JGR Atmospheres

RESEARCH ARTICLE
10.1029/2023JD039621

A New GFSv15 With FV3 Dynamical Core 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, USA, [2]NOAA/National Centers for Environmental Prediction, College Park, MD, USA, [3]ERT Inc, Laurel, MD, USA, [4]NOAA/Physical Sciences Laboratory, Boulder, CO, USA, [5]CIRES, University of Colorado Boulder, Boulder, CO, USA

Key Points:

- A large Atmospheric Model Intercomparison Project ensemble based on NOAA's Global Forecast System version 15 with FV3 dynamical core is created to support attribution of observed climate anomalies at Climate Prediction Center
- The new simulations can replicate the observed climate variability, trends, and extreme seasonal events
- There are some improvements in simulating the extreme events in the new model compared to the older version

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

T. Zhang,
tao.zhang@noaa.gov

Citation:

Zhang, T., Yang, W., Quan, X.-W., Zhu, J., Jha, B., Kumar, A., et al. (2024). A new GFSv15 with FV3 dynamical core based climate model large ensemble and its application to understanding climate variability, and predictability. *Journal of Geophysical Research: Atmospheres*, *129*, e2023JD039621. https://doi.org/10.1029/2023JD039621

Received 11 JULY 2023
Accepted 23 MAR 2024

Abstract NOAA Climate Prediction Center (CPC) has generated a 100-member ensemble of atmospheric model simulations from 1979 to present using the Global Forecast System version 15 (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 Atmospheric Model Intercomparison Project simulations. These simulations are part of CPC's efforts to attribute observed seasonal climate variability to sea surface temperature (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, 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, trends, and 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 fidelity of US surface climate, mainly confined in the Northern and Southeastern US for precipitation and in the east for temperature.

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 to support the attribution of observed seasonal anomalies is reported in this study. We have generated 100-member simulations where each member has identical sea surface temperature (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 Atmospheric Model Intercomparison Project (AMIP)-style simulations, that is, atmosphere-only simulations constrained by time evolving observed sea surface temperatures (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 average of simulations documents the response to SSTs, often referred to as the attributable (or potentially predictable) component of the observed anomalies, or potential for predictions well beyond the limits of when initial atmosphere conditions constrain weather. The

contribution of the atmospheric noise (or unpredictable component) in the 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 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 life and property and the economy of the United States (Changnon, 1999; NOAA, 2017; Philip et al., 2021; Seager et al., 2015). For example, California and Southern States were affected by storms attributed to El Niño 1997–1998. 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 losses and 17,000 jobs in the agriculture sector (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 mortality rates 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 untangle the factors accounting for extreme weather and climate variability on different time scales (Barsugli et al., 2022; Murray et al., 2020). 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 understanding the physical explanation or principal causes for observed climate phenomena and conditions. 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, in August 2004 the first generation Climate Forecast System (CFS), called CFSv1, was implemented into operations at the NCEP and became the first coupled model used for seasonal prediction at NCEP (Saha et al., 2006). Subsequently, the CFSv2 was implemented in March 2011 (Saha et al., 2014), with improvements in forecast model components and the data assimilation. CFSv2 generates a set of 9-month retrospective forecasts with forecasts initialized using the analysis from the corresponding CFS Reanalysis (CFSR) (Saha et al., 2010, 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 understand 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 improvements in computational efficiency, energy conservation, and spatial resolution, 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 decided to be the next dynamical core of NOAA Next Generation Global Prediction System project as an upgrade for the GSM. In recent years, a new global model coupling the FV3 with GFS physical parameterizations, called the finite-volume GFS, or FV3GFS (Zhou et al., 2019), has been developed. The FV3GFS was implemented into the operational GFS 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 FV3GFS 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 analysis is to evaluate the performance of FV3GFS AMIP simulations relative to GFSv2 in replicating observed climate variability and trends for the period of 1979–2021.

This paper is arranged as follows: Section 2 introduces the observational and model data sets as well as analysis methods. Section 3 first presents an assessment of the climatology in the model, and then the simulation of observed trend and Northern Hemisphere (NH) 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. Data Sets and Methods

2.1. Observed and Model Data

The characteristics of observed estimates for land surface climate conditions are based on analysis of 2-m temperature (T2m) from the Global Historical Climatology Network/Climate Anomaly Monitoring System (GHCN/CAMS) (Fan and van den Dool, 2008) and precipitation fields at 1°-by-1° resolution from the Global Precipitation Climatology Center data sets (Schneider et al., 2014). Same as the data used in Zhang et al. (2016), 200-hPa geopotential height fields from the National Centers for Environmental Prediction-National Center for Atmospheric Research reanalysis (Kalnay et al., 1996) are examined to estimate the observed upper-level circulation pattern. 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 observational SST fields are obtained from the Hurrell data set (Hurrell et al., 2008), which are on a 1°-by-1° grid and merge the Hadley Center's SST version 1.1 (HADISST1) with the NOAA Optimal Interpolation SST version 2 (OISSTv2) from November 1981 onward. Global precipitation fields at 2.5°-by-2.5° resolution are from the CPC Merged Analysis of Precipitation (CMAP; Xie & Arkin, 1997).

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 & Lin, 2007) on a cubed-sphere grid (https://www.emc.ncep.noaa.gov/users/meg/fv3gfs/). The GFSv15 uses the RRTMG (Rapid Radiative Transfer Method for General Circulation Models) 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 convective parameterization scheme (Han & 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 and related references are documented 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 FV3GFS model used in our simulations has C96 horizontal resolution (about 100 km) and 64 vertical levels. We use specified observed monthly varying SSTs, sea ice (Hurrell et al., 2008), and carbon dioxide concentrations from the World Data Center for Greenhouse Gases operated by the Japan Meteorological Agency for 1979–2021 to force the model. The aerosols, solar and other greenhouse gases are specified with climatological values. A 100-member ensemble of AMIP simulations is maintained at NOAA's CPC. Each member in the ensemble, although starting from a different initial atmospheric condition, has identical external forcing. The forced response to external forcings is derived from the statistics of 100-member simulations, for example, ensemble average.

To evaluate the robustness in replicating the key features of the observed climate variability and trends by FV3GFS, we also diagnose the AMIP simulations from a 30-member ensemble of the GFSv2 model that spans the same period. The GFSv2 model is the previous version of CPC AMIP simulations and has spectral T126 horizontal resolution and 64 levels in the vertical, and further, is the atmospheric component of the NCEP CFSv2 (Saha et al., 2014).

2.2. Methods

In this study, we follow the methodology of Zhang et al. (2016) to obtain the leading modes of the observed NH winter circulation variability by performing empirical orthogonal function (EOF) analysis with 200-hPa heights for the 42 years of DJF seasonally averaged 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 shown as regressions against the principal component (PC) time series. Here we utilize unrotated EOFs that are built to be both temporally and spatially 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 FV3GFS simulated DJF 200-hPa geopotential heights from each member of AMIP simulations, demonstrating that the model can well capture the observed three leading modes of interannual variability.

Based on a 100-member ensemble of FV3GFS 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 FV3GFS 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 and corresponding expected skill of predictions. The SNR's signal component is the variance of ensemble mean while the noise component is the variance of departure in each member from the ensemble mean (Kumar & Hoerling, 1995). The corresponding expected skill, also called expected value for anomaly correlation (EAC), is a function of SNR (Kumar, 2009). Higher SNR or EAC values indicate larger predictability. The anomaly correlation (AC), computed as the correlation between anomalies for the AMIP ensemble mean and observations, is calculated to complement SNR and EAC analysis. It is expected that larger SNR or EAC would correspond to larger AC (Kumar & Hoerling, 2000). Anomalies are computed relative to a 1991–2020 reference period for observations and AMIP simulations.

3. Results

3.1. 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 & Arkin, 1997).

Figure 1 shows the difference in climatology between JJA and DJF (JJA minus DJF) for observations (left panel) and FV3GFS AMIP ensemble mean (right panel). The largest difference in eddy (zonal mean removed) 200-hPa height is in the NH middle latitude, mainly resulting from the stronger magnitude during DJF (Figure S1 in Supporting Information S1). 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 in both hemispheres precipitation is greater in their respective summers. This feature is realistically reproduced in the model. Surface air temperature difference pattern is also similar between model and observation, with warming in the NH 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 FV3GFS can realistically capture the observed seasonal cycle of climatology. FV3GFS can also simulate the feature of observed climatology of large-scale atmospheric circulations for both DJF and JJA seasons (Figure S1 in Supporting Information S1). Table 1 lists the global mean values of the difference in climatology (JJA minus DJF) for GFSv2 and FV3GFS 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

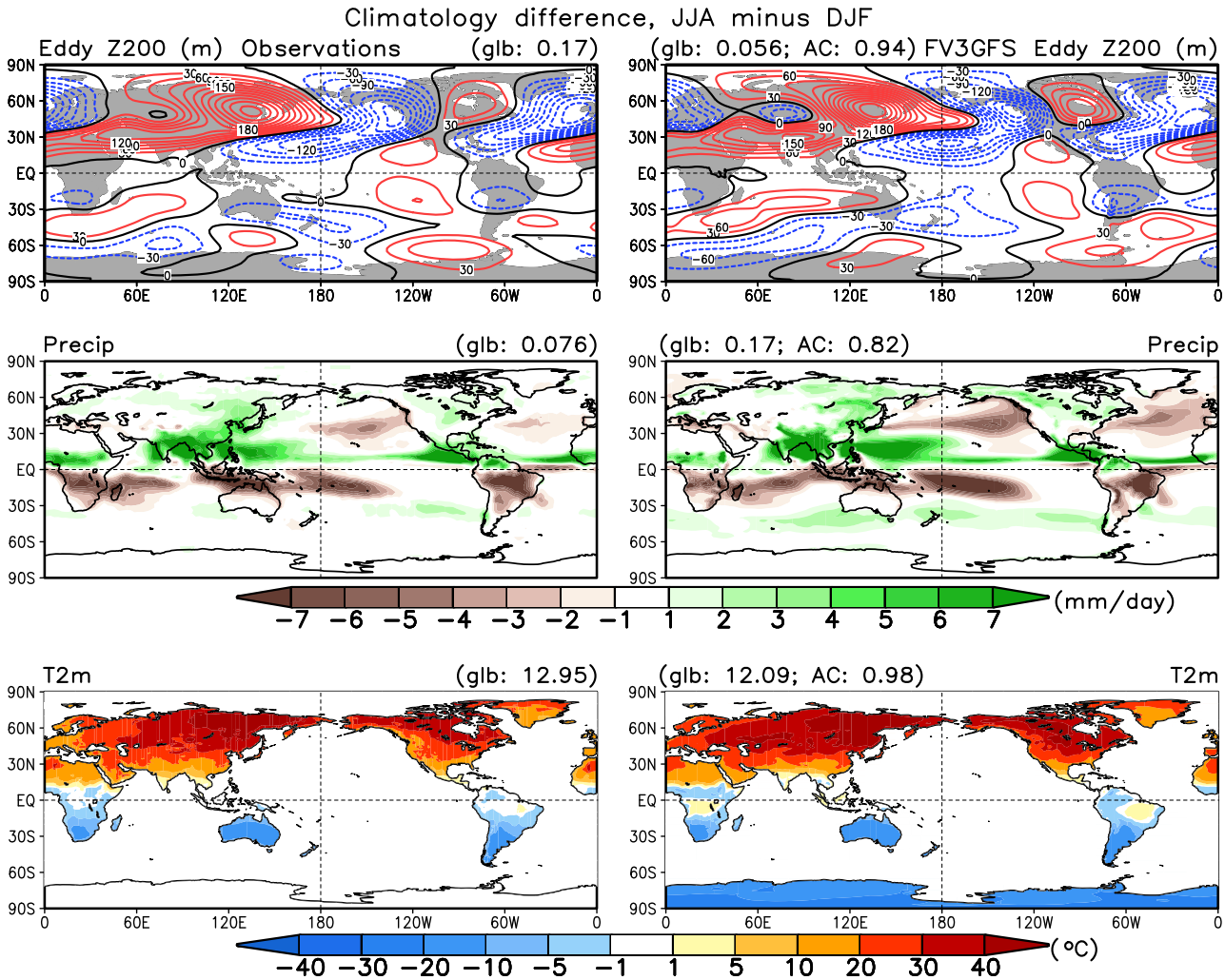

Figure 1. Differences in climatology between JJA and DJF (JJA minus DJF) for (top) eddy 200-hPa height, (middle) precipitation and (bottom) surface air temperature from (left) observations and (right) FV3GFS simulated 100-member Atmospheric Model Intercomparison Project ensemble mean results. The observed and simulated global mean values (the first number) and the pattern correlation values (the second number) are listed in the titles of the plots.

