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,[#, *] Keith Dixon[*]

[+] *University Corporation for Atmospheric Research, Boulder, Colorado*

[*] *Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, New Jersey*

[#] *Princeton University, Princeton, New Jersey*

[@] *Dynamics Research Corporation, Andover, Massachusetts*

[*] *Corresponding author address:* Liwei Jia, Geophysical Fluid Dynamics Laboratory, NOAA, 201 Forrestal Rd., Princeton, NJ 08540.

E-mail: Liwei.Jia@noaa.gov

ABSTRACT

Seasonal climate predictions are of great potential societal value by enabling improved decisions, and are also of inherent scientific value by providing tests to the hypotheses underlying prediction methodologies. Skillful predictions over land are in particular demand due to their influences on such societal factors as agriculture and hydrology. Here we demonstrate skillful seasonal prediction of temperature and precipitation over land in a high-resolution global model using a new statistical optimization technique, and diagnose possible sources of the prediction skill. Specifically, we employ an optimization approach to identify the most predictable components of seasonal 2m air temperature and precipitation over land, and demonstrate the skill of these most predictable components. We then reconstruct new predictions based on the most predictable components, thus removing unpredictable components with the expectation of improving subsequent prediction skill. We find that the two most predictable components of 2m air temperature over global land are characterized by a spatially homogeneous component that is mostly due to changes in external radiative forcing in both boreal winter and summer, and a spatially heterogenous ENSO-related pattern in boreal winter. The most predictable components of precipitation over land in boreal winter and summer are also ENSO-related. These predictable components of temperature and precipitation show significant correlation skill for all leads from 0 to 9 months. Importantly, the reconstructed predictions based only on the leading 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 together with a high-resolution dynamical model leads to significant skill in seasonal predictions of 2m 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 and coauthors 2012; Jung and Coauthors 2012; Kinter and Coauthors 2013; Shaffrey and Coauthors 2009). High-resolution climate models, with the ability to better represent small-scale processes, show advantages simulating many key aspects of climate such as El Niño and Southern Oscillation (ENSO), Indian monsoon (Delworth and coauthors 2012), tropical precipitation, atmospheric circulation, and extratropical cyclones (Jung and Coauthors 2012). At 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 deferent resolutions. At one end of the spectrum is CM2.1 (Delworth and Coauthors 2006) with $1°$ oceanic resolution and approximately 200km atmospheric resolution. At the other extreme sits CM2.5 and CM2.6 (Delworth and coauthors 2012), both with approximately 50km atmosphere and $0.25°$ and $0.1°$ ocean respectively. The simulated climate in the high-resolution CM2.5 showed marked improvements, including a reduction of double intertropical convergence zone, improved simulations of ENSO and Amazonian rainfall, over the coarser-resolution CM2.1 (Delworth and coauthors 2012).

However, high-resolution coupled models are computationally expensive. It is worth exploring which elements of enhanced resolution are critical for each problem of interest. Motivated by the hypothesis that increased atmosphere and land resolution was critical for many of the improvements seen in 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 model: Forecast-oriented Low Ocean Resolution version of CM2.5 (CM2.5_FLOR, called FLOR hereafter). This FLOR model has a high resolution ($\sim$50km) atmosphere and land as that in CM2.5, but a coarser resolution ($1°$) ocean and sea ice as that of CM2.1. FLOR was designed to substantially reduce computing time relative to

CM2.5 to enable the running of large ensembles of simulations needed for climate predictions while still maintaining high-resolution land and atmosphere to allow exploration of regional climate and extremes. FLOR is one of the first high-resolution climate models in the world used for routine seasonal forecasts (Saha and Coauthors 2013; Arribas and Coauthors 2011). The hypothesis underlined the development of FLOR is that atmosphere and land resolution is crucial for seasonal forecasts.

The objective of this paper is to investigate seasonal prediction skill of global 2m air temperature and precipitation over land in the new high-resolution FLOR model using a statistical optimization approach, called average predictability time (APT) (DelSole and Tippett 2009a,b; DelSole et al. 2011). Our hypothesis is that increasing atmosphere and land resolution in the dynamical model together with refined statistical methods can improve seasonal predictions.

Unlike predictions on multi-year 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 is generally about 12 months in length and thereby the externally-forced signals are often overlooked. However, the assessment period for seasonal predictions spans about 20 to 30 years (Saha and Coauthors 2006), indicating the prominence of externally-forced climate signals in seasonal predictions in addition to internal climate variability. Hence, seasonal climate predcitions could be a joint initial-boundary value problem (Doblas-Reyes et al. 2013), similar to decadal predictions. Distinguishing the role of externally-forced changes from internal variability in decadal predictions has been well studied (Smith et al. 2008; Solomon and Coauthors 2011; Yang and Coauthors 2013), but has not been well documented in 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 seasonal predictions.

The rest of the paper is organized as follows: The model and data are introduced in Sec.2.

The methodology is described in Sec.3. Our results are discussed in Sec.4, and summarized in Sec.5.

2. Model and data

The new high-resolution climate model FLOR is a combination of two previously described coupled model configurations. The atmosphere (AM2.5) and land models have an approximately 50km $\times$ 50km spatial resolution, and are those used in GFDL CM2.5 and CM2.6 (Delworth et al. 2012), which have ocean models at $0.25^\circ \times 0.25^\circ$ and $0.1^\circ \times 0.1^\circ$ resolutions respectively. The ocean and sea ice component of the FLOR model are at 1°x1° resolution, based on those of CM2.1. CM2.1 model has atmospheric resolution of 2° latitude x 2.5° longitude, and 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 incorporating a newer, higher order advection scheme; an updated parameterization for eddies (Ferrari et al. 2012); and having a more realistic representation of the solar absorption by the ocean. The initial conditions for the ocean and ice components in FLOR are taken from GFDL's ensemble coupled data assimilation (ECDA) system developed for CM2.1 specifically (Zhang et al. 2007; Zhang and Rosati 2010). The ECDA covers the period 1960 to present and is being updated monthly for GFDL's seasonal-to-decadal experimental forecasts (Yang and Coauthors 2013; Vecchi and Coauthors 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 atmosphere and land components are taken from AM2.5 simulations driven by observed SST. Additional description of aspects of the FLOR model will be described in a series of papers, including Vecchi and Coauthors (2014); Winton and Coauthors (2014) .

Ensemble hindcasts were made from 1980 to present for FLOR and CM2.1. The hindcasts are initialized at the first day of each month, and are run for 12 months. There are 12

(10) ensemble members for each prediction in FLOR (CM2.1). We analyzed hindcasts of 2m air temperature and precipitation over land and global sea surface temperature (SST) during the period of 1980-2012 (1982-2011) in FLOR (CM2.1). A three-month running mean was applied to hindcasts to remove sub-seasonal variability. We also analyzed 2m air temperature, precipitation, SST, meridional and zonal velocity at 925hPa, 850hPa and 200hPa and sea level pressure in FLOR, CM2.1 and CM2.5 1990 control simulations, in which the atmospheric composition (greenhouse gases, aerosols) and external forcing (solar irradiance) are fixed at 1990 levels. The 2m air temperature (1980-2005) from 5-member historical runs of FLOR were used to diagnose externally-forced temperature patterns. The 5 members of FLOR historical runs start from year 101, 141, 181, 221 and 261 of the 1860 control simulation (i.e., the atmospheric composition and external forcing fixed at 1860 levels) respectively.

The observations used in this study are precipitation at 0.5° resolution from National Oceanic and Atmospheric Administration (NOAA)'s precipitation reconstruction over land (Chen et al. 2002); CPC Merged Analysis of Precipitation (CMAP) at 2.5° resolution; GHCN Gridded V2 2m air temperature over land at 0.5° resolution (Fan and van den Dool 2008); Hadley Centre sea ice and sea surface temperature data set (Rayner et al. 2003); NCEP/NCAR Reanalysis-1 surface temperature and precipitation. Velocity and sea level pressure data are from Modern-era Retrospective Analysis for Research and Applications (MERRA). The observed NINO3.4 index was downloaded from NOAA's website at http://www.cpc.ncep.noaa.gov/data/indices/ersst3b.nino.mth.81-10.ascii.

