A New GFSv15 based Climate Model Large Ensemble and Its Application to Understanding Climate Variability, and Predictability

Tao Zhang[1,2], Weiyu Yang[2,3], Xiao-Wei Quan[4,5], Jieshun Zhu[2], Bhaskar Jha[2,3], Arun Kumar[2], Martin P. Hoerling[4,5], Joseph J. Barsugli[4,5], and Wanqiu Wang[2]

[1]*ESSIC, University of Maryland, College Park, MD*

[2]*NOAA/National Centers for Environmental Prediction, College Park, MD*

[3]*ERT Inc Laurel, MD*

[4]*NOAA/Physical Sciences Laboratory, Boulder, CO*

[5]*CIRES, Univ. of Colorado Boulder, Boulder, CO*

Corresponding Author: Tao Zhang, tao.zhang@noaa.gov

Submitted to Journal of Geophysical Research: Atmospheres

Abstract

NOAA Climate Prediction Center (CPC) has generated a 100-member ensemble of Atmospheric Model Intercomparison Project (AMIP) simulations from 1979 to present using the GFSv15 with FV3 dynamical core. The intent of this study is to document a development in an infrastructure capability with a focus to demonstrate the quality of these new simulations is on par with the previous GFSv2 AMIP simulations. These simulations are part of CPC's efforts to attribute observed seasonal climate variability to SST forcings and get updated once a month by available observed SST.

The performance of these simulations in replicating observed climate variability and trends, together with an assessment of climate predictability and the attribution of some climate events is documented. A particular focus of the analysis is on the US climate trend, Northern Hemisphere winter height variability, US climate response to three strong El Niño events, the analysis of signal to noise ratio (SNR), the anomaly correlation for seasonal climate anomalies, and the South Asian flooding of 2022 summer, and thereby samples wide aspects that are important for attributing climate variability. Results indicate that the new model can realistically reproduce observed climate variability and trends as well as extreme events, better capturing the US climate response to extreme El Niño events and the 2022 summer South Asian record-breaking flooding than GFSv2. The new model also shows an improvement in the wintertime simulation skill of US surface climate, mainly confined in the Northern and Southeastern US for precipitation and in the east for temperature.

Key points

- A large AMIP ensemble based on NOAA's GFSv15 with FV3 dynamical core is created to support attribution of observed climate anomalies at CPC.
- The new simulations can replicate the observed climate variability and trends as well as extreme seasonal events.
- There are some improvements in simulating the extreme events in the new model compared to the older version.

Plain Language Summary

To correctly account for extreme weather and climate events such as heatwaves, floods and droughts that have devastating effects on the US economy and human lives, climate model experiments have become a key tool to disentangle numerous responsible factors. A recent development of an updated modeling framework at the National Centers for Environmental Prediction (NCEP) to support the attribution of observed seasonal anomalies is reported in this study. We have generated a 100-member ensemble of simulations in which each member has identical SST forcing but differs only by the initial atmospheric condition. These simulations are updated once a month when the observed SST data becomes available. We use the ensemble mean of these simulations to describe the responses to SST (referred to as the potentially predictable component of observed anomalies) and use the departure of individual members from the ensemble mean to assess the unpredictable component in the atmospheric variability. We document the performance of these simulations in replicating the observed climate variability, trends and extreme events, and find that the new model can realistically reproduce the observed key features and has a better simulation of extreme seasonal events compared to the previous version.

1. Introduction

Needs for understanding climate variability and predictability, understanding of long-lasting climate anomalies, and reasons for success and failures for long-range predictions, can be well served by ensembles of AMIP-style simulations, that is, atmosphere-only simulations that are constrained by the evolution of realistic SSTs and sea ice (Gates et al. 1998). The AMIP approach allows for the isolation of the atmospheric sensitivity to observed and specified evolution of SSTs, though it cannot explain the origin for the SSTs themselves. The ensemble mean of AMIP simulations documents the response to SSTs, often referred to as the potentially predictable (or attributable) component of the observed anomalies, or potential for predictions well beyond the limits of when initial atmosphere conditions constrain weather. The contribution of the unpredictable component in the atmospheric variability can also be assessed from the analysis of the departure of individual AMIP model simulations from the ensemble mean anomalies. In addition, the analysis of individual simulations can evaluate the role of noise in the level of discrepancy of the observed anomalies from the predictable (attributable) component because of the correspondence between observed anomalies and a realization of a single model run (Kumar et al. 2013).

AMIP simulations are well suited to understand causes for extreme weather and climate events including floods, droughts, and heat waves that are known to have devastating effects on human lives and the economy of the United States (Changnon 1999; Seager et al. 2015; NOAA 2017; Philip et al. 2021). For example, southern states and California were plagued by storms attributed to El Niño 1997-98. In addition to the losses of 189 lives, the estimated economic losses nationally were about $4 billion (Changnon 1999). The California drought of 2014 cost California $2.2

billion in damages and 17000 agricultural jobs (Howitt et al. 2014; Seager et al. 2015). The recent heatwave of June 2021, whose temperature records were historically highest in some cities in the Pacific northwest of the U.S. and Canada, caused a sharp increase in sudden deaths and hospital visitations for heat-related illnesses and emergencies (Philip et al. 2021). The key for predicting these events depends strongly on understanding their causal relationship with external drivers (an exercise often referred to as attribution), for example, slowly evolving SST anomalies, decadal variability, and long-term trends. However, as causal relationships seldom explain a large fraction of total variability and are superimposed on the internal variability (e.g., Kumar et al. 2013; Zhang et al. 2018), observations, due to their limited sample, alone are inadequate to fully establish such relationships, particularly on an individual event basis. For this purpose, climate model experiments, for example, the aforementioned AMIP simulations, have become an indispensable tool to disentangle the various factors accounting for extreme weather and climate variability on different time scales (Murray et al. 2020; Barsugli et al. 2022). In this paper, a recent development of such a modeling framework at NCEP in support of the attribution of observed climate anomalies is reported.

Climate attribution is a scientific process for establishing the principal causes or physical explanation for observed climate conditions and phenomena. To date, the attribution efforts at CPC have relied on the current operational seasonal prediction system - the Climate Forecast System v2 (CFSv2) (https://www.cpc.ncep.noaa.gov/products/people/mchen/AttributionAnalysis/). To provide a historical perspective, the first Climate Forecast System (CFS), called CFSv1, was implemented into operations at the NCEP in August 2004 and was the first fully coupled atmosphere–ocean–land model used at NCEP for seasonal prediction (Saha et al. 2006).

Subsequently, the CFSv2 was made operational in March 2011 (Saha et al. 2014), with upgrades to all aspects of the data assimilation and forecast model components. CFSv2 generates a set of 9-month retrospective forecasts with forecasts initialized using the analysis from the corresponding Climate Forecast System Reanalysis (CFSR) (Saha et al. 2014).

In addition to the initialized CFSv2 forecasts, at CPC a large ensemble of AMIP simulations based on GFSv2, the atmospheric component of the CFSv2, updated in real-time, has also been maintained to attribute causes for the observed real-time seasonal climate anomalies by identifying the impacts of anomalous boundary forcing (particularly due to SSTs). The ensemble of AMIP simulations with GFSv2 has been used to diagnose the forced response to observed SSTs, such as the forced atmospheric teleconnections during 1979-2014 (Hartmann 2015; Zhang et al. 2016), the causality of California rains (Seager et al. 2015; Zhang et al. 2018), and US surface climate response associated with El Niño flavors (Zhang et al. 2020).

Despite continued improvements in spatial resolution, energy conservation, and computational efficiency, the hydrostatic spectral dynamical core of the NCEP Global Forecast System (GFS) [Global Spectral Model (GSM)] has not been upgraded since the 1980s. In 2016, the Finite-Volume Cubed-Sphere Dynamical Core (FV3), developed at the NOAA/Geophysical Fluid Dynamics Laboratory (GFDL), was selected as the dynamical core of NOAA Next Generation Global Prediction System (NGGPS) project as an upgrade for the GSM. The advantage of FV3 includes its high efficiency and scalability, run-time switchable nonhydrostatic solver allowing for convective-scale simulation, exact mass and approximate energy conservation, skillful forecasts and adaptability to the present GFS physics and data assimilation system and its robust kinetic energy spectrum (Zhou et al. 2019).

In recent years, a new global model coupling the FV3 with GFS physical parameterizations, called the finite-volume Global Forecast System, or FV3/GFS (Zhou et al. 2019), has been developed. The FV3 GFS was implemented into the operational Global Forecast System as version 15 (GFSv15) in 2019 (https://www.emc.ncep.noaa.gov/users/meg/fv3gfs/).

To continue supporting requirements for the attribution of seasonal climate anomalies and to assess the reasons for the success and failures of operational seasonal forecasts, CPC also upgraded AMIP simulations from the GSM based atmospheric model to one based on FV3GFS. As part of this effort, a large 100-member ensemble of AMIP simulations from 1979 to present using the GFSv15 with FV3 dynamical core has been generated.

The goal of present analysis is to introduce this data set that can be used for understanding various aspects of climate variability, document the performance of these simulations in replicating observed climate variability and trends, development in an infrastructure capability by comparing the quality of FV3 GFS model simulations with those of GFSv2, and give some examples of the assessment of climate predictability and attribution of some climate events. The focus of this study is on the evaluation of the performance of FV3 GFS AMIP simulations relative to GFSv2 in replicating observed climate variability and trends for the period of 1979-2021.

This paper is organized as follows: We introduce the observational and model datasets as well as analysis methods in section 2. Section 3 first presents an assessment of the climatology in the model, and then the simulation of observed trend and northern hemisphere winter height variability. Finally, the simulation of US climate response to ENSO, the assessment of climate predictability and simulation of extreme seasonal events are also presented. Conclusions and discussions are given in section 4.

2. Datasets and methods

a. Observed and model data

The characteristics of observed estimates for land surface climate conditions are based on analysis of the Global Historical Climatology Network/Climate Anomaly Monitoring System (GHCN/CAMS) 2-meter temperature (T2m) (Fan and van den Dool 2008) and gauge-based gridded monthly Global Precipitation Climatology Centre (GPCC) data sets (Schneider et al. 2014), available at 1^{o}-by-1^{o} resolution. Same as the data used in Zhang et al. (2006), observed estimates of the upper-level circulation pattern are based on 200-hPa geopotential height fields using the National Centers for Environmental Prediction –National Center for Atmospheric Research reanalysis (Kalnay et al. 1996). To explore the possible tropical drivers for land surface climate conditions and upper-level circulation patterns, we also analyzed global teleconnection associated with the tropical SST and precipitation variability. The observed SST data, on a 1^{o}-by-1^{o} grid, are from the Hurrell data set (Hurrell et al. 2008), which is a combined version of the Hadley Centre's SST version 1.1 (HADISST1) and the NOAA Optimal Interpolation (OI) SST version 2 (OISSTv2) from November 1981 onward. Global precipitation fields are from the CPC Merged Analysis of Precipitation (CMAP; Xie and Arkin 1997) and are available at 2.5^{o}-by-2.5^{o} resolution.

We utilize an atmospheric model simulation [also referred to as AMIP experiments] based on NOAA's GFSv15 model with the Finite-Volume (FV3) dynamical core (Putman and Lin, 2007) on a cubed-sphere grid (https://www.emc.ncep.noaa.gov/users/meg/fv3gfs/). The GFSv15 uses the Rapid Radiative Transfer Method for General Circulation Models (RRTMG) scheme for shortwave and longwave radiation (Iacono et al. 2008), hybrid eddy-diffusivity mass flux turbulence scheme (Han et al. 2016), GFDL microphysics (Zhou et al. 2019), and scale-aware mass flux convection scheme (Han and Pan 2011). The GFSv15 physics also includes Noah land surface model and a revised bare-soil evaporation scheme. A three-layer thermodynamic sea ice model (Winton 2000) has been coupled to the GFSv15 and it predicts sea ice thickness. Detailed description of parameterization schemes, with associated references, can be found at https://dtcenter.ucar.edu/GMTB/v3.0/sci_doc/GFS_v15_page.html.

A version of this atmospheric model is currently the operational global weather prediction system at NCEP. The FV3 GFS model used in our simulations is run at C96 horizontal resolution with 64 vertical levels and forced with specified observed monthly varying SSTs, sea ice (Hurrell et al. 2008), and carbon dioxide concentrations from the World Data Centre for Greenhouse Gases (WDCGG) operated by the Japan Meteorological Agency (JMA) for 1979–2021. Climatological values are specified for other greenhouse gases, aerosols, solar, and volcanic aerosols. A 100-member ensemble of AMIP simulations is maintained at NOAA's CPC. Each member in the ensemble has identical external forcing but differs only by its initial atmospheric condition. The forced response to external forcings is derived from the statistics of 100-member simulations, e.g., ensemble average.

To assess the robustness of key features in replicating observed climate variability and trends by FV3 GFS, we also diagnose the AMIP simulations from a 30-member ensemble of the GFSv2 model that spans the same period. As the atmospheric component of the NCEP CFSv2 (Saha et al. 2014), the GFSv2 model is the previous version of CPC AMIP simulations and is run at spectral T126 horizontal resolution with 64 vertical levels.

b. Methods

In the present study, we follow the methodology of Zhang et al. (2016) to obtain the observed leading structures of the Northern Hemisphere (NH) wintertime circulation variability by applying empirical orthogonal function (EOF) analysis to DJF seasonally averaged 200-hPa heights for the 42 years of data during 1979-2021 period. The EOF analysis is based on the covariance matrix for 20°N-90°N latitude band and the EOF patterns are presented as regressions against the principal component (PC) time series. Note that unrotated EOFs utilized here are constructed to be both spatially and temporally uncorrelated with each other.

Leading EOF modes of observational variability have contributions both from the atmospheric internal and forced variability. We complement this analysis with the EOF mode analysis of ensemble mean AMIP data to isolate the forced signals. We then provide a comparison of the first three leading modes of variability of the observed and FV3 GFS simulated DJF 200-hPa geopotential heights from the individual members of AMIP simulations, which demonstrates that the model can well capture the observed three leading modes of interannual variability.

Based on a 100-member ensemble of FV3 GFS AMIP simulations, probability density function (PDF) is analyzed to reveal the statistics of US climate trends by examining the frequency distributions of surface climate conditions over two different periods. We also plot the PDF of California rainfall from the large ensemble of FV3 GFS to explore the possible cause of observed failed California rains during the strong 2016 El Niño winter.

Finally, climate predictability in our analysis is further assessed by examining the signal-to-noise ratio (SNR) which quantifies predictable (signal) and unpredictable (noise) components. The signal component in the SNR is the variance of ensemble mean while the noise component is the variance of departure in the individual members from the ensemble mean (Kumar and Hoerling 1995). Higher SNR values indicate larger predictability. The anomaly correlation (AC), defined as the correlation of anomalies between AMIP ensemble means and observations, is calculated to complement SNR analysis. It is expected that larger SNR would correspond to larger AC (Kumar and Hoerling 2000). Anomalies are computed relative to a 1991-2020 reference period for AMIP simulations and observations.

3. Results

a. Assessment of the climatology

Instead of a direct comparison of the climatology between model and observations, we focus on the assessment of the seasonal cycle of climatology because observed estimates of quantities like surface air temperature and rainfall can be problematic (Fan and van den Dool 2008; Xie and Arkin 1997).

Figure 1 shows the difference in climatology between JJA and DJF (JJA minus DJF) for observations (left panel) and FV3 GFS AMIP ensemble mean (right panel). The largest difference in eddy (zonal mean removed) 200-hPa height is in the Northern Hemisphere (NH) middle latitude. The observed positive centers over the Asia and North American Continent and negative centers over the North Pacific and North Atlantic are well captured in the model, with a high pattern correlation of 0.94.

Observed precipitation difference shows that there is increased precipitation in the north of the equator and decreased precipitation in the south of the equator. This feature is realistically reproduced in the model. Surface air temperature difference pattern is also similar between model and observation, with warming in the northern hemisphere land and cooling in the southern hemisphere land. The pattern correlation is 0.98.

The global mean values of the differences in climatology between summer and winter for eddy 200-hPa, precipitation and surface air temperature are also comparable in the model and observations (see the first value in the titles of maps). The results suggest that FV3 GFS can realistically capture the observed seasonal cycle of climatology. Table 1 lists the global mean values of the difference in climatology (JJA minus DJF) for GFSv2 and FV3 GFS and the respective global pattern correlations with observations. For eddy 200-hPa, the global mean values of climatology differences in two models and the pattern correlations with observations are comparable. Compared to GFSv2, the pattern correlation with observations for precipitation and the global mean value for surface air temperature are improved in FV3 GFS model to some extent.

