ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 37, NOVEMBER 2020, 1191–1210

• **Original Paper** •

# Bias Correction and Ensemble Projections of Temperature Changes over Ten Subregions in CORDEX East Asia

Chenwei SHEN, Qingyun DUAN*, Chiyuan MIAO, Chang XING, Xuewei FAN, Yi WU, and Jingya HAN

*State Key Laboratory of Earth Surface Processes and Resource Ecology*, *Faculty of Geographical Science*, *Beijing Normal University*, *Beijing* 100875, *China*

(Received 6 February 2020; revised 1 July 2020; accepted 17 July 2020)

ABSTRACT

Regional climate models (RCMs) participating in the Coordinated Regional Downscaling Experiment (CORDEX) have been widely used for providing detailed climate change information for specific regions under different emissions scenarios. This study assesses the effects of three common bias correction methods and two multi-model averaging methods in calibrating historical (1980–2005) temperature simulations over East Asia. Future (2006–49) temperature trends under the Representative Concentration Pathway (RCP) 4.5 and 8.5 scenarios are projected based on the optimal bias correction and ensemble averaging method. Results show the following: (1) The driving global climate model and RCMs can capture the spatial pattern of annual average temperature but with cold biases over most regions, especially in the Tibetan Plateau region. (2) All bias correction methods can significantly reduce the simulation biases. The quantile mapping method outperforms other bias correction methods in all RCMs, with a maximum relative decrease in root-mean-square error for five RCMs reaching 59.8% (HadGEM3-RA), 63.2% (MM5), 51.3% (RegCM), 80.7% (YSU-RCM) and 62.0% (WRF). (3) The Bayesian model averaging (BMA) method outperforms the simple multi-model averaging (SMA) method in narrowing the uncertainty of bias-corrected results. For the spatial correlation coefficient, the improvement rate of the BMA method ranges from 2% to 31% over the 10 subregions, when compared with individual RCMs. (4) For temperature projections, the warming is significant, ranging from 1.2°C to 3.5°C across the whole domain under the RCP8.5 scenario. (5) The quantile mapping method reduces the uncertainty over all subregions by between 66% and 94%.

**Key words:** CORDEX-EA, bias correction, BMA, temperature projection

**Citation:** Shen, C. W., Q. Y. Duan,, C. Y. Miao, C. Xing, X. W. Fan, Y. Wu, and J. Y. Han, 2020: Bias correction and ensemble projections of temperature changes over ten subregions in CORDEX East Asia. *Adv. Atmos. Sci.*, **37**(11), 1191–1210, https://doi.org/10.1007/s00376-020-0026-6.

**Article Highlights:**

- RCMs have obvious cold biases over the East Asia region, especially in cold seasons.
- Bias correction and BMA methods significantly reduce biases in RCM simulations.
- Temperatures increase between 1.2°C and 3.5°C under the RCP8.5 scenario in 2030–49.
- The warming trend is more remarkable in the northern part of the East Asia region.

## 1. Introduction

Climate change has attracted much attention as its effects on human society and ecological environments have grown over past decades (Grimm et al., 2013; Sun et al., 2015, 2016, 2020; Huang et al., 2016). According to the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report, several independent datasets show that the global average surface temperature increased by 0.65°C–1.06°C from 1880 to 2012 (IPCC, 2013). A set of global climate model (GCM) simulations shows that, relative to current climate (1986–2015), by the end of the 21st century (2081–2100), global average surface temperature further increases by from 0.3°C to 4.8°C under different scenarios [from Representative Concentration Pathway (RCP) 2.6 to RCP8.5] (IPCC, 2014). Accompanied by further increases in surface temperature, obvious impacts on natural and anthropogenic systems around the world have been reported, such as diminishing levels of snow and ice, increasing sea level and more occurrences of climate extremes (e.g., heat waves, droughts and floods) (Mann and Gleick, 2015; Schlaepfer et al., 2017; Kang and Eltahir, 2018). To evaluate climate change impacts and mitigate risks of climate changes, reli-

* Corresponding author: Qingyun DUAN
Email: qyduan@hhu.edu.cn

able projections of future climate change are crucial for decision makers in government, nongovernmental organizations and the general public, especially for regions vulnerable to the adverse effects of climate change. For example, with precipitation and temperature being the main driving variables for hydrological processes, climate change will affect the hydrological cycle (Luo et al., 2018). Since water is an essential resource, hydrological changes will influence flood control policies, hydropower production management and agricultural irrigation adjustment. To investigate the changes in the hydrological cycle, reliable climate change information from regional climate models (RCMs) and GCMs is required to drive the hydrological models (Teutschbein and Seibert, 2010; Mearns et al., 2015).