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 (°C) From Observation, GFSv2, and FV3GFS Ensemble Mean Atmospheric Model Intercomparison Project Simulations, and the Global Pattern Correlations Between Models and Observation for the Corresponding Climatology Difference

	Global mean values			Global pattern correlations with observation	
Variables	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

correlation with observations for precipitation and the global mean value for surface air temperature are improved in FV3GFS 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), since Taylor diagram approach can be used for assessing the climatological mean behavior of model simulations (Heo et al., 2014; Seo et al., 2014; Xiong et al., 2021). Two models generally demonstrated a similar ability to simulate the annual mean precipitation and surface air temperature, featuring the largest pattern correlation (greater than 0.95) for Northern midlatitude temperature and the smallest pattern correlation (about 0.81) for tropical precipitation that has the largest normalized root-mean-square (RMS) error (about 0.65). The RMS error is relatively small (less than 0.5) for surface air temperature. The standard deviation of Northern midlatitude temperature and precipitation and tropical precipitation is somewhat overestimated, and

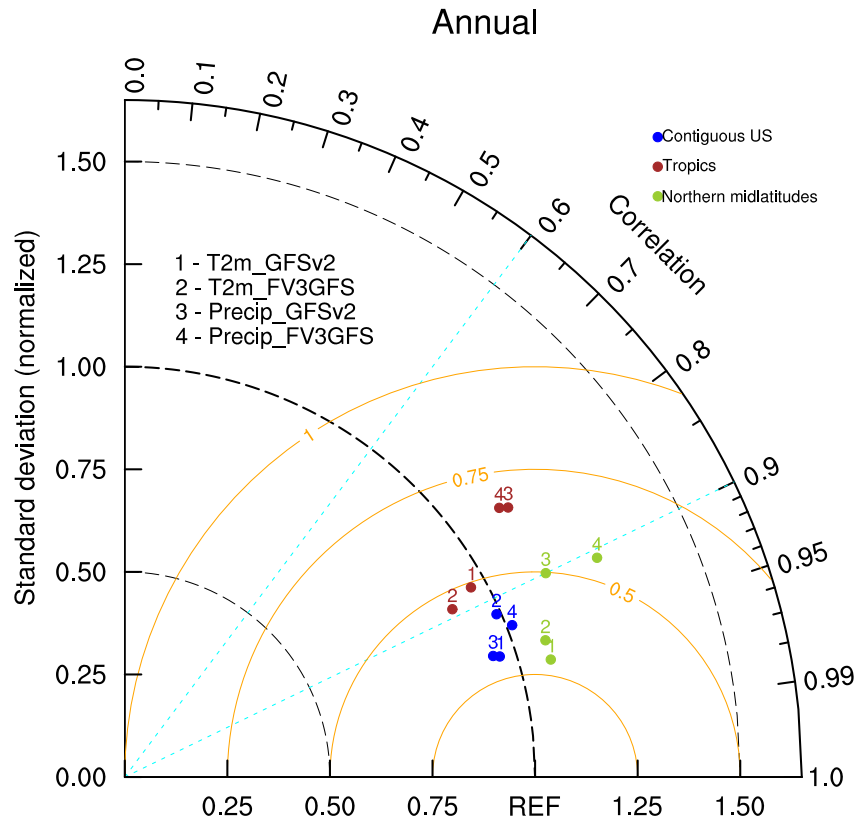

Figure 2. Taylor diagram displaying changes in normalized pattern statistics between GFSv2 and FV3GFS model estimates for the (greenish-yellow dots) Northern midlatitudes (30°N–60°N), (brown dots) tropical (30°S–30°N), and (blue dots) contiguous US pattern of annual mean climatology of precipitation and land surface air temperature. The standard deviations have been normalized by the observed standard deviation of each field.

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 FV3GFS relative to GFSv2. The simulation of observed trends and climate variability are discussed in the following sections.

3.2. Simulation of Observed Trends

A large fraction of recent global land warming trend associated with human activities is communicated to the atmosphere via the indirect influence of trends in SSTs (Compo & Sardeshmukh, 2009; Fahad & Burls, 2022; Hoerling et al., 2006). 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 & Tung, 1998), which can enhance the amplitude of climate feedback in response to anthropogenic activities through positive feedback (Held & 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.

Shown in Figure 3 is 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 FV3GFS 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 FV3GFS are also shown (gray lines). It is clear that observations have

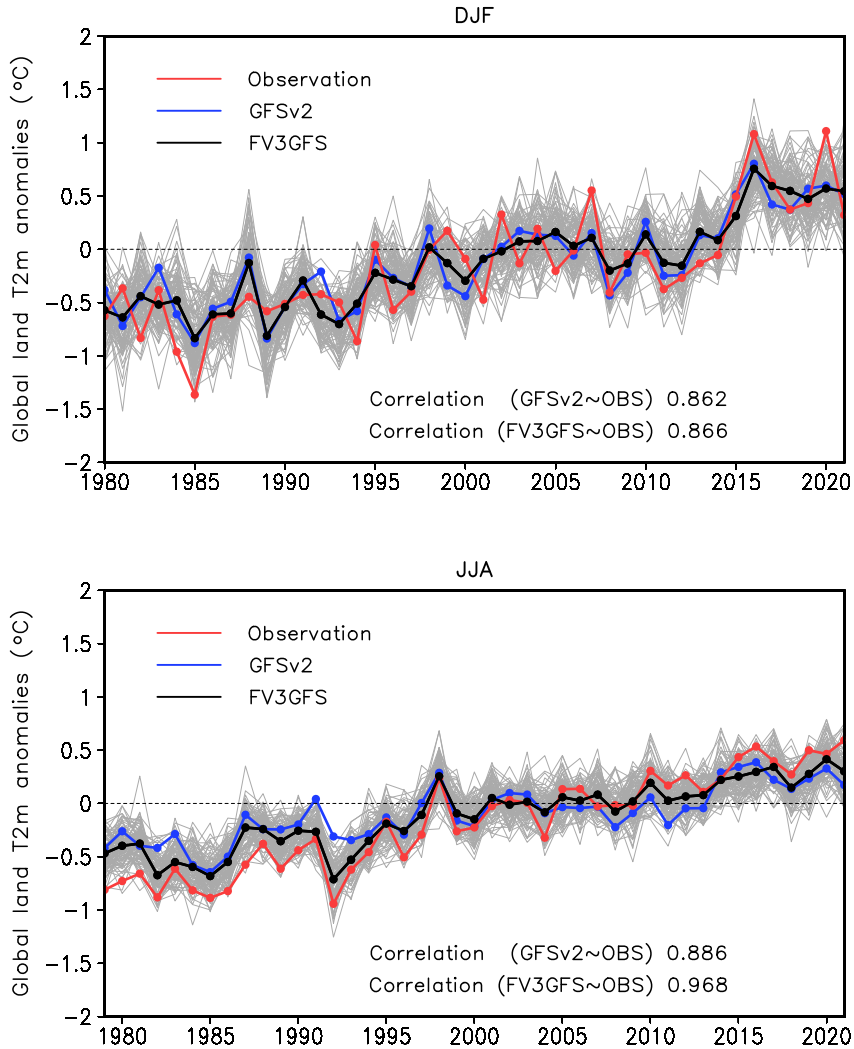

Figure 3. Time series of surface air temperature anomalies for (top) DJF and (bottom) JJA averaged over global land regions from (red line) observations, (blue line) GFSv2 simulated 30-member ensemble mean and (black line) FV3GFS simulated 100-member ensemble mean of Atmospheric Model Intercomparison Project simulations. The gray lines show the spread of individual members of FV3GFS model. The temporal correlations of the global mean land surface air temperature anomalies between model ensemble mean and observations are listed in the plot.

an upward trend of about 1°C since 1979 for both winter and summer. The FV3GFS 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 FV3GFS for winter, but the correlation is somewhat smaller in GFSv2 for summer.

To explore whether FV3GFS model can capture the observed trends, Figure 4 shows the frequency distributions [also called PDFs] 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 assessed by the Kolmogorov-Smirnov test, are based on 1,500 (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

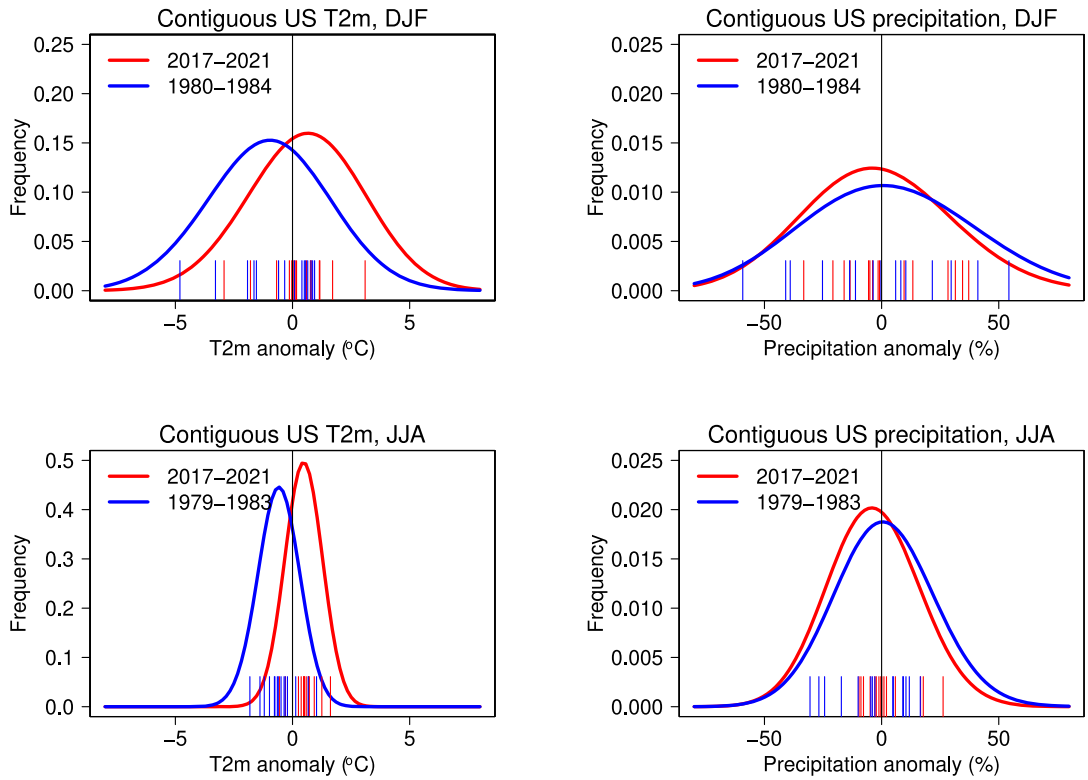

Figure 4. Probability density functions of contiguous U.S. (top) DJF and (bottom) JJA (left) surface air temperature anomalies (°C) and (right) precipitation anomalies (percent departure) for the first (blue curves) and last (red curves) 5-yr periods of 1979–2021. Results are based on 100-member FV3GFS Atmospheric Model Intercomparison Project simulations. Large tick marks at the bottom show observed values for 15 months (3 months × 5-yr) of the first (blue) and last (red) 5-yr periods of 1979–2021.

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 (Figure S2 in Supporting Information S1).

3.3. 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 & Wallace, 1981; Trenberth et al., 1998). Further, understanding the atmospheric response patterns beyond the canonical response to ENSO is also a fundamental problem in quantifying the sources of predictability and attribution of climate variability, and further, may result in improvements in our understanding of seasonal predictability (Barnston et al., 2005; Hoerling and Kumar 2002; Kumar et al., 2005; Zhang et al., 2016). It is thus important to assess the capability of FV3GFS in reproducing the leading modes of climate variability.

Figure 5 presents boreal winter (DJF) 200-hPa height structures for the leading three EOFs of the reanalysis data, explaining a total 56.6% of the height variations north of 20°N. Contours in the left panels and shaded values in the right panels are the observed 200-hPa heights and SSTs regressed against PC time series of each eigenvector given in the middle panels for 1979–2021, respectively.

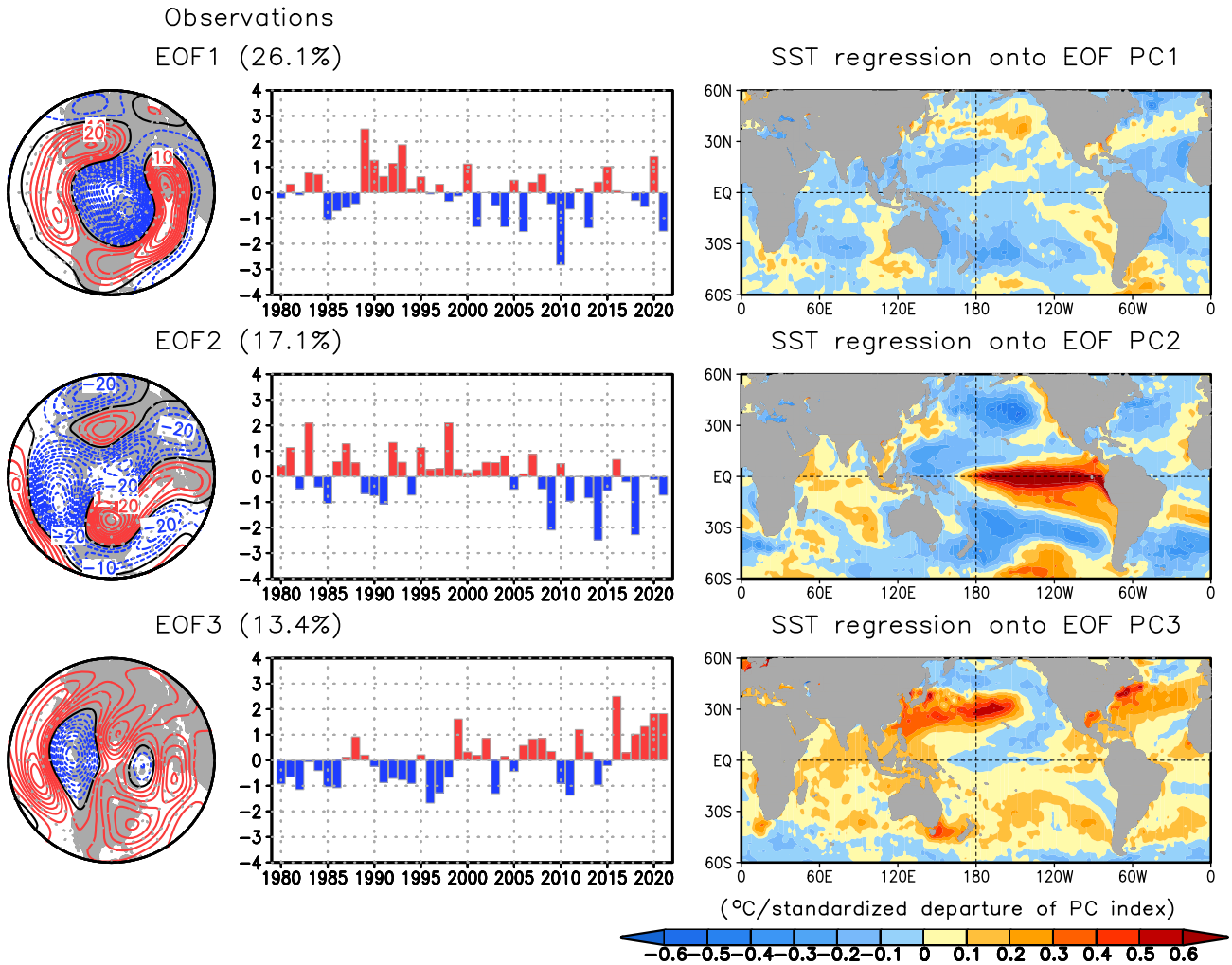

Figure 5. (left) The spatial pattern and (middle) standardized principal component (PC) time series of the leading three empirical orthogonal functions (EOFs) of DJF 200-hPa heights from observations. (right) Regressions of observed DJF sea surface temperature on the PC time series of the leading three EOFs of observed DJF 200-hPa heights. The EOF analysis is computed over the 20°–90°N domain for 1979/1980–2020/2021. The EOF patterns are shown as the regressions of the heights onto the standardized PC time series and drawn at the interval of 5 m for a one standardized departure of PC index.