We further use Taylor diagram (Taylor 2001) to provide a summary of the relative skill with which two models simulate the spatial pattern of annual mean precipitation and surface air temperature over different regions (Figure 2). Two models generally demonstrated a similar ability to simulate the annual mean surface air temperature and precipitation, featuring the largest pattern correlation (greater than 0.98) and the lowest normalized root-mean-square (RMS) error (less than 0.2) for global temperature, and the smallest pattern correlation (about 0.81) and largest RMS error (about 0.65) for tropical precipitation. The standard deviation of global and tropical precipitation is somewhat overestimated, and the standard deviation of tropical temperature is slightly underestimated in the models. Over the contiguous US, the pattern correlations for both temperature and precipitation are greater than 0.90 in two models, while the standard deviations of these two fields are closer to observations in FV3 GFS relative to GFSv2. The simulation of observed trends and climate variability are discussed in the following sections.

b. Simulation of observed trends

Human activities, especially emissions of greenhouse gases, are extremely likely to be the dominant cause of the observed warming trends of global land temperature since the mid-20th century (Wuebbles et al. 2017). A large fraction of these changes is communicated to the atmosphere via the indirect influence of trends in SSTs (Hoerling et al. 2006; Compo and Sardeshmukh 2009; Fahad and Burls 2022). Because the oceans also continuously interact with the atmosphere, SSTs can have considerable effects on global climate variability on different time scales. In addition to the effect of warming oceans on continental temperature trends, increases in SST have also led to an increase in the amount of atmospheric water vapor over the oceans (Yang and Tung 1998). The increased water vapor can enhance the amplitude of climate feedback in response to anthropogenic activities through positive feedback (Held and Soden 2000; Soden et al. 2005).

Because trends contribute to seasonal anomalies especially for temperature-related variables, attribution analysis includes the influence of anthropogenic forcings (either through their direct influence via the radiative forcing or indirect influence via changes in SST that are specified in the AMIP simulations). We therefore document the ability of the AMIP runs to simulate observed trends, particularly in temperature where the influence is most prominent.

Figure 3 shows the time series of DJF (top) and JJA (bottom) surface air temperature anomalies averaged over global land for 1979-2021. The red line indicates observations, and the blue line and black line show the ensemble means of GFSv2 and FV3 GFS AMIP runs, respectively. To compare observations against the individual runs, and to see if the observed variability is within the envelope of model solutions, the time-series of land temperature in the 100 individual runs from FV3 GFS are also shown (gray lines). It is clear that observations have an upward trend of about 1° C since 1979 for both winter and summer. The FV3 GFS model ensemble mean agrees well with the observed trends. In most cases, the observed value is within the envelope of ensemble spread (gray lines) (that, as expected, has larger variability during winter compared to summer). The previous version GFSv2 ensemble mean has a similar temporal correlation with observed trend as FV3 GFS for winter, but the correlation is somewhat smaller in GFSv2 for summer.

To explore whether FV3 GFS model can capture the observed trends, Figure 4 shows the frequency distributions [also called probability density functions (PDF)] of wintertime surface air temperature (top left) and precipitation (top right) from AMIP runs for the first 5-yr period (blue curve) and last 5-yr period (red curve) of the simulations over the contiguous United States. These two curves, which are significantly different according to the Kolmogorov-Smirnov test, are based on 1500 (100 members multiply by 15 months for 5-yr period) model samples. Short tick marks across the bottom indicate 15 observed values during the corresponding 5-yr period.

For the two periods the observed values are located within the spread of model samples for both wintertime temperature and precipitation. A feature to note is that the red curve is shifted toward warmer and drier conditions compared to the blue curve. This indicates that the latter period is warmer and drier than the earlier period for spatial average over the contiguous US. In other words, there is a US warming and drying trend during the winter (Weaver et al. 2014).

The results for summer shown in Figure 4 bottom are similar to those for winter, and two curves are also significantly different through the Kolmogorov-Smirnov test, confirming a US warming and drying trend during the summer as well. Also, as expected, the variability is smaller in summer compared to winter, which is a common feature both in the model and the observation. The US warming trend is also found for both DJF and JJA seasons based on GFSv2 AMIP 30-member ensemble. However, there is no consensus on the precipitation trend for these two seasons in GFSv2 (Fig. S1 in the supplementary material).

c. Simulation of Northern Hemisphere wintertime height variability

The long-lasting climate anomalies are usually related to the leading modes of climate variability (e.g., Hartmann 2015; Zhang et al. 2016). Atmospheric teleconnections associated with ENSO are known to be the underpinnings for North American seasonal climate predictability (Horel and Wallace 1981; Trenberth et al. 1998). Further, understanding the atmospheric response patterns beyond the canonical response to ENSO is also an outstanding problem in quantifying the sources of predictability and attribution of climate variations, and further, may result in improvements in our understanding of seasonal predictability (Hoerling and Kumar 2002; Barnston et al. 2005; Kumar et al. 2005; Zhang et al. 2016). It is thus important to assess the capability of FV3 GFS in reproducing the leading modes of climate variability.

Figure 5 shows wintertime (DJF) 200-hPa height structures based on the leading three EOFs of the reanalysis data, which explain a combined 56.6% of the height variability poleward of 20^{o}N. Contours in the left panels and shaded values in the right panels are the observed 200-hPa heights and SSTs regressed against each eigenvector's PC time series shown in the middle panels for 1979-2021, respectively.

The structure of the first leading mode of the observed variability consists of positive height anomalies in the NH middle latitudes and negative anomalies in the polar regions while the time series for this mode is uncorrelated (the value is -0.026) with Niño-3.4 SST variability. This pattern explains 26.1% of extratropical NH wintertime height variability. Zhang et al. (2016) found a similar mode of observed height variability, though ranked second in its EOF decomposition and explaining a somewhat small fraction of height variance for 1979-2014 period. They further noted that this mode can also be reproduced in a climate simulation having no interannual variability in boundary SSTs or external radiative forcing. It is clear that the observed first mode, therefore, is mainly due to internal atmospheric variability. SST regression map (top right) confirms that this mode, resembling Arctic Oscillation (AO) pattern (Thompson and Wallace 1998), is not related to tropical SST forcing.

Explaining 17.1% of the NH extratropical height variability, the observed second EOF pattern consists of a prominent wave train over the Pacific-North American (PNA) region, resembling the tropical/Northern Hemisphere (TNH) pattern (Mo and Livezey 1986). The time series for the second mode has a moderate correlation (0.58) with Niño-3.4 SST variability. The corresponding SST regression map (middle right) reveals a feature of El Niño SST warming pattern, indicating that the second mode describes the canonical atmospheric teleconnection response associated with ENSO.

The third EOF of the observed variability explains 13.4% of the variance in height variability, whose pattern, temporal variability and the corresponding SST regression (bottom panels) suggest a possible connection with global warming. The EOF3 pattern largely features a same sign hemisphere-wide pattern and the PC3 times series has a distinct upward trend associated with a dominance of SST warming over the global oceans, suggesting a tendency for NH heights (corresponding to a tropospheric warming) to rise since 1979. This observed EOF3 is very similar to the dominant EOF mode in a large ensemble of CMIP simulations in which the only forcing is anthropogenic greenhouse gases (Zhang et al. 2016), supporting the argument that this mode is related to the anthropogenically forced climate change.

We evaluate the model's ability to replicate the leading modes of observed variability. For model simulations, however, the leading EOFs can be computed for each of the 100 individual members.

Further, because of sampling, the spatial pattern and the corresponding PC time series has variations from one ensemble member to another. To quantify the fidelity of leading modes of model variability against observations, one approach is to compute pattern correlations between model and observed EOFs and repeat this process for all 100 individual members. These correlations are shown in Fig. 6 (right panels).

The EOF1 pattern correlation between individual members of the AMIP simulations and observations based on 42 winters ranges from 0.038 to 0.80, and the EOF2 pattern correlation ranges from 0.0076 to 0.72. The corresponding mean value of EOF1 pattern correlations with observation from 100 individual AMIP members is 0.61, much larger than the mean value (0.28) of 100 EOF2 pattern correlations. The range of EOF3 pattern correlation is more scattered, with values ranging from 0.0035 to 0.86, and the corresponding mean value of 0.47. In general, due to sampling variability, there is large uncertainty in the spatial details of EOF structures from one ensemble member to another leading to a similar variability in spatial correlations, especially for the last two modes.

FV3 GFS can reproduce the pattern of observed first three leading modes with moderate to high correlations (see figure captions for correlation values), as is evident from the results of a single member (Figure 6 left) for which the mean correlation for the first three modes with observations is largest. The explained variance for each mode from this run is also very close to observed values. But the correlation of the PC1 time series from this member with the observed PC1 time series is 0.21, much smaller than the PC2 counterpart (0.55). The correlation of PC3 time series between this member and observations (0.44) is roughly double the corresponding value for PC1. We note that the amplitude of these correlations depends on to what extent these modes are a result of atmospheric internal variability and to what extent they are constrained by the evolution of SSTs. If a mode is dominated by the atmospheric internal variability, then even though the spatial pattern of the EOF between observations and model simulation may be the similar, the corresponding time-series could still be uncorrelated.

Since one of the applications of AMIP simulations is to understand the forced response to SSTs, our analysis further explores the forced atmospheric variability during 1979-2021 by using the 100-member ensemble mean of AMIP simulations (Figure 7). The three leading EOFs of the ensemble mean AMIP simulations together explain 84.8% of the total boundary forced ensemble mean height variance.

The height pattern associated with the first mode of forced AMIP response describes a prominent wave train over the PNA region that resembles the TNH pattern. The time series for this leading mode shows a high correlation (0.93) with Niño-3.4 SST variability, featuring positive polarity during warm events (e.g. 1982/83, 1991/92, 1997/98, 2002/03, 2009/10, 2015/16) and negative polarity during cold events (e.g. 1988/89, 1998/99, 1999/2000, 2007/08, 2011/12, 2020/21). The corresponding SST regression against PC1 time series confirms that this mode is clearly related to ENSO, similar to the observed second mode shown in the middle of Figure 5. This forced pattern alone explains 41.5% of the total boundary forced component of extratropical NH wintertime model simulated ensemble mean height response.

Associated with a ubiquitous warming over the global oceans, the second mode of forced AMIP solutions is characterized by a hemisphere-wide increase in heights. This forced mode resembles the observed third mode shown in the bottom of Figure 5 that is strongly related to climate change discussed earlier. The explained variance by this forced mode is 28.0%.

Explaining 15.3% of the total boundary forced height variability over the NH extratropics, the height pattern associated with the third mode of forced AMIP response describes a wave train resembling the classic PNA pattern. Its action centers are in spatial quadrature with the leading forced solution, similar to the second EOF pattern in Zhang et al. (2016). The larger amplitudes in the corresponding PC3 time series tend to occur during ENSO events (e.g. 1982/83, 1997/98, 2015/16 warm events, and 1988/89, 1998/99, 1999/2000, 2007/08, 2011/12 cold events), large projections also occur during several ENSO-neutral years (e.g. 1985/86, 1996/97, 2013/14). Zhang et al. (2016) found that there is a high correlation between PC time series for this forced mode and trans-Niño (TNI) SST index, which measures the evolution of ENSO during its transition phase (Trenberth and Stepaniak 2001). The SST regression map (Figure 7 bottom right) is very close to the SST asymmetry between El Niño and La Niña events (Zhang et al. 2016). Stronger El Niños have larger SST magnitudes in the eastern equatorial Pacific while stronger La Niñas have larger magnitudes in the western equatorial Pacific, causing a positive skew in the Niño-3 index indicative of nonlinearity in SST forcings (Burgers and Stephenson 1999; An and Jin 2004; Zhang et al. 2009; Zhang and Sun 2014). Therefore, the positive phase of the forced third mode is linked to the asymmetry in ENSO teleconnections between their extreme opposite warm and cold phases. Zhang et al. (2016) also argued that for the negative phase of this mode, the SST pattern is analogous to a pattern that is the precursor to El Niño development (e.g., Penland and Sardeshmukh, 1995), featuring warmth in the far western Pacific and coolness in the far eastern Pacific. This may indicate that the negative phase of third mode is shown to be an expression of atmospheric response to a tropical precursor SST for ENSO development that occurs mostly during ENSO-neutral winters.

To assess the robustness of key features of the forced atmospheric variability, we repeat the analysis of Figure 7 by using the 30-member ensemble mean of GFSv2 AMIP simulations (Fig. S2 in the supplementary material). The results are found to be similar, including the EOF ranking and explained variance of three leading forced modes.

The observed leading mode, i.e., the AO pattern is absent among the first three leading modes of AMIP forced solutions. The results lend further support to the previous argument that the observed first mode is very likely attributed to unforced variability.

d. Simulation of US climate response to ENSO

ENSO is the largest source of atmospheric predictability and an important aspect of climate attribution (e.g., Kumar and Hoerling 1998; Goddard and Dilley 2005; Quan et al. 2006), and therefore, it is essential to quantify the fidelity of ENSO response in AMIP simulations.

Figure 8 compares the spatial pattern of the regressions of wintertime 200-hPa height, precipitation and surface air temperature anomalies on the observed Niño-3.4 SST index between FV3 GFS AMIP simulations (right panel) and observations (left panel). The regressions for the model are obtained by first calculating the regressions for individual runs and then averaging 100 regression estimates.

In response to El Niño warming, the observed upper-tropospheric circulation anomaly shows the classic El Niño-related teleconnection pattern consisting of anomalous tropical anticyclones, cyclonic anomalies over the North Pacific and anticyclonic anomalies over the North American continent. The observed precipitation is characterized by reduced convection over the tropical western Pacific and enhanced convection over the tropical Indian Ocean and tropical central and eastern Pacific. The temperature response reveals warming (cooling) over the northern (southern) United States, similar to the observed surface temperature composite during Eastern Pacific (EP) El Niño (Zhang et al. 2020). Appreciable warmth is also observed over Eurasia in the middle latitude.

FV3 GFS AMIP results reproduce the observed key features associated with ENSO. The magnitude of the negative surface temperature anomalies, however, is overestimated over the Southern United States, where the simulated cyclonic anomalies are also stronger. We also note that the magnitude of observed warmth is somewhat underestimated over Eurasia, South Africa, and Australia and overestimated over the Northern South American continent. It should be noted that while the ENSO response in the model simulations is the average of 100 estimates, and therefore, has a higher statistical significance, the observed estimate could be influenced by sampling variability.

Next, we compare the seasonal climate variability for extreme El Niño events, and further, discuss the role of internal variability in shaping the observed anomalies. We also evaluate how well the FV3 GFS model simulates the US climate response to ENSO compared to the previous GFSv2 model that has been used for attribution studies.

Figure 9 shows the wintertime surface air temperature anomalies for three strong El Niño events from observations (left), GFSv2 simulated ensemble mean (middle) and FV3 GFS simulated ensemble mean (right). During the 1982/83 El Niño, maximum warm temperature anomalies are located over the northern United States, but the surface temperature is colder than normal over the southern United States. The above normal anomalies shift gradually from north to south in recent two strong El Niño (1997/98 and 2015/16) events.

Similar to observations, there is a clear southward shift of warm anomalies from 1982/83 El Niño to 2015/16 El Niño for two model ensemble mean results. This is consistent with the US warming trends documented using PDFs (Fig. 4). The models have a moderate (0.4~0.5) pattern correlation with observations in 1982/83 El Niño and a higher pattern correlation (above 0.7) with observations in recent two strong El Niño events. Despite the comparable pattern correlations with observation for two models, there is an improvement in FV3 GFS model relative to GFSv2 in the south-eastern coastal regions of the U.S. where the GFSv2 has too strong cold anomalies but the simulations from FV3 GFS are closer to observations.

Figure 10 shows the corresponding precipitation anomalies for three strong El Niño events from observations and simulations from two models. The observed precipitation patterns for both

1982/83 and 1997/98 are very similar, with wetter anomalies in the west and central US and southern coast. However, negative rainfall anomalies over southern California are observed for 2015/16 winter. It can be seen that the 1997/98 El Niño has the largest wetness in the southwest.

In contrast to the observed anomalies, the ensemble mean precipitation response in two models has a very similar pattern for all three strong El Niño events, characterized by a wetness across the west, central US and southern coast that resembles the observed precipitation responses to 1982/83 and 1997/98 El Niño events. Further, opposite to the observed dryness in Southern California, the model ensemble mean response shows that the Southern California has wet conditions in 2015/16 El Niño, consistent with previous studies (e.g., Chen and Kumar 2018; Zhang et al. 2018). Compared to GFSv2, FV3 GFS model has an increased (more than double) precipitation pattern correlation (0.32 vs. 0.14) with observation during 2015/16 El Niño. Generally, the precipitation response in the models has a high pattern correlation with observations during 1997/98 El Niño. These results indicate that during 2015/16 the observed rainfall anomalies may have been influenced by the atmospheric internal variability.

To explore the role of internal variability in determining the seasonal mean rainfall over California, Figure 11 shows probability density functions (PDFs, estimated as nonparametric fits to the histograms of the raw data) of California winter precipitation during three strong El Niño events based on FV3 GFS AMIP simulations. The long tick marks indicate the corresponding observed values for the three winters. The black PDF, drawn from 100-member ensemble FV3 GFS AMIP simulations of 2015/16, is statistically indistinguishable from the blue PDF drawn from 100-member ensemble FV3 GFS AMIP simulations of 1982/83, and for both, the mean of the PDF is shifted to the right. The results indicate that the most likely California winter precipitation condition is one for wetness in the presence of strong El Niño, with a statistical mode of +42% in 2015/16 runs and +52% in 1982/83 runs. The PDFs also illustrate the fact that even during strong El Niño events, there is also an appreciable probability for California seasonal mean rainfall to be negative. Further, for each PDF since all model simulations that went into its estimation have the same SST forcing, the spread in the PDF is due to atmospheric internal variability. The PDF of California winter precipitation for 1997/98 runs is significantly different from the PDFs for 2015/16 and 1982/83 runs, with a statistical mode of +70%. This is consistent with the observations for which the strongest California rains are for the 1997/98 winter among three extreme El Niño events. The observed California 2015/16 dryness was almost certainly an articulation of unforced variability and is supported by the fact that the observed condition resides within the dry tail of the forced PDF (black curve).

