1 **Validation of a transient dynamical downscaling simulation**
2 **approach with multi-model ensemble-based GCM bias corrections**

3

4 Zhongfeng Xu¹, Ying Han¹, Chi-Yung Tam², Zong-Liang Yang³, Congbin Fu⁴

5

6 1. RCE-TEA, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China

7 2. Earth System Science Programme, The Chinese University of Hong Kong, Hong Kong, China

8 3. Department of Geological Sciences, Jackson School of Geosciences, The University of Texas at
9 Austin, Austin, Texas, USA

10 4. School of Atmospheric Sciences, Nanjing University, 210093, Nanjing, China

11

12 Corresponding author: Zhongfeng Xu (xuzhf@tea.ac.cn)

13

14

15

Abstract: Dynamical downscaling is the most widely used physics-based approach to obtaining fine-scale weather and climate information. However, traditional dynamical downscaling approaches are often degraded by biases in the large-scale forcing. To improve the confidence in future projection of regional climate, we used a novel bias-corrected global climate model (GCM) dataset to drive a regional climate model (RCM) over the period for 1980–2014. The dynamical downscaling simulations driven by the original GCM dataset (MPI-ESM1-2-HR model) (hereafter WRF_GCM), the bias-corrected GCM (hereafter WRF_GCMbc) are validated against that driven by the European Centre for Medium-Range Weather Forecasts Reanalysis 5 dataset (hereafter WRF_ERA5), respectively. The results suggest that, compared with the WRF_GCM, the WRF_GCMbc shows a 50–90% reduction in RMSEs of the climatological mean of downscaled variables (e.g. temperature, precipitation, wind, relative humidity). Similarly, the WRF_GCMbc also shows improved performance in simulating the interannual variability of downscaled variables. The RMSEs of interannual variances of downscaled variables are reduced by 30–60%. An EOF analysis suggests that the WRF_GCMbc can successfully reproduce the dominant tri-pole mode in the interannual summer precipitation variations observed over eastern China as opposed to the mono-pole precipitation pattern simulated by the WRF_GCM. Such improvements are primarily caused by the correct simulation of the location of the western North Pacific subtropical high by the WRF_GCMbc due to the GCM bias correction.

Keywords bias correction, multi-model ensemble mean, dynamical downscaling, future projection of regional climate

Declarations

Funding: The study was supported jointly by the National Natural Science Foundation of China (42075170, 41675105), the National Key Research and Development Program of China (2018YFA0606004, 2017YFA0603803), a Chinese University Direct Grant (No. 4053331) and the Jiangsu Collaborative Innovation Center for Climate Change.

Conflicts of interest/Competing interests: The authors declare no conflicts of interest.

Availability of data and material: The data used in this study is available for open access. The bias corrected GCM dataset is available at ScienceDB <http://www.dx.doi.org/10.11922/sciencedb.00487>. The CMIP6 datasets are available at <https://esgf-data.dkrz.de/projects/esgf-dkrz/>. The ERA5 reanalysis dataset was provided by the European Centre for Medium-Range Weather Forecasts (<https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>)

Code availability: Not applicable

Authors' contributions: ZX designed the study and carried out the WRF simulations. ZX and YH performed the analysis and validation of dynamical downscaling simulations. All co-authors contributed to writing and commented on the manuscript.

1 Introduction

Dynamical downscaling is an important approach to generate a full set of physically consistent high-resolution climate information, which is crucial for climate-related studies, e.g. climate extremes, air quality, hydrology, agriculture, and wind power. The traditional dynamical downscaling (TDD) approach employs a regional climate model (RCM) nested within a global climate model (GCM) to generate high-resolution climate information (Giorgi and Mearns 1991). The TDD approach has been widely used in regional climate-related studies (e.g., Jacob and Podzun 1997; Leung et al. 2003; Gao et al. 2011; Tang et al. 2016). The TDD approach has also been used in previous and ongoing regional climate assessment projects such as the Regional Climate Model Intercomparison Project for Asia (RMIP; Fu et al. 2005), the North American Regional Climate Change Assessment Program (NARCCAP) (Mearns et al. 2009), and the Coordinated Regional Climate Downscaling Experiment (CORDEX) (Gutowski et al. 2016). The ensemble simulations generated through the combination of multi-GCM and multi-RCM can provide us a good opportunity to measure the uncertainty of regional climate projections. However, the ensemble simulations require enormous computing resources. The significant biases in GCM and RCM models could also lead to an overestimation of the uncertainty in future projections.

In previous studies using the TDD approach, GCM biases were passed into RCMs through the lateral boundary conditions (LBC), which degraded the RCM simulations and causing large uncertainties in future projections (e.g., Wu and Lynch 2000; Sato et al. 2007; Xu and Yang 2012; Ji et al. 2019). To remove the GCM biases, some studies employed a “pseudo global warming” method in which the LBCs was constructed by adding GCM-projected mean climate changes onto reanalysis dataset (e.g., Wu and Lynch 2000; Sato et al. 2007; Kawase et al. 2009; Chen et al. 2020). Over the past decade, many dynamical downscaling approaches were developed by directly correcting GCM biases, which proved to be effective in improving the dynamical downscaling simulations (Xu et al. 2019; Adachi et al. 2020). Previous GCM bias correction methods primarily include mean bias correction (Holland et al.

2010), mean and variance bias corrections (Xu and Yang 2012, 2015), quantile-quantile correction (Colette et al. 2012), nested bias correction (Rocheta et al. 2017), trend-preserving bias correction (Hoffmann et al. 2016), and multi-model ensemble (MME) bias correction (Dai et al. 2020). These studies suggested that GCM bias correction can significantly improve the dynamical downscaling simulation. For example, RCM can reasonably reproduce the Atlantic tropical cyclones simulation after considering the GCM mean bias correction (Done et al. 2013; Bruyère et al. 2014). The GCM mean and variance bias correction can further improve the precipitation and temperature extreme compared with GCM mean bias correction (Xu and Yang 2012). The MME-based bias correction is expected to generate a more reliable projection of future climate background relative to the bias correction of a single GCM (Dai et al. 2020).