In past decades, GCMs have been widely used to investigate the mechanisms of climate changes (Gulizia and Camilloni, 2015; Eyring et al., 2016; Soden et al., 2018). Several studies show that GCMs can satisfactorily reproduce the spatial pattern and variability of historical temperature (Miao et al., 2014; McSweeney et al., 2015; Tang et al., 2016; Ashfaq et al., 2017; Ruan et al., 2019). However, GCMs are less skillful in regional climate factor simulations due to their coarse resolution (ranging from 100 to 200 km) (Bao et al., 2015; Li et al., 2018a). The dynamical downscaling of outputs from GCMs by RCMs can reproduce regional climate information that is more detailed and reliable. Compared with GCMs, RCMs can provide higher-resolution, more-detailed local information such as topography, land-use categories and soil moisture data, which are important in simulating regional climate information (Fulakeza et al., 2002; Salzmann et al., 2007; Akhtar et al., 2009; Jones and Brunsell, 2009; Halder et al., 2016). Although RCMs can alleviate the deficiencies in GCMs, the simulations of RCMs also greatly depend on (1) the quality of initial conditions, (2) lateral and boundary conditions provided by GCMs used to drive the RCMs and (3) simplified physical parameterization schemes for subgrid-scale physical processes (Au-Yeung and Chan, 2012; Rocheta et al., 2017). As a result, various techniques, such as bias correction and ensemble post-processing methods, have been developed to remove the systemic biases in RCM simulations and give better projections of future climate. The Coordinated Regional Downscaling Experiment (CORDEX) is a program sponsored by the World Climate Research Programme (Giorgi et al., 2009). It was established to provide an international coordinated downscaling framework for advancing RCM development, evaluation and applications (Gutowski et al., 2016). Within CORDEX, RCM ensembles have been created for multiple regions throughout the world (through dynamic downscaling driven by CMIP5 GCMs). Since the influence of climate change varies across regions and time scales, regional climate studies are necessary for correctly detecting climate change signals. East Asia is one of the most vulnerable regions to climate change, since it is a large domain that comprises diverse terrestrial features and complex climate systems (Li et al., 2018b; Miao et al., 2019; Zheng et al., 2019). The main climate system of East Asia is a monsoonal system, and the monsoons are always accompanied by extreme events such as heat waves, typhoons, droughts and floods (Ding and Chan, 2005; Chang et al., 2012; Lee et al., 2017). Observations show that since the middle of the 20th century, the average surface temperatures and frequencies of heat waves have increased for most regions of East Asia (Hijioka et al., 2014; Zhou et al., 2016).

Focusing on the CORDEX East Asia domain (CORDEX-EA), five RCMs have been employed to provide ensemble simulations of regional climate information. Several studies have examined the performance of RCMs in simulating temperature and the projections of future temperature over CORDEX-EA. Kim et al. (2016) analyzed the spatial pattern of projected temperature data from RCMs participating in studies of the CORDEX-EA domain. The results showed that there will be a warming of 1°C–3°C over the whole domain by the year 2050 and the temperature will increase more at high latitude. Park et al. (2016) focused on the performance of RCMs participating in CORDEX-EA research in simulating summer temperature means and extremes. They found that, compared to the Asian Precipitation-Highly-Resolved Observational Data Integration Towards Evaluation dataset, the RCMs show systematic biases in seasonal means and the simulations of temperature means are more accurate than those of the extremes. Gu et al. (2018) showed that the RCMs have improved model performances as compared to the raw GCM outputs, and the projected trends of the RCM temperatures are increasing from the ensemble mean by around 1°C $yr^{-1}$ over the entire domain. Further, there have also been studies focused on the bias correction and ensemble calibration of RCM simulations over CORDEX-EA. Ngai et al. (2017) showed that RCM bias is comparable to the bias of the driving GCMs, and the bias correction methods can substantially reduce the bias in simulating historical temperature. Kim and Suh (2013) used a Bayesian model averaging (BMA) method to calibrate the probabilistic predictions of temperature. They found that BMA outperforms the equal-weighted method and other ensemble calibration methods in calibrating seasonal means and distributions of the simulations. In addition, the BMA forecasts outperform the single-RCM forecasts.

Bias correction methods and ensemble calibration methods have been used in the aforementioned studies to improve the simulation of historical temperature and improve the reliability of temperature projections. However, little attention has been paid to the performance of different bias correction methods in improving the accuracy of RCM simulations, and comparison of their ability in narrowing future projections over the CORDEX-EA region is insufficient as well.

In this study, we evaluated the effects of three common bias correction methods on improving historical temperature simulations, and then used the most effective bias correction method together with an ensemble method to narrow the uncertainties in temperature projections. The changes in

future climate were analyzed based on the most reliable projection of temperatures. The paper is organized as follows: Section 2 describes the observational data, the climate model data, and the bias correction and ensemble averaging methods. Section 3 presents the design of the bias correction experiment, including the selection of calibration and validation cases and the evaluation of bias correction methods, as well as the results of ensemble calibration. Finally, the future temperature changes under the RCP4.5 and RCP8.5 scenarios are also discussed, based on the projection period (2030–49) and reference period (1980–99). A summary and conclusions are provided in section 4.

## 2. Data and methods

### 2.1. *Observations and model setup*

The reference data of monthly temperature over CORDEX-EA were taken from the Climatic Research Unit Time-Series (CRU TS) 4.03 dataset developed by the University of East Anglia (available at http://data.ceda.ac.uk/badc/cru/data/cru_ts/cru_ts_4.03/data/tmp), with a resolution of 0.5° (Harris et al., 2014) and a time frame of 1901-2018. To conduct bias correction, only the data from 1980 to 2005 were used in the study, to coincide with the historic period of CORDEX-EA.