To further understand the cause for observed Southern California failed rains, we calculate the 2015/16 winter precipitation pattern correlation with observation from 100 individual members and make composites for the four runs that had the best or the worst correlation among the sample of 100 (Figure 12). The analysis approach follows that of Kumar et al. (2013).

The analysis based on individual model simulations indicates that on an individual run basis the observed dryness over Southern California can be replicated. This is evident from the composite of best four runs for which the anomaly correlation is the largest positive (left panels). For the composite of four runs that have the largest negative anomaly correlation, the simulated rainfall anomaly is opposite to the observed rainfall over California, and further, the wet condition over

Southern California is similar to that in SST forced signal in ensemble mean results (top right panel). Thus, the internal atmospheric variability, rather than a boundary-forced signal, was the likely cause for the failed Southern California rains in 2016 even in the presence of one of the largest El Niño. In summary, the FV3 GFS model can realistically capture the observed US climate variability associated with ENSO.

e. Assessment of climate predictability

Predictability of seasonal atmospheric climate variability depends on the fraction of total variability that is related to boundary conditions (referred to as the external, or potentially predictable variability) and the fraction of variability unrelated to external forcings (referred to as the internal, or unpredictable variability). Extensive efforts have been made in the past several decades to quantify potential predictability of seasonal mean climate variability by using either AMIP simulations or initialized coupled forecast systems (Kumar and Hoerling 1995; Kumar et al. 2007; Jha et al. 2019). The purpose of the analysis in this section is to assess the climate predictability based on a large ensemble FV3 GFS AMIP simulations and to quantify how the predictability measured by signal-to-noise ratio (SNR) is changing as the modeling systems are being improved.

We start our analysis by comparing the total variance of observed and FV3 GFS simulated DJF 200-hPa height anomalies over 1979-2021 period (Figure 13 right panel). It is evident that the model can realistically reproduce the observed total variance of upper-tropospheric circulation anomaly during winter that is characterized by the small variability in the tropical regions and a larger variability in the extratropical regions. The observed maximum centers of variability over

Aleutian and Greenland in the northern hemisphere and those over the southern higher latitude are also well captured in the model.

Shown in the left panel of Figure 13 is the two components of the simulated total variance, predictable (top) and unpredictable (bottom), which are derived from the variance of ensemble mean and the variance of departure in the individual members from the ensemble mean, respectively (Kumar and Hoerling 1995). The external variance for DJF 200-hPa height simulated by FV3 GFS is mainly located in the tropical eastern Pacific, the North Pacific and North American continent, similar to previous findings based on different periods (Kumar et al. 2007; Jha et al. 2019). This is to be expected since the ensemble mean variance is dominated by SST-forced atmospheric variability and its spatial structure is in agreement with the atmospheric response to ENSO (Trenberth et al. 1998; see also Fig. 7 and associated discussion). The simulated internal variance is largest in the middle and high latitudes, especially in the North American continent and the northern Asia and is similar to the best estimate of the internal variance of observed winter 200-hPa height using multiple models as noted in previous studies (Kumar et al. 2007; Jha et al. 2019).

Next, we calculate the ratio of the external and the internal variance in dimensionless units as signal-to-noise ratio to assess potential predictability, the results of which is given in Figure 14 that shows FV3 GFS simulated winter (left) and summer (right) SNR pattern for 200-hPa height (top), precipitation (middle) and surface air temperature (bottom). It is found that the largest SNR values for DJF 200-hPa height reside in the tropics and decrease gradually polewards due to an increase in atmospheric internal variability from tropics to extratropics (Figure 13), consistent with the previous findings that the predictability is larger in the tropics than the extratropics (e.g., Quan et al. 2004). The summer SNR pattern is very similar to the winter pattern, while the SNR values of heights are somewhat stronger in the tropical Atlantic. This difference is likely associated with the stronger height trend in the model over the tropical Atlantic for the summer compared to the winter. The analysis of previous GFSv2 AMIP runs also indicates that there is a somewhat larger height trend over the tropical Atlantic during summer relative to winter (not shown).

It should be noted that there is little consensus in the scientific community on the difference of seasonal predictability of 200-hPa height between winter and summer. Based on the NMME coupled forecast system, Jha et al. (2019) showed that SNR values for summer are lower than SNR for winter because of the weaker SST forcing during summer. However, Kumar et al. (2003) argued that due to a reduction in the internal variability, the magnitude of seasonal predictability for winter and summer are quite similar by using two atmospheric general circulation model (AGCM) simulations. A close examination of their results also reveals that the seasonal predictability is slightly stronger over the tropical Atlantic for summer than for winter, consistent with our findings.

SNR pattern for precipitation is also quite similar for two seasons, with larger values located in the tropical oceans. By comparing the results with those in the coupled forecast system (Jha et al. 2019), we notice that the improvement of seasonal predictability of precipitation is evident in the tropical Indian Ocean and tropical Atlantic. Similar to the spatial structure of SNR for 200-hPa height, the precipitation SNR also shows a large decline from the tropics to the extratropics and the relatively larger values over the tropical Atlantic in summer than in winter.

Consistent with SNR for 200-hPa height and precipitation, SNR of surface air temperature is also confined within the tropical land for both seasons, having larger values over North Africa, the Middle East, Asia, Northern Mexico and South America. Except for Northern Mexico, SNR values are larger over other regions in summer than in winter. The larger SNR value for winter over Northern Mexico could be related to the amplitude of ENSO SST variability. Consistent with increased SNR for precipitation over the northern Indian Ocean, SNR for surface air temperature is higher over North Africa, the Middle East, and Asia in summer compared to winter.

If the SNR estimates based on the AMIP simulations are correct estimates of corresponding predictability in observations, then generally larger SNR values imply a higher skill for seasonal prediction (Kumar and Hoerling 2000). To assess this, the corresponding maps of anomaly correlation (AC) (Figure 15), the value of which at each grid point is computed between AMIP ensemble mean and observed anomaly over the analysis period, confirms this relationship and is consistent with the theoretical analysis and model results in previous studies (Kumar et al. 2007; Jha et al. 2019). The stronger AC values for 200-hPa height over the tropical Atlantic extend northward in summer compared to winter, in agreement with larger SNR values there. The increase in AC values for surface air temperature over the Middle East and Asia in summer relative to winter is also in line with the increase of AC for precipitation over the Northern Indian Ocean.

A close look of the wintertime US prediction skill reveals that the stronger AC values for precipitation in FV3 GFS are located in the northwest, western coast and southern coast, where the 200-hPa AC values are higher (Figure 16 left). The surface air temperature AC pattern is characterized by maximum values in the west and the east and minimum values in the central US, largely consistent with the 200-hPa AC distribution.

We also calculated the AC values in GFSv2 and made the difference in the AC for two models to examine the changes in prediction skill. Some US regions experience an increase of prediction skill, as indicated by the shaded regions shown in Figure 16 right. Compared to GFSv2, FV3 GFS shows an increase of precipitation AC values over the north and the southeast, where the increase of 200-hPa AC values is also visible. The obvious improvement of surface air temperature prediction skill is located in the east, consistent with the improvement of surface air temperature response to three strong El Niño events over these regions as seen in Figure 9.

f. Simulation of extreme events—2022 summer South Asia flooding

In this section, we will evaluate the model's capability in simulating the extreme events by considering a case study for a specific extreme event in 2022. Most regions of Pakistan experienced record-breaking monsoonal rainfall from mid-June until the end of August 2022 that resulted in considerable losses of human lives and the economy of Pakistan (https://www.worldweatherattribution.org/analysis/rainfall/). Observed SST anomaly for 2022 summer had a La Niña condition in tropical Pacific and warm condition in the eastern Indian Ocean and coastal regions (not shown). How well does the FV3 GFS model simulate the observed South Asia flooding for 2022 summer compared to previous GFSv2 model?

Figure 17 shows the spatial map of precipitation anomalies for JJA 2022 from observations (top), FV3 GFS (middle) and GFSv2 (bottom) ensemble mean AMIP simulations. Observations show a large increase in south Asia flooding shown in black box region which includes Pakistan and northwest India. The ensemble mean results from FV3 GFS can reproduce the observed wet condition in South Asia, but the magnitude is somewhat weaker. This is to be expected when comparing ensemble mean anomalies with observations that are equivalent to a single model realization. Opposite to observations and FV3 GFS simulations, the previous version GFSv2 did not replicate the observed wet condition but indicates dry conditions. It is clear that the FV3 GFS model has a better simulation of South Asian flooding compared to old version GFSv2.

We also examined the FV3 GFS individual members to better understand the ensemble mean results. The top panel of Figure 18 shows the precipitation anomaly averaged over the box region from 100 individual members. The black line is the observed value, and the green line is the model ensemble mean value. Among 100 members, only 5 members produce the dry condition. This suggests that the observed SSTs specified as the forcing favor wet conditions over South Asia for 2022 summer. It can be seen from the bar plot that the magnitude of a single member (member 100) is very close to observations. Examining the spatial map of this member (bottom right panel) confirms that the model is capable of realistically simulating both the magnitude and spatial structure of observed wet conditions over South Asia for 2022 summer (bottom left panel). The results suggest that the FV3 GFS model can serve as a suitable tool to examine the causality of the extreme events.

4. Summary and discussion

The intent of this paper is to document an update in infrastructure of AMIP simulations that are used for real-time attribution of seasonal climate anomalies at CPC, i.e., based on the FV3 GFS.

These simulations are updated in real-time as SST observations become available. We evaluated the performance of these simulations in reproducing observed climate trends and variability and assessed climate predictability and the model's capability in capturing the extreme events.

We demonstrated that the FV3 GFS model can realistically capture the observed temperature trend over global land that has an upward trend of about 1° C since 1979 for both winter and summer. Associated with the warming trend over the global land, there is also a US warming and drying climate trend for DJF and JJA seasons as demonstrated by the frequency distributions of a large ensemble of model samples. The observed feature of the larger variability during winter compared to summer is also replicated in the model.

Observed three leading modes of variability for the period of 1979-2021 were identified based on the EOF analysis of wintertime extratropical Northern Hemisphere 200-hPa heights. The observed leading mode describes the extratropical atmospheric circulation pattern associated with Arctic Oscillation (AO) variability and is found to be independent of tropical SST variability. The observed second mode describes the canonical atmospheric teleconnection associated with ENSO resembling the tropical/Northern Hemisphere (TNH) pattern and the third mode features a hemisphere-scale increasing trend in heights associated with global warming. The FV3 GFS model is able to replicate the three primary modes of observed variability but there is an appreciable variability in the detailed EOF structures among individual members associated with sampling resulting in a large scatter in the spatial correlation with observations, especially for the last two modes.

Forced atmospheric teleconnections during 1979-2021 were examined using the 100-member ensemble mean of AMIP simulations. The leading mode of the forced variability is similar to the observed second mode that describes the TNH pattern associated with ENSO. The second forced mode resembles the observed third mode which is related to anthropogenically forced climate change. The forced third mode describes a wave train resembling the PNA pattern resulting from atmospheric sensitivity to ENSO asymmetry and from sensitivity to a tropical precursor SST for ENSO development. Our results are similar to the three primary forced modes of Zhang et al. (2016) except for a different EOF ranking for the latter two modes, implying that the forced primary modes do not depend on the selection of a particular model but are more determined by the nature of boundary forcing used.

FV3 GFS AMIP simulations realistically captured the observed key features including US climate variability associated with ENSO. Consistent with the US warming trend, a gradual southward shift of stronger warm anomalies from early strong El Niño (1982/83) to recent strong El Niño (2015/16) was evident in both observations and model ensemble mean results. There was an improvement in FV3 GFS relative to previous GFSv2 in simulating the surface temperature response to extreme El Niño events over the US southeastern coastal regions. The observed US precipitation pattern featured wetness in the west and central US and southern coast for both 1982/83 and 1997/98 El Niño events but had dry conditions over Southern California for 2015/16 winter. The ensemble mean precipitation response in the model was similar for three strong El Niño events, having a wetness across the west and central US and southern coast. We explored the role of internal variability in determining the seasonal mean rainfall in California based on a large ensemble of FV3 GFS AMIP simulations. The results indicate that the observed California

2015/16 dryness was likely an articulation of unforced variability (the internal atmospheric noise), rather than a boundary-forced signal, in agreement with previous findings (Zhang et al. 2018; Kumar and Chen 2020).

The climate predictability measured by SNR was also assessed based on FV3 GFS AMIP simulations. The SNR pattern, in general, was similar to each other between winter and summer, with largest values in the tropical regions and a decrease in the SNR towards high latitudes. The SNR values for 200-hPa heights and precipitation over the tropical Atlantic and those for surface air temperature over North Africa, the Middle East, Asia and South America were larger in summer than in winter. It is noted that there was an improvement of seasonal predictability of precipitation over the tropical Indian Ocean and tropical Atlantic, compared to the results in the initialized forecast system (Jha et al. 2019). The seasonal prediction skill measured by AC generally follows the SNR pattern, supporting previous theoretical analysis and model results (Kumar et al. 2007; Jha et al. 2019). The comparison of the wintertime AC pattern of US surface climate between FV3 GFS and GFSv2 reveals that the prediction skill of precipitation over the Northern and Southeastern US, and that of surface air temperature over the eastern US are somewhat improved in the new model relative to the old version.

The model's capability in simulating the extreme events was evaluated for a case study for the 2022 summer South Asia record-breaking flooding. The ensemble mean results from FV3 GFS reproduced the observed wet conditions but those from previous GFSv2 indicated dry conditions in South Asia, suggesting a better simulation of extreme events in the new model relative to the old model. Analysis of FV3 GFS individual runs further confirmed that the model could replicate the magnitude and spatial pattern of South Asian flooding and indicates that the 2022 summer flooding over that region could have been driven by the observed SST forcing.

In summary, the FV3 GFS model can realistically replicate the observed climate variability and trends as well as extreme events. We also plan to use the same infrastructure for other sensitivity studies to understand various aspects of climate variability, e.g., atmospheric responses to Central Pacific (CP) vs. Eastern Pacific (EP) events, role of SST anomalies in different ocean basins etc. In future, counterfactual simulations in which an estimate of the observed long-term changes in the SST due to anthropogenic forcing is removed will also be conducted in parallel with the current AMIP simulations to understand the influence of climate change on extreme events.

Acknowledgments

This work was supported by NOAA's Climate Program Office.

Open Research

Data Availability Statement

The version of GFSv15 is available from https://www.emc.ncep.noaa.gov/users/meg/fv3gfs/. The data were analyzed with the NCAR Command Language version 6.6.2 (National Center for Atmospheric Research, 2019) available from https://www.ncl.ucar.edu/ and the Grid Analysis and Display System (GrADS) Version 2.1.1.b0 available from http://cola.gmu.edu/grads/. The Data and scripts for analysis are publicly available at (Zhang et al., 2023) https://doi.org/10.5281/zenodo.8023560.

REFERENCES

An, S-I., and F-F. Jin, 2004: Nonlinearity and asymmetry of ENSO. *J. Climate*, 17, 2399–2412.

Barnston, A. G., A. Kumar, L. Goddard, and M. P. Hoerling, 2005: Improving seasonal predictions practices through attribution of climate variability. *Bull. Ame. Meteor. Soc.*, **85**, 59-72.

Barsugli, J. J., D. M. Easterling, D. S. Arndt, D. A. Coates, T. L. Delworth, M. P. Hoerling, N. Johnson, S. B. Kapnick, A. Kumar, K. E. Kunkel, C. J. Schreck, R. S. Vose, T. Zhang, 2022: Development of a Rapid Response Capability to Evaluate Causes of Extreme Temperature and Drought Events in the United States. *Bull. Amer. Met. Soc.*,103, DOI: 10.1175 /BAMS-D-21-0237.1. https://www.ametsoc.org/ams/index.cfm/publications/bulletin-of-the-american-meteorological-society-bams/explaining-extreme-events-from-a-climate-perspective/

Burgers, G., and D. B. Stephenson, 1999: The "normality" of El Niño. *Geophys. Res. Lett.*, 26, 1027–1030.

Changnon, S. A., 1999: Impacts of 1997–98 El Niño generated weather in the United States. Bull. Amer. Meteor. Soc., 80, 1819–1827, https://doi.org/10.1175/1520-0477(1999)080,1819: IOENOG.2.0.CO;2.

Chen, M., and A. Kumar, 2018: Winter 2015/16 atmospheric and terrestrial anomalies over North America: El Nino response and the role of noise. *Mon. Wea. Rev.*, **146**, 909-927.