The structure of the first leading mode of the observed variability consists of positive (negative) height anomalies in the NH middle latitudes (polar regions) while the time series for this mode is uncorrelated (the value is −0.026) with Niño-3.4 SSTs. This pattern explains 26.1% variability in extratropical NH wintertime heights. 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 replicated in a climate simulation without interannual variability in 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 & 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 & 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 depicts the canonical atmospheric teleconnection response linked to 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 time varying external forcing is anthropogenic greenhouse gases (Zhang et al., 2016), supporting the argument that this mode is associated with 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 have 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. In general, due to the sampling variability in the EOF rankings, 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.

FV3GFS 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 model simulation and observations may be similar, the corresponding time series could still be uncorrelated. The corresponding SST regression map (Figure 6, right) is very similar to observations (Figure 5, right), supporting the physical mechanism associated with the observed three leading modes.

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 explain a combined 84.8% of the variance in ensemble mean height variability that is externally forced.

The height pattern related to the first mode of the AMIP forced response exhibits a prominent wave train over the PNA region that resembles the TNH pattern. The temporal variability for this mode has a high correlation (0.93) with Niño-3.4 SST variability, featuring positive polarity during warm events (e.g., 1982/1983, 1991/1992, 1997/1998, 2002/2003, 2009/2010, 2015/2016) and negative polarity during cold events (e.g., 1988/1989, 1998/1999, 1999/2000, 2007/2008, 2011/2012, 2020/2021). 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 variance in the total boundary forced component of 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. Having a moderate spatial correlation of 0.69, this forced mode resembles the observed third mode shown in the bottom of Figure 5 that is strongly associated with climate change discussed earlier. The explained variance by this forced mode is 28.0%.

Explaining 15.3% of the total boundary forced component of the NH extratropical height variability, the height pattern related to the third mode of forced AMIP response depicts a classic PNA-like wave train pattern. Its action centers are in spatial quadrature with the leading mode of forced solutions, resembling the second EOF pattern in Zhang et al. (2016). The corresponding PC3 time series shows the larger amplitudes during both ENSO events (e.g., 1982/1983, 1997/1998, 2015/2016 El Niño events, and 1988/1989, 1998/1999, 1999/2000, 2007/2008, 2011/2012 La Niña events) and some ENSO-neutral years (e.g., 1985/1986, 1996/1997, 2013/2014). Zhang et al. (2016) found that there is a high correlation between PC time series for this forced mode and SST trans-Niño

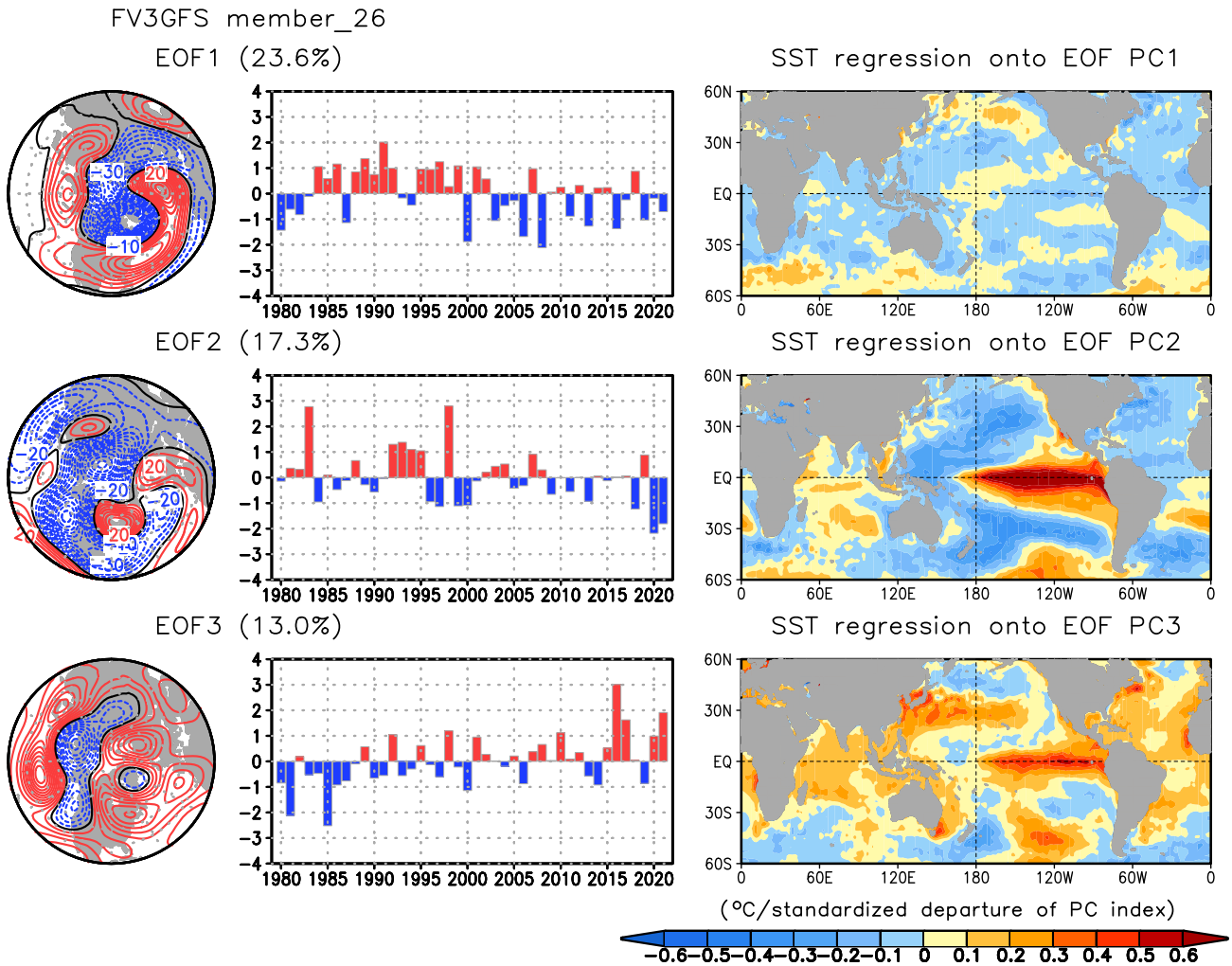

Figure 6. (left) The spatial pattern and (middle) standardized principal component (PC) time series of the leading three EOFs of DJF 200-hPa heights from a single member (member 26) of FV3GFS Atmospheric Model Intercomparison Project simulations. (right) Regressions of observed DJF sea surface temperature on the PC time series of the leading three EOFs of DJF 200-hPa heights from this member (member 26). This single member is selected among members that resemble observations, subject to the largest mean value of three leading empirical orthogonal function (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 20°–90°N domain for 1979/1980–2020/2021. The EOF patterns are shown as the regressions of the heights onto the standardized PC time series and drawn at the interval of 5 m for a one standardized departure of PC index.

index (Trenberth & Stepaniak, 2001), which measures the gradient in SST anomalies across the equatorial Pacific and describes the feature of ENSO evolution in the transition phase. The SST regression map (Figure 7 bottom right) is very close to the pattern of SST asymmetry between two phases of ENSO (Zhang et al., 2016). The SST asymmetry results from the fact that the magnitude of SST anomalies is larger over the eastern equatorial Pacific during El Niño while it is larger over the western equatorial Pacific during La Niña, causing a positive SST skewness in the Niño-3 region (An & Jin, 2004; Burgers & Stephenson, 1999; Zhang & Sun, 2014; Zhang et al., 2009) and indicating nonlinearity in tropical SST forcings. Therefore, the positive phase of the forced third mode is linked to the asymmetry in ENSO teleconnections between warm and cold phases. Zhang et al. (2016) also argued that for the negative phase of this mode, the associated tropical SST pattern resembles the El Niño precursor pattern (e.g., Penland & Sardeshmukh, 1995), featuring warmth (coolness) in the far western (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.

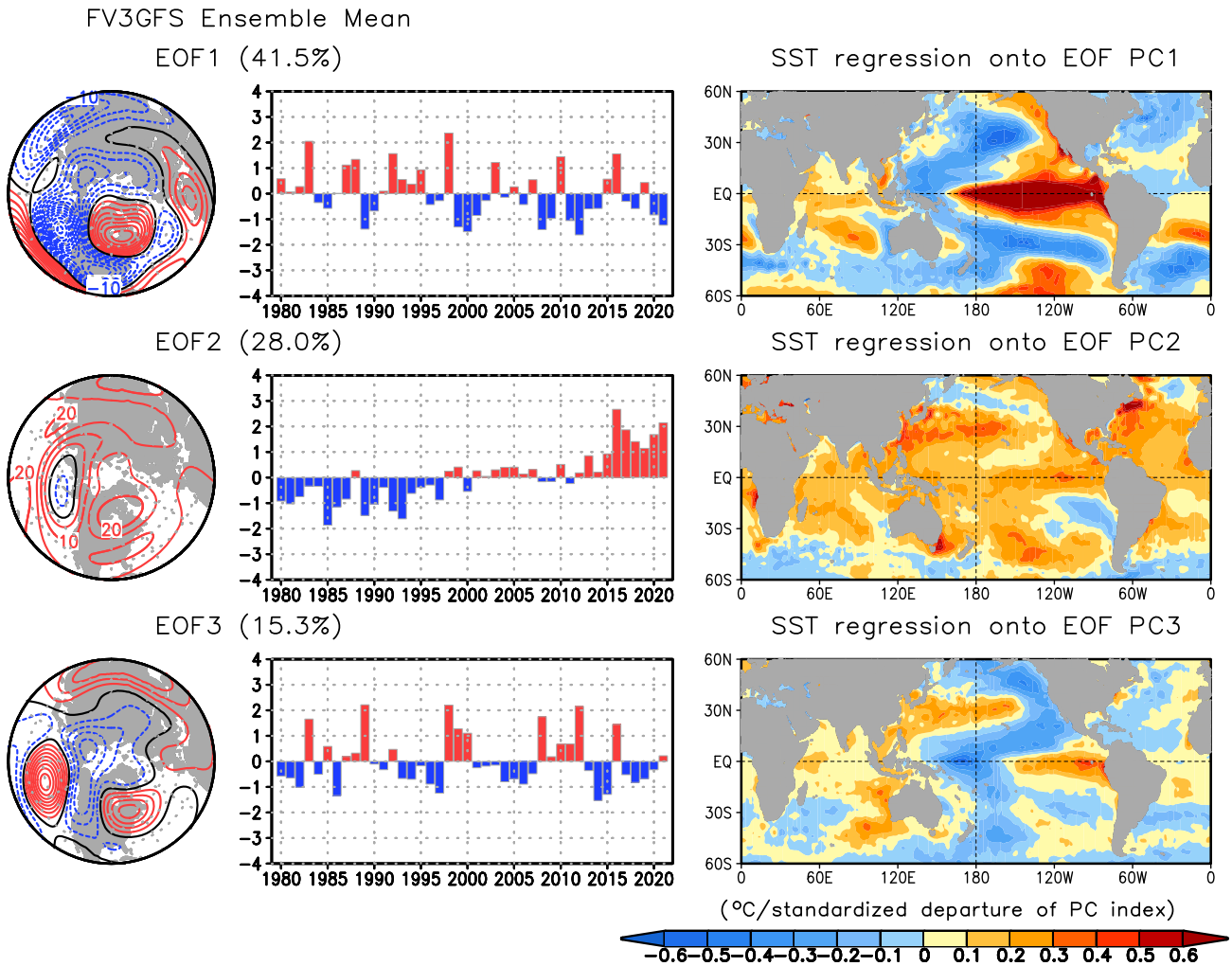

Figure 7. (left) The spatial pattern and (middle) standardized principal component (PC) time series of the leading three empirical orthogonal functions (EOFs) of FV3GFS simulated 100-member ensemble mean DJF 200-hPa heights. (right) Regressions of observed DJF sea surface temperature on the PC time series of the leading three EOFs of FV3GFS simulated 100-member ensemble mean DJF 200-hPa heights. The EOF analysis is computed over the 20°–90°N domain for 1979/1980–2020/2021. The EOF patterns are shown as the regressions of the heights onto the standardized PC time series and drawn at the interval of 5 m for a one standardized departure of PC index.

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 (Figure S3 in Supporting Information S1). The results are found to be similar, including the EOF ranking and explained variance of three leading forced modes.

The observed leading mode, that is, 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.

To explore whether the leading three EOF modes of FV3GFS AMIP ensemble mean DJF 200-hPa height anomalies are predictable signals, the observed DJF 200-hPa height anomalies are first projected onto the three leading EOF patterns separately, and the time series of the projections are then compared with the corresponding PC time series of the leading three EOFs of FV3GFS ensemble mean DJF 200-hPa height (Figure S4 in Supporting Information S1). Our analysis implies that the leading two modes are more predictable than the third mode.