Five RCMs were used in the CORDEX-EA experiment: the Hadley Centre Global Environmental Model, version 3, with Regional Atmosphere configurations (HadGEM3-RA); the Fifth-generation Pennsylvania State-National Center for Atmospheric Research Mesoscale Model (MM5); the Weather Research and Forecasting (WRF) model; the Regional Climate Model, version 4 (RegCM4); and the Yonsei University Regional Climate Model (YSU-RCM). The selected RCMs include three non-hydrostatic models (HadGEM3-RA, MM5 and WRF) and two hydrostatic models (RegCM4 and YSU-RCM) (von Storch et al., 2000; Cha et al., 2008; Giorgi et al., 2012; Baek et al., 2013; Wang et al., 2013). Table 1 lists the detailed configurations of the five RCMs, including dynamics processes, physical parameterization schemes, and spectral nudging (Gu et al., 2018). The CORDEX-EA domain covers East Asia, India, South Asia, and the northern part of Australia (Fig. 1), and the spatial resolution is 50 km (except HadGEM3-RA, whose resolution is 0.44°). The historical experiment (1980-2005) and the future projections under the RCP4.5 and RCP8.5 scenarios (2006–49) are driven by the outputs of HadGEM2-AO (Hadley Centre Global Environmental Model, version 2, with Atmosphere and Ocean and sea ice configurations), whose horizontal resolution is 1.875° × 1.25°. Several studies have confirmed the good performance of HadGEM2-AO for simulating East Asia's climatology (Martin et al., 2011; Baek et al., 2013; Sperber et al., 2013).

To further correct the biases at smaller spatial scales, 10 subregions were selected: northwestern China (NW; 36°–43°N, 75°–103°E); the Tibetan Plateau (TP; 28°–35°N, 75°–103°E); northeastern China (NE; 30°–42°N, 104°–121°E), northern China (NC; 42°–55°N, 113°–132°E); southern China (SC; 18°–30°N, 104°–122°E); the Korean Peninsula and Japan (KJ; 43°–51°N, 91°–112°E); Mongolia (MG; 30°–42°N, 125°–141°E); India (5°–27°N, 69°–91°E), Indochina (InC; 8°–28°N, 92°–110°E), and Southeast Asia (SEA; 10°S–6°N, 95°–151°E) (Zou and Zhou, 2016; Zhou et al., 2016; Li et al., 2018a; Tang et al., 2018). The reference period (1980–99) and projected period (2030–49) are analyzed to explore future climate change.

### 2.2. *Bias correction methods*

As mentioned above, due to the deficiencies in RCMs, large biases can be found in simulations when compared to the observations. Thus, bias correction methods based on the observations can be implemented to improve the performance of RCMs. Considering the relatively short time sequence of data, here three common stationary temperature bias correction methods are used from the R packages "hyfo" and "DownscaleR"; namely, variance scaling, additive scaling, and quantile mapping based on empirical distribution (Wilcke et al., 2013). To facilitate bias correction and ensemble calibration, temperatures from RCMs were interpolated to a common grid of 0.5° × 0.5° latitude/longitude, following CRU, using bilinear interpolation.

**Table 1**. Configurations of the five RCMs in the CORDEX-EA region (after Gu et al., 2018).

| Name | HadGEM3-RA | RegCM4 | MM5 | WRF | YSU-RCM |
|---|---|---|---|---|---|
| Resolution | 0.44° | 50 km | 50 km | 50 km | 50 km |
| Dynamics process | Non-hydrostatic | Hydrostatic | Non-hydrostatic | Non-hydrostatic | Hydrostatic |
| Convective scheme | Revised mass flux scheme | MIT-Emanuel | Kain–Fritsch II | Kain–Fritsch II | Simplified Arakawa–Schubert |
| Land-surface parameterization | MOSES2 | CLM3 | CLM3 | NOAH | NOAH |
| Planetary boundary layer | MOSES2 nonlocal | Holtslag | YSU | YSU | YSU |
| Spectral nudging | No | Yes | Yes | Yes | Yes |
| Research center | Met Office Hadley Centre | International Centre for Theoretical Physics | Seoul National University | NCAR's Mesoscale and Microscale Meteorology Laboratory | Climate Limited-area Modelling Community |

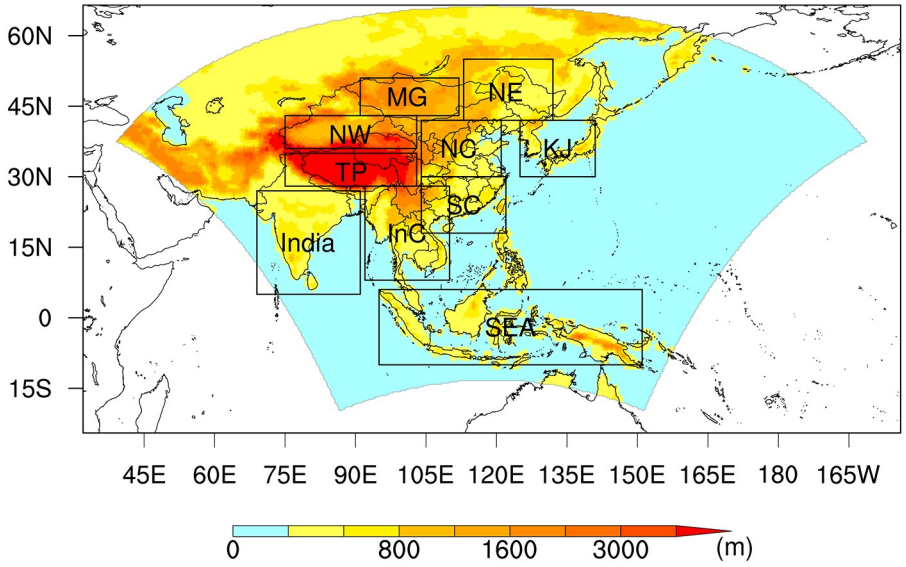


**Fig. 1**. Simulation domain and topology of CORDEX-EA and the 10 selected subregions: northwestern China (NW), Tibetan Plateau (TP), northeastern China (NE), northern China (NC), southern China (SC), Korean Peninsula and Japan (KJ), Mongolia (MG), India, Indochina (InC), and Southeast Asia (SEA).