3. Review of Methodology

To identify predictable components, we employed the technique of average predictability time (APT). Following DelSole and Tippett (2009b), APT is defined as

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

where σ_τ^2 denotes the forecast variance at lead time τ, and σ_{clim}^2 denotes the climatological variance. 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 linear combination of variables, $\mathbf{q}^T\mathbf{x}(\tau, t, e)$, that maximizes APT. Let the weights of the linear combination be specified by vector $\mathbf{q}$. The forecast variance of $\mathbf{q}^T\mathbf{x}(\tau, t, e)$ at lead time τ is

$$\sigma_\tau^2 = \mathbf{q}^T\overline{(\mathbf{x}(\tau,t,e) - \langle\mathbf{x}(\tau,t,e)\rangle)(\mathbf{x}(\tau,t,e) - \langle\mathbf{x}(\tau,t,e)\rangle)^T}\mathbf{q} = \mathbf{q}^T\hat{\mathbf{\Sigma}}_\tau\mathbf{q}, \tag{2}$$

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

$$\sigma_{clim}^2 = \mathbf{q}^T\left(\mathbf{x}(\tau,t,e) - \overline{\mathbf{x}(\tau,t,e)}\right)\left(\mathbf{x}(\tau,t,e) - \overline{\mathbf{x}(\tau,t,e)}\right)^T\mathbf{q} = \mathbf{q}^T\hat{\mathbf{\Sigma}}_{clim}\mathbf{q}, \tag{3}$$

where the overline denotes the average over all start times, lead times and ensemble members, $\hat{\mathbf{\Sigma}}_{clim}$ is the estimated climatological covariance matrix. Substituting (2) and (3) into (1) gives

$$APT = 2\sum_{\tau=1}^{\infty}\left(\frac{\mathbf{q}^T(\hat{\mathbf{\Sigma}}_{clim} - \hat{\mathbf{\Sigma}}_\tau)\mathbf{q}}{\mathbf{q}^T\hat{\mathbf{\Sigma}}_{clim}\mathbf{q}}\right). \tag{4}$$

It can be shown that maximizing (4) leads to a generalized eigenvalue problem

$$2\sum_{\tau=1}^{\infty}\left(\hat{\mathbf{\Sigma}}_{clim} - \hat{\mathbf{\Sigma}}_\tau\right)\mathbf{q} = \lambda\hat{\mathbf{\Sigma}}_{clim}\mathbf{q}. \tag{5}$$

The eigenvalue λ gives APT value, and each eigenvector $\mathbf{q}$ corresponds to a component. Each component is uncorrelated with one another due to the symmetry characteristic of $\hat{\mathbf{\Sigma}}_\tau$ and $\hat{\mathbf{\Sigma}}_{clim}$. We order the eigenvalues and their associated eigenvectors by decreasing order, such that the first eigenvector maximizes APT, the second maximizes APT subject being uncorrelated with the first, and so on. More details of this APT technique is found in DelSole and Tippett (2009a,b), and the application of this technique can be found in Jia and DelSole (2011, 2012); Yang and Coauthors (2013).

For typical global forecast data, the number of grid points exceeds the number of samples, so the above covariance matrixes are singular and the eigenvalue problem cannot be solved. A standard approach is to project the data onto the leading principal components (PCs) of the predictand, and then to maximize APT only in the subspace spanned by the leading PCs. In this paper, we chose 30 PCs for 2m air temperature and precipitation. The sensitivity of results to the number of PCs have been tested, and are not sensitive when using more than 20 PCs. Taking into account that leading PCs principally project on grids with large variances, we normalize original precipitation hindcasts by dividing the standard deviation at each grid. The normalized precipitation is able to capture large-scale precipitation structure.

The squared error skill score (SESS) is used to measure skill, which is defined as

$$SESS = 1 - \frac{\sum_n (O_n - P_n)^2}{\sum_n (O_n - \overline{O})^2}, \tag{6}$$

where O_n is the observation at time n, P_n is the prediction of O_n, and $\overline{O}$ is the time mean of O_n for all years. The value of SESS is one for perfect forecast, and is negative if a prediction has a mean squared error larger than a prediction based on the climatological mean.

4. Results

a. Climate mean state, variability and prediction skill

Fig. 1 shows annual mean precipitation and 2m air temperature in observations and 1990 control simulations of FLOR and CM2.1, as well as the bias in FLOR and CM2.1. Compared to observed mean precipitation, FLOR and CM2.1 show a dry bias in most of South America, although FLOR shows less bias than CM2.1. FLOR also simulates mean precipitation better than CM2.1 in tropical Africa, eastern China and the southeastern Unites States. As for annual mean 2m air temperature, the pattern is well simulated in FLOR and CM2.1, but CM2.1 shows a cold bias in most areas. FLOR shows less bias along the Andes and in equatorial Africa than does CM2.1.

Besides annual mean temperature and precipitation, we also examined seasonal climate and variability of several important variables. Fig. 2 shows the scatter plot of pattern correlation of seasonal mean climate and standard deviation between observation and model simulations for FLOR vs. CM2.1 and FLOR vs. CM2.5. The pattern correlations between model simulations and observations are higher in FLOR than in CM2.1 for nearly all variables and seasons, both in mean climate and standard deviation. In contrast, the pattern correlations in FLOR are comparable to those in CM2.5 for seasonal climate, and are slightly higher than CM2.5 for standard deviation, although the ocean resolution is increased in CM2.5. The fact that mean climate and climate variability improved considerably in FLOR over CM2.1, but moderately in CM2.5 over FLOR, supports the hypothesis that atmospheric and land resolution is critical to the improvements in these quantities, and this is the focus of our study.

Another phenomenon worth examining is ENSO, and its teleconnections to remote regions. Fig. 3 shows the patterns of correlation between NINO3 sea surface temperature anomalies (SSTA) and the global anomalies of surface temperature and precipitation, for observations, a 280-yr control simulation of FLOR, CM2.1 and a 260-yr control simulation of CM2.5. In the observations, NINO3 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 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, southern Africa and positive correlations over the equatorial Pacific and Indian Oceans, contiguous U.S., 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 atmospheric resolution (going from CM2.1 to FLOR) than by the increase in ocean resolution (going from FLOR to CM2.5).

To examine ENSO prediction skill, we show in Fig. 4 the SESS values and anomaly correlations of NINO3.4 index as a function of initial months and target months in FLOR and CM2.1. Both FLOR and CM2.1 show very high correlation skills. FLOR shows higher correlation than CM2.1 at short leads for initial months from August to December. The SESS values in FLOR are much larger (i.e., higher skill) than CM2.1 in boreal winter and spring of the target month. Such skill improvements in SESS are not seen in anomaly correlation, indicating that conditional bias are reduced in FLOR. We also found that the SESS values at long leads initialized in November, December, January are lower in FLOR than those in CM2.1. It will be shown shortly that the low SESS at long leads in FLOR might account for the low skill of the most predictable precipitation pattern at long leads in FLOR.

We now identify predictable components of global precipitation over land in FLOR and CM2.1 hindcasts using APT analysis. The most predictable components of precipitation are shown in Fig. 5b, c, for FLOR and CM2.1 respectively. The most predictable components in the two models are all significantly correlated with the NINO3.4 index. Thus, to evaluate the patterns emerging from the model, the predictable patterns are compared to the observed precipitation regression pattern with NINO3.4 index (Fig. 5a). The most predictable pattern in FLOR is much closer to the observed precipitation teleconnection pattern to ENSO than CM2.1, particularly in South America, southern United States, eastern China, Australia and southern 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 well captured in FLOR. 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.

The improvements of retrospective prediction skill of the most predictable pattern in FLOR over CM2.1 are striking (Fig. 5e, f). The SESS values of ensemble mean prediction of the most predictable precipitation pattern are much larger in FLOR than CM2.1 in nearly all initial and target months, especially for target months from October to the following March. To gain insight into the skill difference between these two models, we compute the standard deviation of the time series of the most predictable pattern and normalize it relative to that from observation. A resulting value close to one implies that the predicted variability is close to 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). Fig. 5d shows the normalized standard deviation as a function of target 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 in FLOR than 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 global 2m air temperature over land (not shown). As described in Sec. 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 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 after the spring barrier period 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 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-Feburary (DJF) and June-August (JJA) are discussed in this section.

The spatial patterns of the leading predictable component in DJF and the first two predictable components in JJA are show in Fig. 6 a, b, e, as these components are well separated from the others. The first predictable pattern in DJF shows wet anomalies in eastern China, southern North and South America, southeast Africa, the Andes, and dry anomalies in northern Australia, southern Africa, northeastern South America. In JJA, the leading 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 in previous studies (Ropelewski and Halpert 1996; Yang and DelSole 2012). In fact, the time series of these components are significantly correlated with NINO3.4 index, and the SST regression pattern on these components displays a classic ENSO structure (not shown). Therefore, these predictable components of precipitation over land are likely ENSO-related.

