Serrano, F. Lienert, A. P. Biescas, and L. R. L. Rodrigues, 2013: Seasonal climate predictability and forecasting: Status and prospects. *Wiley Interdiscip. Rev.: Climate Change*, **4**, 245–268, doi:10.1002/wcc.217.
- 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.
- Goddard, L., and Coauthors, 2013: A verification framework for interannual-to-decadal predictions experiments. *Climate Dyn.*, **40**, 245–272, doi:10.1007/s00382-012-1481-2.
- Jia, L., 2011: Robust multi-year predictability on continental scales. Ph.D. dissertation, George Mason University, 101 pp. [Available online at <http://hdl.handle.net/1920/6326>.]
- , and T. DelSole, 2011: Diagnosis of multiyear predictability on continental scales. *J. Climate*, **24**, 5108–5124, doi:10.1175/2011JCLI4098.1.
- , and —, 2012: Multi-year predictability of temperature and precipitation in multiple climate models. *Geophys. Res. Lett.*, **39**, L17705, doi:10.1029/2012GL052778.
- , —, and M. K. Tippett, 2014: Can optimal projection improve dynamical model forecasts? *J. Climate*, **27**, 2643–2655, doi:10.1175/JCLI-D-13-00232.1.
- Jung, T., and Coauthors, 2012: High-resolution global climate simulations with the ECMWF model in Project Athena: Experimental design, model climate, and seasonal forecast skill. *J. Climate*, **25**, 3155–3172, doi:10.1175/JCLI-D-11-00265.1.
- Kapnick, S., and T. Delworth, 2013: Controls of global snow under a changed climate. *J. Climate*, **26**, 5537–5562, doi:10.1175/JCLI-D-12-00528.1.
- Kinter, J. L., and Coauthors, 2013: Revolutionizing climate modeling with Project Athena: A multi-institutional, international collaboration. *Bull. Amer. Meteor. Soc.*, **94**, 231–245, doi:10.1175/BAMS-D-11-00043.1.
- Knutson, T. R., and Coauthors, 2006: Assessment of twentieth-century regional surface temperature trends using the GFDL CM2 coupled models. *J. Climate*, **19**, 1624–1651, doi:10.1175/JCLI3709.1.
- 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:JOLSIO>2.0.CO;2.
- Meehl, G. A., and Coauthors, 2009: Decadal prediction: Can it be skillful? *Bull. Amer. Meteor. Soc.*, **90**, 1467–1485, doi:10.1175/2009BAMS2778.1.

- Meinshausen, M., and Coauthors, 2011: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, **109**, 213–241, doi:10.1007/s10584-011-0156-z.
- Msadek, R., G. A. Vecchi, M. Winton, and R. G. Gudgel, 2014: Importance of initial conditions in seasonal predictions of Arctic sea ice extent. *Geophys. Res. Lett.*, **41**, 5208–5215, doi:10.1002/2014GL060799.
- Paolino, D. A., J. L. Kinter III, B. P. Kirtman, D. Min, and D. M. Straus, 2012: The impact of land surface and atmospheric initialization on seasonal forecasts with CCSM. *J. Climate*, **25**, 1007–1021, doi:10.1175/2011JCLI3934.1.
- Rayner, N. A., D. E. Parker, E. B. 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**, 4407, doi:10.1029/2002JD002670.
- Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.
- Ropelewski, C. F., and M. S. Halpert, 1996: Quantifying Southern Oscillation–precipitation relationships. *J. Climate*, **9**, 1043–1059, doi:10.1175/1520-0442(1996)009<1043:QSOPR>2.0.CO;2.
- Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, doi:10.1175/JCLI-D-12-00823.1.
- Shaffrey, L. C., and Coauthors, 2009: U.K. HiGEM: The new U.K. High-Resolution Global Environment Model—Model description and basic evaluation. *J. Climate*, **22**, 1861–1896, doi:10.1175/2008JCLI2508.1.
- Smith, D. M., S. Cusack, A. W. Colman, C. K. Folland, G. R. Harris, and J. M. Murphy, 2007: Improved surface temperature prediction for the coming decade from a global climate model. *Science*, **317**, 796–799, doi:10.1126/science.1139540.
- Solomon, A., and Coauthors, 2011: Distinguishing the roles of natural and anthropogenically forced decadal climate variability: Implications for prediction. *Bull. Amer. Meteor. Soc.*, **92**, 141–156, doi:10.1175/2010BAMS2962.1.
- Sutton, R. T., and D. L. R. Hodson, 2005: Atlantic Ocean forcing of North American and European summer climate. *Science*, **309**, 115–118, doi:10.1126/science.1109496.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.
- Ting, M., Y. Kushnir, R. Seager, and C. Li, 2009: Forced and internal twentieth-century SST trends in the North Atlantic. *J. Climate*, **22**, 1469–1481, doi:10.1175/2008JCLI2561.1.
- Vecchi, G. A., and Coauthors, 2013: Multiyear predictions of North Atlantic hurricane frequency: Promise and limitations. *J. Climate*, **26**, 5337–5357, doi:10.1175/JCLI-D-12-00464.1.
- , and Coauthors, 2014: On the seasonal forecasting of regional tropical cyclone activity. *J. Climate*, **27**, 7994–8016, doi:10.1175/JCLI-D-14-00158.1.
- Winton, M., W. G. Anderson, T. L. Delworth, S. M. Griffies, W. J. Hurlin, and A. Rosati, 2014: Has coarse ocean resolution biased simulations of transient climate sensitivity? *Geophys. Res. Lett.*, **41**, 8522–8529, doi:10.1002/2014GL061523.
- Yang, X., and T. DelSole, 2012: Systematic comparison of ENSO teleconnection patterns between models and observations. *J. Climate*, **25**, 425–446, doi:10.1175/JCLI-D-11-00175.1.
- , and Coauthors, 2014: A predictable AMO-like pattern in the GFDL fully coupled ensemble initialization and decadal forecasting system. *J. Climate*, **26**, 650–661, doi:10.1175/JCLI-D-12-00231.1.
- , and Coauthors, 2015: Seasonal predictability of extratropical storm tracks in GFDL's high-resolution climate prediction model. *J. Climate*, in press.
- Zhang, S., and A. Rosati, 2010: An inflated ensemble filter for ocean data assimilation with a biased coupled GCM. *Mon. Wea. Rev.*, **138**, 3905–3931, doi:10.1175/2010MWR3326.1.
- , M. J. Harrison, A. Rosati, and A. T. 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.
- Zhang, T., M. P. Hoerling, J. Perlwitz, D.-Z. Sun, and D. Murray, 2011: Physics of U.S. surface temperature response to ENSO. *J. Climate*, **24**, 4874–4887, doi:10.1175/2011JCLI3944.1.