#### 2.2.1. *Additive scaling*

The additive scaling method performs bias correction based on the bias between the average simulation and average observation during the calibration period. It is expressed as

$$T_{\rm bc}(t) = T_{\rm val}(t) - \overline{T}_{\rm cal} + \overline{T}_{\rm obs-cal}\,, \tag{1}$$

where $T_{\rm val}$ and $T_{\rm bc}$ denote the validation temperature sequence before and after bias correction, respectively, and $\overline{T}_{\rm cal}$ and $\overline{T}_{\rm obs-cal}$ are the average of the simulation and observation temperatures in the calibration, respectively.

#### 2.2.2. *Variance scaling*

Building on the additive scaling method, the variance scaling method further corrects the variance of the temperature (Terink et al., 2010) using the following procedure. First, we apply the additive scaling method, which corrects the average of the temperature:

$$T^{*1}_{\rm val}(t) = T_{\rm val}(t) - \overline{T}_{\rm cal} + \overline{T}_{\rm obs-cal}\,. \tag{2.1}$$

Then, we convert the average temperature to 0, using

$$T^{*2}_{\rm val}(t) = T^{*1}_{\rm val}(t) - \overline{T^{*1}_{\rm val}}\,. \tag{2.2}$$

Next, we scale the variance of temperature according to the ratio of the temperature variance between the calibration and validation periods:

$$T^{*3}_{\rm val}(t) = T^{*2}_{\rm val}(t)\left[\frac{\sigma_{\rm m}(T_{\rm obs-cal})}{\sigma_{\rm m}\left(T^{*2}_{\rm cal}\right)}\right]. \tag{2.3}$$

And finally, we add the temperature in step 3 to the temperature in step 1:

$$T^{*}_{\rm val}(t) = T^{*3}_{\rm val}(t) + \overline{T^{*1}_{\rm val}}\,. \tag{2.4}$$

#### 2.2.3. *Quantile mapping*

The quantile mapping (QM) method derives from the empirical transformation developed by Panofsky and Brier (1968). It has been widely used in the bias correction of both GCMs and RCMs (Wilcke et al., 2013; Miao et al., 2016). By contrast with the methods mentioned above, the QM method focuses not only on the mean of the distribution but also on correcting the quantiles of the distribution. The QM method estimates the cumulative distribution function (CDF) from the simulation data in the calibration period and then finds the corresponding percentile values of the model projections. The corrected projections can be derived through inverse CDFs of the observations. The transfer function is shown as follows:

$$T_{\rm bc}(t) = F^{-1}_{\rm O}\left(F_{\rm mc}\left[T_{\rm val}(t)\right]\right)\,, \tag{3}$$

where the subscripts O and mc denote the observation and model calibration periods, respectively. $F_{\rm mc}$ is the CDF from simulation data in the calibration period, and $F^{-1}_{\rm O}$ is the inverse CDF of observations.

### 2.3. *Simple multi-model averaging method*

Simple multiple-model averaging (SMA), which gives each member in the ensemble equal weight, is the most common ensemble post-processing method. For an ensemble of $n$ members, the weight of each member is $1/n$ (Arsenault et al., 2015).

### 2.4. *BMA method*

The BMA method was introduced by Raftery et al. (2005) to combine different model forecasts into an

ensemble and calibrate the under-dispersion during the ensemble forecasts (Duan and Phillips, 2010; Miao et al., 2013). BMA can be viewed as a post-processing method for producing the forecast probability density function (PDF) of output variables, which is a weighted average of the bias-corrected PDF of each individual ensemble member. The weights reflect the relative performance of each member in the ensemble during the training period. Following the notation in Raftery et al. (2005), the BMA-weighted forecast PDF of variable $y$ is

$$p(y|f_1,\ldots,f_k,O) = \sum_{k=1}^{K} w_k g_k(y|f_k,O) \ , \tag{4}$$

where $w_k$ is the weight of the $k$th forecast $f_k$ over a training period and all $w_k$ must sum to 1. $K$ is the number of forecasts, and $O$ denotes the observations during the training period. Further, $g_k(y|f_k,O)$ is the conditional PDF of $y$ given the $k$th forecast $f_k$ and observation $O$. We considered the temperature forecast in our case to be a normal PDF with mean $f_{k,\mathrm{bc}}$ (bias-corrected forecast from bias correction methods described above) and standard deviation $\sigma^2$. Therefore, the conditional PDF is expressed as

$$y|(f_k,O) \sim N(f_{k,\mathrm{bc}},\sigma^2) \ . \tag{5}$$

The mean of the BMA-weighted forecast can be interpreted as a weighted sum of the normal distributions with equal variance but centered at the bias-corrected forecast $f_{k,\mathrm{bc}}$:

$$E[y|f_1,\ldots,f_k] = \sum_{k=1}^{K} w_k f_{k,\mathrm{bc}} \ . \tag{6}$$

The BMA weights $w_k$ and variance $\sigma^2$ are estimated through the maximum likelihood method (Raftery et al., 2005). The expectation maximization algorithm is adopted in this study to calculate the BMA weights and variance. For a detailed description of the BMA method, see Raftery et al. (2005).