An important question is whether the components diagnosed in the dynamical model exist in the real world. We show in Fig. 6c, d, f the anomaly correlations between the time series of the predictable components and those from observations. The observed time series were derived by projecting predictable components on observations. The correlations are statistically significant at 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 below zero at March initial condition in both seasons for the first component. The negative skill of predictions initialized in March is presumably due to the spring barrier of ENSO prediction (Barnston et al. 2012). The SESS values of the second component in JJA are positive for all initial months. We emphasize that our statistical optimization method is able to identify components that are physically meaningful (e.g., ENSO-related). Also, it is impressive that the predictable components diagnosed in a dynamical model can be predicted with significant skill in the real world.

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

The spatial patterns of the first two predictable components of 2m air temperature in DJF and JJA and their associated correlation skill and SESS are shown in Fig. 7. The leading 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 in 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. As the associated time series of the leading predictable component exhibits an increasing trend in both seasons (not shown), the leading predictable component indicates a multi-decadal warming signal. To explore the mechanism of the most predictable component, we diagnosed externally-forced patterns of 2m air temperature over land in DJF and JJA from FLOR historical runs, using signal-to-total maximizing

EOF method (Ting et al. 2009). As shown in Fig. 8, the externally-forced patterns in two seasons bear great similarity with the most predictable patterns, which implies that the most predictable components are the responses to external forcings.

The spatial pattern of the second predictable component in DJF (Fig. 7e) shows dipole structures in North America, South America, Africa and positive sign in Australia and southern Asia, negative sign in mid to high latitudes of Eurasia. Again, the time series of the second component in DJF is significantly correlated with NINO3.4 index (cc=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 a temperature teleconnection pattern with ENSO (Yang and DelSole 2012; Zhang et al. 2011). In JJA, relatively weak amplitudes of the second component are found compared to those in DJF (Fig. 7f). Unlike in DJF, the correlation between the time series of the second predictable component and NINO3.4 index is not statistically significant in JJA, suggesting that other processes than ENSO contribute to the predictability of JJA temperature. The mechanism of this component remains to be studied.

The anomaly correlations of the first two predictable components, shown in Fig.7 c, d, g, h, demonstrate very high correlations in both seasons. The correlations are nearly unchanged with initial months. The SESS values are higher than 0.8 in most cases for the first component in both seasons. As for the second component, the SESS values are smaller, and are negative in JJA. A close scrutiny of the associated time series of the second component in JJA reveals that the negative SESS values are due to the overestimation of the predictable pattern in model compared to observations (not shown).

The above analysis based upon APT isolated predictable components with different time scales and mechanisms, i.e. the externally-forced trend component on multi-decadal scales and the ENSO-related component on interannual scales. The trend component explains 7% (6.4%) of total variance in DJF (JJA), and the ENSO-related component in DJF explains

4.8% of total variance. It is noteworthy that both forced and unforced (i.e., ENSO) internal variability contribute to seasonal predictions of temperature.

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 leading few predictable components. We hypothesize that since the reconstructed predictions filter out unpredictable components in the model, they will be more likely to yield higher skill when compared with observations than the raw predictions directly from 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 initial months are computed for predictions constructed from the leading few predictable components (as those shown in Figs. 6 and 7) of temperature and precipitation respectively, and are compared with raw predictions from FLOR in DJF and JJA (Fig. 9). The reason for averaging SESS over initial months is that the geographical distribution of SESS among different initial months are very close. Overall, the actual SESS values of reconstructed and raw predictions are larger in temperature than precipitation. The map of SESS difference (Fig. 9 far right column), defined as 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 improved skill in reconstructed predictions for temperature and precipitation in both seasons. These results are impressive in that reconstructed predictions using only 1-2 predictable components beat raw predictions. The improvements in precipitation predictions are generally higher than those in temperature as indicated by the darker color in the difference map of precipitation. Note that the skill improvements in precipitation are mostly over areas with negative SESS values in raw predictions. And, a large amount 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 reconstructed and raw predictions are small (Fig. 10). Only moderate improvements are found, and even decreases in correlation skill are seen in certain areas. The fact that moderate to no improvements in correlation skill but significant improvements in SESS over nearly the whole globe reveals that reconstructing predictions based on predictable components substantially reduces conditional biases.

To further compare reconstructed and raw predictions, we show in Fig. 11 the percentage of grid points in each bin (interval of 0.04) for SESS and anomaly correlation of reconstructed vs. raw predictions. For example, a value of 0.5 indicate 0.5% of total grid points in that bin. Values above the diagonal line imply that SESS/correlation of reconstructed predictions is higher than that from 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 grids with negative SESS in raw predictions, consistent with the results from geographical distribution maps shown in Fig. 9. The improvements in correlation are smaller than those in SESS. As SESS takes into account conditional biases, the higher improvements in SESS than correlation again implies reduced conditional bias in reconstructed predictions.

5. Summary and discussion

This study investigated seasonal prediction skill of 2m 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 mean climate and variability (including ENSO teleconnection patterns) better than the lower resolution model CM2.1. A further increase in ocean resolution (CM2.5) does not

In addition, FLOR exhibits higher skill in predicting the NINO3.4 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 skill of temperature and precipitation over land in FLOR for DJF and JJA separately. It is shown that the two most predictable components for 2m air temperature over land are characterized by an externally-forced multi-decadal 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 at March initial condition might be related to the spring barrier of ENSO prediction.

The reconstructed predictions based on the first few predictable components were compared to 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 reconstructed predictions. A question might be raised as 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 of PCs with small variances (leading 30PCs were used in APT analysis). To address this question, we examined the SESS of predictions that were reconstructed based on the leading 30PCs and without any optimal filtering. The resulting SESS values were lower than the reconstructed predictions from the first few predictable components (not shown). Therefore, optimal filtering of unpredictable components does contribute to the skill improvements of temperature and precipitation over land.

Our results suggest that a high-resolution dynamical model and refined statistical optimization techniques improve seasonal predictions of 2m air temperature and precipitation over land. The increased resolution in FLOR leads to better simulation of mean climate and variability, and improved predictions of ENSO, 2m air temperature and precipitation over land. Further improvements in skill are expected when the data assimilation system is available for FLOR. The statistical optimization method (APT) is able to isolate predictable components on different time scales that associated with different physical mechanisms. It is noteworthy that both externally-forced multi-decadal trend component and the internal ENSO-related component contribute to seasonal predictions of 2m air temperature. Reconstructing predictions based on predictable components provides a strategy to improve seasonal predictions. 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, and Timothy DelSole 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 National Oceanic and Atmospheric Administration's Climate Program Office.

REFERENCES

Arribas, A. and Coauthors, 2011: The GloSea4 ensemble prediction system for seasonal forecasting. *Mon. Wea. Rev.*, **139**, 1891–1910.

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. Mereor. Soc.*, **93**, 631–651.

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. *Clim. Dyn.*, **40**, 775–803.

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. Hydrometeor.*, **3**, 249–266.

DelSole, T. and M. K. Tippett, 2009a: Average predictability time. part i: Theory. *J. Atmos. Sci.*, **66**, 1172–1187.

DelSole, T. and M. K. Tippett, 2009b: Average predictability time. Part II: Seamless diagnoses of predictability on multiple time scales. *J. Atmos. Sci.*, **66**, 1188–1204.

DelSole, T., M. K. Tippett, and J. Shukla, 2011: A significant component of unforced multidecadal variability in the recent acceleration of global warming. *J. Climate*, **24**, 909–926.

Delworth, T. L. and Coauthors, 2006: GFDL's CM2 global coupled climate models. part I: Formulation and simulation characteristics. *J. Climate*, **19**, 643–674.

Delworth, T. L. and coauthors, 2012: Simulated climate and climate change in the GFDL CM2.5 high-resolution coupled climate model. *J. Climate*, **25**, 2755–2781.

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. *WIREs Clim 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. Geo. Res.*, **113**, D01 103, doi:10.1029/2007JD008 470.

Jia, L. and T. DelSole, 2011: Diagnosis of multiyear predictability on continetal scales. *J. Climate*, **24**, 5108–5124.

Jia, L. and T. DelSole, 2012: Multi-year predictability of temperature and precipitation in multiple climate models. *Geophys. Res. Lett.*, **39**, doi:10.1029/2012GL052 778.