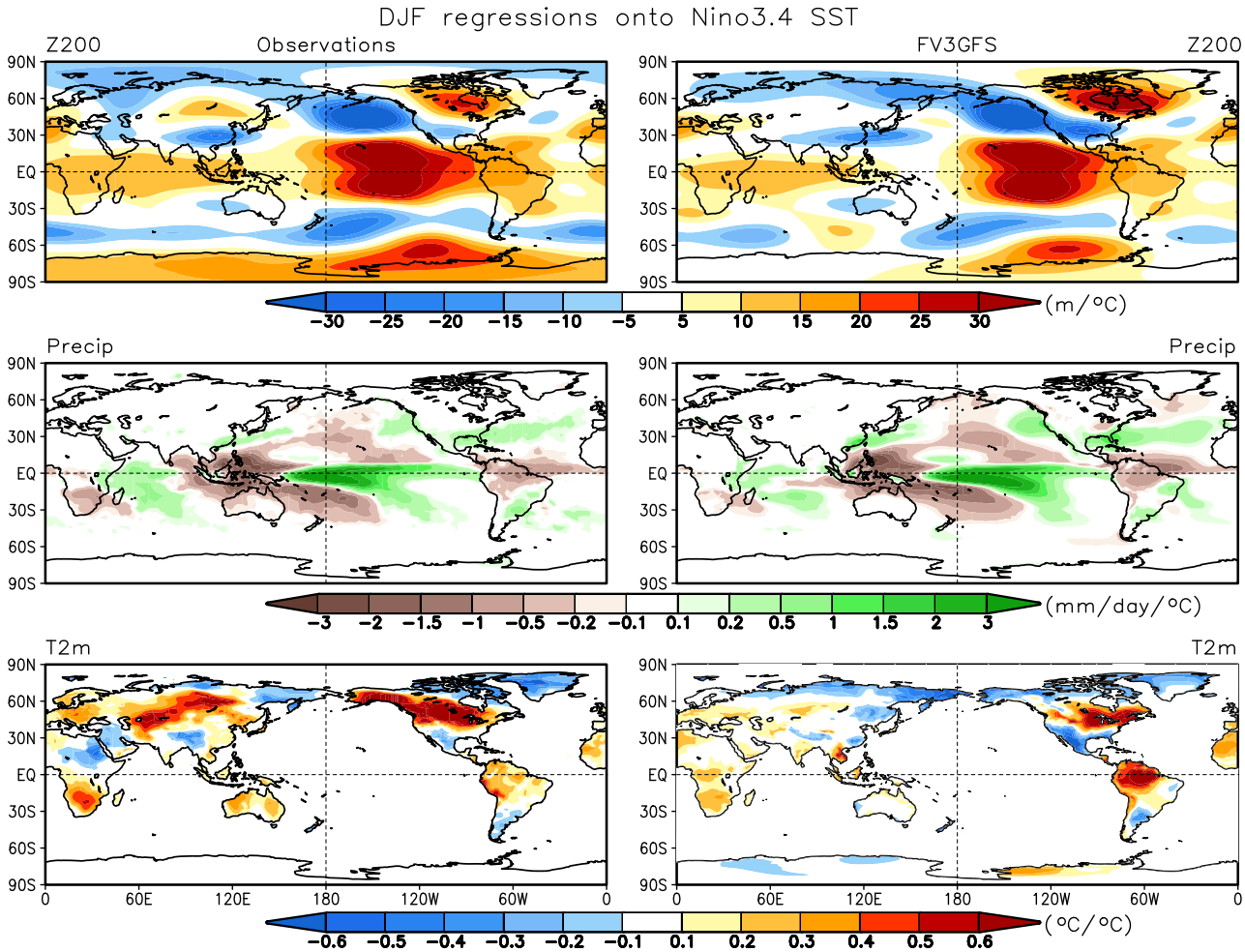

Figure 8. The spatial pattern of regressions of DJF (top) 200-hPa height, (middle) precipitation and (bottom) surface air temperature anomalies on the observed Niño3.4 sea surface temperature index from (left) observations and (right) FV3GFS Atmospheric Model Intercomparison Project simulations. We first calculate the regressions from individual runs and then average 100 regression estimates as the regressions for the model.

3.4. Simulation of US Climate Response to ENSO

ENSO is the largest source of atmospheric predictability and an important aspect of climate attribution (e.g., Goddard & Dilley, 2005; Kumar & Hoerling, 1998; 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 FV3GFS 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 the tropical anticyclonic anomalies, the North Pacific cyclonic anomalies 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 (EP). The temperature response reveals warming (cooling) over the northern (southern) United States, similar to the observed surface temperature composite during EP El Niño (Zhang et al., 2020). Appreciable warmth is also observed over Eurasia in the middle latitude.

FV3GFS AMIP results reproduce the observed key features associated with ENSO. The simulated rainfall response, however, is smaller than observations over the western Indian Ocean, an area which is an important source of teleconnections in the early-winter phase of ENSO episodes (e.g., Abid et al., 2021). The magnitude of the negative surface temperature anomalies 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 severely underestimated over Eurasia, South Africa, and Australia and overestimated over the Northern South American continent. FV3GFS AMIP simulations also capture the observed responses to ENSO for JJA season featuring weakened tropical anticyclonic anomalies and the North Pacific cyclonic anomalies, reduced convection over the tropical Indian Ocean and northward expansion of enhanced convection over the tropical western Pacific, and cooling prevailing over most regions of the globe. However, the observed cyclonic anomalies over the western U. S., the negative precipitation response over the eastern Indian Ocean, and the coolness over the Eastern Europe and the warmth over the Southern South American continent are underestimated (Figure S5 in Supporting Information S1). 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 FV3GFS 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 FV3GFS simulated ensemble mean (right). During the 1982/1983 El Niño, stronger warm anomalies dominate the northern US, but the cold anomalies prevail over the southern US. The above normal anomalies shift gradually from north to south in recent two strong El Niño (1997/1998 and 2015/2016) events.

Similar to observations, there is a clear southward shift of warm anomalies from 1982/1983 El Niño to 2015/2016 El Niño for two model ensemble mean results. This is consistent with the US warming trends documented using PDFs (Figure 4). The models have a moderate (0.4–0.5) pattern correlation with observations in 1982/1983 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 FV3GFS 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 FV3GFS are closer to observations. However, FV3GFS fails to capture the observed El Niño warming over the Pacific Northwest, with erroneous cold signal for 1982/1983 and 1997/1998 El Niño and weaker warming magnitude for 2015/2016 El Niño. By contrast, GFSv2 can reproduce the observed Pacific Northwest warming during three strong El Niño winters.

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 1982/1983 and 1997/1998 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 2015/2016 winter. It can be seen that the 1997/1998 El Niño has the largest wetness in the southwest.

Compared 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 1982/1983 and 1997/1998 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 2015/2016 El Niño, consistent with previous studies (e.g., Chen & Kumar, 2018; Zhang et al., 2018). Compared to GFSv2, FV3GFS model has an increased (more than double) precipitation pattern correlation (0.32 vs. 0.14) with observation during 2015/2016 El Niño. Generally, the precipitation response in the models has a high pattern correlation with observations during 1997/1998 El Niño. These results indicate that during 2015/2016 the observed rainfall anomalies may have been influenced by the atmospheric internal variability.

To explore the role of atmospheric noise in determining the seasonal mean rainfall over California, Figure 11 presents PDFs of California wintertime rainfall during three strong El Niño events based on FV3GFS AMIP

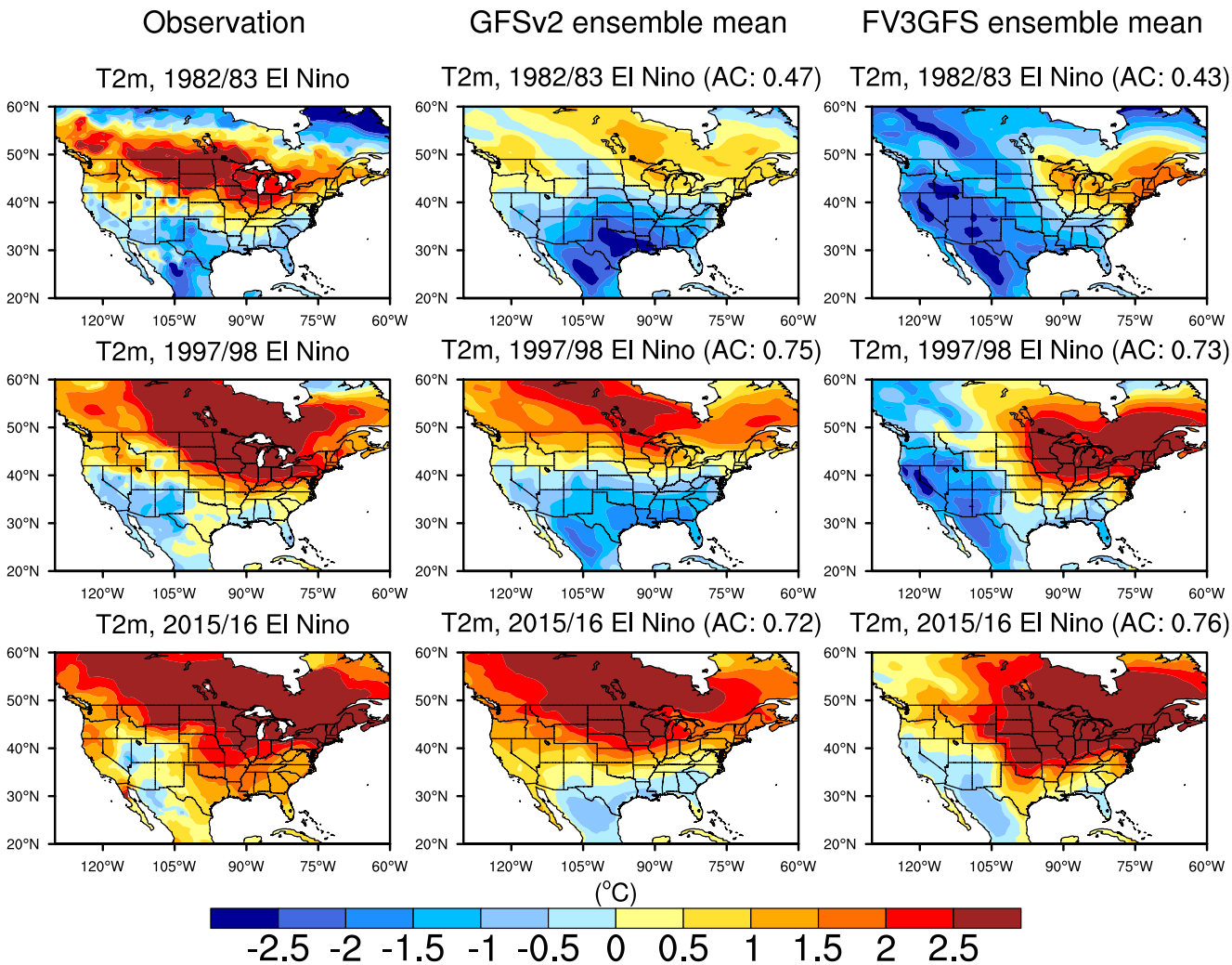

Figure 9. Surface air temperature anomalies for (top) 1982/1983 DJF, (middle) 1997/1998 DJF and (bottom) 2015/2016 DJF from (left) observations, (middle) GFSv2 simulated 30-member ensemble mean, and (right) FV3GFS simulated 100-member ensemble mean. The pattern correlations between models and observations are listed in the titles of the plots.

simulations. The long tick marks show the observed values for the three El Niño winters. The black PDF, drawn from 100-member ensemble FV3GFS AMIP simulations of 2015/2016, is statistically indistinguishable from the blue PDF drawn from 100-member ensemble FV3GFS AMIP simulations of 1982/1983, and for both, the mean of the PDF is shifted to the right. The results indicate that when strong El Niño is present, the most likely California wintertime rainfall condition is one for wetness with a statistical mode of +42% in 2015/2016 runs and +52% in 1982/1983 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 1997/1998 runs is significantly different from the PDFs for 2015/2016 and 1982/1983 runs, with a statistical mode of +70%. This is consistent with the observations for which the strongest California rains are for the 1997/1998 winter among three extreme El Niño events. The observed California 2015/2016 dryness is consistent with the range of unforced variability simulated by the FV3GFS model in this year and is supported by the evidence that the observed value 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 2015/2016 winter precipitation pattern correlation with observation from 100 individual members and make composites for the four