### 2.5. *Model evaluation in spatial simulation*

In this study, we conducted the bias correction for each subregion and for each month of the year. In addition, considering the time dependence of the model biases and the relatively short time sequence of historical simulations (only 26 years, from 1980 to 2006), cross validation (Miao et al., 2016) was used to calibrate and validate the performance of the RCMs. For calibration, 20 years were randomly selected out of the 26 years and then three bias correction methods were applied to get the bias correction factors. The remaining 6 years were used for validation, and the bias correction factors were applied. The sampling method used in cross validation was simple random sampling without replacement, which guarantees that each member has an equal opportunity to be selected. Validation was conducted by comparing the corrected temperatures with CRU data. We repeated the whole cross-validation process 30 times to overcome the limitation of insufficient sample sizes and enhance the robustness of the validation results. The performance of different bias correction methods was also evaluated based on the cross-validation results.

Based on the cross-validation results, for each pixel in each subregion, 30 results were acquired for each method and each month. If at least 27 corrected results agreed well with the CRU data, the corresponding method was considered to be effective. For all subregions and RCMs, we summed up the number of all effective pixels and calculated their percentage out of all pixels over the 30 validations. In this way, we analyzed the performance of each method across all regions and RCMs. Further, the temporal average (averaging over 12 months and 6 years) of validated and observed temperature was used to calculate the relative decrease in root-mean-square error (RMSE) ($\mathrm{RMSE}_{\mathrm{cal},i}$, $i = 1,\ldots,3$). The subscript $i$ represents different bias correction methods. In this way, the relative decrease in RMSE was derived with reference to the RMSE between the raw simulation and observations ($\mathrm{RMSE}_{\mathrm{raw}}$). The relative decrease in RMSE can be expressed as $\mathrm{RMSE}_{\mathrm{dec},i} = 100(\mathrm{RMSE}_{\mathrm{raw}} - \mathrm{RMSE}_{\mathrm{cal},i})/\mathrm{RMSE}_{\mathrm{raw}}$ (unit: %). The $\mathrm{RMSE}_{\mathrm{dec},i}$ reflects the effectiveness of the bias correction methods. Similarly, the relative decrease in mean average error (MAE) for each pixel was calculated using the equation $\mathrm{MAE}_{\mathrm{dec},i} = 100(\mathrm{MAE}_{\mathrm{raw}} - \mathrm{MAE}_{\mathrm{cal},i})/\mathrm{MAE}_{\mathrm{raw}}$ (unit: %). The performance of the bias correction methods over the five RCMs and 10 subregions was analyzed based on the results. If a particular bias correction method was most effective (with most effective pixels in the months and quantiles, or the most relative decreases in RMSE or MAE) among all subregions and RCMs, it is considered to be the most suitable bias correction method and is used in the future projection period.

While the most effective bias correction method was being obtained, the historical temperature simulations (1980-2005) were corrected for all subregions. Then, the ensemble averaging methods were applied to the bias-corrected temperatures. To validate the performance of ensemble averaging methods, monthly temperature distributions, interannual variability and Taylor diagrams were used. The Taylor diagram is especially useful in evaluating multiple aspects of complex models (IPCC, 2001). It incorporates three evaluation terms–spatial correlation, centered RMSE and standard deviations–and graphically measures how closely the simulation patterns match the observations (Taylor, 2001).

## 3. Results

### 3.1. *Bias correction evaluation*

Figure 2 illustrates the spatial pattern of annual average (1980-2005) temperatures of CRU and the driving GCM (HadGEM2-AO) as well as the annual average temperature biases of the driving GCM and the five RCMs. In addition, the annual average temperatures of the five RCMs are shown in Fig. S1 in the electronic supplementary material (ESM). We can see that both the GCM and the RCMs cap-