Compo, G., and P. Sardeshmukh, 2009: Oceanic influence on recent continental warming. *Clim. Dyn.*, **32**, 333-342. https://doi.org/10.1007/s00382-008-0448-9

Gates, W. L., J. Boyle, C. Covey, C. Dease, C. Doutriaux, R. Drach, M. Fiorino, P. Gleckler, J. Hnilo, S. Marlais, T. Phillips, G. Potter, B. Santer, K. Sperber, K. Taylor and D. Williams, 1998: An Overview of the Results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.*, 73, 1962-1970.

Fahad, A, and N. Burls, 2022: The influence of direct radiative forcing versus indirect sea surface temperature warming in southern hemisphere subtropical anticyclones under global warming. *Clim. Dyn*., 58, 2333-2350, https://doi.org/s00382-021-06006-1.

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 M. Dilley, 2005: El Niño: Catastrophe or opportunity? *J. Climate*, 18, 651–665, https://doi.org/10.1175/JCLI-3277.1.

Han, J., and Pan, H.-L. (2011). Revision of convection and vertical diffusion schemes in the NCEP Global Forecast System. *Weather and Forecasting*, 26(4), 520–533. https://doi.org/10.1175/WAF-D-10-05038.1

Han, J., Witek, M. L., Teixeira, J., Sun, R., Pan, H.-L., Fletcher, J. K., & Bretherton, C. S. (2016). Implementation in the NCEP GFS of a hybrid eddy-diffusivity mass-flux (EDMF) boundary layer parameterization with dissipative heating and modified stable boundary layer mixing. *Weather and Forecasting*, 31(1), 341–352. https://doi.org/10.1175/WAF-D-15-0053.1

Hartmann, D. L., 2015: Pacific sea surface temperature and the winter of 2014. *Geophys. Res. Lett.*, 42, 1894–1902, doi:10.1002/2015GL063083.

Held, I. M., and B. J. Soden, 2000: Water Vapor Feedback and Global Warming. *Annual Review of Energy and the Environment,* 25, 441-475, https://doi.org/10.1146/annurev.energy.25.1.441.

Hoerling, M. P., and A. Kumar, 2002: Atmospheric response patterns associated with tropical forcing. *J. Climate*, 15, 2184–2203, doi:10.1175/1520-0442(2002)015<2184:ARPAWT>2.0.CO;2.

Hoerling, M., T. Xu, G. Bates, A. Kumar, and B. Jha, 2006: Warm oceans raise land temperatures. *EOS*, Transactions American Geophysical Union, **87** (19), 189-193.

Horel, J. D., and J. M. Wallace, 1981: Planetary-scale atmospheric phenomena associated with the Southern Oscillation. *Mon. Wea. Rev.*, 109, 813–829, doi:10.1175/1520-0493(1981)109<0813:PSAPAW>2.0.CO;2.

Howitt, R. E., J. Medellin-Azuara, D. MacEwan, J. R. Lund, and D. A. Summer, 2014: Economic analysis of the 2014 drought for California agriculture. Tech. Rep., Center for Watershed Sciences, University of California, Davis, 20 pp. [Available online at https://watershed.ucdavis.edu/files/content/news/Economic_Impact_of_the_2014_California_Water_Drought.pdf.]

Hurrell, J., J. Hack, D. Shea, J. Caron, and J. Rosinski, 2008: A new sea surface temperature and sea ice boundary dataset for the Community Atmosphere Model. *J. Climate*, 21, 5145–5153, doi:10.1175/2008JCLI2292.1.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, 113(D13). https://doi.org/10.1029/2008JD009944

Jha, B., A. Kumar, and Z.-Z. Hu, 2019: An update on the estimate of predictability of seasonal mean atmospheric variability using North American Multi-Model Ensemble. *Clim Dyn.,* 53, 7397–7409. https://doi.org/10.1007/s00382-016-3217-1

Kalnay, E., et al., 1996: The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, 77, 437–471.

Kumar, A., and M. P. Hoerling, 1995: Prospects and limitations of seasonal atmospheric GCM predictions. *Bull. Amer. Meteor. Soc.*, 76 , 335–345.

Kumar, A., and M. P. Hoerling, 1998: Annual cycle of Pacific-North American seasonal predictability associated with different phases of ENSO. *J. Climate*, 11, 3295-3308.

Kumar, A., and M. P. Hoerling, 2000: Analysis of conceptual model of seasonal climate variability and implication for seasonal predictions. *Bull. Amer. Meteor. Soc.*, 81, 255–264, https://doi.org/10.1175/1520-0477(2000)081,0255:AOACMO.2.3.CO;2.

Kumar, A., S. D. Schubert, and M. S. Suarez, 2003: Variability and predictability of 200-mb seasonal mean heights during summer and winter. *J. Geophys. Res.* 108:4169. doi:10.1029/2002JD002728.

Kumar, A., Q. Zhang, P. Peng, and B. Jha, 2005: SST-forced atmospheric variability in an atmospheric general circulation model. *J. Climate*, 18, 3953–3967, doi:10.1175/JCLI3483.1.

Kumar, A., B. Jha, Q. Zhang, and L. Bounoua, 2007: A new methodology for estimating the unpredictable component of seasonal atmospheric variability. *J. Climate*, 20, 3888–3901.

Kumar, A., M. Chen, M. Hoerling, and J. Eischeid, 2013: Do extreme climate events require extreme forcings? Geophys. Res. Lett., 40, 3440-3445. doi:10.1002/grl.50657.

Kumar, A., and M. Chen, 2020: Understanding skill of seasonal mean precipitation prediction over California during boreal winter and role of predictability limits. *J. Climate*, **33**, 6141–6163, https://doi.org/10.1175/JCLI-D-19-0275.1.

Mo, K. C., and R. E. Livezey, 1986: Tropical–extratropical geopotential height teleconnections during the Northern Hemisphere winter. *Mon. Wea. Rev.*, 114, 2488–2515, doi:10.1175/1520-0493(1986)114<2488:TEGHTD>2.0.CO;2.

Murray, D., A. Hoell, M. Hoerling, J. Perlwitz, X.-W. Quan, D. Allured, T. Zhang, J. Eischeid, C. Smith, J. Barusgli, J. McWhirter, C. Kreutzer and R. S. Webb, 2020: Facility for Weather and Climate Assessments (FACTS): A Community Resource for Assessing Weather and Climate Variability. *Bull. Amer. Meteor. Soc.*, 101, E1214–E1224, https://doi.org/10.1175/BAMS-D-19-0224.1.

NOAA 2017: Billion dollar weather and climate disasters: Table of events. NOAA, accessed 7 June 2017, https://www.ncdc.noaa. gov/billions/events/US/2015.

Penland, C., and P. D. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, 8, 1999–2024.

Philip, S. Y., Kew, S. F., van Oldenborgh, G. J., Anslow, F. S., Seneviratne, S. I., Vautard, R., Coumou, D., Ebi, K. L., Arrighi, J., Singh, R., van Aalst, M., Pereira Marghidan, C., Wehner, M., Yang, W., Li, S., Schumacher, D. L., Hauser, M., Bonnet, R., Luu, L. N., Lehner, F., Gillett, N., Tradowsky, J., Vecchi, G. A., Rodell, C., Stull, R. B., Howard, R., and Otto, F. E. L., 2021: Rapid attribution analysis of the extraordinary heatwave on the Pacific Coast of the US and Canada June 2021, Earth Syst. Dynam. Discuss., https://doi.org/10.5194/esd-2021-90.

Putman, W. M., and Lin, S.-J., 2007: Finite-volume transport on various cubed-sphere grids. *Journal of Computational Physics*, 227(1), 55–78. https://doi.org/10.1016/j.jcp.2007.07.022

Quan, X. W., P. J. Webster, A. M. Moore, and H.-R. Chang, 2004: Seasonality in SST forced atmospheric short-term climate predictability. *J. Climate*, **17**, 3090-3180.

Quan, X., M. Hoerling, J. Whitaker, G. Bates, and T. Xu, 2006: Diagnosing sources of U.S. seasonal forecast skill. *J. Climate*, 19, 3279–3293, https://doi.org/10.1175/JCLI3789.1.

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

Saha, S., et al., 2010: The NCEP climate forecast system reanalysis, *Bull. Amer. Meteor. Soc.*, 91, 1015–1057.

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.

Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, and B. Rudolf, 2014: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theor. Appl. Climatol.*, 115, 15–40, doi:10.1007/s00704-013-0860-x.

Seager, R., M. Hoerling, S. Schubert, H. Wang, B. Lyon, A. Kumar, J. Nakamura, and N. Henderson, 2015: Causes of the 2011 to 2014 California drought. J. Climate, 28, 6997–7024, https://doi.org/10.1175/JCLI-D-14-00860.1.

Soden, B. J., Jackson, D. L., Ramaswamy, V., Schwarzkopf, M. D., & Huang, X., 2005: The radiative signature of upper tropospheric moistening. *Science*, 310, 841–844. https://doi.org/10.1126/science.1115602.

Taylor, K.E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106**, 7183-7192. https://doi.org/10.1029/2000JD900719.

Thompson, D.W.J., and J.M. Wallace 1998: The Arctic Oscillation signature in wintertime geopotential height and temperature fields. *Geophys. Res. Lett.* 25, 1297-1300.

Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res.*, 103, 14291–14324, doi:10.1029/97JC01444.

Trenberth, K. E., and D. P. Stepaniak, 2001: Indices of El Niño Evolution. *J. Climate*, 14, 1697–1701.

Weaver, S., A. Kumar., and M. Chen: 2014: Recent Increases in Extreme Temperature Occurrence over Land. *Geophys. Rea. Lett*., DOI: 10.1002/2014GL060300.

Winton, M., 2000: A reformulated three-layer sea ice model. *J. Atmos. Oceanic Tech.*, **17**, 525–531.

Wuebbles, D.J., D.W. Fahey, K.A. Hibbard, B. DeAngelo, S. Doherty, K. Hayhoe, R. Horton, J.P. Kossin, P.C. Taylor, A.M. Waple, and C.P. Weaver, 2017: Executive summary. In: Climate Science Special Report: Fourth National Climate Assessment, Volume I [Wuebbles, D.J., D.W. Fahey, K.A. Hibbard, D.J. Dokken, B.C. Stewart, and T.K. Maycock (eds.)]. U.S. Global Change Research Program, Washington, DC, USA, pp.12-34, doi: 10.7930/J0DJ5CTG.

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

Yang, H., and K. K. Tung, 1998: Water vapor, surface temperature and the greenhouse effect—a statistical analysis of tropical-mean data. *J. Climate*, 11, 2686–2697.

Zhang, T., D.-Z. Sun, R. Neale, and P. J. Rasch, 2009: An evaluation of ENSO asymmetry in the Community Climate System Models: A view from the subsurface. *J. Climate*, 22, 5933–5961.

Zhang, T., and D.-Z. Sun, 2014: ENSO Asymmetry in CMIP5 Models. *J. Climate*, 27, 4070-4093, doi:10.1175/JCLI-D-13-00454.1.

Zhang, T., M. P. Hoerling, J. Perlwitz, and T. Xu, 2016: Forced Atmospheric Teleconnections During 1979-2014. *J. Climate*, 29, 2333-2357, DOI: http://dx.doi.org/10.1175/JCLI-D-15-0226.1.

Zhang, T., M. P. Hoerling, K. Wolter, J. Eischeid, L. Cheng, A. Hoell, J. Perlwitz, X. Quan, and J. Barsugli, 2018: Predictability and prediction of the Southern California rains during strong El Niño events: A focus on the failed 2016 winter rains. *J. Climate*, 31, 555–574, https://doi.org/10.1175/JCLI-D-17-0396.1.

Zhang, T., M. P. Hoerling, A. Hoell, J. Perlwitz, and J. Eischeid, 2020: Confirmation for and Predictability of Distinct U.S. Impacts of El Niño Flavors. *J. Climate*, 33, 5971–5991, https://doi.org/10.1175/JCLI-D-19-0802.1.

Zhou, L., S.-J. Lin, J.-H. Chen, L. M. Harris, X. Chen, and S. L. Rees, 2019: Toward Convective-Scale Prediction within the Next Generation Global Prediction System, *Bull. Amer. Meteor. Soc.*, 100, 1225-1243. https://doi.org/10.1175/BAMS-D-17-0246.1

Table 1. The global mean values of the difference in climatology between JJA and DJF (JJA minus DJF) for eddy 200-hPa height (m), precipitation (mm/day) and land surface air temperature (ºC) from observation, GFSv2 and FV3 GFS ensemble mean AMIP simulations, and the global pattern correlations between models and observation for the corresponding climatology difference.

Variables	Global mean values			Global pattern correlations with observation	
	Observation	GFSv2	FV3GFS	GFSv2	FV3GFS
Eddy Z200	0.17	0.079	0.056	0.96	0.94
Precip	0.076	0.11	0.17	0.72	0.82
T2m	12.95	10.73	12.09	0.98	0.98

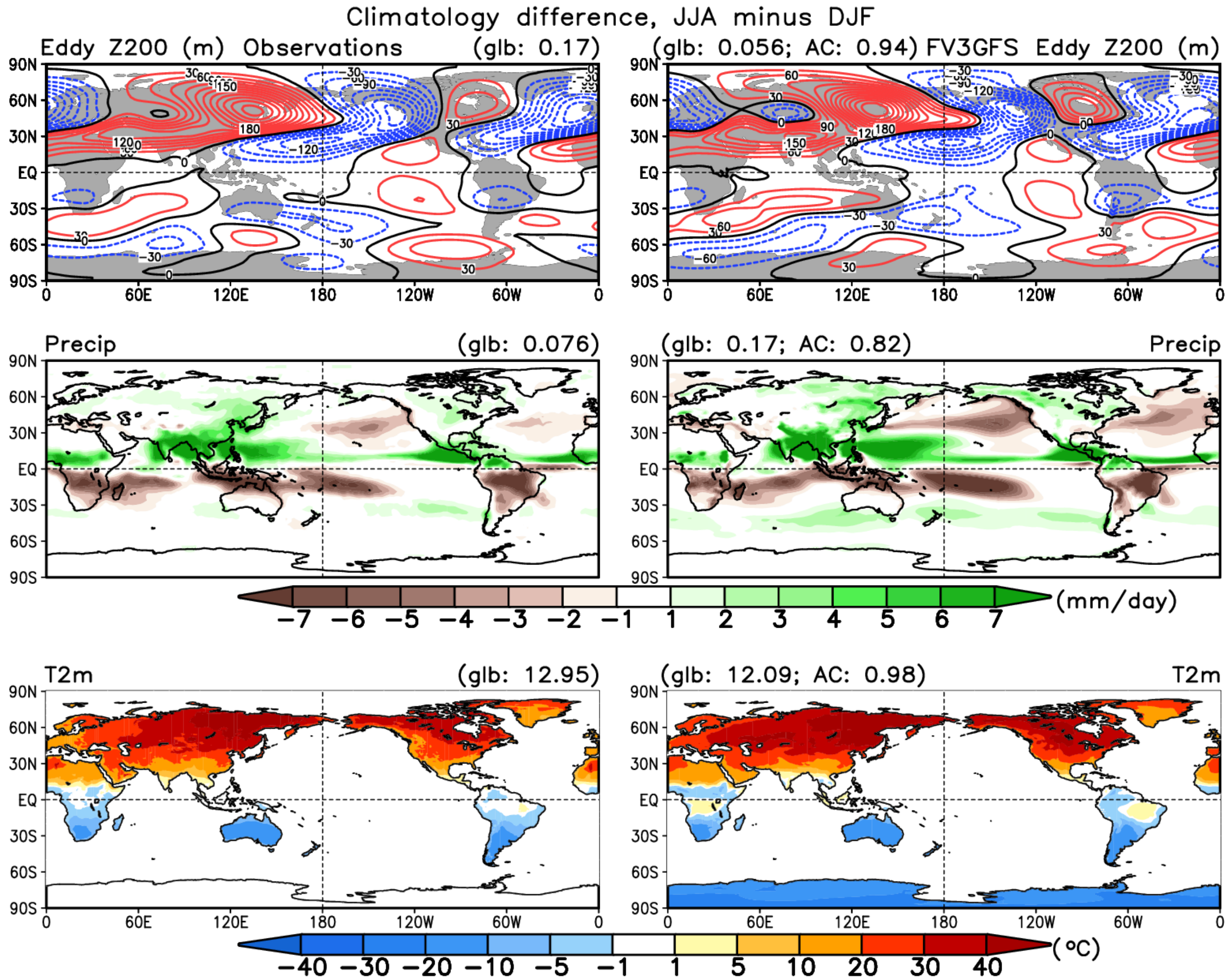

Figure 1. Differences in climatology between JJA and DJF (JJA minus DJF) for (top) eddy 200-hPa height, (middle) precipitation and (bottom) surface air temperature from (left) observations and (right) FV3 GFS simulated 100-member AMIP ensemble mean results. The observed and simulated global mean values (the first number) and the pattern correlation values (the second number) are listed in the titles of the plots.