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.

Kinter, J. L. and Coauthors, 2013: Revolutionizing climate modeling with project Athena: A multi-institutional, international collaboration. *Bull. Amer. Mereor. Soc.*, **94**, 231–245.

Meehl, G. A., et al., 2009: Decadal prediction: Can it be skillful? *BAMS*, **90**, 1467–1485.

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. Geo. Res.*, **108**, NO. D14, 4407, doi:10.1029/2002JD002 670.

Ropelewski, C. F. and M. S. Halpert, 1996: Quantifying Southern Oscillation-precipitation relationships. *J. Climate*, **9**, 1043–1059.

Saha and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483–3517.

Saha, S. and Coauthors, 2013: The NCEP Climate Forecast System Version 2. *J. Climate (early online release)*, doi:10.1175/JCLI–D–12–00 823.1.

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.

Smith, D. M., S. Cusack, A. W. Colman, C. K. Folland, G. R. Harris, and J. M. Murphy, 2008: Improved surface temperature prediction for the coming decade from a global climate model. *Science*, **317**, 796–799.

Solomon, A. and Coauthors, 2011: Distinguishing the roles of natural and anthropogenically forced decadal climate variability. *Bull. Amer. Mereor. Soc.*, **92**, 141–156.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experimental design. *Bull. Am. Meteor. Soc.*, **93**, 485–498.

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.

Vecchi, G. and Coauthors, 2014: On the seasonal prediction of regional tropical cyclone frequency. *J. Climate*, to be submitted.

Vecchi, G. A. and Coauthors, 2013: Multiyear predictions of North Atlantic hurricane frequency: Promise and limitations. *J. Climate*, **26**, 5337–5357.

Winton, M. and Coauthors, 2014: Eddy-scale ocean resolution and the simulation of transient climate sensitivity. *Geophys. Res. Lett.*, to be submitted.

Yang, X. and Coauthors, 2013: A predictable AMO-like pattern in the GFDL fully coupled ensemble initialization and decadal forecasting system. *J. Climate*, **26**, 650–661.

Yang, X. and T. DelSole, 2012: Systematic comparison of ENSO teleconnection patterns between models and observations. *J. Climate*, **25**, 425–446.

Zhang, S., 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.

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.

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.

List of Figures

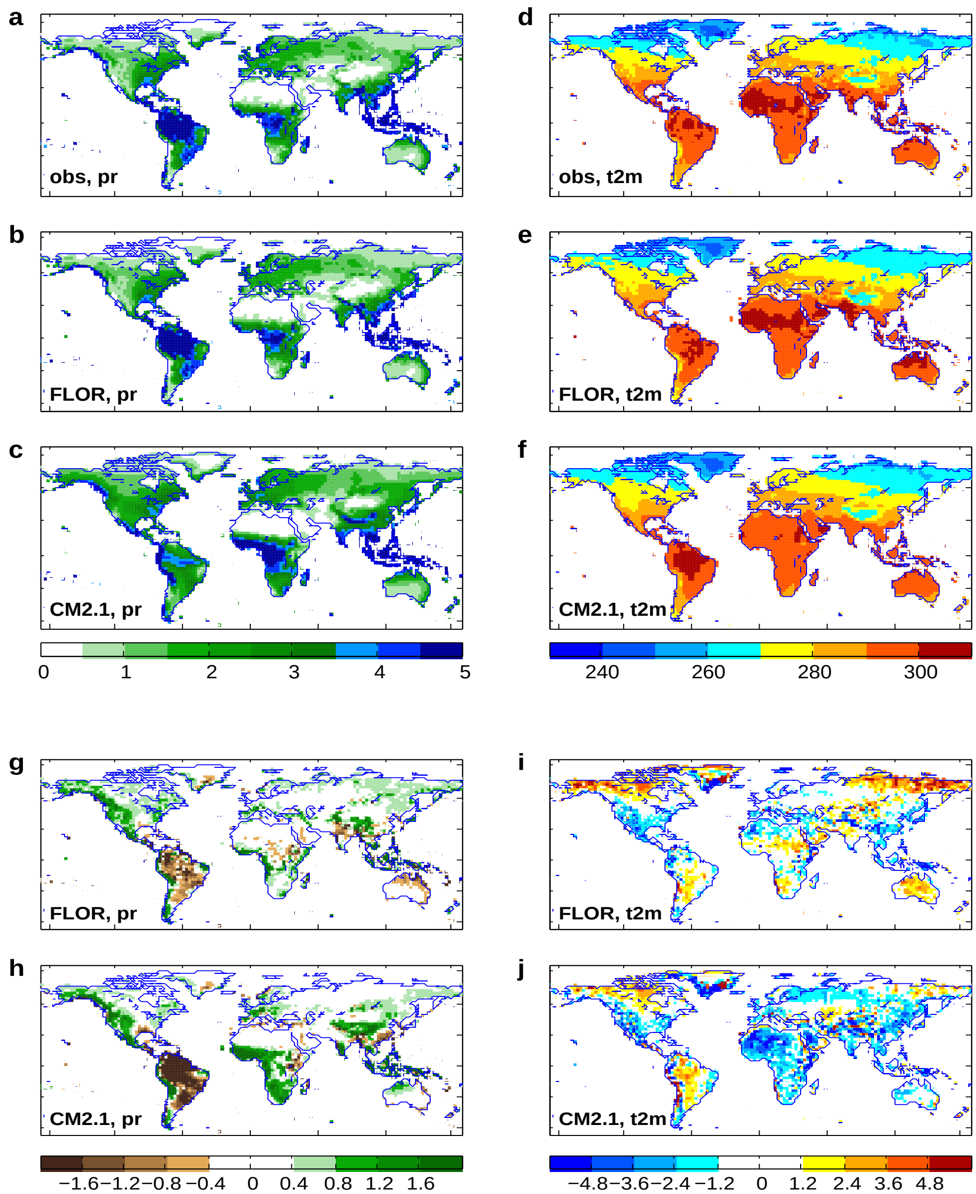

Figure 1: Annual mean precipitation (a,b,c) and 2m air temperature (d,e,f) in observations (1981-2010), FLOR (601-1200) and CM2.1 (101-300) control simulations; and the bias of annual mean precipitation (g,h) and 2m air temperature (i,j) in FLOR and CM2.1. The units of precipitation is mmday^{-1}. The units of temperature is Kelvin.