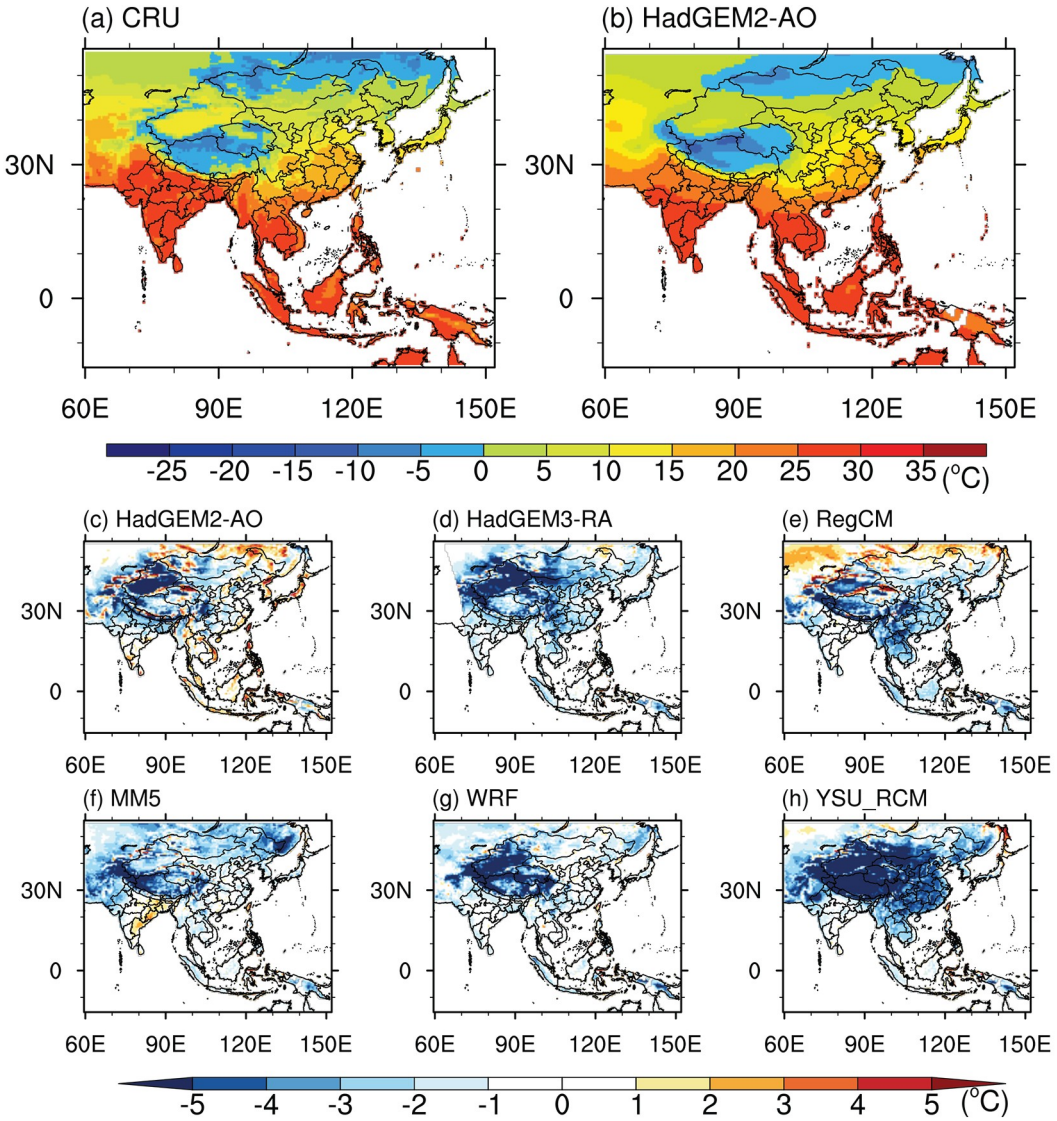


**Fig. 2**. Spatial distribution of annual average temperatures according to (a) CRU and (b) the HadGEM2-AO GCM. (c–h) Annual temperature biases of (c) the HadGEM2-AO GCM and (d–h) five RCMs during the years 1980–2005.

tured the spatial pattern of the observations with a decreasing south to north temperature gradient. However, both the driving GCM and the RCMs generally underestimated the annual average temperature in most regions, especially in the TP region, where the greatest bias exceeded −8°C. The cold biases in the simulation came mainly from December–January–February (DJF), and the biases in June-July-August (JJA) were small (Fig. S2). Several studies have also reported similar annual average and seasonal temperature bias patterns (Ham et al., 2016; Guo et al., 2018; Li et al., 2018a; Hui et al., 2019). The cold biases in DJF were remarkable in most RCMs, with the largest bias exceeding −8°C in most RCMs. The only exception was RegCM, which had warm biases exceeding 4°C at high latitudes in CORDEX-EA, which is consistent with previous studies (Gao and Giorgi, 2017). The RCMs' performance also varied in JJA temperature simulations. For example, the YSU-RCM model presented a large cold bias in NE and the TP (exceeding −8°C), while the bias was small in other RCMs.

In addition, compared to the driving GCM, some RCMs had improved their skill in simulating the temperature in several regions. For example, WRF and MM5 reduced the biases over the NC, SC and KJ regions, and RegCM and YSU-RCM improved their performance in the India region. However, for the other regions (e.g., TP and NW), the improvement of the RCMs was less and the cold biases were even larger than those of the driving GCM. Increased resolution does not always lead to improvement in simulations (Prömmel et al., 2010; Gu et al., 2018), and there are several reasons for this phenomenon. For instance, due to their simplified physical parameterization schemes, longwave radiation is underestimated in RCMs, which limits the heating of the low-level atmosphere (Hui et al., 2019). Also, the overestimation of albedo in RCMs (in the lower boundary conditions) may also lead to cold biases (Meng et al., 2018).

Given the large biases in simulating annual and seasonal average temperatures, bias correction is necessary to improve the RCMs' performance. We calculated the percentage of effective pixels for each method over 30 rounds of validation (Fig. 3). We found that all the methods can effectively improve RCM performance. The percentage of effective pixels at the 25% quantile was higher than at the 50% and 75% quantiles. The QM method outperformed the other two methods for the 25%, 50% and 75% quantiles. This may be due to the fact that, although all bias correction methods corrected the biases in the RCMs, their foci are different. For example, the additive scaling method focuses on the mean difference between the calibration model and observation data (Teutschbein and Seibert, 2012; Ezéchiel et al., 2016), while the variance scaling method focuses on adjustment of variance (Luo et al., 2018). The two methods can adjust the monthly mean values, but they neglect the cumulative distribution of the temperature, and thus they cannot adjust the quantiles of the simulation. However, in contrast with these two methods, the QM method constructs CDFs of RCMs and adjusts the distributions according to the corresponding distribution of observation data (Bennett et al., 2014; Singh et al., 2017; Ayugi et al., 2020). Therefore, the QM method can adjust not only the mean value of temperature but also the quantile values.