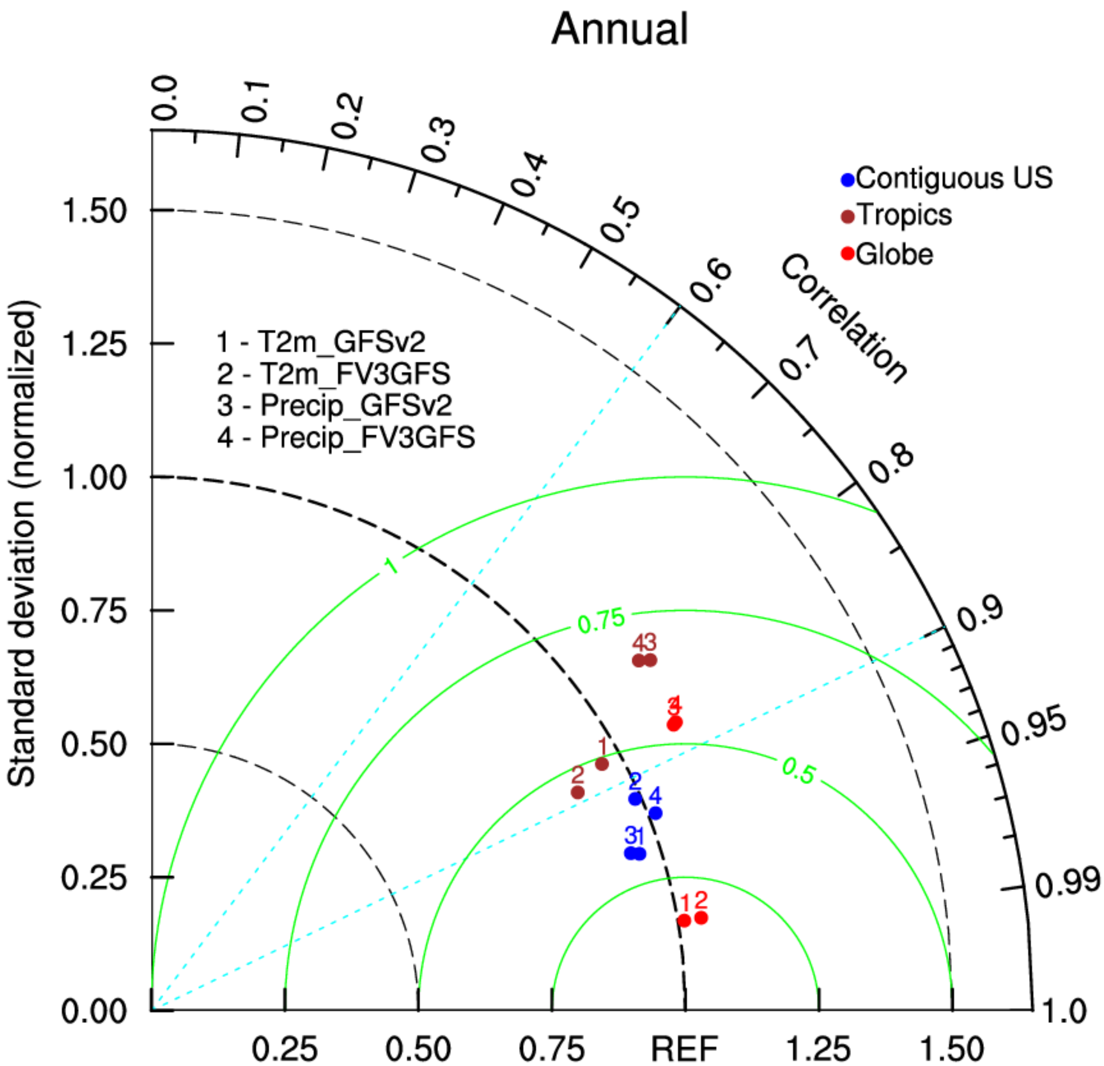

Figure 2. Taylor diagram displaying changes in normalized pattern statistics between GFSv2 and FV3 GFS model estimates for the (red dots) global, (brown dots) tropical, and (blue dots) contiguous US pattern of annual mean climatology of precipitation and land surface air temperature. The standard deviations have been normalized by the observed standard deviation of each field.

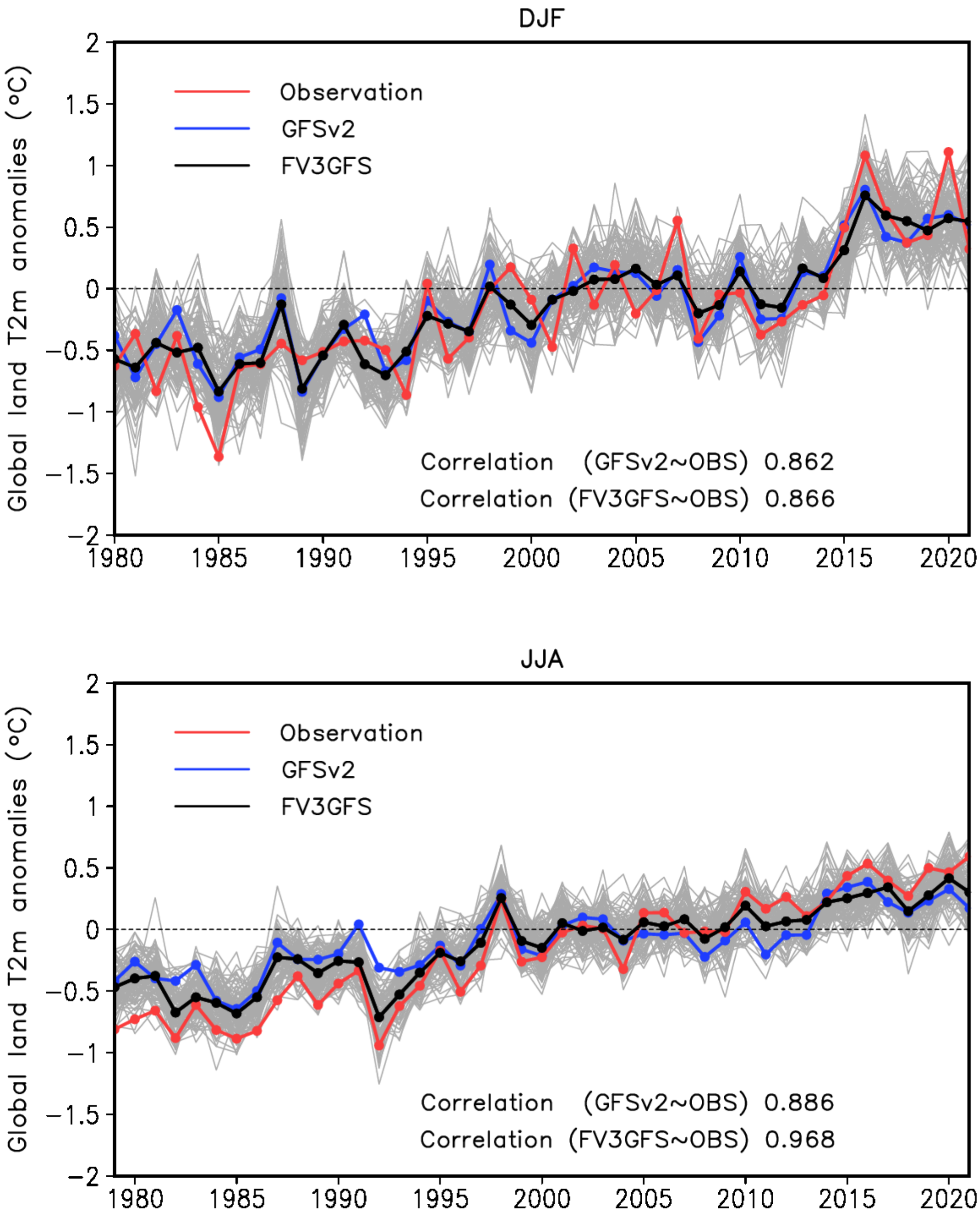

Figure 3. Time series of surface air temperature anomalies for (top) DJF and (bottom) JJA averaged over global land regions from (red line) observations, (blue line) GFSv2 simulated 30-member ensemble mean and (black line) FV3 GFS simulated 100-member ensemble mean of AMIP simulations. The gray lines show the spread of individual members of FV3 GFS model. The temporal correlations of the global mean land surface air temperature anomalies between model ensemble mean and observations are listed in the plot.

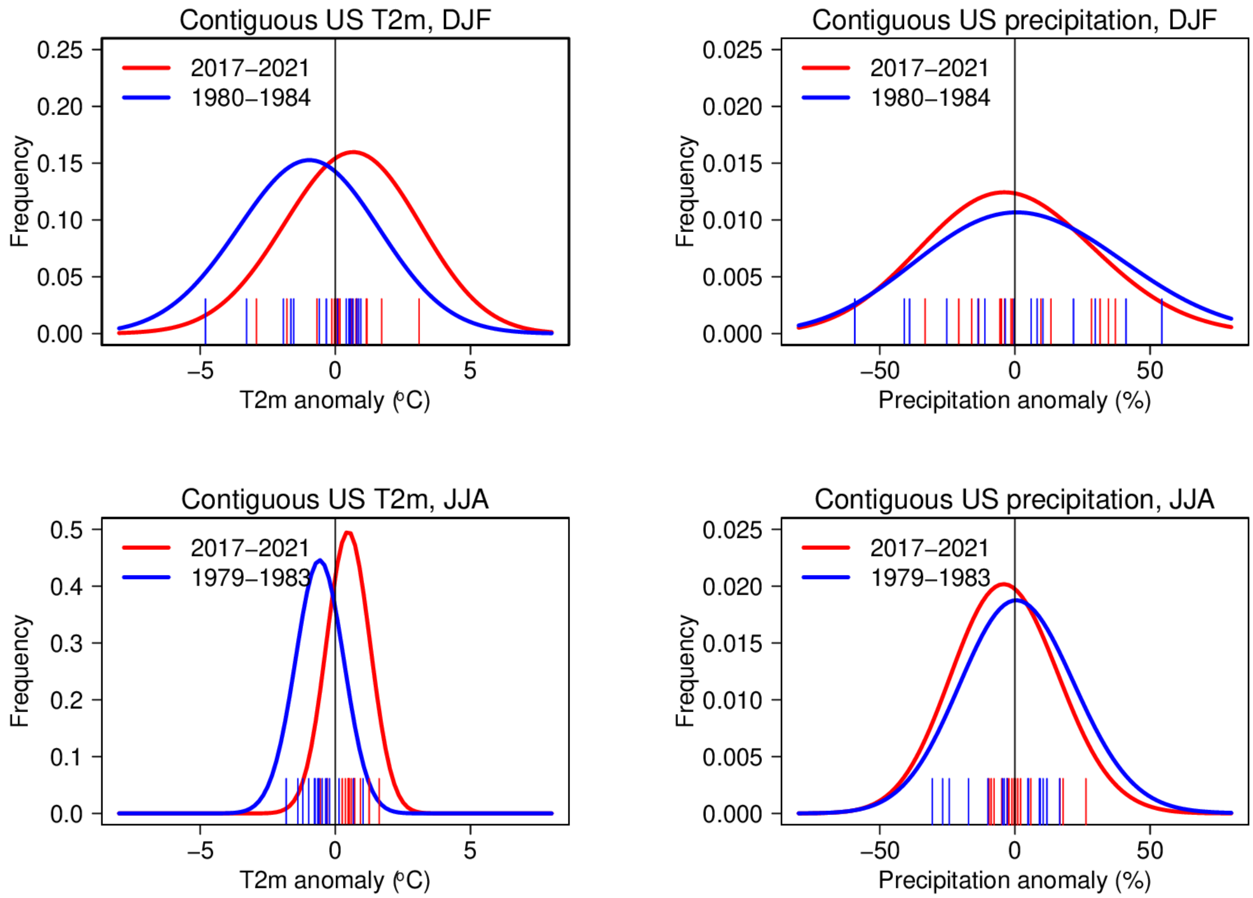

Figure 4. PDFs of contiguous U.S. (top) DJF and (bottom) JJA (left) surface air temperature anomalies (°C) and (right) precipitation anomalies (percent departure) for the first (blue curves) and last (red curves) 5-yr periods of 1979-2021. Results are based on 100-member FV3 GFS AMIP simulations. Large tick marks at the bottom show observed values for 15 months of the first (blue) and last (red) 5-yr periods of 1979-2021.

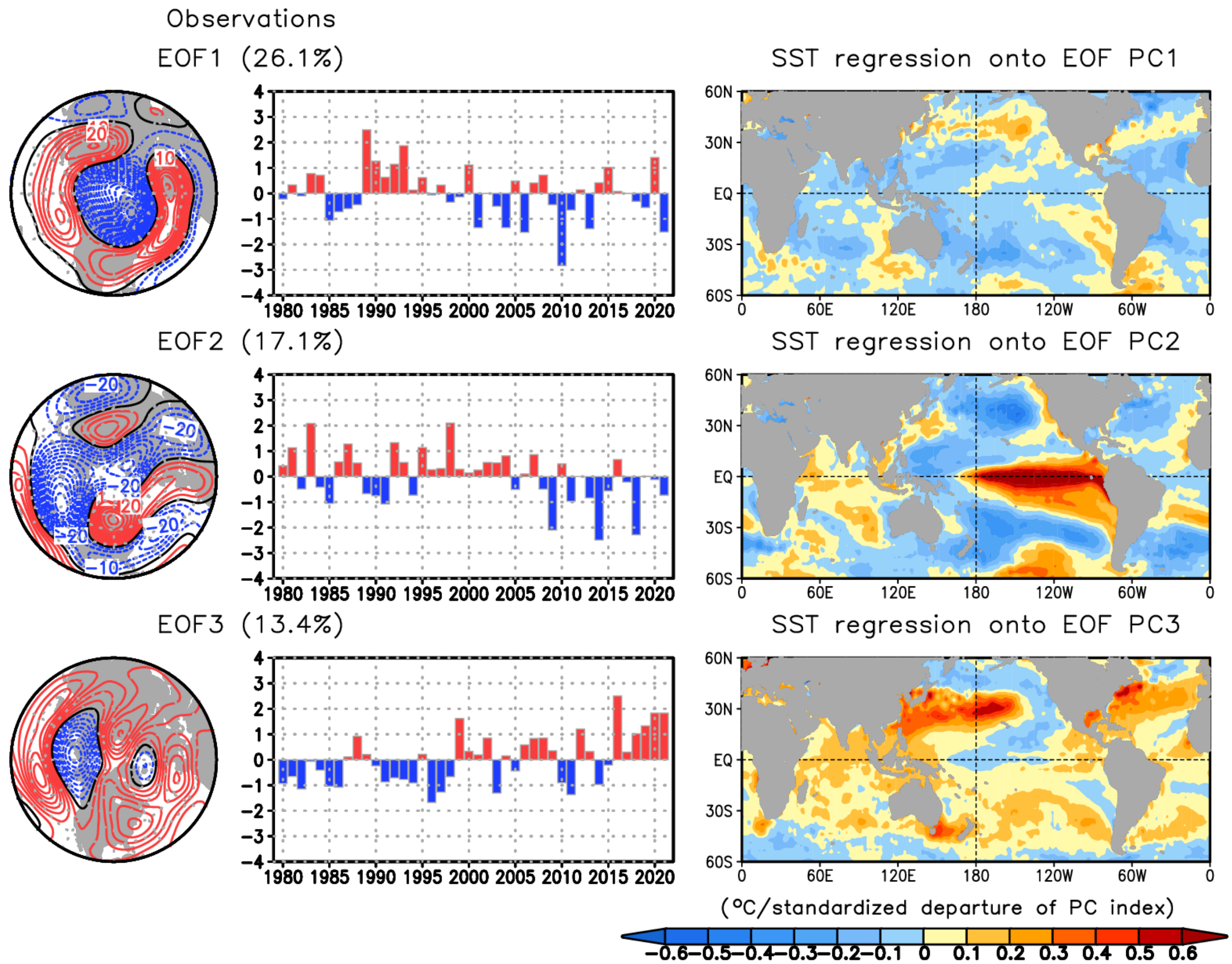

Figure 5: (left) The spatial pattern and (middle) standardized PC time series of the leading three EOFs of DJF 200-hPa heights from observations. (right) Regressions of observed DJF SST on the PC time series of the leading three EOFs of observed DJF 200-hPa heights. The EOF analysis is computed over the 20º–90ºN domain for 1979/80–2020/21. The EOF patterns are shown as the regressions of the heights onto the standardized PC time series and drawn at the interval of 5m for a 1 standardized departure of PC index.