However, previous GCM bias correction methods also have limitations. For instance, the GCM variance bias correction may modify GCM long-term trends inappropriately. To overcome this limitation, Hoffmann et al. (2016) developed a trend-reserving bias correction method through correcting the variance bias of detrended time series. The quantile-quantile correction failed to preserve the intervariable dependencies and introduced spurious precipitation variability (White and Toumi 2013). The MME-based bias correction only corrected GCM mean bias and left other biases untouched (Dai et al. 2020).

With the advantages and limitations of previous GCM bias correction methods in mind, we recently developed a GCM bias correction method (hereafter termed MVT), which corrected GCM mean, variance, and long-term trend biases (Xu et al. 2021). The MVT bias correction integrates the advantage of previous GCM bias correction methods and has the potential to generate more reliable climate projections. In this paper, we aim to evaluate the performance of the dynamical downscaling approach with MVT bias correction in simulating climatological mean and interannual variability. We briefly introduce the dataset, the MVT bias correction method, and experimental design in section 2. Section 3 presents the validation result. Discussion and conclusion are given in section 4.

2 Data, methods and experimental design

2.1 Data

The dataset used include the European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5) dataset, historical simulation derived from the MPI-M earth system model (MPI-ESM1-2-HR), and a bias-corrected GCM dataset from 1980 to 2014. The ERA5 was produced by the European Centre for Medium-Range Weather Forecasts Reanalysis using 4D-var data assimilation (Hersbach et al. 2020). ERA5 dataset assimilated more observations and had a higher grid spacing of 31 km compared with ERA-Interim. ERA5 shows considerable improvement in tropospheric temperature, wind, and humidity, but not in the stratosphere relative to the ERA-Interim (Hersbach et al. 2020). The historical simulation generated by the MPI-ESM1-2-HR was also used (Müller et al. 2018). MPI-ESM1-2-HR is one of GCMs in the Coupled Model Intercomparison Project Phase 6 (CMIP6). MPI-ESM1-2-HR and its predecessor, MPI-ESM-MR, show generally good performance in the simulation of SST and atmospheric circulation among the CMIP5 and CMIP6 models (Huang et al. 2019; Zhang et al. 2021). Both the ERA5 and MPI-ESM1-2-HR dataset were regridded to a horizontal grid spacing of 1.25° longitude $\times$ 1.25° latitude by bilinear interpolation.

We also used a bias-corrected GCM dataset generated in our previous study (Xu et al. 2021), which was constructed using historical simulations of 18 CMIP6 models (Eyring 2016) and ERA5 over the period for 1979–2014. The GCM mean and variance biases were corrected based on the ERA5. Moreover, the individual GCM (MPI-ESM1-2-HR) long-term non-linear trend was replaced with that derived from multi-model ensemble mean (MME) to reduce the uncertainty of future projection. Consequently, the bias-corrected data showed a better quality than any individual CMIP6 models in terms of the climatological mean, interannual variance, and climate extremes. The detailed validation of the bias-corrected GCM data and related bias correction method can be found in Xu et al. (2021). We briefly interpreted the bias correction method in the following section for the readers' convenience.

2.2 GCM bias correction method

The GCM and ERA5 data can be broken down into a long-term non-linear trend plus an interannual perturbation term:

$$\text{GCM} = \text{GCM}_{LT} + \text{GCM}' \quad (1)$$

$$\text{ERA} = \text{ERA}_{LT} + \text{ERA}' \quad (2)$$

The bias-corrected GCM data (GCM_{mvt}) was constructed as follows:

$$\text{GCM}_{mvt} = \text{MME}_{LT} - (\overline{\text{MME}_{LT|H}} - \overline{\text{ERA}_{LT|H}}) + \text{GCM}' \times r_s \quad (3)$$

where MME_{LT} is the long-term non-linear trend of MME. $\overline{\text{MME}_{LT|H}}$ and $\overline{\text{ERA}_{LT|H}}$ are the climatological mean of the long-term trend derived from the MME and ERA5 data over the historical period (1979–2014), respectively. The second term on the right-hand side of Eq. (3) represents the mean bias of GCM and we assumed it remains the same in the historical and future periods. In the third term on the right-hand side of Eq. (1), the GCM variance bias was corrected by multiplying the perturbation term with a scaling factor, $r_s = \sigma_{\text{ERA}}/\sigma_{\text{GCM}}$. Here, r_s is defined by the ratio of the standard deviation of detrended reanalysis data to that of detrended GCM data over the historical period. We also assumed that the ratio remains the same in the historical and future periods. Eq. (3) can be rearranged as:

$$\text{GCM}_{mvt} = \overline{\text{ERA}_{LT|H}} + (\text{MME}_{LT} - \overline{\text{MME}_{LT|H}}) + \text{GCM}' \times r_s \quad (4)$$

Thus, the bias-corrected GCM data has a reanalysis-based mean climate ($\overline{\text{ERA}_{LT|H}}$), a long-term non-linear trend of MME ($\text{MME}_{LT} - \overline{\text{MME}_{LT|H}}$), and a bias-corrected interannual variance derived from individual GCM ($\text{GCM}' \times r_s$). We used 18 CMIP6 models to compute the long-term trend of MME, which is expected to be more reliable than the individual GCM projected trend. The GCM' was computed using MPI-ESM1-2-HR model output. Compared with the previous MME-based bias correction method (Dai et al. 2020), our method is particularly suitable for dynamical downscaling simulations of transient climate.

2.3 Model and experimental setup

We used the Weather Research and Forecasting (WRF) model with the Advanced Research WRF (ARW) dynamical core version 4.2 (Skamarock et al. 2008).

WRF model is a non-hydrostatic model developed and maintained by NCAR. The model domain is centered at 28°N, 118°E with 469×319 horizontal grid points and 33 vertical levels (Fig.1). Based on the evaluation of multiple parameterization schemes over East Asian domain, the WRF model was configured with the Kain-Fritsch convective parameterization scheme (Kain 2004), RRTMG shortwave and longwave radiation schemes (Iacono et al. 2008), Lin microphysics scheme (Lin et al. 1983), the Noah-MP land surface model (Niu et al. 2011; Yang et al. 2011), and the Yonsei University planetary boundary layer scheme (Hong et al. 2006).