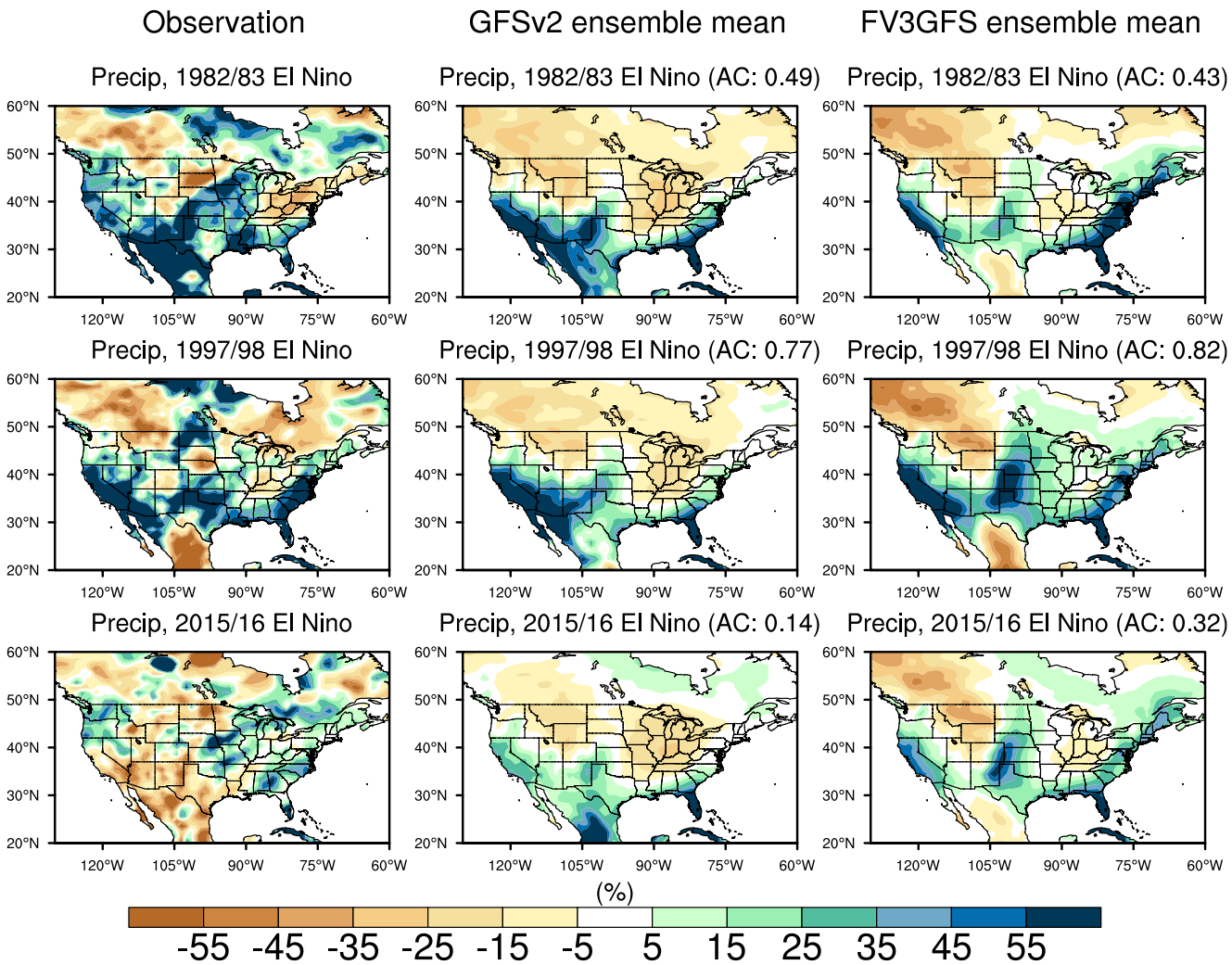

Figure 10. Precipitation anomalies (percent departures) for (top) 1982/1983 DJF, (middle) 1997/1998 DJF and (bottom) 2015/2016 DJF from (left) observations, (middle) GFSv2 simulated 30-member ensemble mean, and (right) FV3GFS simulated 100-member ensemble mean. The pattern correlations between models and observations are listed in the titles of the plots.

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 AC is the largest positive (left panels). For the composite of four runs that have the largest negative AC, 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, instead of a boundary-forced signal, the internal atmospheric variability was the likely cause for the 2016 failed Southern California rains even in the presence of one of the largest El Niño. In summary, the FV3GFS model can realistically capture the observed US climate variability associated with ENSO.

3.5. Assessment of Climate Predictability

Predictability of seasonally averaged atmospheric climate variability depends on the component 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 unforced variability that cannot be predicted in a multi-decadal AMIP run). Extensive efforts have been made in the past several decades to

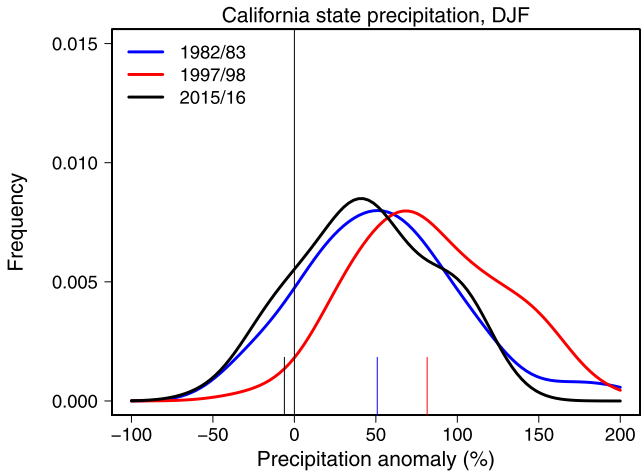

Figure 11. Probability density functions of California state precipitation anomalies (percent departure) for 1982/1983 DJF (blue curve), 1997/1998 DJF (red curve) and 2015/2016 DJF (black curve). Results are based on 100-member FV3GFS Atmospheric Model Intercomparison Project simulations. Large tick marks at the bottom show corresponding observed values.

quantify potential predictability of seasonal mean climate variability by using either AMIP simulations or initialized coupled forecast systems (Jha et al., 2019; Kumar & Hoerling, 1995; Kumar et al., 2007). The purpose of the analysis in this section is to assess the climate predictability based on a large ensemble FV3GFS 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 begin our analysis by comparing the variance of observed and FV3GFS 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 smaller tropical variability and a larger variability in the extratropical regions. The observed maximum centers of variability over Aleutian and Greenland in the NH 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, external (predictable, top left) and internal (unpredictable, bottom left), which are derived from the variance of ensemble mean and the variance of differences in the individual members from the ensemble mean, respectively (Kumar & Hoerling, 1995). The external variance for DJF 200-hPa height simulated by FV3GFS is mainly located in the tropical eastern Pacific, the North Pacific and North American continent, similar to previous findings based on different periods (Jha et al., 2019; Kumar et al., 2007). This is to be expected since the ensemble mean variance is dominated by atmospheric variability forced by SSTs and its spatial structure is in agreement with the atmospheric response to ENSO (Trenberth et al., 1998; see also Figure 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 estimate of the internal variability of observed winter 200-hPa height using multiple models as noted in previous studies (Jha et al., 2019; Kumar et al., 2007).

To assess potential predictability, we have derived the signal-to-noise ratio (SNR) of FV3GFS simulated 200-hPa height, precipitation, and surface air temperature based on the calculation of the ratio of the external and the internal variance (Figure S6 in Supporting Information S1). The largest SNR values for DJF 200-hPa height reside in the tropics and decrease gradually poleward due to an increase in atmospheric internal variability from tropics to extratropics (Figure 13). The summer SNR pattern is very similar to the winter pattern, while the SNR values of heights are somewhat stronger in the tropical Atlantic. Note that there is little consensus on the difference of seasonal predictability of 200-hPa height between winter and summer (Jha et al., 2019; Kumar et al., 2003). SNR pattern for precipitation is also quite similar for two seasons, with larger values located in the tropical oceans, 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 expected value of the skill of seasonal climate predictions, which measures the theoretical limit of predictability, is determined by SNR (Kumar, 2009). Thus, we calculate the square root of the ratio of SNR and SNR+1 as expected value for AC (EAC) to assess the expected skill measure, the results of which are given in Figure 14 that shows FV3GFS simulated winter (left) and summer (right) EAC pattern for 200-hPa height (top), precipitation (middle) and surface air temperature (bottom). By definition, the EAC pattern resembles well the SNR pattern. Consistent with the previous findings that the predictability is larger in the tropics than the extratropics (e.g., Quan et al., 2004), the largest EAC values of DJF 200-hPa height are located in the tropics and decline in the extratropics. Despite appreciable similarity between two seasons, the EAC values of heights are somewhat stronger in the tropical western Pacific and northeastern Pacific and weaker in the tropical Atlantic for the winter compared to the summer. The strong EAC values of precipitation also reside in the tropical ocean, where the expected predictability of precipitation is more evident over the Northern Indian Ocean and tropical

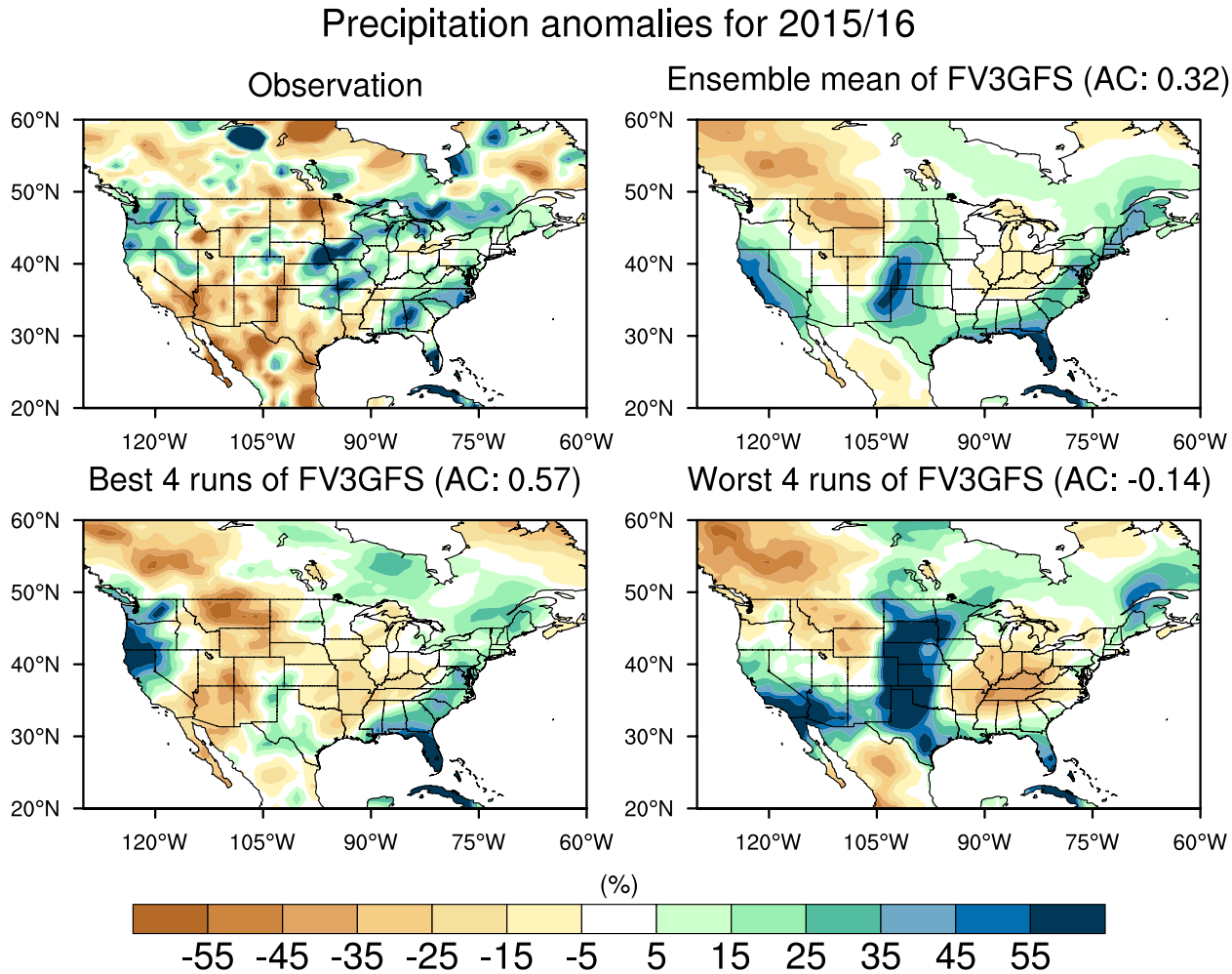

Figure 12. Precipitation anomalies (percent departures) for 2015/2016 DJF from (top left) observations, (top right) FV3GFS simulated 100-member ensemble mean, (bottom left) the composite of four best runs, and (bottom right) the composite of four worst runs among 100-member FV3GFS Atmospheric Model Intercomparison Project simulations. The pattern correlations between the model and observation are listed in the titles of the plots.

Atlantic in the summer relative to the winter. Similar to SNR of surface air temperature, EAC values of temperature are larger over North Africa, the Middle East and Asia in the summer than in the winter.

If the SNR or associated EAC estimates based on the AMIP simulations are correct estimates of corresponding predictability in observations, then generally larger SNR or EAC values imply a higher seasonal prediction skill (Kumar, 2009; Kumar & Hoerling, 2000). To assess this, the corresponding map of AC (Figure 15), the value of which at each location is computed between the observed and the AMIP ensemble mean anomaly over the analysis period, confirms this relationship and is consistent with the theoretical analysis and model results in previous studies (Jha et al., 2019; Kumar et al., 2007). The stronger AC values for 200-hPa height over the tropical Atlantic extend northward in summer compared to winter, in agreement with larger SNR and EAC 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. The trend in the data can contribute in part to the increase of AC values in 200-hPa height and surface air temperature, particularly in the summer (not shown).

A close look of the wintertime US prediction skill reveals that the stronger AC values for precipitation in FV3GFS are located in the northwest, western coast and southern coast, where the 200-hPa height 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 height AC distribution.