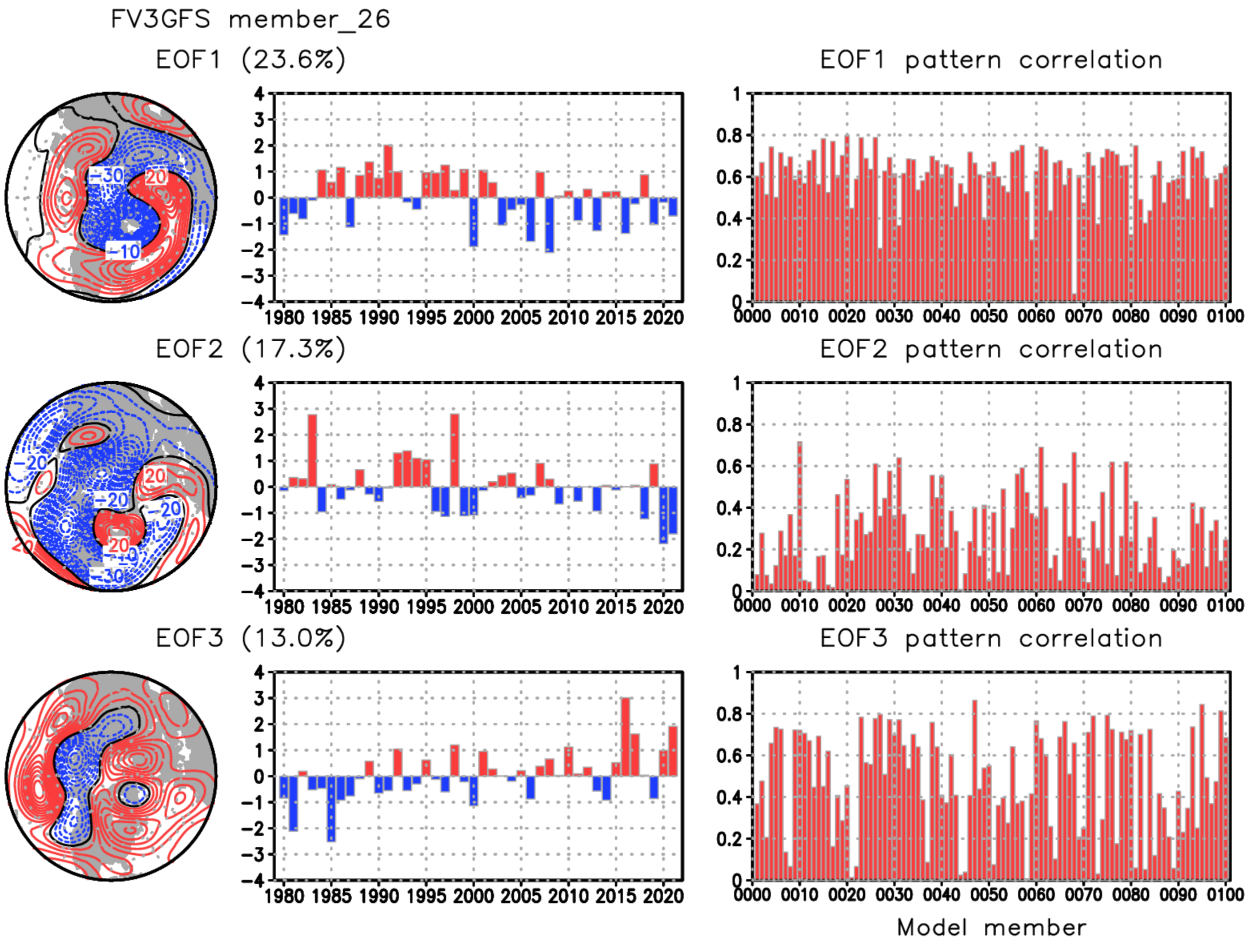

Figure 6: (left) The spatial pattern and (middle) standardized PC time series of the leading three EOFs of DJF 200-hPa heights from a single member (member 26) of FV3 GFS AMIP simulations. (right) The respective EOF pattern correlation values of the leading three EOFs of DJF 200-hPa heights with observations from 100 individual members of FV3 GFS AMIP simulations. This single member is selected among members that resemble observations, subject to the largest mean value of three leading EOF pattern correlations of DJF 200-hPa heights with observations, with EOF1 pattern correlation of 0.79, EOF2 pattern correlation of 0.61, and EOF3 pattern correlation of 0.78. The EOF analysis is computed over the 20º–90ºN domain for 1979/80–2020/21. The EOF patterns are shown as the regressions of the heights onto the standardized PC time series and drawn at the interval of 5m for a 1 standardized departure of PC index.

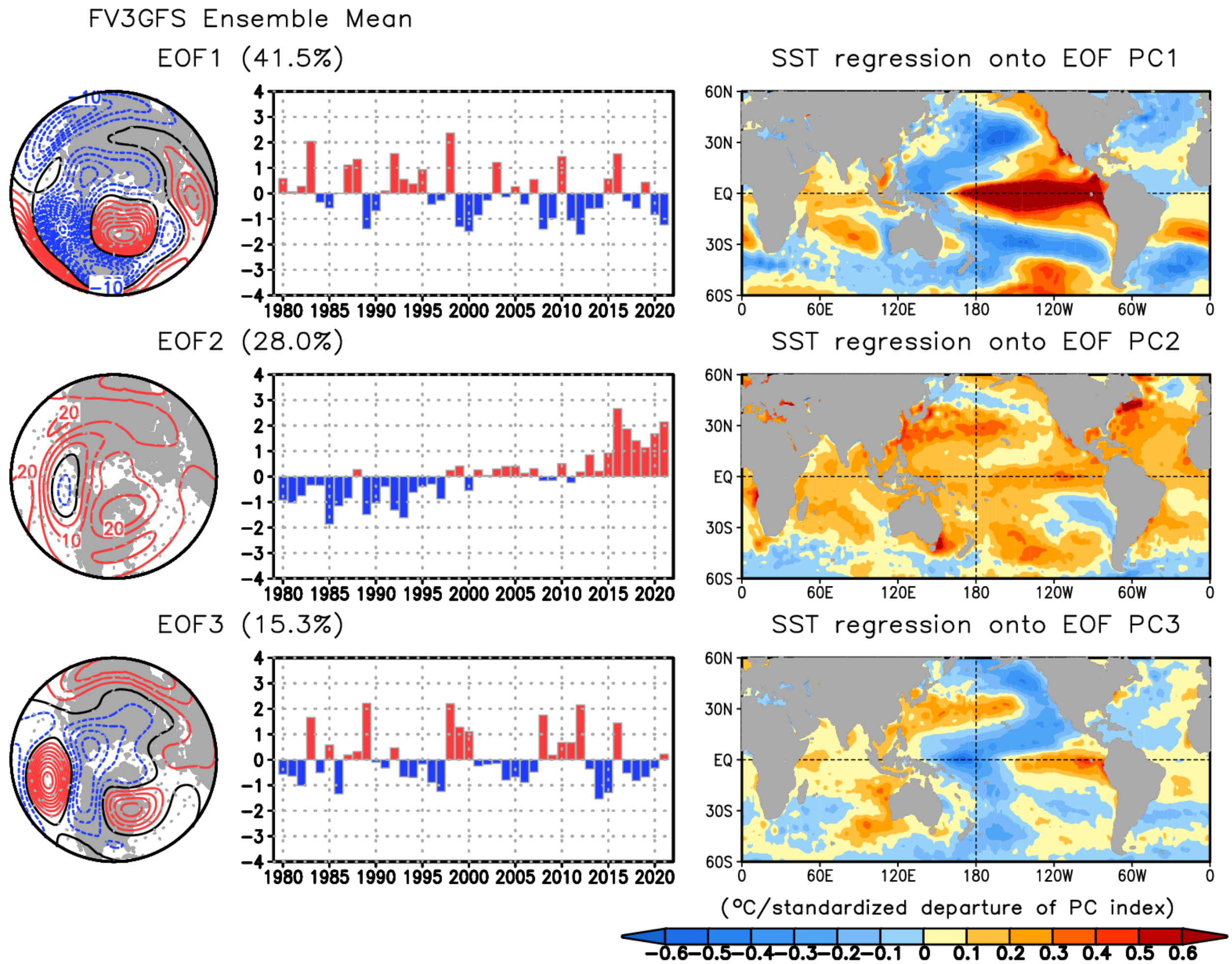

Figure 7: (left) The spatial pattern and (middle) standardized PC time series of the leading three EOFs of FV3 GFS simulated 100-member ensemble mean DJF 200-hPa heights. (right) Regressions of observed DJF SST on the PC time series of the leading three EOFs of FV3 GFS simulated 100-member ensemble mean DJF 200-hPa heights. The EOF analysis is computed over the 20°–90°N domain for 1979/80–2020/21. The EOF patterns are shown as the regressions of the heights onto the standardized PC time series and drawn at the interval of 5m for a 1 standardized departure of PC index.

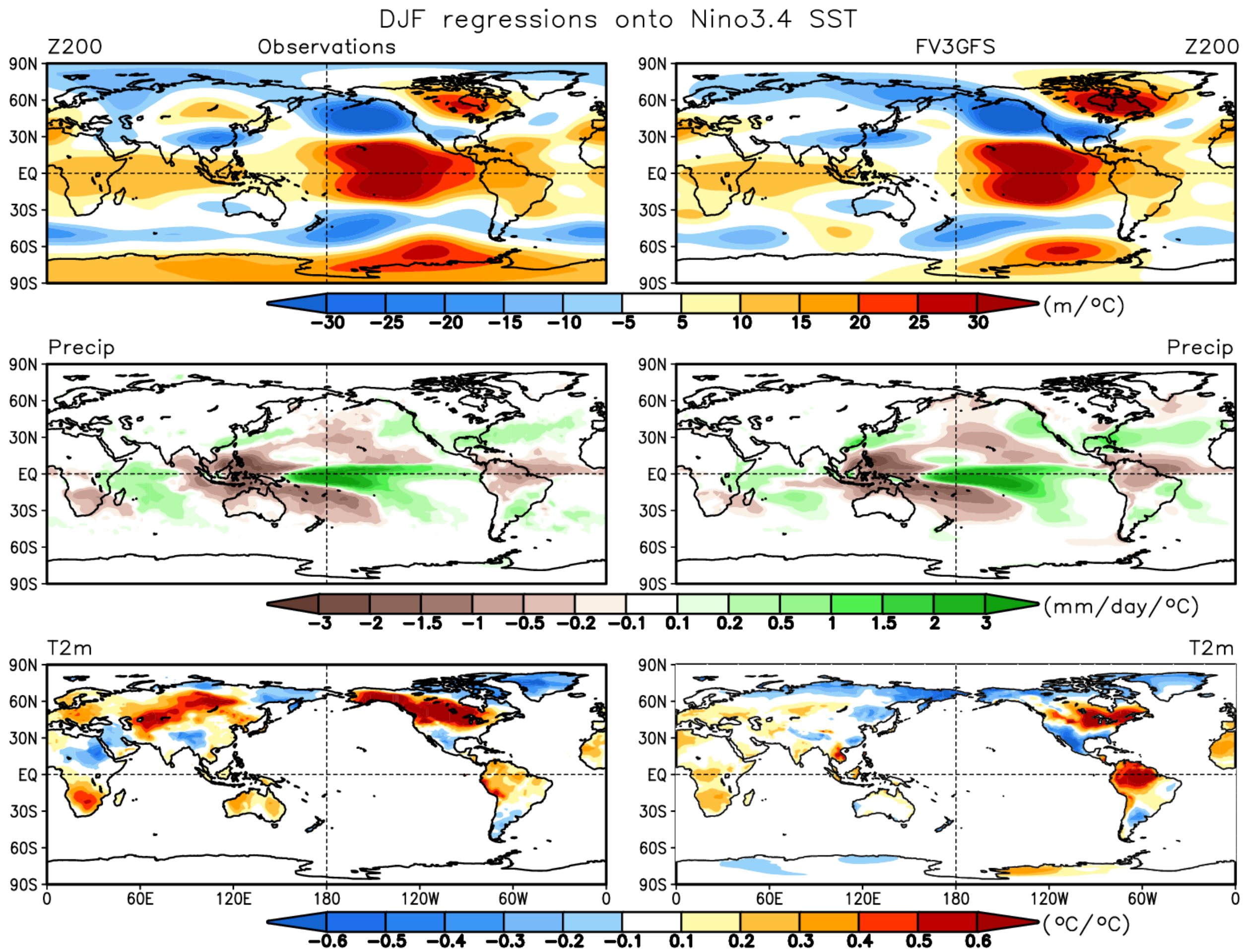

Figure 8: The spatial pattern of regressions of DJF (top) 200-hPa height, (middle) precipitation and (bottom) surface air temperature anomalies on the observed Niño3.4 SST index from (left) observations and (right) FV3 GFS AMIP simulations. We first calculate the regressions from individual runs and then average 100 regression estimates as the regressions for the model.

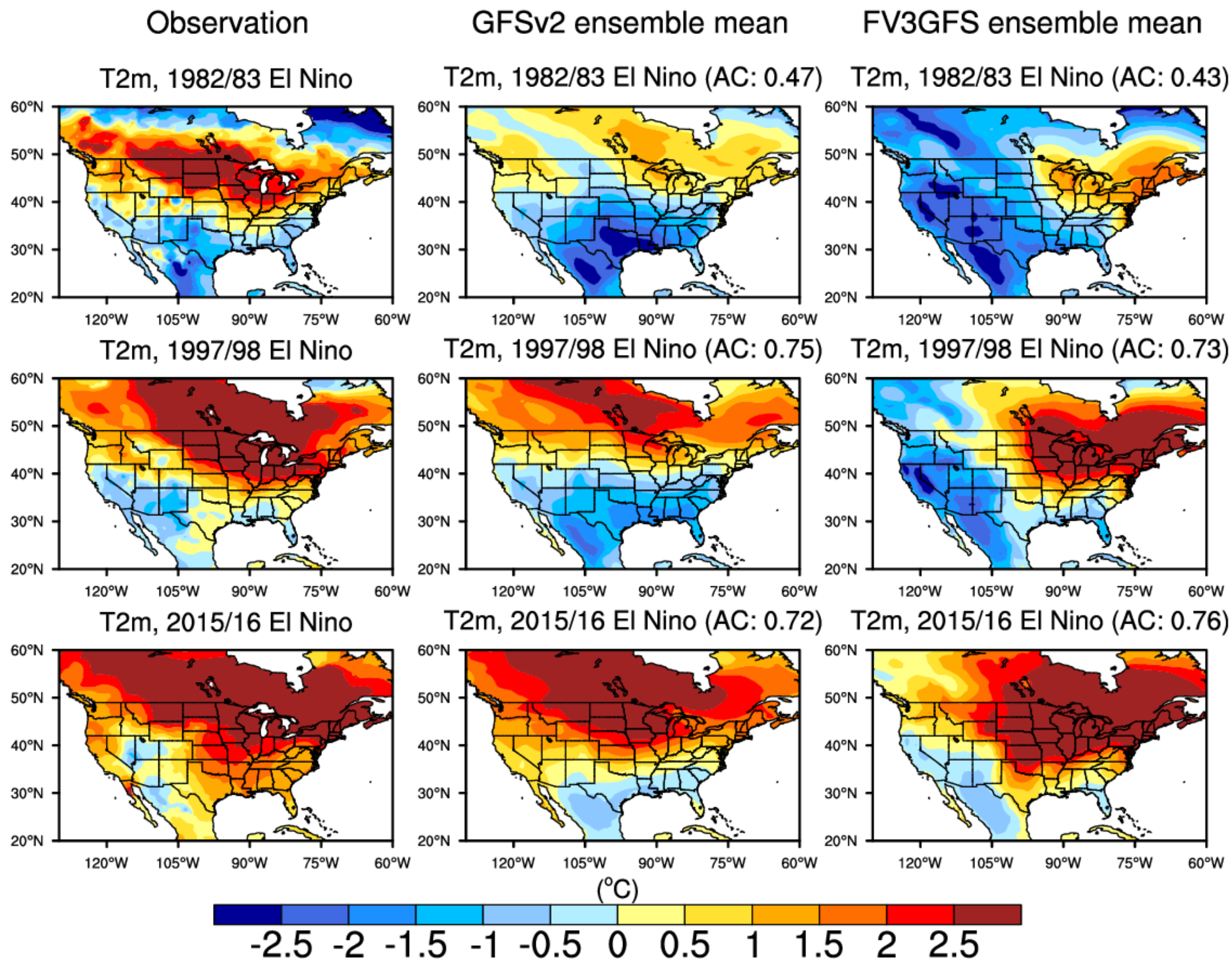

Figure 9. Surface air temperature anomalies for (top) 1982/83 DJF, (middle) 1997/98 DJF and (bottom) 2015/16 DJF from (left) observations, (middle) GFSv2 simulated 30-member ensemble mean, and (right) FV3 GFS simulated 100-member ensemble mean. The pattern correlations between models and observations are listed in the titles of the plots.