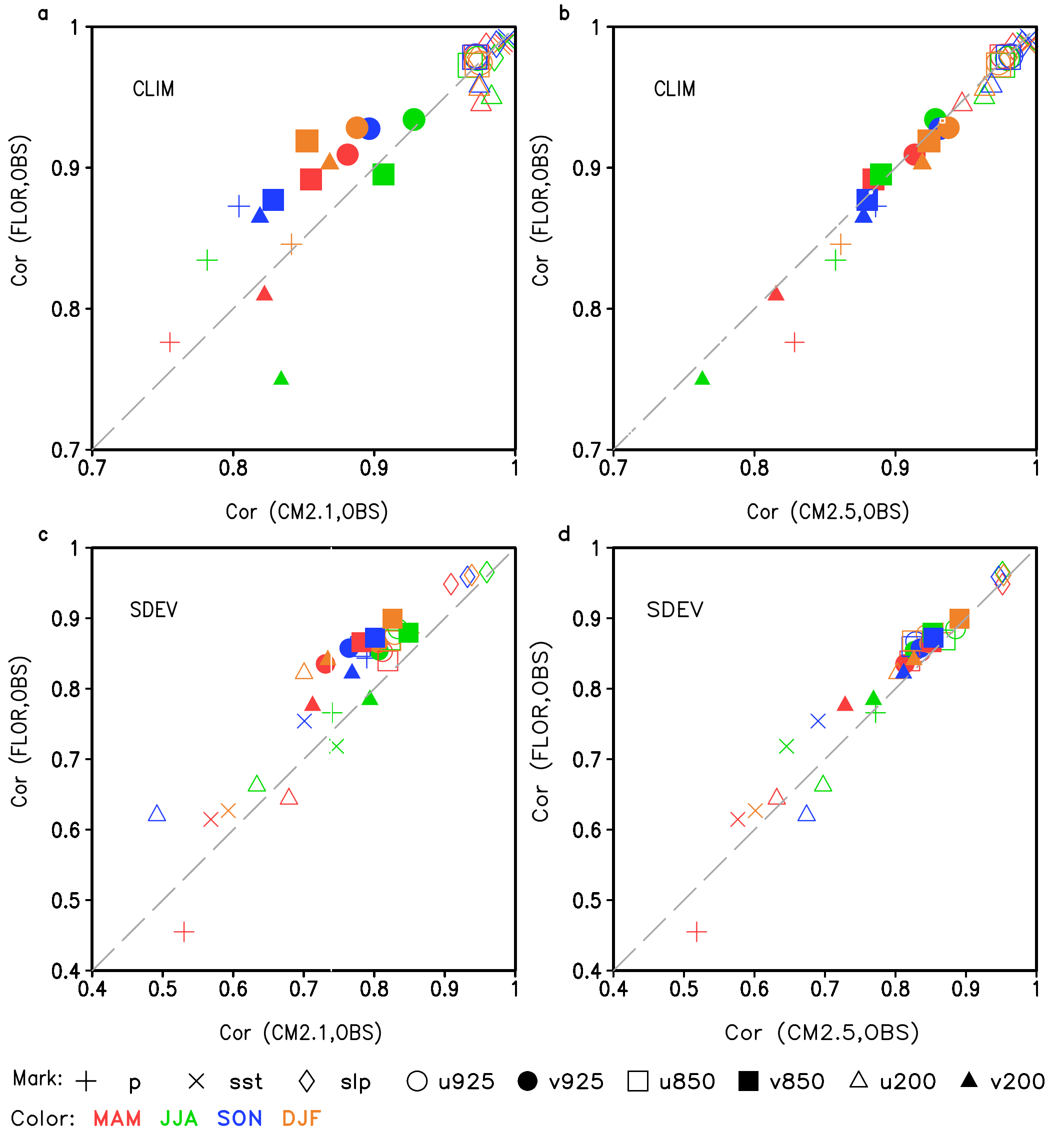

Figure 2: Scatter plot of pattern correlation between CM2.1(101-300) and observation(1982-2000) (x-axis) and FLOR (601-1200) and observation (y-axis) for seasonal mean climate (a) and standard deviation (c); and between CM2.5 (1-100) and observation (x-axis) and FLOR and observation (y-axis) for seasonal mean climate (b) and standard deviation (d) for precipitation, sea surface temperature, sea level pressure, zonal and meridional velocity at 925hPa, 850hPa and 200hPa. Different colors indicate different seasons. Each symbol represents a particular variable.

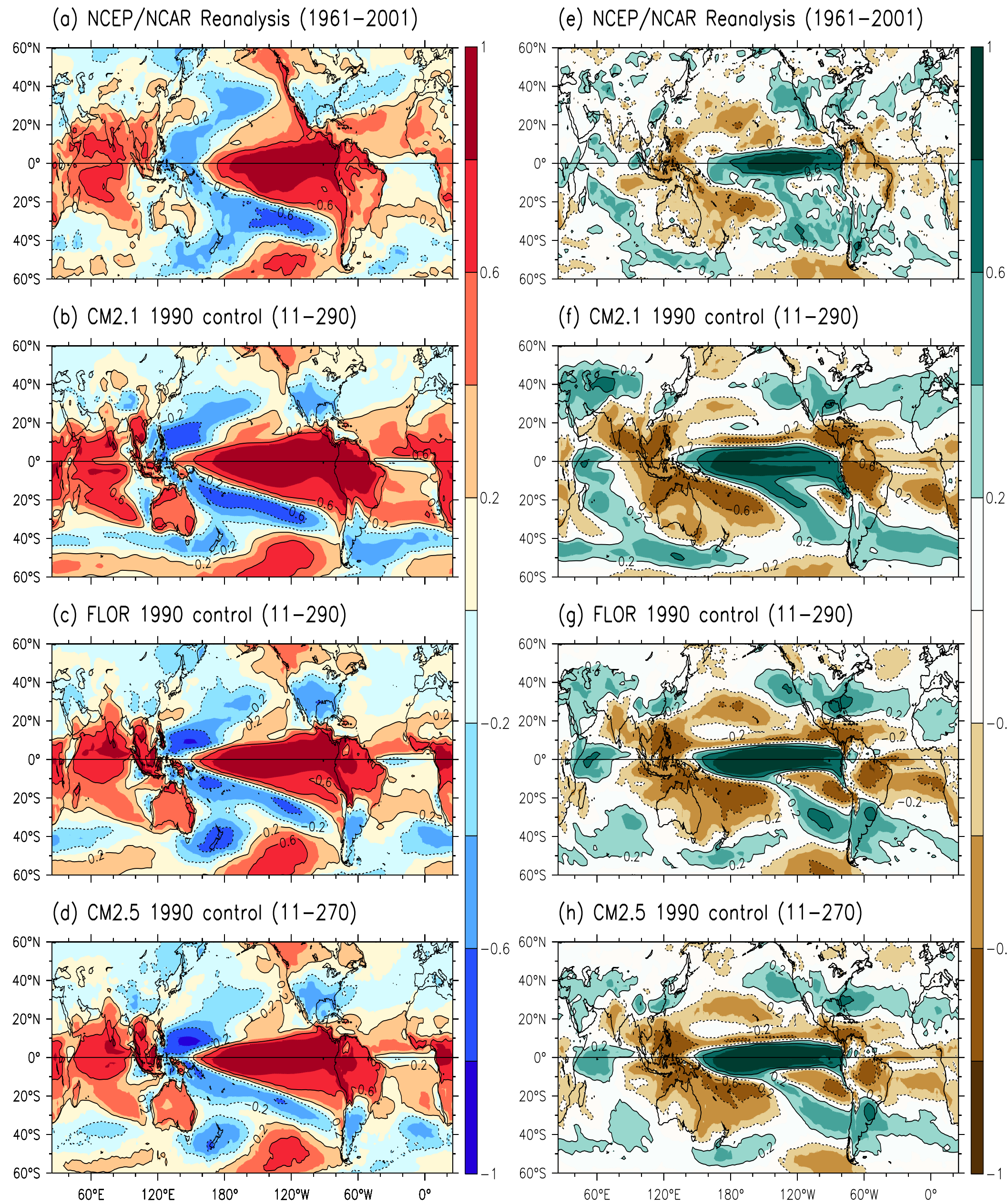

Figure 3: Correlation between local surface temperature (left) and precipitation(right) anomalies and NINO3 (150W-90W, 5S-5N) SST anomalies, for annual (June-May) means from (a,e) NCEP/NCAR Reanalysis-1 (1951-2001) observations, and 1990 control runs from (b,f) CM2.1 (11-290); (c,g) FLOR (11-290) and (d,h) CM2.5 (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 (25,50,75)% of the amplitude at periods of (25, 33, 49) yr, also truncates the initial and final decades from the anomaly time series.