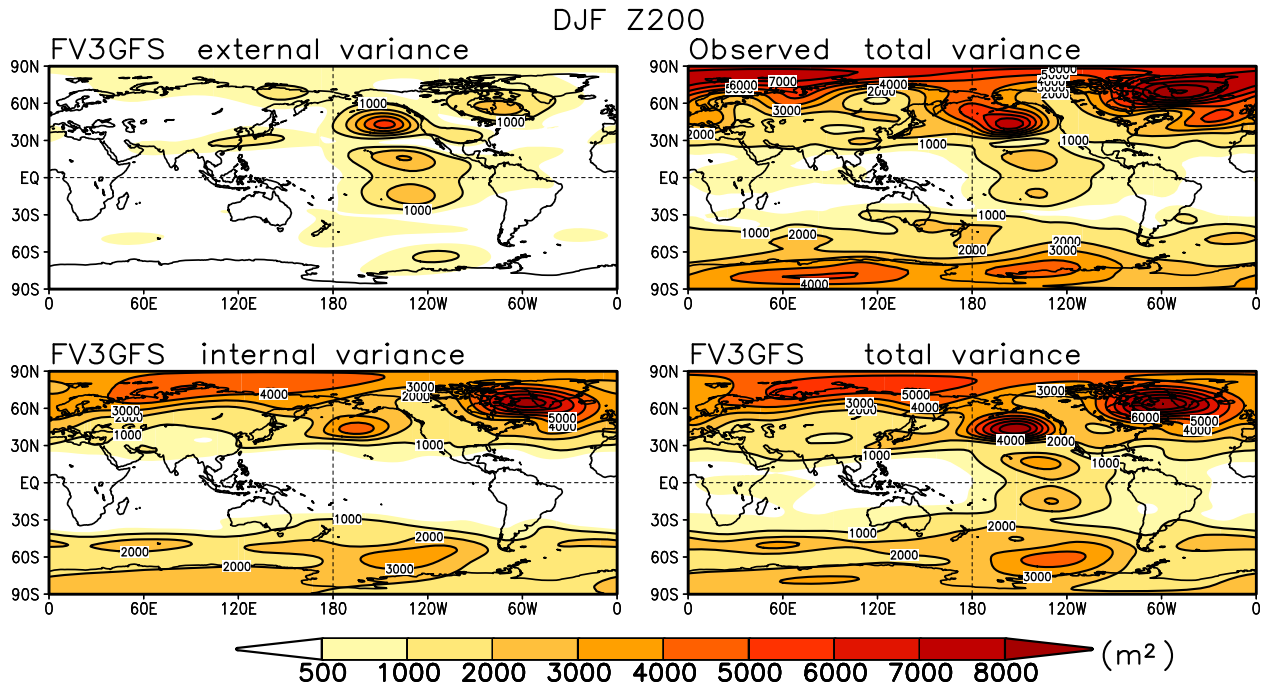

Figure 13. The spatial pattern of FV3GFS simulated (top left) external variance, (bottom left) internal variance, (top right) observed total variance and (bottom right) FV3GFS simulated total variance of DJF 200-hPa height anomaly. Model results are based on 100-member FV3GFS Atmospheric Model Intercomparison Project simulations.

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, FV3GFS shows an increase of precipitation AC values over the north and the southeast, where the increase of 200-hPa height 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.

3.6. 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). We have quantified the realism of simulated historical South Asian Summer Monsoon (SASM) rainfall-ENSO linkage to evaluate the model's ability to reproduce SASM precipitation variability (Figure S7 in Supporting Information S1). Consistent with previous findings (e.g., Zhang et al., 2022), there is a weak negative correlation between observed SASM precipitation anomalies and ENSO. Contrary to observations, the previous version GFSv2 shows a weak positive correlation with ENSO. Clearly, FV3GFS can realistically replicate the historical relationship between SASM precipitation and ENSO. How well does the FV3GFS 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), FV3GFS (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 FV3GFS can reproduce the observed wet condition in South Asia, but the magnitude is somewhat weaker, particularly over Pakistan, where the floods actually occurred. This is to be expected when comparing ensemble mean anomalies with observations that are equivalent to a single model realization. Opposite to observations and

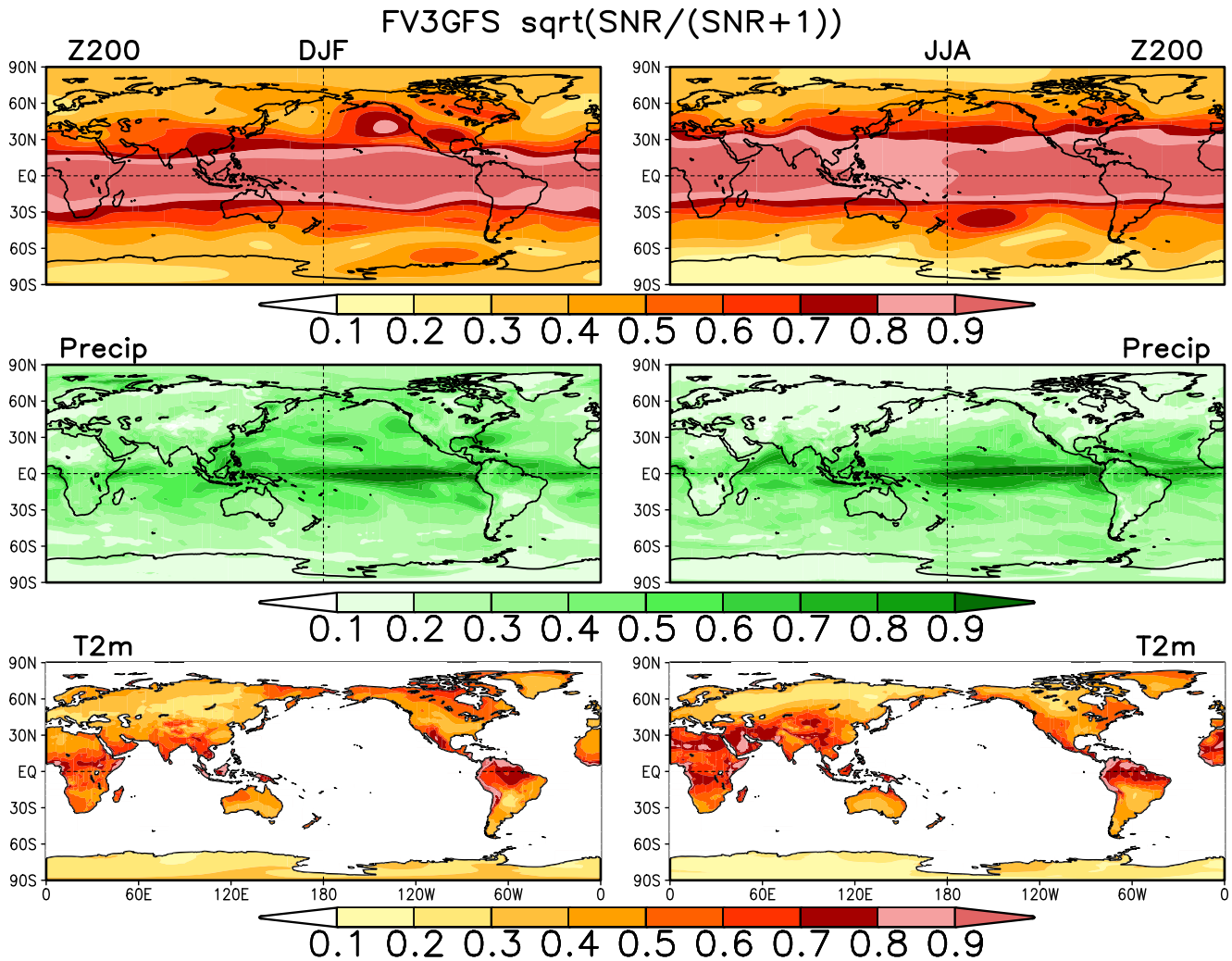

Figure 14. The spatial pattern of (left) DJF and (right) JJA expected skill represented by the square root of signal-to-noise ratio (SNR)/(SNR+1) of (top) 200-hPa height, (middle) precipitation and (bottom) surface air temperature anomalies. The SNR estimate is computed as the ratio of external-to-internal variance in 100-member FV3GFS Atmospheric Model Intercomparison Project simulations.

FV3GFS simulations, the previous version GFSv2 did not replicate the observed wet condition but indicated dry conditions. It is clear that the FV3GFS model has a better simulation of South Asian flooding compared to old version GFSv2.

We also examined the FV3GFS 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 (0.89 mm/day) that is roughly two times as large as the historical ensemble mean JJA precipitation standard deviation (0.43 mm/day) over this region for 1979–2021. The simulated precipitation standard deviation is very close to the observed value of 0.49. Among 100 members, only five members produce the dry condition and many members produce larger wet conditions from normal (greater than 2 times standard deviation). 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 FV3GFS model can serve as a suitable tool to examine the causality of the extreme events.

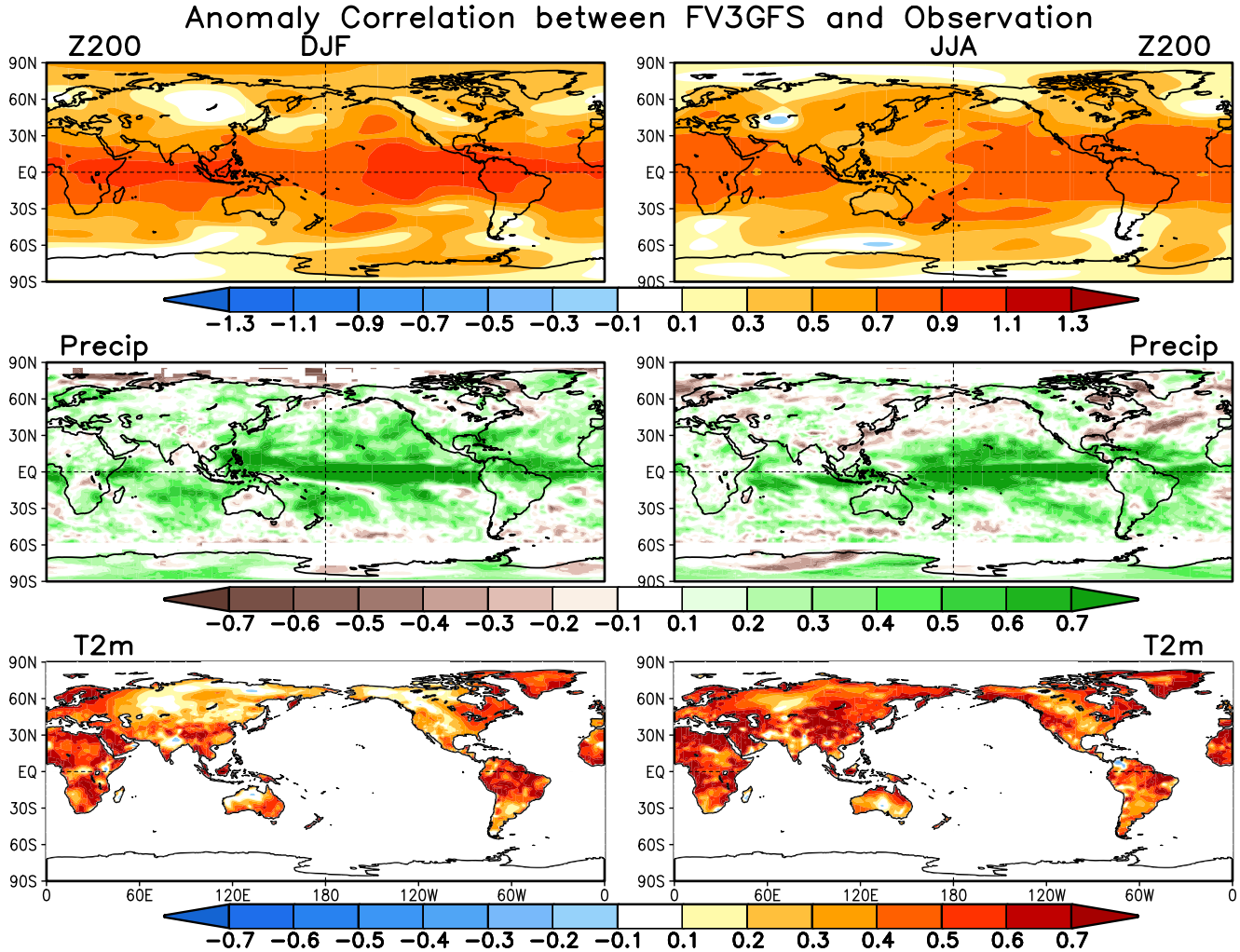

Figure 15. The spatial pattern of (left) DJF and (right) JJA anomaly correlation of (top) 200-hPa height, (middle) precipitation, and (bottom) surface air temperature between observations and FV3GFS simulated 100-member ensemble mean.

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, that is, based on the FV3GFS. 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 FV3GFS 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 NH 200-hPa heights. The observed leading mode describes the extratropical atmospheric circulation pattern associated with AO variability and is found to be independent of tropical SST variability. The second mode of the observed variability describes the canonical atmospheric teleconnection related to ENSO and resembles the TNH pattern and the third mode features a hemisphere-scale increasing trend in heights associated with global warming. The FV3GFS model is able to replicate the three primary modes of

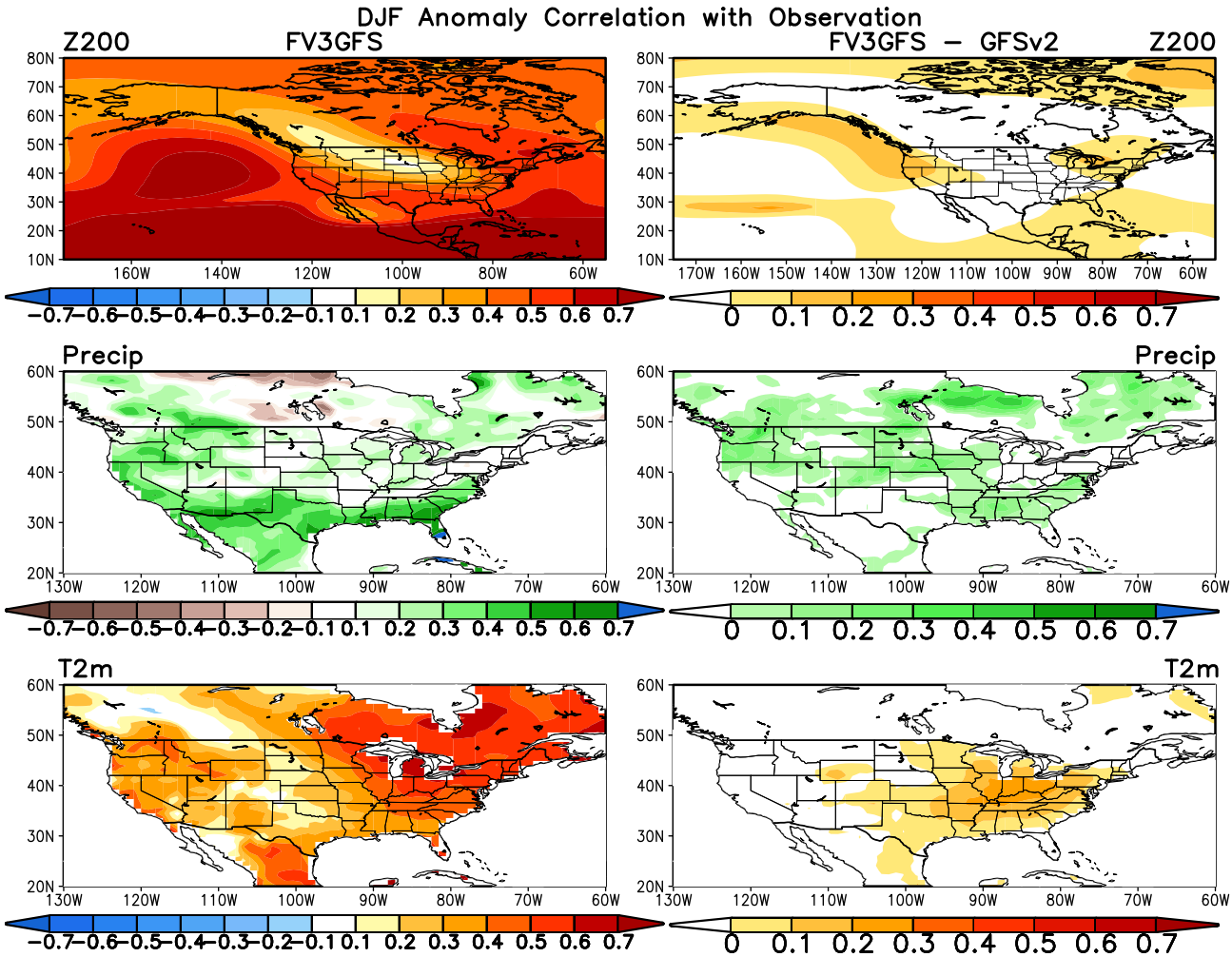

Figure 16. The spatial pattern of DJF anomaly correlation (AC) with observations of (top) 200-hPa height, (middle) precipitation, and (bottom) surface air temperature from (left) FV3GFS Atmospheric Model Intercomparison Project (AMIP) ensemble mean and (right) the difference in AC with observations between FV3GFS and GFSv2 AMIP ensemble mean.