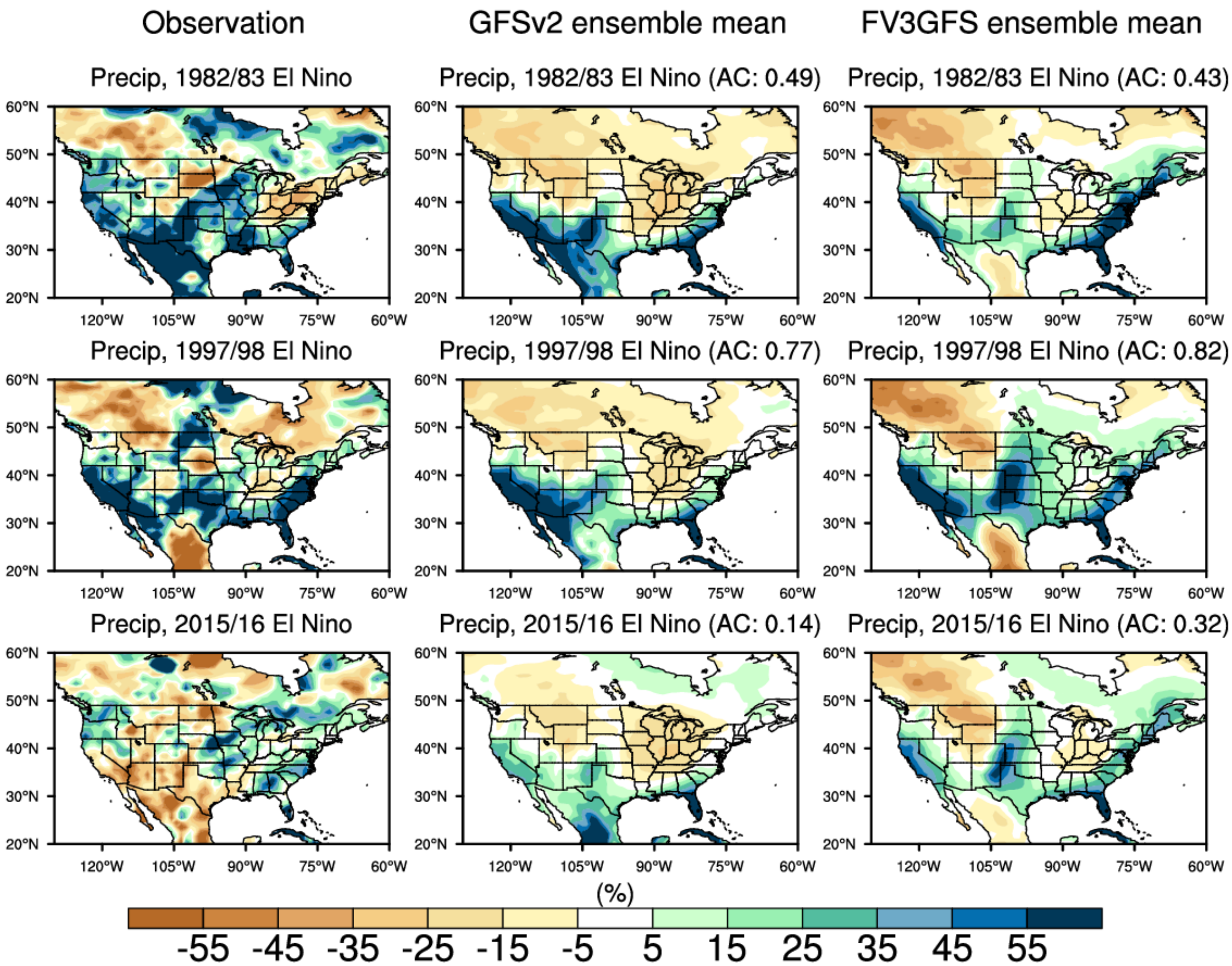

Figure 10. Precipitation anomalies (percent departures) for (top) 1982/83 DJF, (middle) 1997/98 DJF and (bottom) 2015/16 DJF from (left) observations, (middle) GFSv2 simulated 30-member ensemble mean, and (right) FV3 GFS simulated 100-member ensemble mean. The pattern correlations between models and observations are listed in the titles of the plots.

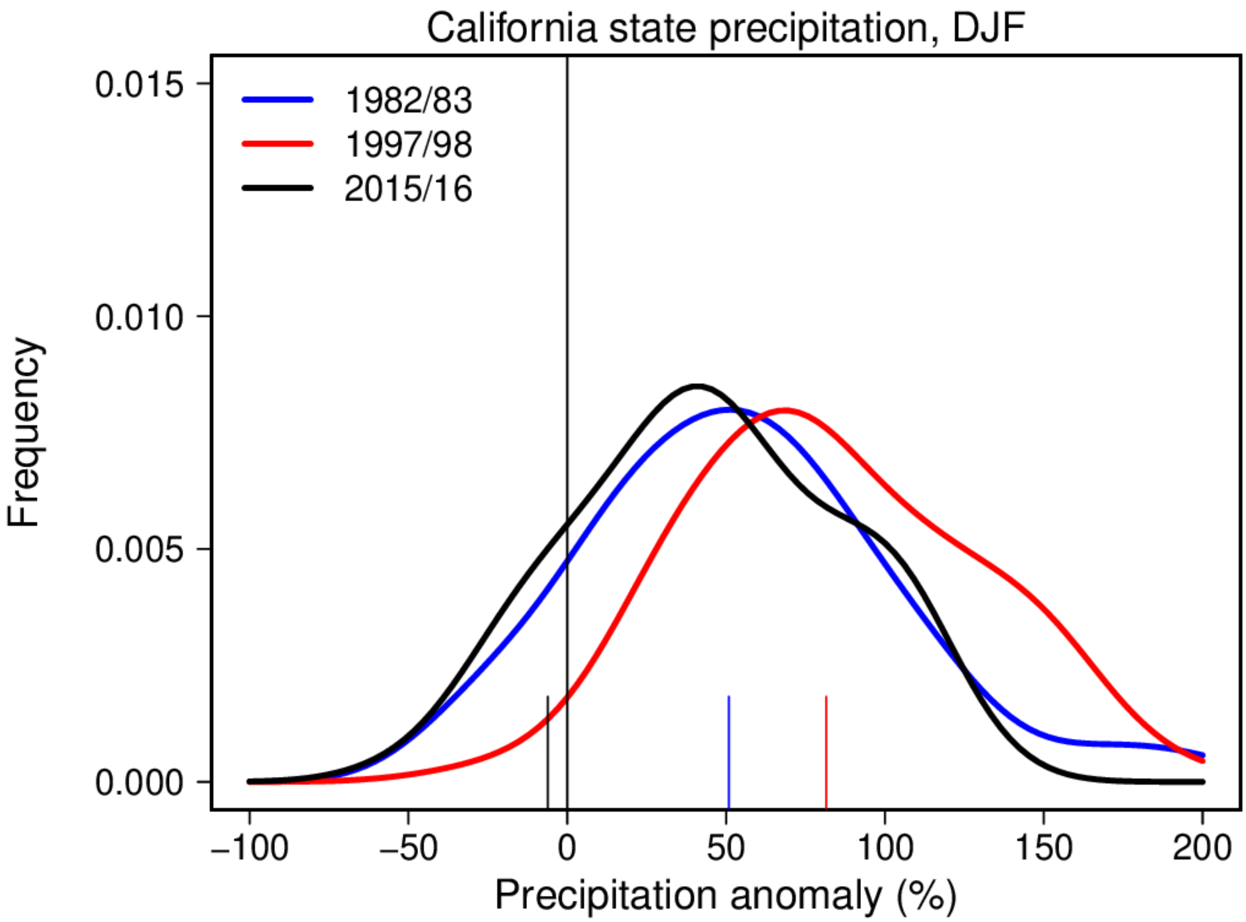

Figure 11. PDFs of California state precipitation anomalies (percent departure) for 1982/83 DJF (blue curve), 1997/98 DJF (red curve) and 2015/16 DJF (black curve). Results are based on 100-member FV3 GFS AMIP simulations. Large tick marks at the bottom show corresponding observed values.

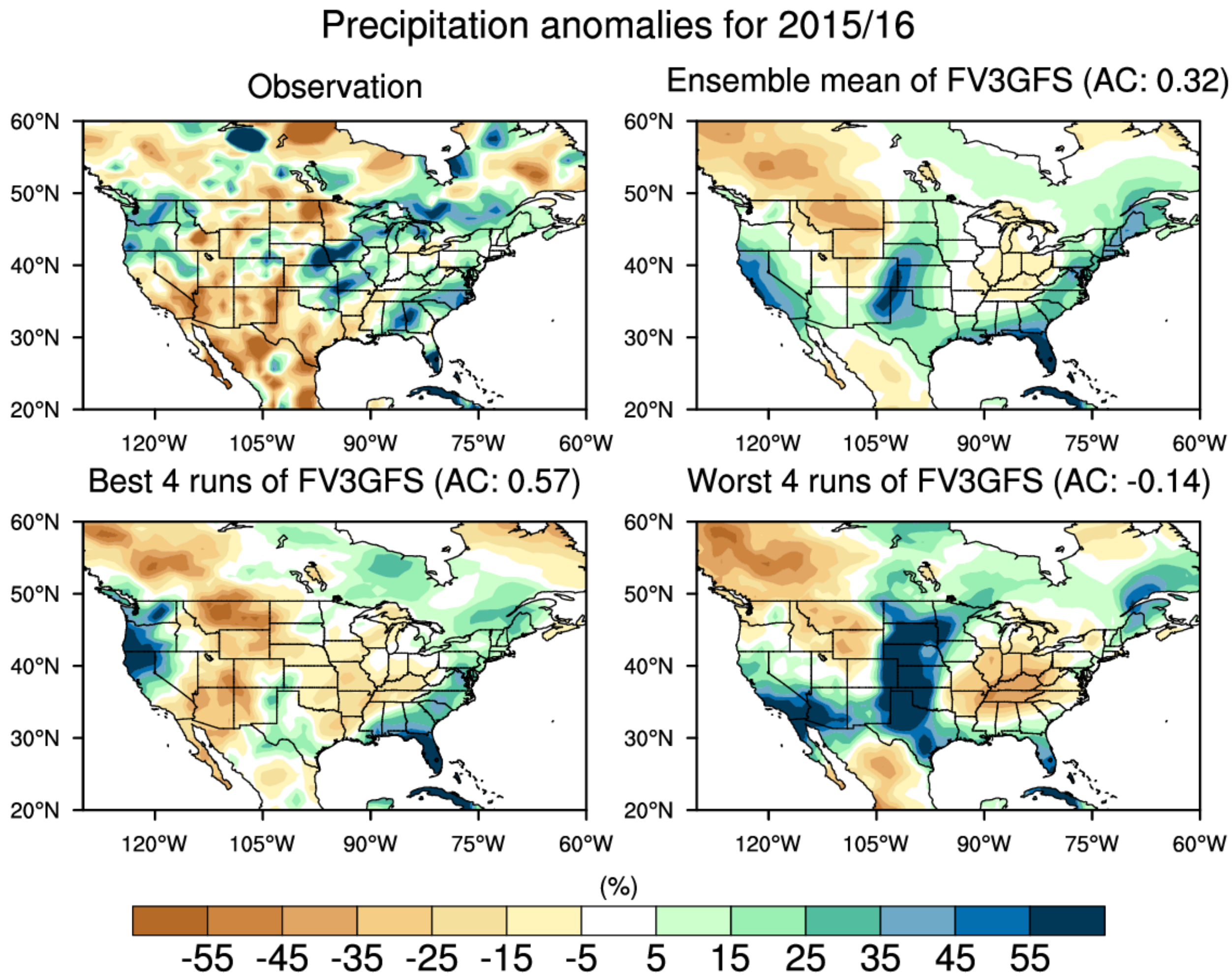

Figure 12. Precipitation anomalies (percent departures) for 2015/16 DJF from (top left) observations, (top right) FV3 GFS simulated 100-member ensemble mean, (bottom left) the composite of 4 best runs, and (bottom right) the composite of 4 worst runs among 100-member FV3 GFS AMIP simulations. The pattern correlations between the model and observation are listed in the titles of the plots.

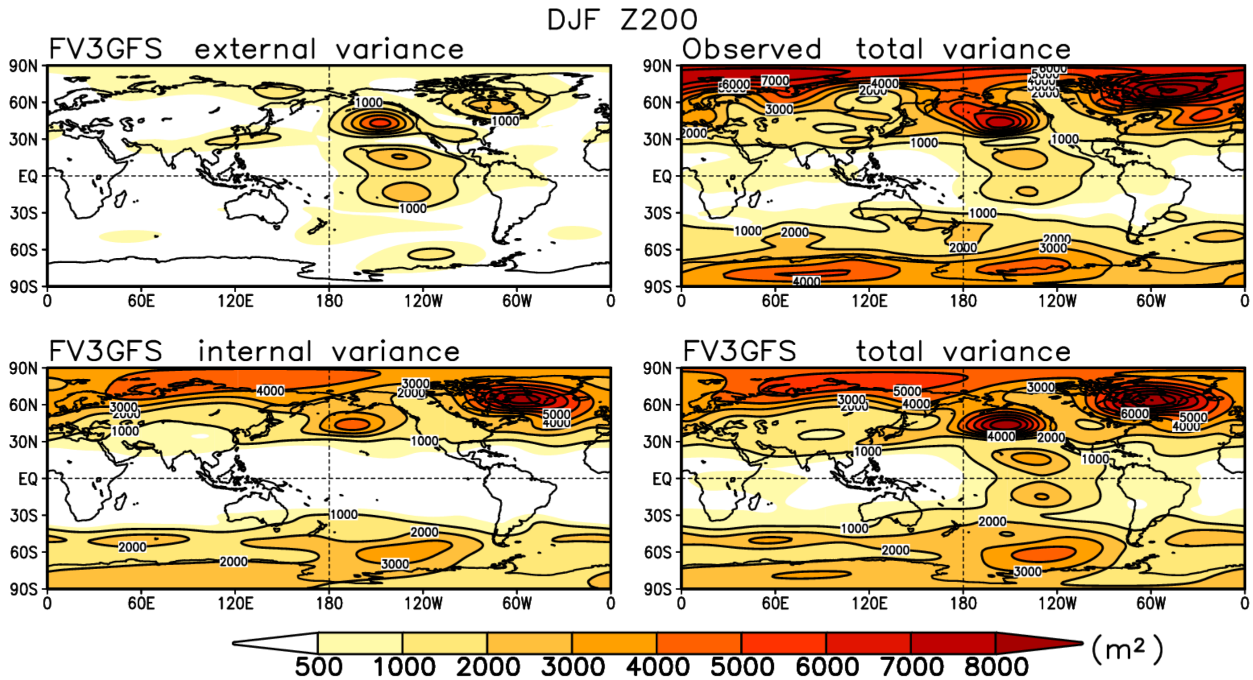

Figure 13: The spatial pattern of FV3 GFS simulated (top left) external variance, (bottom left) internal variance, (top right) observed total variance and (bottom right) FV3 GFS simulated total variance of DJF 200-hPa height anomaly. Model results are based on 100-member FV3 GFS AMIP simulations.

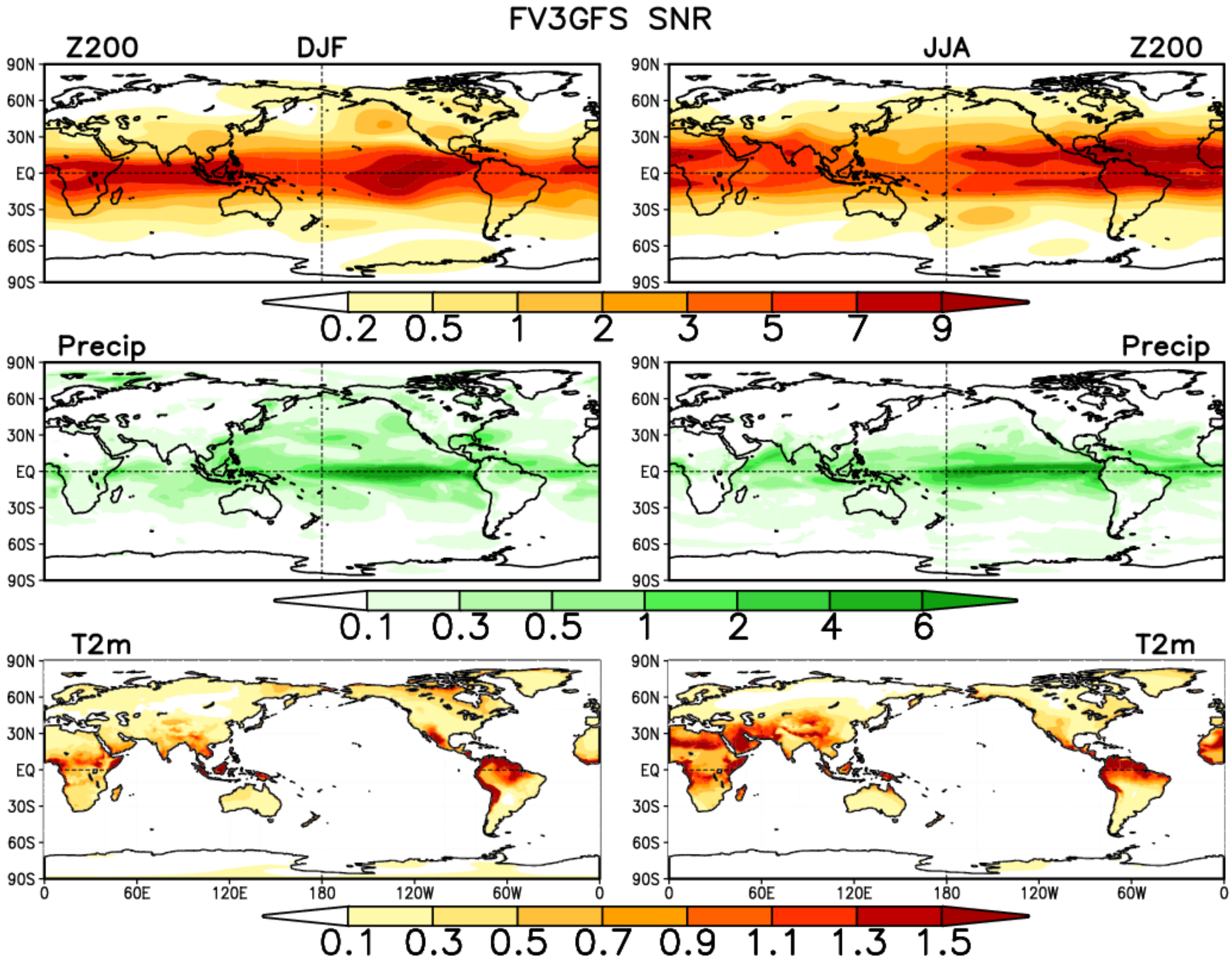

Figure 14: The spatial pattern of (left) DJF and (right) JJA signal-to-noise ratio (SNR) estimate of (top) 200-hPa height, (middle) precipitation and (bottom) surface air temperature anomalies computed as the ratio of external-to-internal variance in 100-member FV3 GFS AMIP simulations.