Three 25-km WRF simulations were carried out throughout 1980–2014 with the MPI-ESM1-2-HR (WRF_GCM), bias-corrected GCM data (WRF_GCMbc), and the ERA5 (WRF_ERA5) as large-scale forcing data, respectively. The first two years were discarded for spin-up.

2.4 Validation metrics

It is known that GCMs and RCMs contain systematic biases and both biases can degrade the dynamical downscaling simulations (Xu and Yang 2015). In this study, we only focus on the impact of GCM biases on dynamical downscaling simulations. We compared the WRF_GCM and WRF_GCMbc against the WRF_ERA5 to examine the impact of GCM bias corrections and exclude impacts of the RCM biases on dynamical downscaling simulations. Assuming the ERA5 can provide perfect initial condition and LBCs, the difference between the WRF_GCM (WRF_GCMbc) and WRF_ERA5 represents the impacts of biases from the SST and LBCs.

We measure the amplitude of interannual variation of a scalar variable (M) using standard deviation:

$$S = \sqrt{\frac{1}{N} \sum_{i=1}^N (M_i - \bar{M})^2} \quad (5)$$

where N is the number of years. $\bar{M}$ is the climatological mean of M_i . Similarly, we can define the standard deviation of a vector variable (Xu et al. 2016):

$$S_v = \sqrt{\frac{1}{N} \sum_{i=1}^N ((u_i - \bar{u})^2 + (v_i - \bar{v})^2)} \quad (6)$$

$\bar{u}$ and $\bar{v}$ represent the climatological mean of zonal and meridional components of vector wind, respectively. S_v takes both the vector length and vector direction into account, which can measure the amplitude of interannual variation of vector wind.

To validate the WRF simulations we used mean error (ME), root-mean-square error (RMSE), and root-mean-square vector error (RMSVE):

$$ME = \frac{1}{N} \sum_{i=1}^N M_i - O_i \quad (7)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (M_i - O_i)^2} \quad (8)$$

$$RMSVE = \sqrt{\frac{1}{N} \sum_{i=1}^N (M_{u,i} - O_{u,i})^2 + (M_{v,i} - O_{v,i})^2} \quad (9)$$

where M and O are scalar variables derived from the GCM-driven WRF simulation and the ERA5-driven WRF simulation, respectively. N is the sample size. RMSVE is the same as RMSE except for vector fields. The subscripts u and v denote the zonal and meridional components of vector wind, respectively. ME measures the mean difference between two simulations. RMSE and RMSVE measure the overall difference between two simulations in terms of scalar variable and vector variable, respectively. RMSE (RMSVE) summarizes multiple statistics of a climate model because it is a function of the mean error, pattern correlation, and variance (Xu et al. 2016; Xu and Han 2019). These statistical metrics were computed within the verification zone as shown in Fig.1 using the 33-year WRF simulations from 1982 to 2014.

3 Results

3.1 Climatological mean

3.1.1 Surface variables

To examine the impact of GCM bias corrections on the dynamical downscaling simulation, Figure 2 illustrates the MEs of climatological mean precipitation and the surface air temperature (SAT) derived from the WRF_GCM and WRF_GCMbc, respectively. The WRF_GCM significantly overestimates (underestimates) precipitation and SAT over the western North Pacific to the south (north) of 20°N (Figs. 2a, 2c). The error of SST in the large-scale forcing dataset shows a similar

spatial pattern as that in the WRF_GCM, which indicates that the errors in precipitation and SAT primarily result from the SST bias over the western North Pacific (Figure not shown). A warmer SST leads to higher SAT and more precipitation and vice versa. In contrast, the WRF_GCM significantly underestimates the precipitation but overestimates the SAT over the Indian and Indochina peninsula (Figs. 2a, 2c). The warm bias in SAT is closely related to the dry bias in precipitation over the land area. The decrease in precipitation is accompanied by more downward solar radiation and reduces evapotranspiration from the land surface, which in turn leads to a higher SAT. The errors in downscaled precipitation and SAT are related to the errors in upper-air variables, which will be further addressed in the next section. In the WRF_GCMbc, the errors in precipitation and SAT are greatly removed relative to that in the WRF_GCM due to the GCM bias correction (Fig. 4). The RMSE of climatological mean precipitation (SAT) is reduced from 3.37 mm d⁻¹ (0.74°C) in the WRF_GCM to 0.95 mm d⁻¹ (0.19°C) in the WRF_GCMbc (Table 1).

In addition to the precipitation and SAT, other surface variables are also evaluated and summarized in Table 1. Compared with the WRF_CAM, the WRF_CAMbc shows much smaller RMSEs in terms of the climatological mean SAT, 2-m specific humidity, precipitation, snow depth, sensible heat flux, latent heat flux, planetary boundary height, and 10-m wind speed. Generally, the RMSEs of the surface variables are reduced by 50–80%. The improvement appears to be more significant in summer than in winter for SAT, Snowh, SH, LH, PLBH, and Wind10, which is generally consistent with the previous study (Xu and Yang 2012). For example, the RMSE of SAT is reduced from 0.74 in the WRF_GCM to 0.19 in the WRF_GCMbc in summer. In contrast, the RMSE of SAT is reduced from 0.86 in the WRF_GCM to 0.41 in the WRF_GCMbc in winter. The errors in SAT primarily occur over mid-high latitudes and the Tibetan plateau in winter, which is likely related to the errors in snow simulation.

3.1.2 Upper air variables

We also computed the RMSEs of upper-air temperature, geopotential height, vector wind, and relative humidity using the outputs of the WRF_GCM and

WRF_GCMbc, respectively. The WRF_GCM shows the largest RMSEs in the air temperature, geopotential height, and vector wind in the upper troposphere. Whereas, the largest RMSE of relative humidity occurs in the mid-troposphere (Fig. 3). Compared with the WRF_GCM, the WRF_GCMbc show much smaller RMSEs of upper-air variables in all vertical levels in summer (Fig. 3). Generally, the RMSEs of downscaled upper-air variables are reduced by approximately 70–90% due to the GCM bias corrections. Similarly, the WRF_GCMbc also shows very significant improvement in the simulation of upper-air variables in spring, autumn, and winter (Figure not shown).