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 for 1979–2021 were analyzed based on 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 analogous to the PNA pattern caused by atmospheric sensitivity to ENSO asymmetry and by sensitivity to a tropical ENSO precursor SST. 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 rely on the analysis of a specific model but are more governed by the nature of boundary forcing used.

FV3GFS 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 (1982/1983) to recent strong El Niño (2015/2016) was evident in both observations and model ensemble mean results. There was an improvement in FV3GFS relative to previous

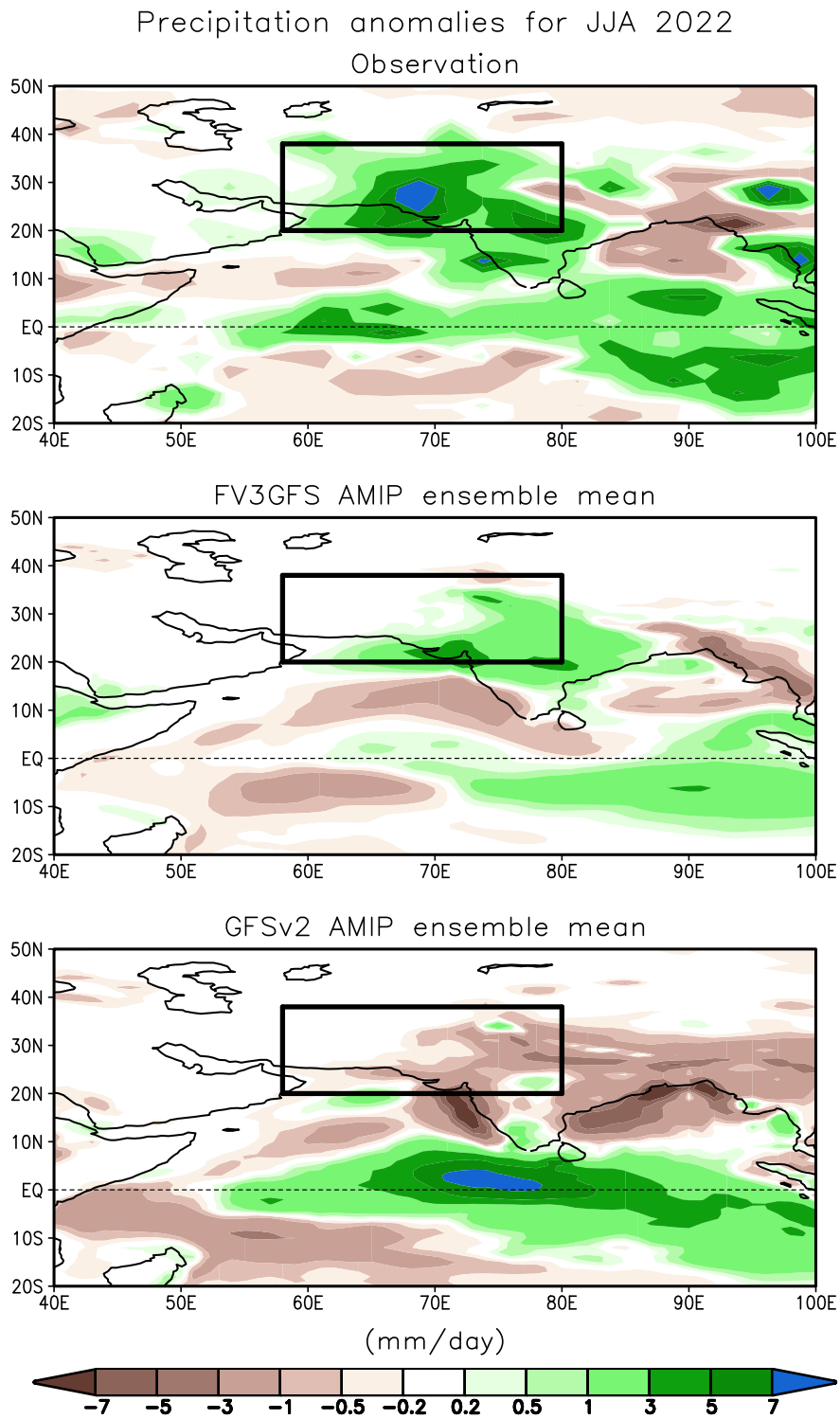

Figure 17. Precipitation anomalies for 2022 JJA from (top) observations, (middle) FV3GFS simulated 100-member ensemble mean, and (bottom) GFSv2 simulated 30-member ensemble mean. The outlined box shows the South Asia region bounded by 58°–80°E, 20°–38°N.

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 1982/1983 and 1997/1998 El Niño events but had dry conditions over Southern California for 2015/2016

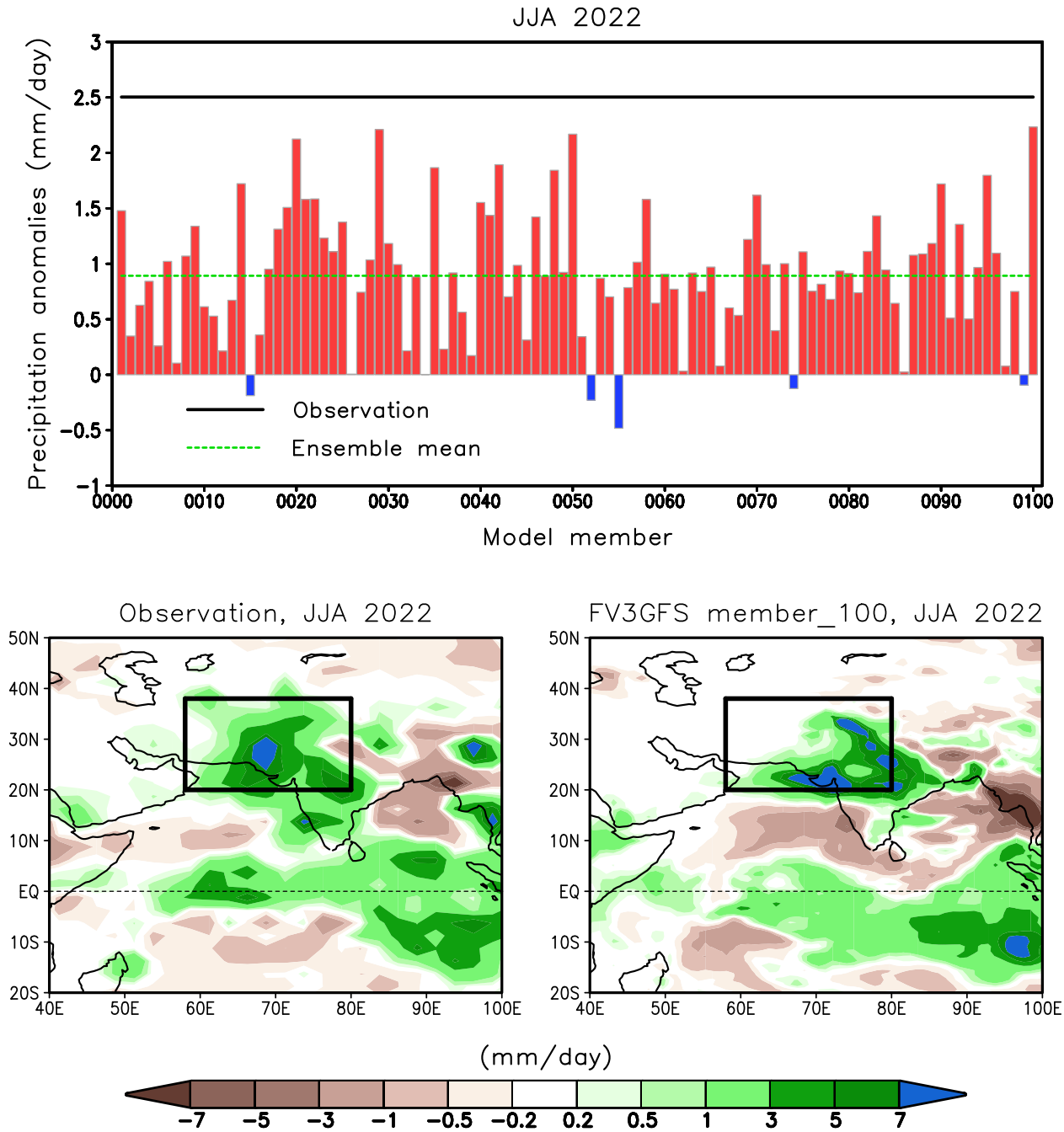

Figure 18. (top) Precipitation anomalies for 2022 JJA averaged over the South Asia region (58°–80°E, 20°–38°N) from observations (black line), FV3GFS simulated 100 individual members (red and blue bars) and 100-member ensemble mean (green line), and (bottom) the comparison of the spatial pattern of precipitation anomalies for 2022 JJA between (left) observations and (right) a single member (100th member) from FV3GFS Atmospheric Model Intercomparison Project simulations. The outlined box shows the South Asia region bounded by 58°–80°E, 20°–38°N.

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 FV3GFS AMIP simulations. The results indicate that the observed California 2015/2016 dryness was likely an articulation of unforced variability (the internal atmospheric variability) and was not a boundary-forced signal, in agreement with previous findings (Kumar & Chen, 2020; Zhang et al., 2018).

The climate predictability measured by SNR and EAC was also assessed based on FV3GFS 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 toward high latitudes. The SNR values for 200-hPa heights and precipitation over the tropical Atlantic and those for surface air temperature over North Africa, Asia, the Middle East, 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 Atlantic and the Indian Ocean, compared to the results in the initialized forecast system (Jha et al., 2019). As the expected skill of seasonal climate prediction is determined by SNR (Kumar, 2009), the EAC pattern mimics the SNR pattern. The seasonal prediction skill measured by AC generally follows the SNR and EAC patterns, supporting previous theoretical analysis and model results (Jha et al., 2019; Kumar et al., 2007). The comparison of the wintertime AC pattern of US surface climate between FV3GFS 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.

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 FV3GFS 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 FV3GFS individual runs further confirmed that the model could replicate the magnitude and spatial pattern of South Asian flooding and indicated that the 2022 summer flooding over that region could have been driven by the observed SST forcing.

In summary, the FV3GFS 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, for example, atmospheric responses to Central Pacific versus Eastern Pacific El Niño 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.

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 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 (Zhang et al., 2023).