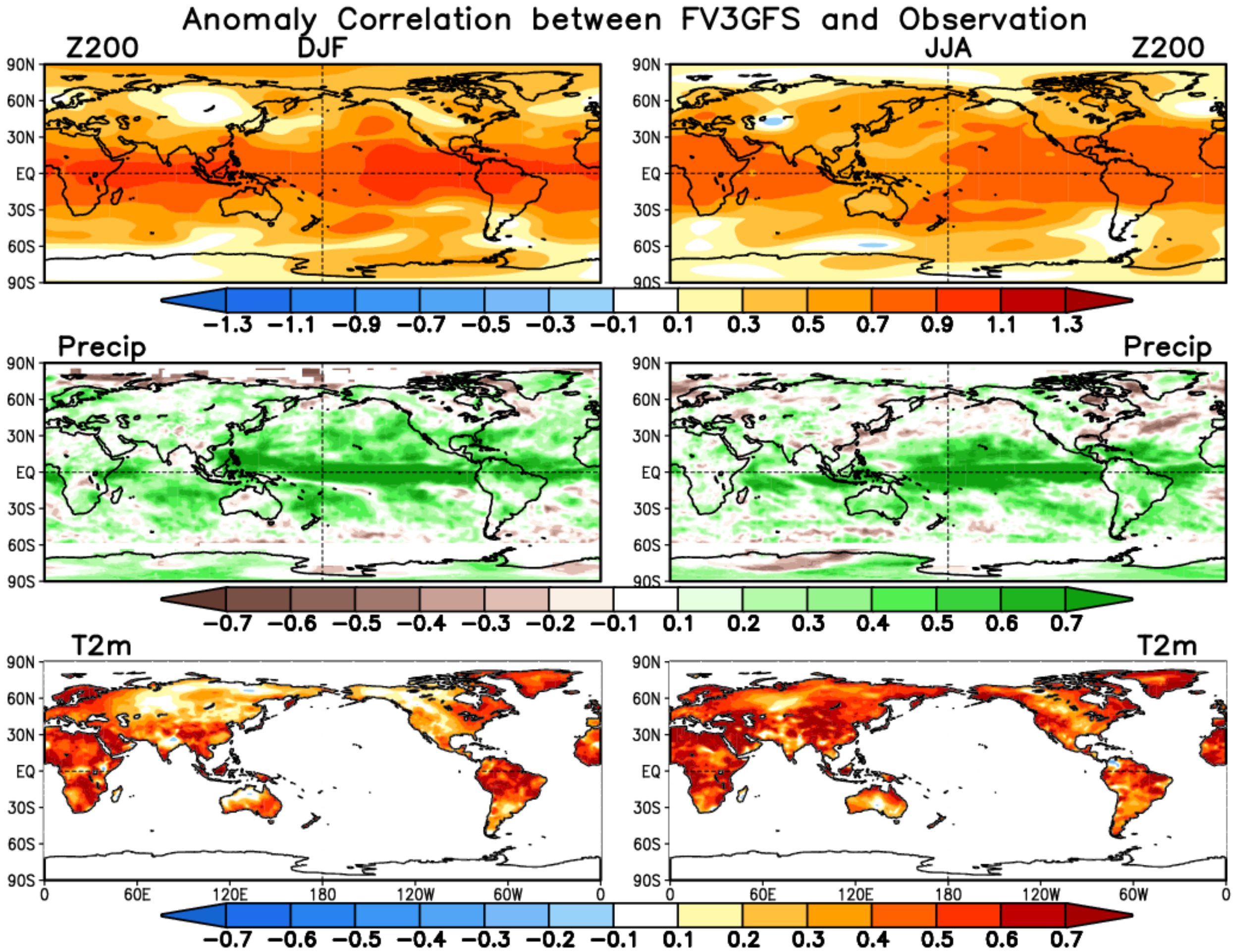

Figure 15: The spatial pattern of (left) DJF and (right) JJA anomaly correlation of (top) 200-hPa height, (middle) precipitation and (bottom) surface air temperature between observations and FV3 GFS simulated 100-member ensemble mean.

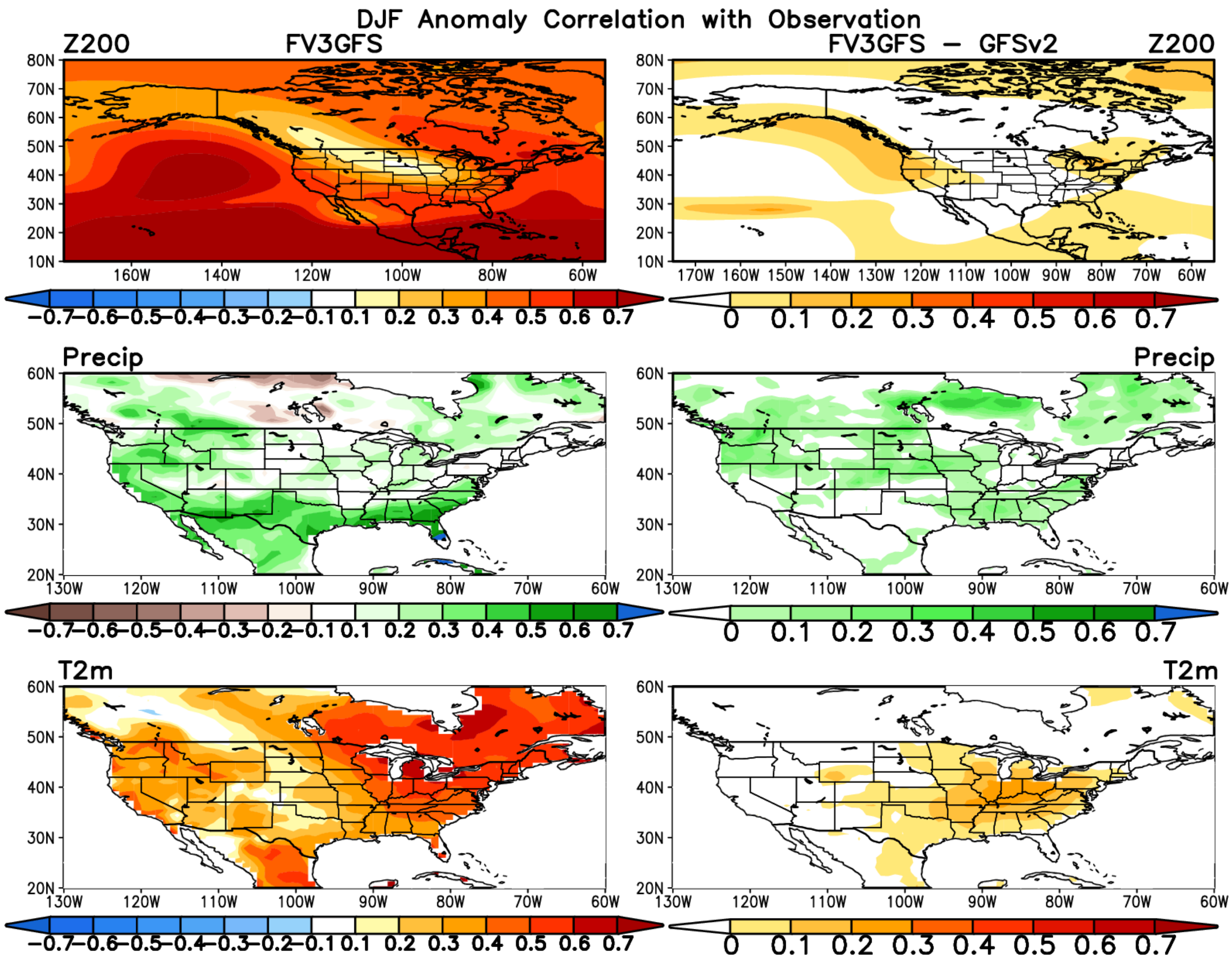

Figure 16: The spatial pattern of DJF anomaly correlation with observations of (top) 200-hPa height, (middle) precipitation and (bottom) surface air temperature from (left) FV3 GFS AMIP ensemble mean and (right) the difference in anomaly correlation with observations between FV3 GFS and GFSv2 AMIP ensemble mean.

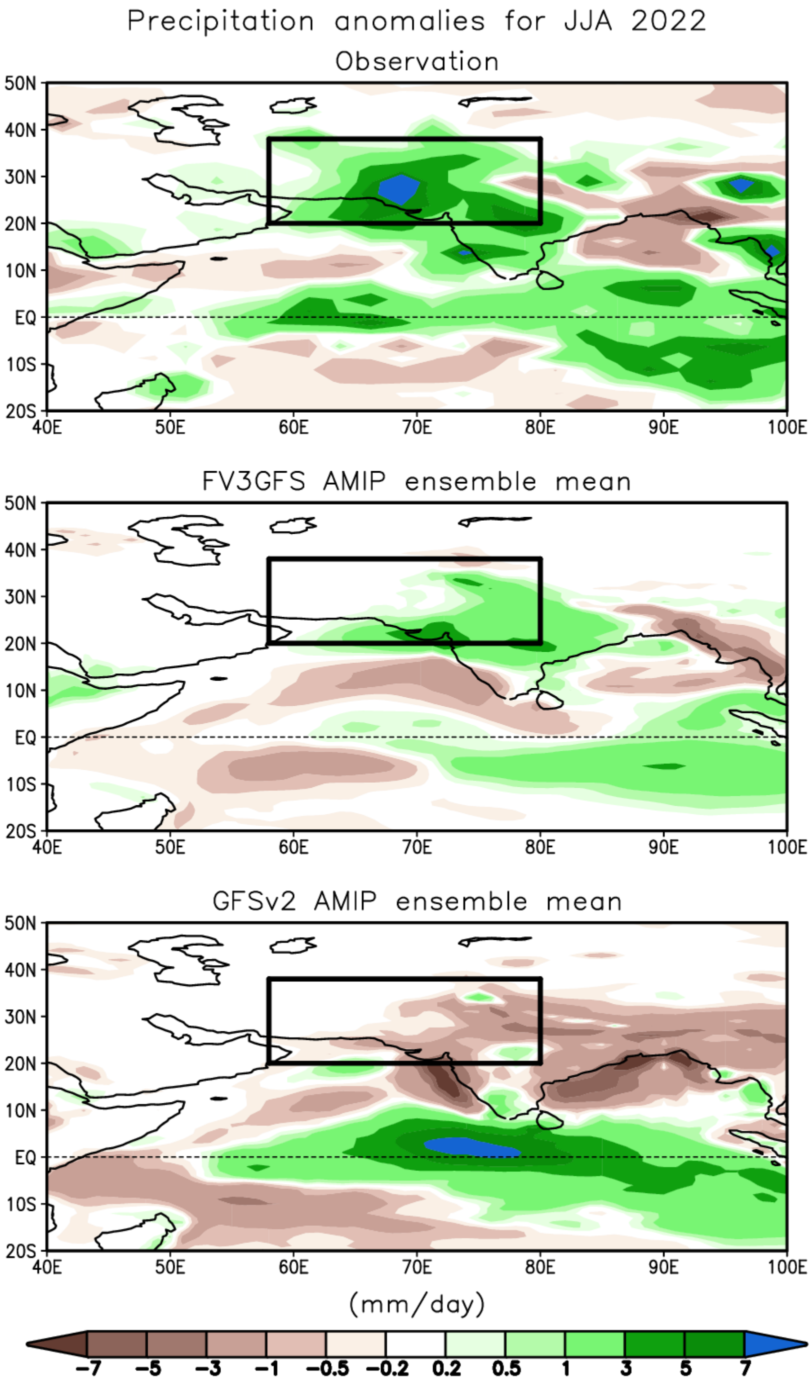

Figure 17. Precipitation anomalies for 2022 JJA from (top) observations, (middle) FV3 GFS simulated 100-member ensemble mean, and (bottom) GFSv2 simulated 30-member ensemble mean. The outlined box shows the South Asia region bounded by 58°-80°E, 20°-38°N.

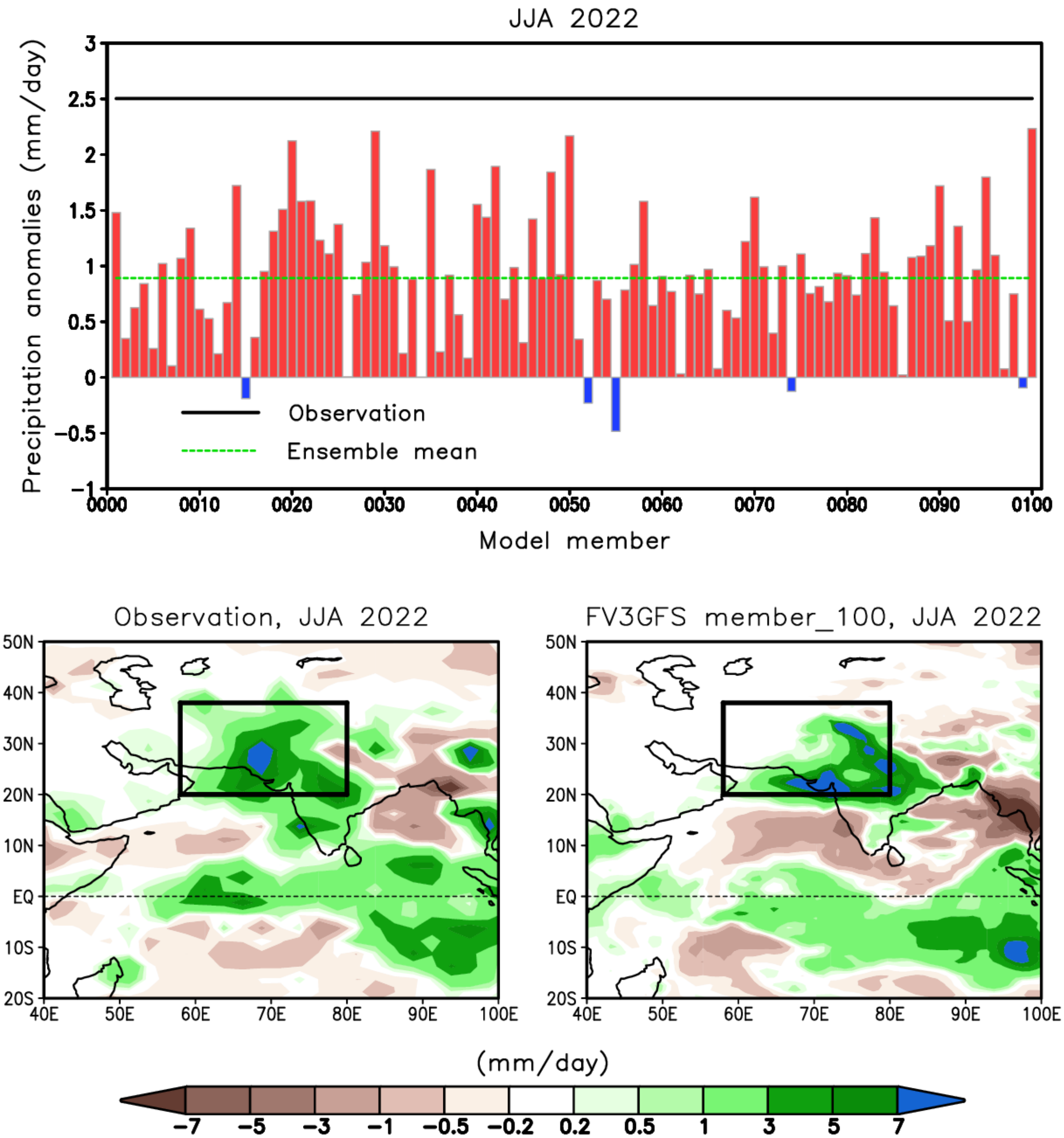

Figure 18. (top) precipitation anomalies for 2022 JJA averaged over the South Asia region (58^o-80^oE, 20^o-38^oN) from observations (black line), FV3 GFS simulated 100 individual members (red and blue bars) and 100-member ensemble mean (green line), and (bottom) the comparison of the spatial pattern of precipitation anomalies for 2022 JJA between (left) observations and (right) a single member (100th member) from FV3 GFS AMIP simulations. The outlined box shows the South Asia region bounded by 58^o-80^oE, 20^o-38^oN.