Figure 4 illustrates the spatial pattern of climatological mean errors in 850-hPa geopotential height, vector wind, and relative humidity in summer. The WRF_GCM significantly underestimates the geopotential height over the western North Pacific between 5°N and 30°N but overestimates it over the Asian continent and western North Pacific to the north of 30°N. In association with the geopotential height errors, the 850-hPa vector wind shows a significant cyclonic error in the western North Pacific between 5°N and 30°N and an anticyclonic error around the Tibetan Plateau (Fig. 4a). The anticyclonic error in 850-hPa vector wind and the negative error in relative humidity lead to an underestimation of precipitation and overestimation of SAT over India and Bay of Bengal (Figs. 2a, 2c, 4a, 4c). Similarly, the cyclonic error and positive error in relative humidity tend to generate more precipitation over the western North Pacific in the WRF_GCM (Figs. 2a, 4a, 4c). The WRF_GCM shows very similar spatial patterns of the errors in 850-hPa geopotential height and vector wind as that in the large-scale forcing dataset (i.e., MPI-ESM1-2-HR) (Figure not shown). This indicates that the errors in the downscaled 850-hPa circulation are largely inherited from the large-scale forcing data. In terms of the simulation of 850-hPa relative humidity, the WRF_GCM also shows significant errors characterized by a significant overestimation over the western North Pacific and equatorial Indian Ocean and a significant underestimation over India and Bay of Bengal due to the error in SST (Fig. 4c). The errors in circulation, relative humidity, precipitation, and SAT

are also largely removed in the WRF_GCMbc experiment due to the GCM bias corrections (Figs. 2, 4).

Note that the GCM bias correction leads to a more significant improvement in the dynamical downscaling simulations compared with that presented in [Xu and Yang \(2012\)](#) especially in the lower troposphere. In the previous study, the dynamical downscaling simulations were forced by observed SST data and the GCM bias corrections were only applied to the upper-air variables ([Xu and Yang 2012](#)). In contrast, we simultaneously corrected the upper-air variable biases and SST biases in this study. The additional bias corrections to SST lead to a greater improvement in the dynamical downscaling simulations, especially for the lower tropospheric variables.

3.2 Interannual variability

3.2.1 Surface variables

In addition to the climatological means, interannual variability is also of great importance in the future projection of regional climate. An enhanced interannual variability can often cause more damages to agriculture, economy and affects food and water security. Here, the interannual variability is measured by the standard deviation of a scalar or vector variable (Eqs. 5, 6). Compared with the WRF_ERA5, the WRF_CAM significantly underestimates the interannual variability of precipitation over India, the Bay of Bengal, and the South China Sea and overestimates the precipitation variability over the western North Pacific to the south of 20°N. The error in precipitation variability shows a similar spatial pattern as that in climatological mean precipitation. In the WRF_GCMbc, the errors in precipitation variability were significantly reduced (Figs. 5a, 5b). The RMSE of summer precipitation variability is reduced from 1.95 mm d⁻¹ in the WRF_CAM to 0.93 mm d⁻¹ in the WRF_GCMbc (Table 1). The WRF_GCM underestimates the interannual variance of summer SAT in northwest India and eastern China but overestimates it over the Bay of Bengal, Indochina peninsula, and northwest Pacific. These errors in the interannual variance of downscaled SAT are also greatly removed in the WRF_GCMbc due to the GCM bias correction.

Similar to precipitation and SAT, other surface variables also show significantly improved interannual variability in the WRF_GCMbc relative to the WRF_GCM. The RMSEs in the interannual standard deviation of various variables were reduced by approximately 30–60% due to the GCM bias corrections. Among these surface variables, 2-m specific humidity shows relatively greater improvement characterized by a greater reduction in RMSE. Overall, the improvement in the downscaled interannual variability is relatively small compared with that in the climatological means (Table 1).

3.2.2 Upper air variables

The comparison of the WRF_GCMbc with the WRF_GCM suggests that the GCM bias correction significantly improves the interannual variability of upper air temperature, geopotential height, vector wind, and relative humidity at all vertical levels (Fig. 6). The RMSEs in the interannual variability are reduced by 20–70%. Similarly, the WRF_GCMbc also shows significant improvement in the interannual variability of upper-air variables in spring, autumn, and winter (Figure not shown). Note that the improvements in the interannual variability of upper-air variables are smaller than that in climatological means (Figs. 3, 6). This indicates that it is more challenging to improve the downscaled interannual variability than the climatological means through correcting GCM biases.

The WRF_GCM significantly underestimates the interannual variability of 850-hPa geopotential height over the Eastern China-East China Sea and overestimates the variability over the maritime continent and western North Pacific to the east of 140°E (Fig. 7a). This suggests that the western North Pacific subtropical high (WNPSH) shows a weaker interannual variability over East Asia to the south of 40°N in the WRF_GCM than the WRF_ERA5. In this region, the 850-hPa vector wind also shows a weaker interannual variability over eastern China and the South China Sea, which likely accounts for the underestimated summer precipitation variability in the WRF_GCM (Figs. 5a, 7c). Conversely, the WRF_GCM overestimates the vector wind variability over the western North Pacific to the south of 20°N where also shows an overestimated precipitation variability (Figs. 5a, 7c). It appears that the 850-hPa

vector wind variability can partly account for the precipitation variability. In the WRF_GCMbc, the errors in geopotential height and vector wind are largely removed due to the GCM bias corrections. Consequently, the interannual variability of downscaled precipitation is also improved in the WRF_GCMbc.