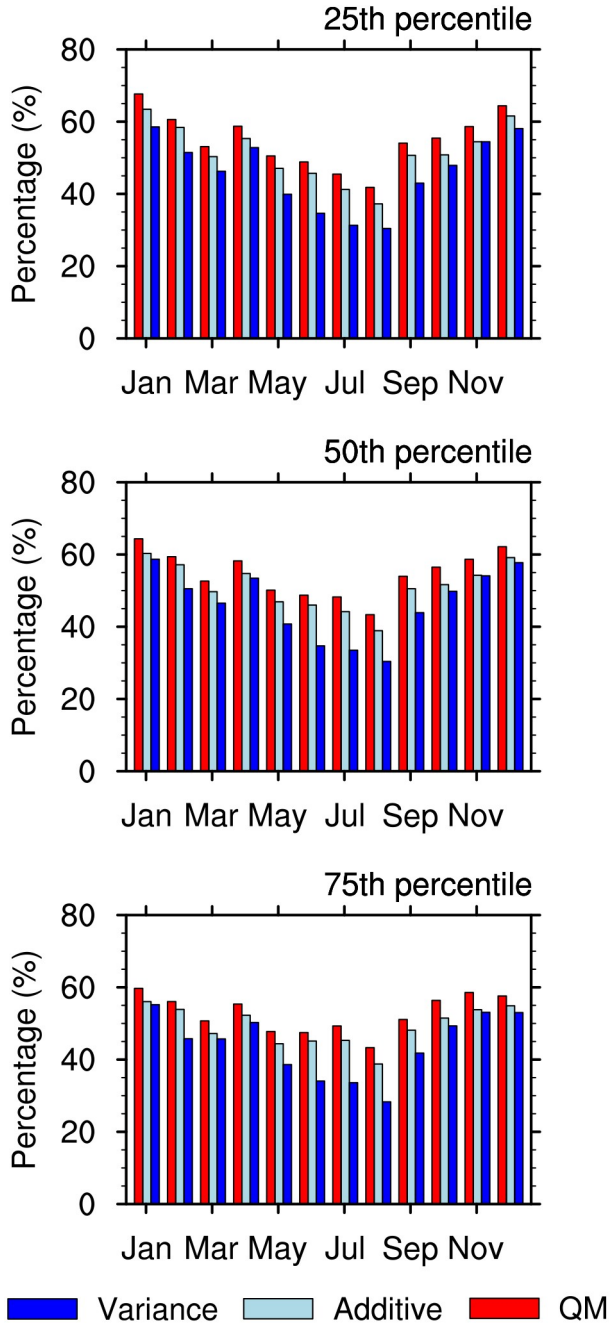


**Fig. 3**. Comparisons of different bias correction methods at the 25th, 50th and 75th quantiles.

We calculated the RMSE for both raw and bias-corrected temperature data, then the relative decrease in the RMSE for the 10 subregions and five RCMs was calculated based on the 30 rounds of cross validation (Fig. 4). Results showed that all the bias correction methods effectively reduced biases for all RCMs. The QM method was most effective among the methods, with the maximum relative decrease in the RMSE reaching 59.8% (HadGEM3-RA), 63.2% (MM5), 51.3% (RegCM), 80.7% (YSU-RCM) and 62.0% (WRF). For subregions, although all the bias correction methods significantly reduced the biases over most subregions, results varied. For example, in the SEA region, MM5, bias-corrected by the additive scaling method, was worse than the raw simulation. We analyzed the results (Fig. S3) and found that this may be due to the fact that the additive scaling method can only adjust the mean difference between the model values and observations, while for extreme values, the additive scaling method failed (Fang et al., 2015). In the SEA region, for MM5, the additive scaling method narrowed the biases for high temperature but amplified the bias for low temperature when compared to the raw simulations. However, the QM-corrected results almost perfectly fitted the CRU distribution. Several previous studies have also shown that bias correction methods are region-dependent, and the scaling method may have adverse effects in some regions (Berg et al., 2012; Ayugi et al., 2020). Moreover, for almost all subregions (except the MG region as simulated by WRF and MM5), the QM method outperformed the other two methods. We also found that, for the TP and NW regions, where the bias was largest among the subregions, the bias reduction was significant, with the maximum relative decrease in the RMSE reaching 61.5% and 80.7%, respectively (both for YSU-RCM).

Figure 5 gives the spatial distribution of the relative decrease in the MAE for the 10 subregions (the QM method performed best, so only the results of QM are shown here). The spatial distribution results are consistent with results in Fig. 4. The bias correction was effective in most regions, especially in the TP, NW, SC and NC subregions, where the MAE was reduced more than 50% compared to the MAE of the raw model output. For the YSU-RCM model, which had the largest biases in the simulation, the bias reductions were remarkable: the relative decrease in the MAE was more