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

References

Abid, M. A., Kucharski, F., Molteni, F., Kang, I., Tompkins, A. M., & Almazroui, M. (2021). Separating the Indian and Pacific Ocean impacts on the euro-Atlantic response to ENSO and its transition from early to late winter. *Journal of Climate*, *34*(4), 1531–1548. https://doi.org/10.1175/JCLI-D-20-0075.1

An, S.-I., & Jin, F.-F. (2004). Nonlinearity and asymmetry of ENSO. *Journal of Climate*, *17*(12), 2399–2412. https://doi.org/10.1175/1520-0442(2004)017<2399:naaoe>2.0.co;2

Barnston, A. G., Kumar, A., Goddard, L., & Hoerling, M. P. (2005). Improving seasonal predictions practices through attribution of climate variability. *Bulletin of the American Meteorological Society*, *85*(1), 59–72. https://doi.org/10.1175/bams-86-1-59

Barsugli, J. J., Easterling, D. M., Arndt, D. S., Coates, D. A., Delworth, T. L., Hoerling, M. P., et al. (2022). Development of a rapid response capability to evaluate causes of extreme temperature and drought events in the United States. *Bulletin of the American Meteorological Society*, *103*(3), S14–S20. https://doi.org/10.1175/BAMS-D-21-0237.1

Burgers, G., & Stephenson, D. B. (1999). The "normality" of El Niño. *Geophysical Research Letters*, *26*(8), 1027–1030. https://doi.org/10.1029/1999gl900161

Changnon, S. A. (1999). Impacts of 1997–98 El Niño generated weather in the United States. *Bulletin America Meteorology Social*, *80*(9), 1819–1827. https://doi.org/10.1175/1520-0477(1999)080,1819:IOENOG.2.0.CO;2

Chen, M., & Kumar, A. (2018). Winter 2015/16 atmospheric and terrestrial anomalies over North America: El Nino response and the role of noise. *Monthly Weather Review*, *146*(3), 909–927. https://doi.org/10.1175/mwr-d-17-0116.1

Compo, G., & Sardeshmukh, P. (2009). Oceanic influence on recent continental warming. *Climate Dynamics*, *32*(2–3), 333–342. https://doi.org/10.1007/s00382-008-0448-9

Fahad, A., & Burls, N. (2022). The influence of direct radiative forcing versus indirect sea surface temperature warming in southern hemisphere subtropical anticyclones under global warming. *Climate Dynamics*, *58*(9–10), 2333–2350. https://doi.org/10.1007/s00382-021-06006-1

Fan, Y., & van den Dool, H. (2008). A global monthly land surface air temperature analysis for 1948-present. *Journal of Geophysical Research*, *113*(D1), D01103. https://doi.org/10.1029/2007JD008470

Gates, W. L., Boyle, J., Covey, C., Dease, C., Doutriaux, C., Drach, R., et al. (1998). An overview of the results of the atmospheric model Intercomparison project (AMIP I). *Bulletin America Meteorology Social*, *73*(12), 1962–1970. https://doi.org/10.1175/1520-0477(1992)073<1962:atamip>2.0.co;2

Goddard, L., & Dilley, M. (2005). El Niño: Catastrophe or opportunity? *Journal of Climate*, *18*(5), 651–665. https://doi.org/10.1175/JCLI-3277.1
Han, J., & 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. *Geophysical Research Letters*, *42*(6), 1894–1902. https://doi.org/10.1002/2015GL063083
Held, I. M., & Soden, B. J. (2000). Water vapor feedback and global warming. *Annual Review of Energy and the Environment*, *25*(1), 441–475. https://doi.org/10.1146/annurev.energy.25.1.441
Heo, K. Y., Ha, K. J., Yun, K. S., Lee, S. S., Kim, H. J., & Wang, B. (2014). Methods for uncertainty assessment of climate models and model predictions over East Asia. *International Journal of Climatology*, *34*(2), 377–390. https://doi.org/10.1002/joc.3692
Hoerling, M. P., & Kumar, A. (2002). Atmospheric response patterns associated with tropical forcing. *Journal of Climate*, *15*(16), 2184–2203. https://doi.org/10.1175/1520-0442(2002)015<2184:ARPAWT>2.0.CO;2
Hoerling, M. P., Xu, T., Bates, G., Kumar, A., & Jha, B. (2006). Warm oceans raise land temperatures. *EOS, Transactions American Geophysical Union*, *87*(19), 189–193. https://doi.org/10.1029/2006eo190003
Horel, J. D., & Wallace, J. M. (1981). Planetary-scale atmospheric phenomena associated with the Southern Oscillation. *Monthly Weather Review*, *109*(4), 813–829. https://doi.org/10.1175/1520-0493(1981)109<0813:PSAPAW>2.0.CO;2
Howitt, R. E., Medellin-Azuara, J., MacEwan, D., Lund, J. R., & Summer, D. A. (2014). *Economic analysis of the 2014 drought for California agriculture. Technical report* (p. 20). Center for Watershed Sciences, University of California, Davis. Retrieved from https://watershed.ucdavis.edu/files/content/news/Economic_Impact_of_the_2014_California_Water_Drought.pdf
Hurrell, J., Hack, J., Shea, D., Caron, J., & Rosinski, J. (2008). A new sea surface temperature and sea ice boundary dataset for the Community Atmosphere Model. *Journal of Climate*, *21*(19), 5145–5153. https://doi.org/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), 4147. https://doi.org/10.1029/2008JD009944
Jha, B., Kumar, A., & Hu, Z.-Z. (2019). An update on the estimate of predictability of seasonal mean atmospheric variability using North American Multi-Model Ensemble. *Climate Dynamics*, *53*(12), 7397–7409. https://doi.org/10.1007/s00382-016-3217-1
Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*(3), 437–471. https://doi.org/10.1175/1520-0477(1996)077<0437:tnyrp>2.0.co;2
Kumar, A. (2009). Finite samples and uncertainty estimates for skill measures for seasonal prediction. *Monthly Weather Review*, *137*(8), 2622–2631. https://doi.org/10.1175/2009MWR2814.1
Kumar, A., & Chen, M. (2020). Understanding skill of seasonal mean precipitation prediction over California during boreal winter and role of predictability limits. *Journal of Climate*, *33*(14), 6141–6163. https://doi.org/10.1175/JCLI-D-19-0275.1
Kumar, A., Chen, M., Hoerling, M., & Eischeid, J. (2013). Do extreme climate events require extreme forcings? *Geophysical Research Letters*, *40*(13), 3440–3445. https://doi.org/10.1002/grl.50657
Kumar, A., & Hoerling, M. P. (1995). Prospects and limitations of seasonal atmospheric GCM predictions. *Bulletin America Meteorology Social*, *76*(3), 335–345. https://doi.org/10.1175/1520-0477(1995)076<0335:palosa>2.0.co;2
Kumar, A., & Hoerling, M. P. (1998). Annual cycle of Pacific-North American seasonal predictability associated with different phases of ENSO. *Journal of Climate*, *11*(12), 3295–3308. https://doi.org/10.1175/1520-0442(1998)011<3295:acopna>2.0.co;2
Kumar, A., & Hoerling, M. P. (2000). Analysis of conceptual model of seasonal climate variability and implication for seasonal predictions. *Bulletin America Meteorology Social*, *81*(2), 255–264. https://doi.org/10.1175/1520-0477(2000)081,0255:AOACMO.2.3.CO;2
Kumar, A., Jha, B., Zhang, Q., & Bounoua, L. (2007). A new methodology for estimating the unpredictable component of seasonal atmospheric variability. *Journal of Climate*, *20*(15), 3888–3901. https://doi.org/10.1175/jcli4216.1
Kumar, A., Schubert, S. D., & Suarez, M. S. (2003). Variability and predictability of 200-mb seasonal mean heights during summer and winter. *Journal of Geophysical Research*, *108*(D5), 4169. https://doi.org/10.1029/2002JD002728
Kumar, A., Zhang, Q., Peng, P., & Jha, B. (2005). SST-forced atmospheric variability in an atmospheric general circulation model. *Journal of Climate*, *18*(19), 3953–3967. https://doi.org/10.1175/JCLI3483.1
Mo, K. C., & Livezey, R. E. (1986). Tropical–extratropical geopotential height teleconnections during the Northern Hemisphere winter. *Monthly Weather Review*, *114*(12), 2488–2515. https://doi.org/10.1175/1520-0493(1986)114<2488:TEGHTD>2.0.CO;2
Murray, D., Hoell, A., Hoerling, M., Perlwitz, J., Quan, X.-W., Allured, D., et al. (2020). Facility for weather and climate assessments (FACTS): A community resource for assessing weather and climate variability. *Bulletin America Meteorology Social*, *101*(7), E1214–E1224. https://doi.org/10.1175/BAMS-D-19-0224.1
NOAA. (2017). *Billion dollar weather and climate disasters: Table of events*. NOAA. Retrieved from https://www.ncdc.noaa.gov/billions/events/US/2015
Penland, C., & Sardeshmukh, P. D. (1995). The optimal growth of tropical sea surface temperature anomalies. *Journal of Climate*, *8*, 1999–2024. https://doi.org/10.1175/1520-0442(1995)008<1999:togots>2.0.co;2
Philip, S. Y., Kew, S. F., van Oldenborgh, G. J., Anslow, F. S., Seneviratne, S. I., Vautard, R., et al. (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., & 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., Hoerling, M., Whitaker, J., Bates, G., & Xu, T. (2006). Diagnosing sources of U.S. seasonal forecast skill. *Journal of Climate*, *19*(13), 3279–3293. https://doi.org/10.1175/JCLI3789.1
Quan, X. W., Webster, P. J., Moore, A. M., & Chang, H.-R. (2004). Seasonality in SST forced atmospheric short-term climate predictability. *Journal of Climate*, *17*(16), 3090–3180. https://doi.org/10.1175/1520-0442(2004)017<3090:sisasc>2.0.co;2
Saha, S., Moorthi, S., Pan, H. L., Wu, X., Wang, J., Nadiga, S., et al. (2010). The NCEP climate forecast system reanalysis. *Bulletin America Meteorology Social*, *91*(8), 1015–1057. https://doi.org/10.1175/2010bams3001.1
Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., et al. (2014). The NCEP climate forecast system version 2. *Journal of Climate*, *27*(6), 2185–2208. https://doi.org/10.1175/JCLI-D-12-00823.1
Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang, W., Zhang, Q., et al. (2006). The NCEP climate forecast system. *Journal of Climate*, *19*(15), 3483–3517. https://doi.org/10.1175/jcli3812.1

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., & Rudolf, B. (2014). GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, *115*(1–2), 15–40. https://doi.org/10.1007/s00704-013-0860-x

Seager, R., Hoerling, M., Schubert, S., Wang, H., Lyon, B., Kumar, A., et al. (2015). Causes of the 2011 to 2014 California drought. *Journal of Climate*, *28*(18), 6997–7024. https://doi.org/10.1175/JCLI-D-14-00860.1

Seo, Y.-W., Kim, H., Yun, K.-S., Lee, J.-Y., Ha, K.-J., & Moon, J.-Y. (2014). Future change of extreme temperature climate indices over East Asia with uncertainties estimation in the CMIP5. *Asia-Pacific Journal of Atmospheric Science*, *50*(S1), 609–624. https://doi.org/10.1007/s13143-014-0050-5

Soden, B. J., Jackson, D. L., Ramaswamy, V., Schwarzkopf, M. D., & Huang, X. (2005). The radiative signature of upper tropospheric moistening. *Science*, *310*(5749), 841–844. https://doi.org/10.1126/science.1115602

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000JD900719

Thompson, D. W. J., & Wallace, J. M. (1998). The Arctic Oscillation signature in wintertime geopotential height and temperature fields. *Geophysical Research Letters*, *25*(9), 1297–1300. https://doi.org/10.1029/98gl00950

Trenberth, K. E., Branstator, G. W., Karoly, D., Kumar, A., Lau, N.-C., & Ropelewski, C. (1998). Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *Journal of Geophysical Research*, *103*(C7), 14291–14324. https://doi.org/10.1029/97JC01444

Trenberth, K. E., & Stepaniak, D. P. (2001). Indices of El Niño evolution. *Journal of Climate*, *14*(8), 1697–1701. https://doi.org/10.1175/1520-0442(2001)014<1697:lioeno>2.0.co;2

Weaver, S., Kumar, A., & Chen, M. (2014). Recent increases in extreme temperature occurrence over land. *Geophysical Research Letters*, *41*(13), 4669–4675. https://doi.org/10.1002/2014GL060300

Winton, M. (2000). A reformulated three-layer sea ice model. *Journal of Atmospheric and Oceanic Technology*, *17*(4), 525–531. https://doi.org/10.1175/1520-0426(2000)017<0525:artlsi>2.0.co;2

Xie, P., & Arkin, P. A. (1997). Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin America Meteorology Social*, *78*(11), 2539–2558. https://doi.org/10.1175/1520-0477(1997)078<2539:gpayma>2.0.co;2

Xiong, Y., Ta, Z., Gan, M., Yang, M., Chen, X., Yu, R., et al. (2021). Evaluation of CMIP5 climate models using historical surface air temperatures in central Asia. *Atmosphere*, *12*(3), 308. https://doi.org/10.3390/atmos12030308

Yang, H., & Tung, K. K. (1998). Water vapor, surface temperature and the greenhouse effect—A statistical analysis of tropical-mean data. *Journal of Climate*, *11*(10), 2686–2697. https://doi.org/10.1175/1520-0442(1998)011<2686:wvstat>2.0.co;2

Zhang, T., Hoerling, M. P., Hoell, A., Perlwitz, J., & Eischeid, J. (2020). Confirmation for and predictability of distinct U.S. Impacts of El Niño flavors. *Journal of Climate*, *33*(14), 5971–5991. https://doi.org/10.1175/JCLI-D-19-0802.1

Zhang, T., Hoerling, M. P., Perlwitz, J., & Xu, T. (2016). Forced atmospheric teleconnections during 1979–2014. *Journal of Climate*, *29*(7), 2333–2357. https://doi.org/10.1175/JCLI-D-15-0226.1

Zhang, T., Hoerling, M. P., Wolter, K., Eischeid, J., Cheng, L., Hoell, A., et al. (2018). Predictability and prediction of the southern California rains during strong El Niño events: A focus on the failed 2016 winter rains. *Journal of Climate*, *31*(2), 555–574. https://doi.org/10.1175/JCLI-D-17-0396.1

Zhang, T., Jiang, X., Yang, S., Chen, J., & Li, Z. (2022). A predictable prospect of the South Asian summer monsoon. *Nature Communications*, *13*(1), 7080. https://doi.org/10.1038/s41467-022-34881-7

Zhang, T., & Sun, D.-Z. (2014). ENSO asymmetry in CMIP5 models. *Journal of Climate*, *27*(11), 4070–4093. https://doi.org/10.1175/JCLI-D-13-00454.1

Zhang, T., Sun, D.-Z., Neale, R., & Rasch, P. J. (2009). An evaluation of ENSO asymmetry in the community climate system models: A view from the subsurface. *Journal of Climate*, *22*, 5933–5961. https://doi.org/10.1175/2009jcli2933.1

Zhang, T., Yang, W., Quan, X., Zhu, J., Jha, B., Kumar, A., et al. (2023). A new GFSv15 based climate model large ensemble and its application to understanding climate variability, and predictability [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.8023560

Zhou, L., Lin, S.-J., Chen, J.-H., Harris, L. M., Chen, X., & Rees, S. L. (2019). Toward convective-scale prediction within the next generation global prediction system. *Bulletin America Meteorology Social*, *100*(7), 1225–1243. https://doi.org/10.1175/BAMS-D-17-0246.1