3.3 Leading EOF mode of summer precipitation over Eastern China

It's known that the leading Empirical Orthogonal Function (EOF) of observed summer precipitation shows a tri-pole mode characterized by a dry-wet-dry pattern over the North China-Yangtze River Basin-South China at interannual and interdecadal time scales (Chang et al. 2000; Ding et al. 2008). The WRF_ERA5 reasonably reproduces the tri-pole pattern (-,+, -) of observed summer precipitation anomaly over eastern China. The leading EOF shows a positive precipitation anomaly over the Yangtze River basin and a negative precipitation anomaly over North China and South China (Fig. 8a). The first leading mode explains 20.5% of the total variance. However, the WRF_GCM fails to reproduce the spatial pattern of leading EOF of observed summer precipitation anomaly and shows a monopole pattern over eastern China, instead (Fig. 8b). This indicates that summer precipitation varies homogeneously over eastern China in the WRF_GCM. In contrast, the WRF_GCMbc successfully reproduces the tri-pole pattern of summer precipitation anomaly due to the GCM bias corrections (Fig. 8c). The first leading EOF explains 25.8% of the total variance in the WRF_GCMbc. In addition to the first EOF mode, the second and third EOF modes in the WRF_GCMbc are also consistent with that in the WRF_ERA5 characterized by higher pattern correlation than in the WRF_GCM (Table 2). The first three EOF modes explain 52.3%, 43.1%, and 51.4% of the total variance in the WRF_ERA5, WRF_GCM, and WRF_GCMbc, respectively.

Note that the WRF_GCMbc also well reproduces the decline trend of the time series of the leading EOF, which is likely due to a correct response of precipitation to the external forcing (Fig. 8). However, one should not expect the WRF_GCMbc can reproduce the year-to-year variation as that in the WRF_ERA5. The reason is that the bias-corrected GCM data inherits the temporal variation of GCM (Xu et al. 2021).

The GCM is a land-atmosphere-ocean coupled model and its temporal variation is largely an unforced variability varying independent from that in ERA5. Likewise, the WRF_GCMbc cannot reproduce the year-to-year variation in the WRF_ERA5, either, because RCM largely inherit the temporal variation from the large-scale forcing data.

As discussed in section 3.2 the interannual variability of precipitation is related to that of atmospheric circulation. Therefore, we computed the difference of the 850-hPa geopotential height and vector wind between the positive- and negative-phase years of the leading EOF time series. The positive (negative) phase years are defined as the time series of EOF greater (smaller) than 0.5 (−0.5) standard deviation. In the WRF_ERA5 and WRF_GCM, the positive phase of lead EOF is corresponding to a wet anomaly in the Yangtze River basin and dry anomalies in North and South China, and vice versa (Figs. 8a, 8c). In the positive-phase years, the 850-hPa geopotential height shows a positive (negative) anomaly to the south (north) of 30°N indicating a southward (northward) shift of the WNPSH in the WRF_ERA5 (Fig. 9a). In association with the WNPSH anomaly, 850-hPa circulation shows an anticyclonic (cyclonic) anomaly to the south (north) of 30°N in the positive-phase years. Thus, the anticyclonic anomaly results in a dry anomaly over South China. The Yangtze River basin is dominated by a southwesterly anomaly, which brings abundant moisture and leads to a wet anomaly there. The southward shift of WNPSH hinders the northward transfer of moisture and results in a dry anomaly over North China (Fig. 9a). The spatial patterns of the 850-hPa geopotential height and vector wind anomalies in the WRF_GCMbc are very similar to those in the WRF_ERA5 (Figs. 9a, 9c). Consequently, the WRF_GCMbc well reproduce the leading EOF of the summer precipitation anomaly. In contrast, the WRF_GCM generates a positive anomaly in the 850-hPa geopotential height over East Asia between 20°N and 40°N in the positive-phase years relative to the negative-phase years (Fig. 9b). In turn, the enhanced WNPSH causes a dry anomaly over eastern China and vice versa. Clearly, the WRF_GCMbc can well reproduce the leading EOF mode of summer precipitation anomaly over eastern China and the corresponding circulation patterns relative to the WRF_GCM.

As the WNPSH plays a crucial role in modulating summer precipitation patterns over eastern China, Figure 10 further illustrates the 1500-gpm and 1450-gpm contours of the summer WNPSH in the WRF_ERA5, WRF_GCM, and WRF_GCMbc during 1982–2014. The 1500-gpm contour can indicate the location of the WNPSH. Based on the geostrophic balance relationship, the 1450-gpm contour can roughly represent the prevailing wind over eastern China. In the WRF_ERA5, the WNPSH extends southwestward in the positive-phase years and results in a positive (negative) geopotential height anomaly to the south (north) of 30°N (Figs. 9a, 10a). The prevailing wind flows from the Indo-China peninsula to the East China Sea as indicated by the 1450-gpm contour in the positive-phase years, which generates more precipitation over the Yangtze River basin and less precipitation over South China and North China (Figs. 8a, 9a, red lines in 10d). In the negative years, the WNPSH extends northward and the 1450-gpm contour shows a south-north orientation, which brings more moisture and precipitation to North China (blue lines in Fig. 10b). The WRF_GCM shows a clear error in the ridge of the WNPSH characterized by a roughly 5-degree northward migration relative to that in the WRF_ERA5 (Figs. 10a, 10b). The 1450-gpm contour extends westward in the positive-phase year relative to the negative-phase years over eastern China between 20°N and 40°N and leads to a monopole pattern of summer precipitation anomaly (Figs. 8b, 10e). In the WRF_GCMbc, the error in the ridge of WNPSH is greatly removed due to the GCM bias corrections. The 1450-gpm contours also show very similar spatial patterns between 20°N and 40°N as that in the WRF_ERA5, i.e. a southwest-northeast orientation in the positive-phase years and a south-north orientation in the negative phase years (Figs. 10a, 10f). Such a spatial pattern of geopotential height is crucial to generate the leading EOF pattern of observed summer precipitation anomaly over eastern China.