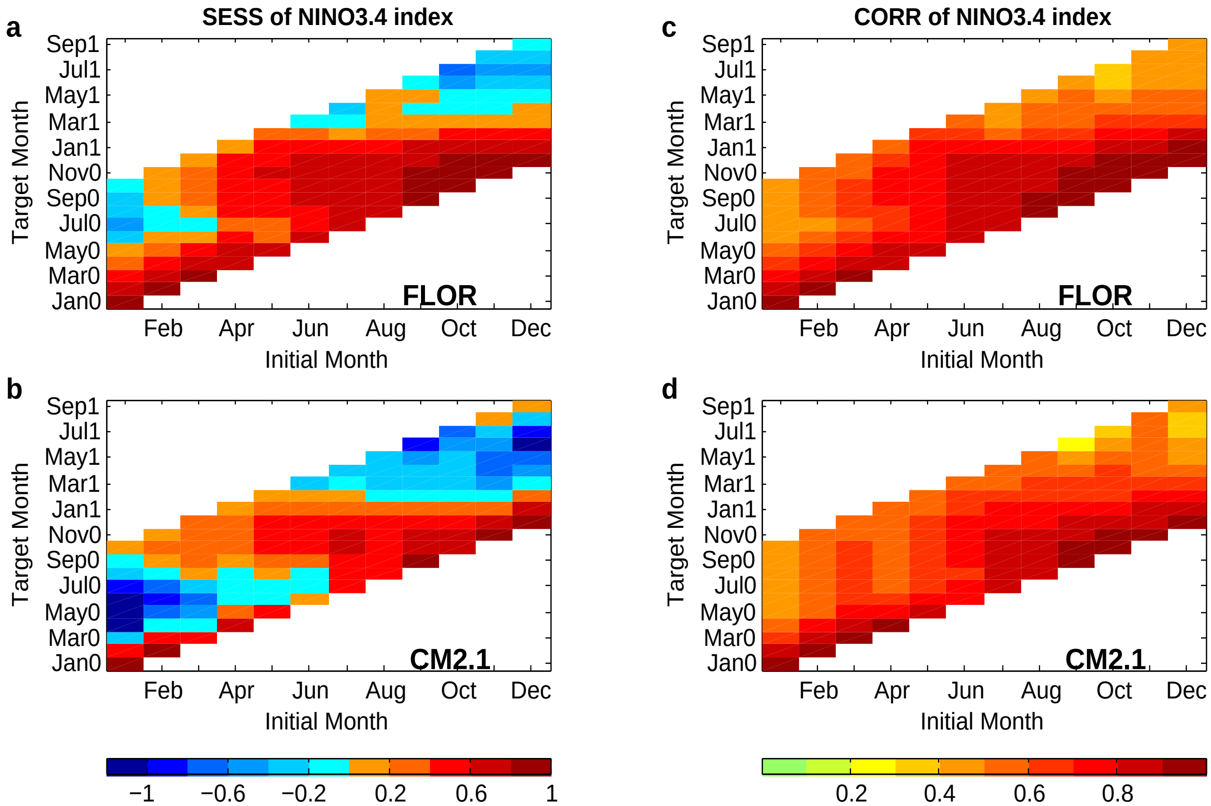

Figure 4: Squared error skill score (a,b) and anomaly correlation (c,d) of NINO3.4 index in FLOR and CM2.1 for each initial month and target month during 1981-2010 (1983-2010) in FLOR (CM2.1). Each target month indicates a 3-month mean (e.g., target month Jan. denotes Jan.-Mar. mean.)

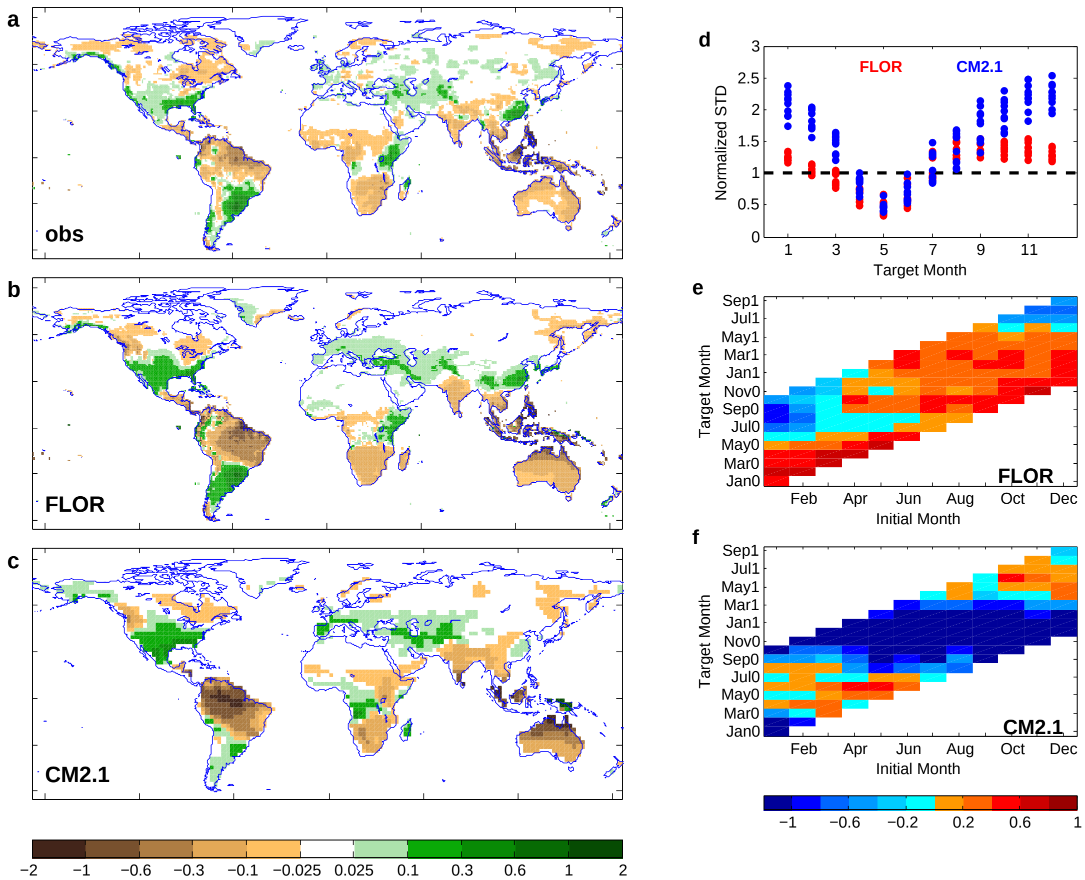

Figure 5: Observed precipitation teleconnection pattern to ENSO (in mmday^{-1} per unit variate) (a); Spatial structure of the most predictable component of precipitation over land (in mmday^{-1} per unit variate) from FLOR (b) and CM2.1 (c); Standard deviation of time series of the most predictable pattern for different target months and initial months, normalized relative to the observations (d). The squared error skill score for each initial month and target month in FLOR (e) and CM2.1(f).

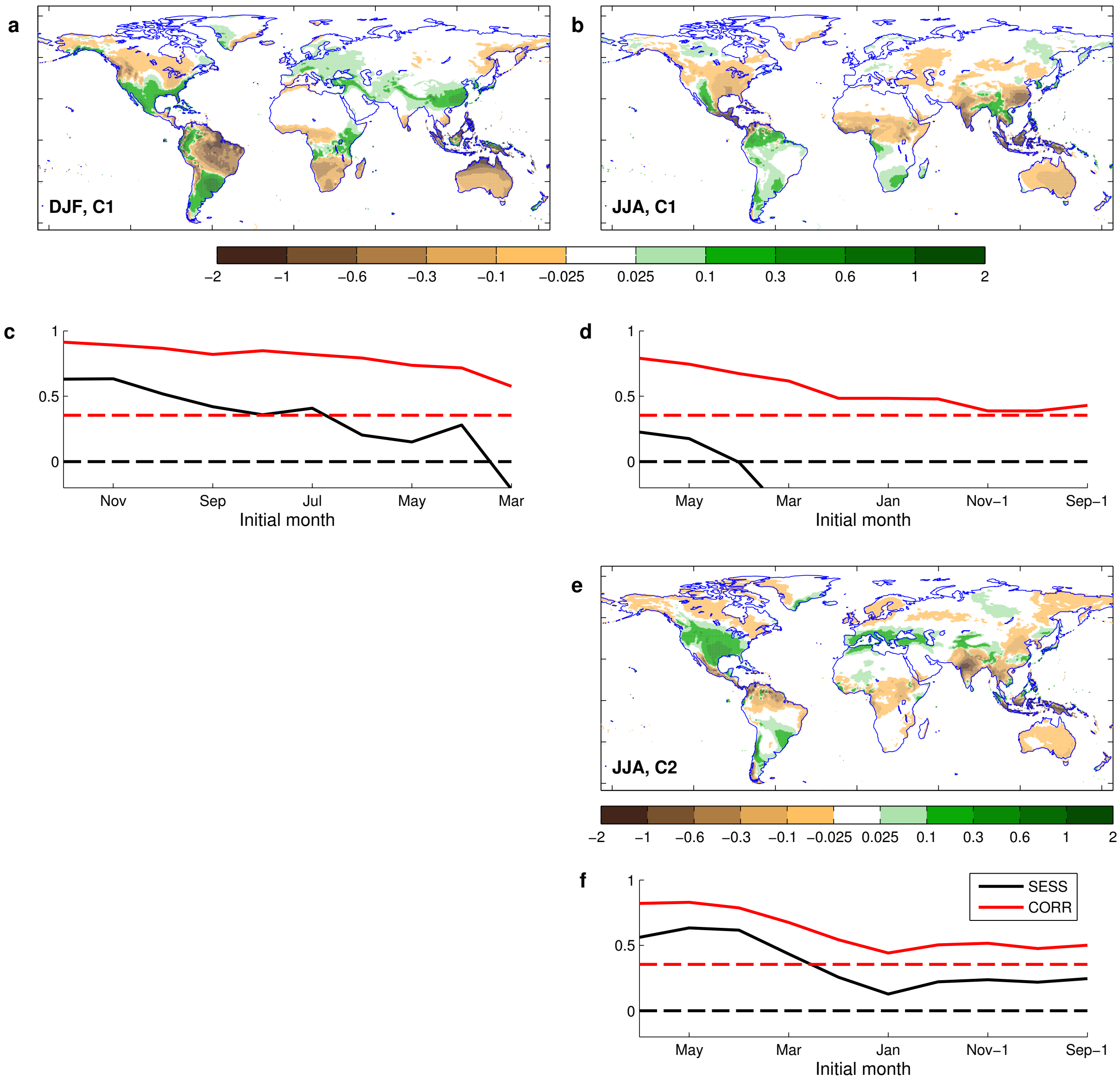

Figure 6: Spatial structure of the most predictable component of precipitation over land (in mmday^{-1} per unit variate) in DJF (a) and the first two predictable components in JJA (b, e); The corresponding SESS (black solid), anomaly correlation (red solid) skill as a function of initial month. The red dash lines indicate the 5% significance level for anomaly correlation. The black dash lines indicate zero SESS.