4 Discussion and conclusions

It is well known that the large errors in the large-scaling forcing data can be propagated into RCM through the lateral boundary conditions and degraded RCM

simulations. This is often referred to as the “garbage in, garbage out” issue (Giorgi and Mearns 1999; Hall 2014; Giorgi 2019). Therefore, we corrected GCM mean, variance, and long-term trend (MVT) bias before the dynamical downscaling simulations to reduce the “garbage” in the large-scale forcing data. The MVT bias correction integrates the advantage of previous GCM bias correction methods and the MME-based long-term trend. It has the potential to generate more reliable climate projections. Our evaluation based on three long-term RCM simulations suggests that the dynamical downscaling approach with MVT bias corrections can significantly improve the RCM simulations in terms of the climatological mean, interannual variability of multiple variables. The RMSEs of the climatological means are reduced by 70–90% for the upper-air variables (air temperature, geopotential height, vector wind, relative humidity) and 50–80% for the surface variables (e.g., SAT, precipitation, sensible heat flux, 10-m wind speed). The RMSEs of interannual variances are reduced by 20–70% for the downscaled upper-air and surface variables due to the GCM bias corrections.

More interestingly, the dynamical downscaling simulation with MVT bias corrections can significantly improve the leading EOF mode of summer precipitation anomaly over eastern China compared with the TDD approach. The GCM bias correction generates a more accurate large-scale circulation, e.g. the correct location of WNPSH, which in turn improves the downscaled precipitation variability. This is one of the important merits of GCM bias corrections before dynamical downscaling simulation relative to the bias correction after dynamical downscaling simulation. Correcting the mean bias in the downscaled precipitation does not affect the EOF mode because the EOF analysis uses an anomalous field as input. In contrast, the leading EOF of downscaled precipitation anomaly can be improved if we corrected GCM biases in large-scale circulation and SST before the dynamical downscaling simulation. The GCM bias correction can generate a more accurate large-scale circulation, which in turn improves the dynamics in RCM and generate more reasonable precipitation variability.

Our validation suggests that the dynamical downscaling approach with MVT bias corrections shows greatly improved performance in regional climate simulation in the Asian-western North Pacific domain. Based on the results presented in this paper, we expect the dynamical downscaling approach can also generate improved simulations in other regions of the world. It would be interesting to validate and compare the dynamical downscaling simulation with MVT bias corrections in different domains, especially for the validation of the spatial-temporal variability of precipitation. In addition to the GCM biases, RCM also contains significant biases which are not considered in this study yet. One may consider using GCM bias corrections together with spectral nudging during the dynamical downscaling simulations to constrain both the GCM and RCM biases (Xu and Yang 2015).

Acknowledgements We acknowledge the climate modeling groups for producing and making available their model outputs of CMIP6 experiments. The ERA5 reanalysis dataset was provided by the European Centre for Medium-Range Weather Forecasts (<https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>). ZX is supported jointly by the National Natural Science Foundation of China (42075170) and the National Key Research and Development Program of China (2018YFA0606004). YH is supported jointly by the National Key Research and Development Program of China (2017YFA0603803) and the National Science Foundation of China (41675105). CYT is supported by a Chinese University Direct Grant (No. 4053331). This work was also supported by the Jiangsu Collaborative Innovation Center for Climate Change.

References

- Adachi SA, Tomita H (2020) Methodology of the constraint condition in dynamical downscaling for regional climate evaluation: A review. *J Geophys Res Atmos* 125(11). DOI:10.1029/2019JD032166.
- Bruyère CL, Done JM, Holland GJ, Fredrick S (2014) Bias corrections of global models for regional climate simulations of high-impact weather, *Clim Dyn* 43: 1847–1856. DOI:10.1007/s00382-013-2011-6
- Chang C P, Zhang Y, Li T (2000) Interannual and Interdecadal Variations of the East Asian Summer Monsoon and Tropical Pacific SSTs. Part II: Meridional Structure of the Monsoon. *J Climate* 13: 4326-4340.
- Chen J, Wang Z, Tam C-Y, Lau N-C, Lau D, Mok HY (2020) Investigating climate change impacts on western North Pacific tropical cyclones and induced storm surges over the Pearl River Delta region based on pseudo-global warming experiments, *Sci Rep* 10, 1965. <https://doi.org/10.1038>
- Colette A, Vautard R, Vrac M (2012) Regional climate downscaling with prior statistical correction of the global climate forcing. *Geophys Res Lett* 39, L13707. doi:10.1029/2012GL052258.
- Dai AG, Rasmussen RM, Ikeda K, Liu CH (2020) A new approach to construct representative future forcing data for dynamic downscaling. *Clim Dyn*, 55: 315–323. doi: 10.1007/s00382-017-3708-8.
- Ding Y, Wang Z, Sun Y (2008) Inter-decadal variation of the summer precipitation in East China and its association with decreasing Asian summer monsoon. Part I: Observed evidences. *Int J Climatol* 28: 1139–1161. doi:10.1002/joc.1615.
- Done JM, Holland GJ, Bruyère CL, Leung LR, Suzuki-Parker A (2013) Modeling high-impact weather and climate: Lessons from a tropical cyclone perspective, *Clim Change* 129(3-4): 381–395. doi:10.1007/s10584-013-0954-6.
- Eyring, V. Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, et al (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci Model Dev* 9(5): 1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>.

528 Fu C, Wang S, Xiong Z, Gutowski WJ, Lee D-K, McGregor JL, et al (2005) Regional
 529 Climate Model Intercomparison Project for Asia. *Bull Amer Meteorol Soc* 86: 257–
 530 266.

531 Gao X J, Shi Y, Giorgi F. (2011) A high resolution simulation of climate change over
 532 China. *Sci China Earth Sci*, 54: 462–472.

533 Giorgi F, Mearns LO (1991) Approaches to the simulation of regional climate change:
 534 a review. *Rev Geophys*, 29: 191–216.

535 Giorgi F and Mearns L O (1999) Introduction to special section: regional climate
 536 modeling revisited *J Geophys Res Atmos* 104: 6335–6352.

537 Giorgi F (2019) Thirty years of regional climate modeling: Where are we and where
 538 are we going next? *J Geophys Res Atmos* 124: 5696–5723.

539 Gutowski WJ, Giorgi F, Timbal B, Frigon A, Jacob D, Kang H-S, et al (2016) WCRP
 540 COordinated Regional Downscaling EXperiment (CORDEX): a diagnostic MIP for
 541 CMIP6. *Geosci Model Dev* 9: 4087–4095.
 542 <https://doi.org/10.5194/gmd-9-4087-2016>

543 Hall A (2014) Projecting regional change. *Science* 346: 1461–1462. DOI:
 544 10.1126/science.aaa0629

545 Jacob D, Podzun R (1997) Sensitivity studies with the regional climate model REMO.
 546 *Meteorol Atmos Phys* 63: 119–129.

547 Ji F, Evans JP, Luca AD, Jiang N, Olson R, Fita L, et al (2019) Projected change in
 548 characteristics of near surface temperature inversions for southeast Australia. *Clim*
 549 *Dyn* 52: 1487–1503. <https://doi.org/10.1007/s00382-018-4214-3>

550 Kawase H, Yoshikane T, Hara M, Kimura F, Yasunari T, Ailikun B (2009)
 551 Intermodel variability of future changes in the Baiu rainband estimated by the
 552 pseudo global warming downscaling method. *J Geophys Res* 114, D24110.
 553 doi:10.1029/2009JD011803.

554 Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, et al
 555 (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological*
 556 *Society* 146: 1999–2049.

557 Hoffmann P, Katzfey JJ, McGregor JL, Thatcher M (2016) Bias and variance
 558 correction of sea surface temperatures used for dynamical downscaling, *J Geophys*
 559 *Res Atmos* 121(21): 12877–12890. doi:10.1002/2016JD025383.

560 Holland GJ, Done J, Bruyère C, Cooper C, Suzuki A (2010) Model investigations of
 561 the effects of climate variability and change on future Gulf of Mexico tropical
 562 cyclone activity. *Proc. Offshore Technology Conf.*, Houston, TX, ASCE, OTC
 563 20690.

564 Hong S-Y, Noh Y, Dudhia J (2006) A new vertical diffusion package with an explicit
 565 treatment of entrainment processes, *Mon Weather Rev* 134: 2318–2341.

566 Huang F, Xu Z, Guo W (2019) Evaluating vector winds in the Asian-Australian
 567 monsoon region simulated by 37 CMIP5 models. *Clim Dyn* 53: 491–507.
 568 <https://doi.org/10.1007/s00382-018-4599-z>.

569 Iacono MJ, Delamere JS, Mlawer EJ, Shephard MW, Clough SA, Collins WD (2008)
 570 Radiative forcing by long-lived greenhouse gases: Calculations with the AER
 571 radiative transfer models. *J Geophys Res* 113, D13103. doi:10.1029/2008JD009944

572 Kain JS (2004) The Kain–Fritsch convective parameterization: An update. *J Appl*
 573 *Meteor* 43: 170–181.

574 Leung LR, Qian Y, Bian X (2003) Hydroclimate of the western United States based
 575 on observations and regional climate simulation of 1981–2000. Part I: Seasonal
 576 statistics. *J Clim* 16: 1892–1911.

577 Lin Y-L, Farley RD, Orville HD (1983) Bulk Parameterization of the Snow Field in a
 578 Cloud Model. *Journal of Applied Meteorology and Climatology* 22: 1065–1092.

579 Müller WA, Jungclaus JH, Mauritsen T, Baehr J, Bittner M, Budich R, et al (2018) A
 580 high-resolution version of the Max Planck Institute Earth System Model
 581 MPI-ESM1.2-HR. *J Adv Model Earth Syst* 10: 1383–1413.

582 Niu G-Y, Yang Z-L, Mitchell KE, Chen F, Ek MB, Barlage M, et al (2011) The
 583 community Noah land surface model with multiparameterization options
 584 (Noah-MP): 1. Model description and evaluation with local-scale measurements, *J*
 585 *Geophys Res* 116, D12109, doi:10.1029/2010JD015139

586 Rocheta E, Evans JP, Sharma A (2017) Can Bias Correction of Regional Climate
587 Model Lateral Boundary Conditions Improve Low-Frequency Rainfall Variability?
588 J Climate 30: 9785–9806.

589 Sato T, Kimura F, Kitoh A (2007) Projection of global warming onto regional
590 precipitation over Mongolia using a regional climate model. J Hydrol 333: 144–
591 154.

592 Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, et al (2008) A
593 description of the Advanced Research WRF version 3, NCAR Tech. Note
594 NCAR/TN-4751STR, 113 pp.

595 Tang J, Li Q, Wang S, Lee D-K, Hui P, Niu X, et al (2016) Building Asian climate
596 change scenario by multiregional climate models ensemble. Part I: Surface air
597 temperature. Int J Climatol 36: 4241–4252.

598 White RH, Toumi R (2013) The limitations of bias correcting regional climate model
599 inputs. Geophys Res Lett 40: 2907–2912. doi:10.1002/grl.50612.

600 Wu W, Lynch AH (2000) Response of the seasonal carbon cycle in high latitudes to
601 climate anomalies. J Geophys Res 105 (D18): 22897–22908.

602 Xu Z, Yang Z-L (2012) An Improved Dynamical Downscaling Method with GCM
603 Bias Corrections and Its Validation with 30 Years of Climate Simulations. J
604 Climate 25: 6271–6286.

605 Xu Z, Yang Z-L (2015) A new dynamical downscaling approach with GCM bias
606 corrections and spectral nudging. J Geophys Res Atmos 120(8): 3063–3084
607 doi:10.1002/2014JD022958

608 Xu Z, Han Y (2019) Short communication comments on ‘DISO: A rethink of Taylor
609 diagram’. Int J Climatology 40(4): 2506–2510 doi:10.1002/joc.6359

610 Xu Z, Hou Z, Han Y, Guo W (2016) A diagram for evaluating multiple aspects of
611 model performance in simulating vector fields, Geosci Model Dev 9: 4365–4380.