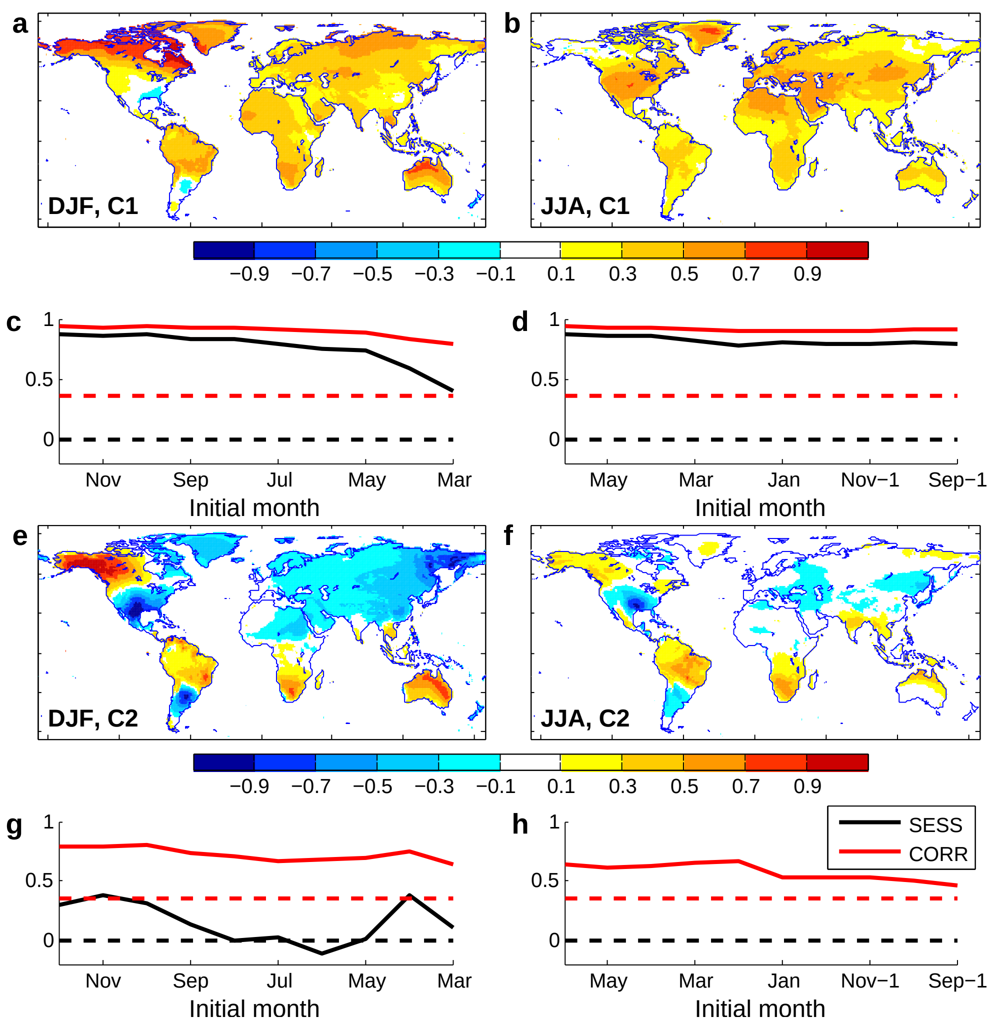

Figure 7: Spatial structure of the first two predictable components of 2m air temperature (in degree kelvin per unit variate) in DJF (a, e) and JJA (b, f) and the corresponding SESS (black solid) and anomaly correlation (red solid) skill as a function of initial month. The red dash lines indicate the 5% significance level for anomaly correlation. The black dash lines indicate zero SESS.

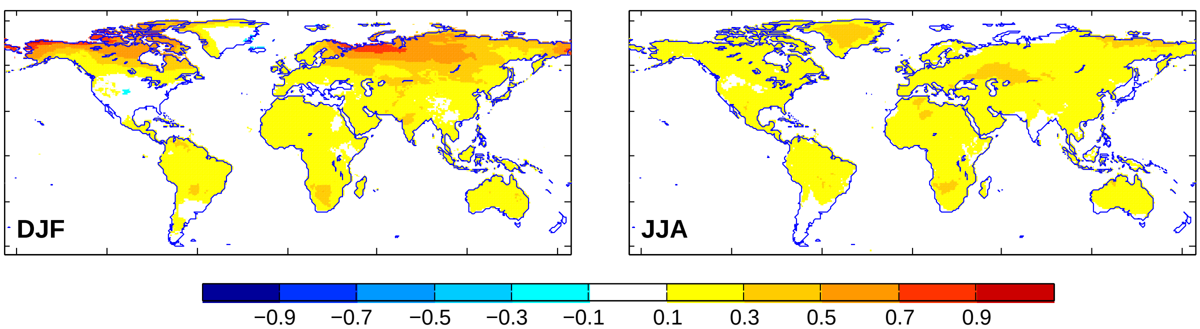

Figure 8: Externally-forced pattern of 2m air temperature (in degree kelvin per unit variate) over land in DJF (left) and JJA (right) derived from 5-member historical runs of FLOR from 1980 to 2005.

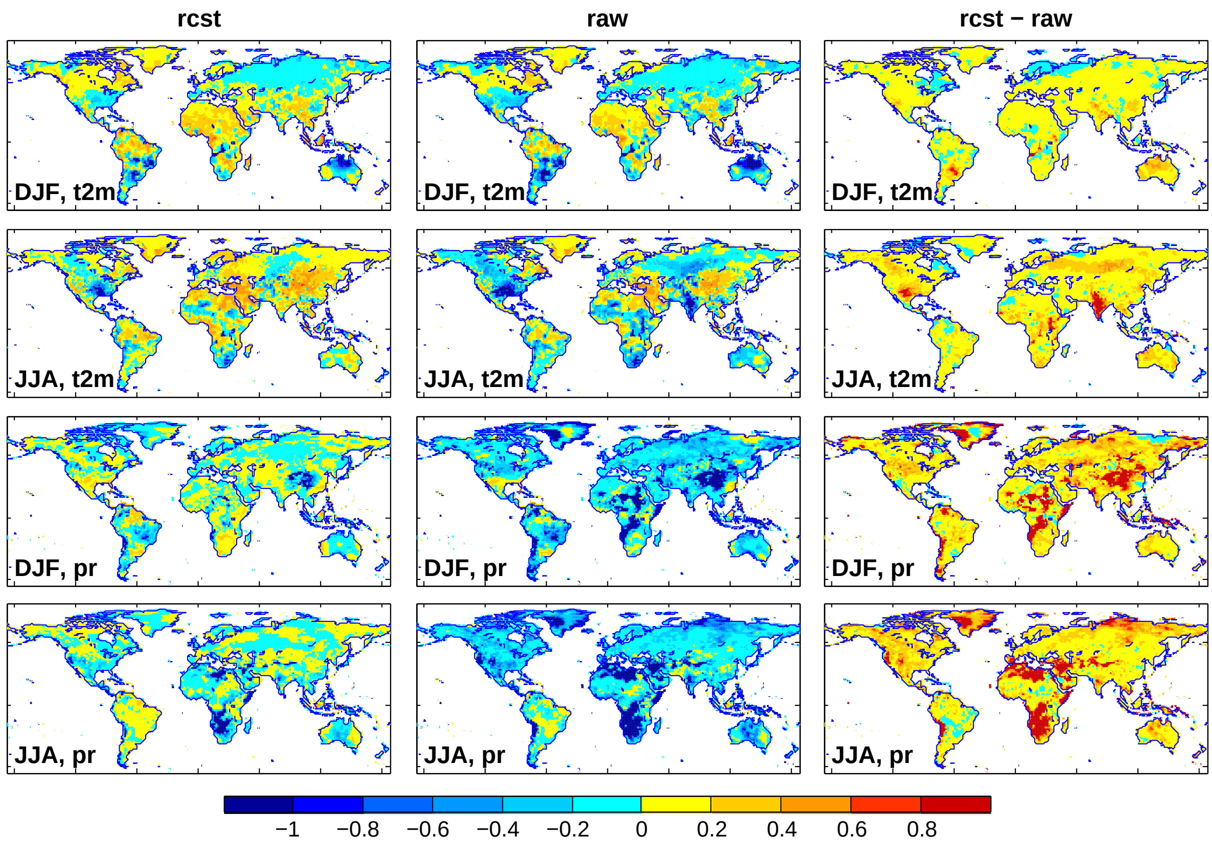

Figure 9: SESS of reconstructed predictions of 2m air temperature and precipitation from the leading predictable components(far left column), raw predictions directly from FLOR(middle column), and SESS of reconstructed predictions minus SESS of raw predictions (far right column). The SESS is averaged over lead times from 0 to 9 months.

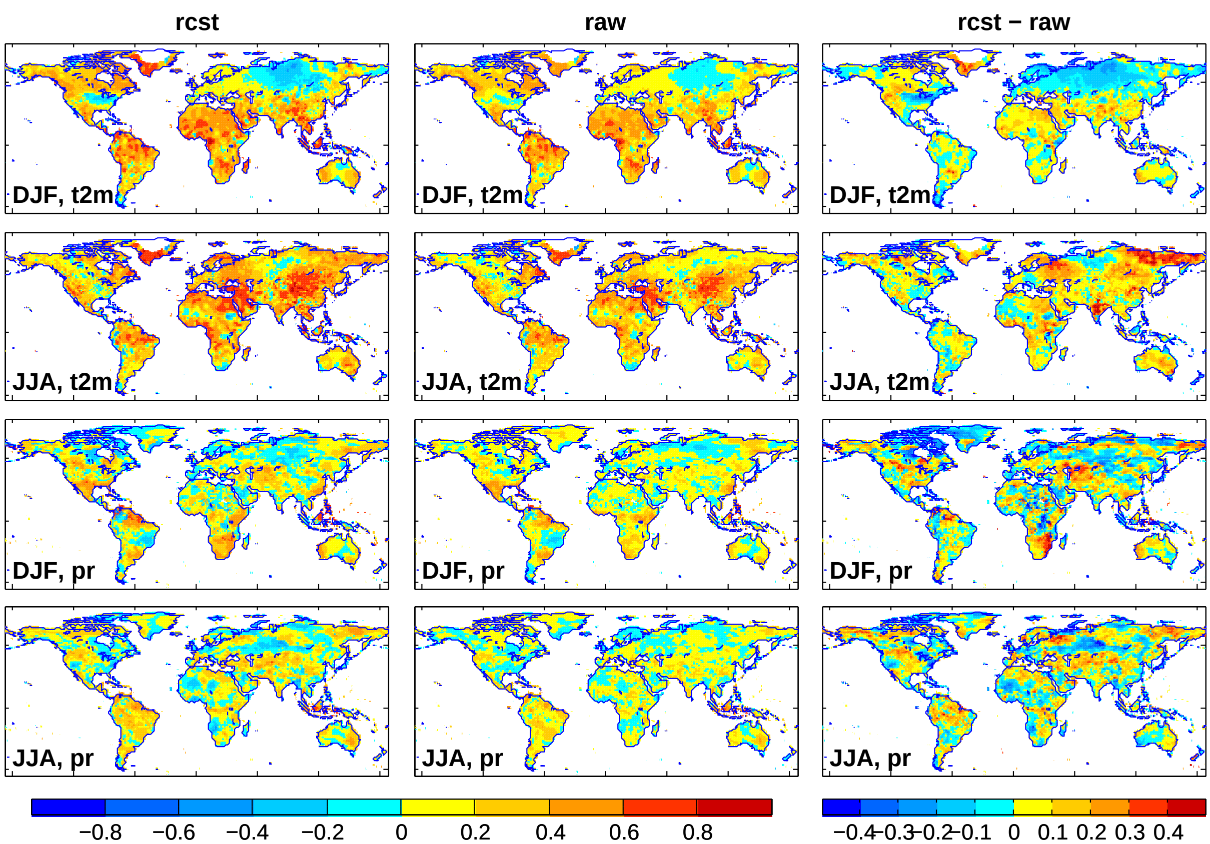

Figure 10: Correlation of reconstructed predictions of 2m air temperature and precipitation from the leading predictable components(far left column), raw predictions directly from FLOR(middle column), and correlation of reconstructed predictions minus correlation of raw predictions (far right column). The correlation is averaged over lead times from 0 to 9 months

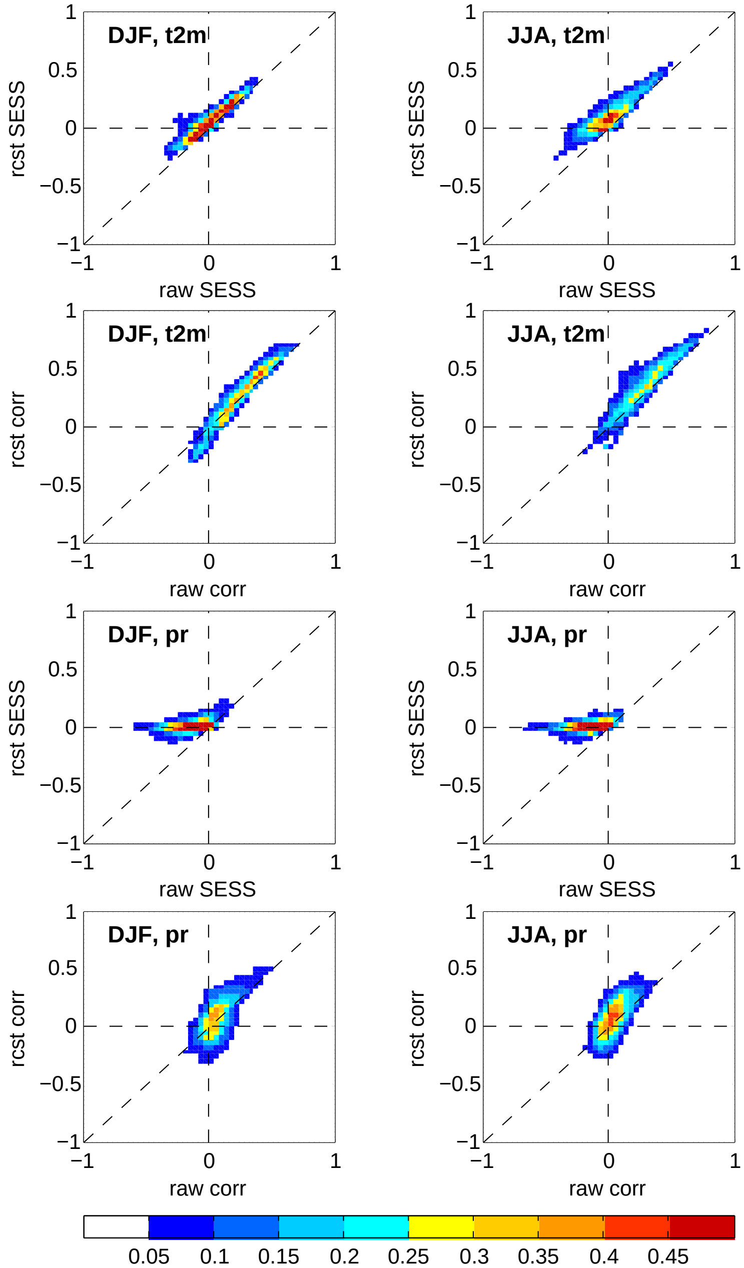

Figure 11: Percentage of grid points in each bin from -1 to 1 (interval of 0.04) for reconstructed vs. raw SESS and anomaly correlation of 2m air temperature and precipitation in DJF (left), JJA (right).