612 Xu Z, Han Y, Yang ZL (2019) Dynamical downscaling of regional climate: a review
613 of methods and limitations. Sci China Earth Sci 62: 365–375.
614 <https://doi.org/10.1007/s11430-018-9261-5>

Xu Z, Han Y, Tam C-Y, Yang Z-L, Fu C (2021) Bias-corrected CMIP6 global dataset for dynamical downscaling of the historical and future climate (1979–2100). Scientific Data DOI: 10.21203/rs.3.rs-424811/v1. Under review

Yang Z-L, Niu G-Y, Mitchell KE, Chen F, Ek MB, Barlage M, et al (2011) The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins, J Geophys Res 116, D12110, doi:10.1029/2010JD015140

Zhang M-Z, Xu Z, Han Y, Guo W (2021) Evaluation of CMIP6 Models toward dynamical downscaling over eight CORDEX domains. Clim Dyn. DOI:10.21203/rs.3.rs-911581/v1. Under review.

Table 1 RMSEs of the climatological mean and interannual variance in different seasons. The REMEs are calculated with the WRF_GCM/WRF_GCMbc against the WRF_ERA5 over the validation domain illustrated in Fig.1. The variables evaluated include the surface air temperature (SAT), 2-m specific humidity (Q2), precipitation (Precip), snow depth (Snowh), sensible heat flux (SH), latent heat flux (LH), planetary boundary height (PLBH), and 10-m wind speed (Wind10).

	SAT (°C)	Q2 (g kg ⁻¹)	Precip (mm d ⁻¹)	Snowh (cm)	SH (W s ⁻²)	LH (W s ⁻²)	PBLH (m)	Wind10 (m s ⁻¹)
<u>Climatological mean</u> (WRF_GCM/WRF_GCMbc)								
Spring	0.87/0.18	0.69/0.13	3.15/0.95	4.46/1.62	6.07/2.23	15.34/5.25	42.49/16.99	0.42/0.14
Summer	0.74/0.19	0.69/0.13	3.37/1.05	2.61/0.70	6.77/1.63	21.94/4.53	53.11/14.65	0.69/0.16
Autumn	0.59/0.17	0.59/0.12	3.25/1.05	2.74/0.95	5.48/2.60	16.99/5.77	40.72/15.64	0.62/0.15
Winter	0.86/0.41	0.74/0.11	3.30/0.52	3.78/1.72	10.2/6.91	20.27/6.58	46.84/23.50	0.50/0.16
<u>Interannual standard deviation</u> (WRF_GCM/WRF_GCMbc)								
Spring	0.34/0.23	0.26/0.11	1.79/1.00	1.76/0.86	3.77/1.96	6.90/3.52	22.36/12.93	0.19/0.08
Summer	0.29/0.13	0.27/0.09	1.95/0.93	1.25/0.54	3.20/1.20	7.01/3.82	20.02/11.12	0.21/0.11
Autumn	0.28/0.11	0.30/0.10	2.06/1.04	1.24/0.72	3.37/2.03	8.32/3.21	20.09/10.44	0.16/0.09
Winter	0.30/0.19	0.17/0.07	1.53/0.68	1.32/0.91	3.38/1.69	6.57/3.03	15.51/10.07	0.15/0.08

Table 2 Pattern correlation between the WRF_GCM (WRF_GCMbc) and WRF_ERA5 for the first three EOF patterns of summer precipitation over eastern China (20°–45°N, 105°–125°E). The number in the bracket indicates the percentage of total variance explained by each EOF mode.

	EOF1	EOF2	EOF3
WRF_GCM	0.26 (18.2%)	0.54 (14.5%)	0.25 (10.4%)
WRF_GCMbc	0.88 (25.8%)	0.85 (16.6%)	0.43(9.0%)
WRF_ERA5	(20.5%)	(19.3%)	(12.5%)

Figures

Fig. 1 WRF model Domain (shaded area) and terrain height (m). The thick red line outlines the verification region.

645

646

Fig. 2 Mean errors of the precipitation and surface air temperature in summer during the period of 1982–2014. The shaded area indicates the significance level of 0.05.

647

648

649

650

Fig. 3 RMSE of climatological mean air temperature, geopotential height, wind vector and relative humidity derived from the WRF_GCM and WRF_GCMbc against the WRF_ERA5 in summer (June-July-August).

651

652

653

Fig. 4 Mean errors of 850-hPa geopotential height (m), vector wind, and relative humidity in summer (June-July-August) during the period 1982–2014. Shaded areas represent that the errors reach the significance level of 0.05. The red, blue and black arrows indicate that the zonal wind, meridional wind and both of them reach the significance level of 0.05, respectively.

654

655

656

657

658

Fig. 5 Same as in Fig.2 except for standard deviation of interannual variability.

659

660

661

Fig. 6 Same as in Fig. 3 except for the RMSEs of interannual variance.

662

663

664

665

Fig. 7 Mean errors of interannual variability of 850-hPa geopotential height (a, b) and vector wind (c, d) in summer. The interannual variability is measured by the standard deviation over the period of 1982–2014. The hatched area indicates the interannual variances between two experiments reach the significance level of 0.05 based on F test.

666

667

Fig. 8 Leading EOF mode and PC time series of summer precipitation anomaly derived from the WRF_ERA5, ERA_GCM, and WRF_GCMbc, respectively. The thick black line indicates the linear trend of the time series in the lower panels. The variance explained is shown on the top of each panel.

670

671

Fig. 9 Difference in 850-hPa circulation between the positive phase and negative phase years. Contours: geopotential height (gpm), Arrows: vector wind (m s^{-1}). The colour of the arrow indicates the difference in moisture anomaly (g kg^{-1}). The positive year is defined by the time series of first EOF (Fig.8) greater than its 0.5 standard deviation.

672

673

Fig. 10 1500-gpm and 1450-gpm contours of geopotential heights in the WRF_ERA5 (a, d), WRF_GCM (b, e), and WRF_GCMbc (c, f) during the period of 1982–2014. The thin red, blue, and grey contours indicate the geopotential height in each positive-phase, negative-phase, and normal years, respectively. The positive (negative) years are defined as the time series of EOF greater (smaller) than 0.5 (–0.5) standard deviation. The thick red and blue contours represent the geopotential height averaged over the positive-phase and negative-phase years, respectively.