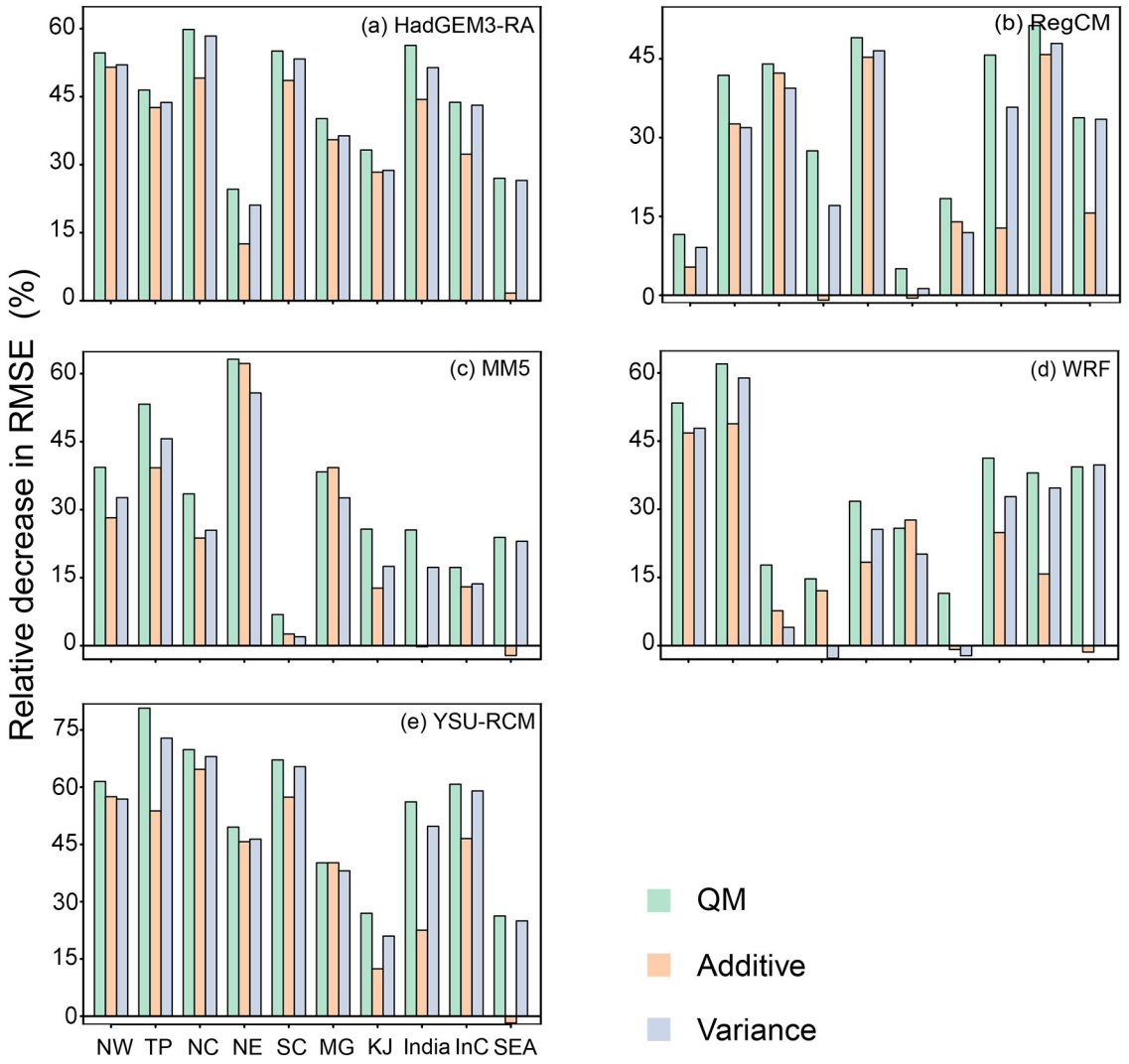


**Fig. 4**. Relative decrease in RMSE (%) of three bias correction methods for the 10 subregions for the five RCMs: (a) HadGEM3-RA, (b) RegCM, (c) MM5, (d) WRF, and (e) YSU-RCM.

than 60% for most regions and more than 70% in most parts of China. The spatial distribution results for winter (DJF) and summer (JJA) (Fig. S4) also showed that the bias correction method effectively reduced the bias in seasonal simulations, especially in winter. The correction for summer focused mainly on the southern part of the CORDEX region, including India, InC and SC. In summary, all bias correction methods effectively reduced the biases in the simulations, and the QM method was most effective for almost all subregions and all RCMs. Therefore, the QM method was chosen as the most suitable method to correct the temperature simulations in the following text. Figure S5 provides the annual average temperature error distribution of the bias-corrected data using the QM method. The results show that, compared with the raw RCM simulations, the errors were removed remarkably well by the bias correction process, especially for the NE, NC, KJ, SC, INC, SEA and India regions, where the errors were small. However, for the NW, TP and MG regions, although bias correction did decrease the cold biases, some cold biases remained. Furthermore, due to the remarkable cold biases in the NW, TP and MG regions, the bias correction tended to overcorrect cold biases in these regions, especially for the TP and NW regions in the YSU-RCM model.

### 3.2. *Multi-model averaging based on bias correction*

In addition to the bias correction method, BMA and SMA were also used to further narrow the uncertainty in the corrected RCMs. Figure 6 gives the seasonal distribution of CRU, raw simulated, bias-corrected, SMA-weighted [bias-corrected (BC)] and BMA-weighted (BC) temperatures. Results show that, compared to the CRU data, the raw simulation had significant cold biases in winter (DJF) and biases were small in summer (JJA) over the 10 subregions. This is consistent with the results in Fig. S2 and several previous studies, where the cold biases in winter were large and in summer were small (Ham et al., 2016; Hui et al., 2019). Moreover, RCM performance was region-dependent; more

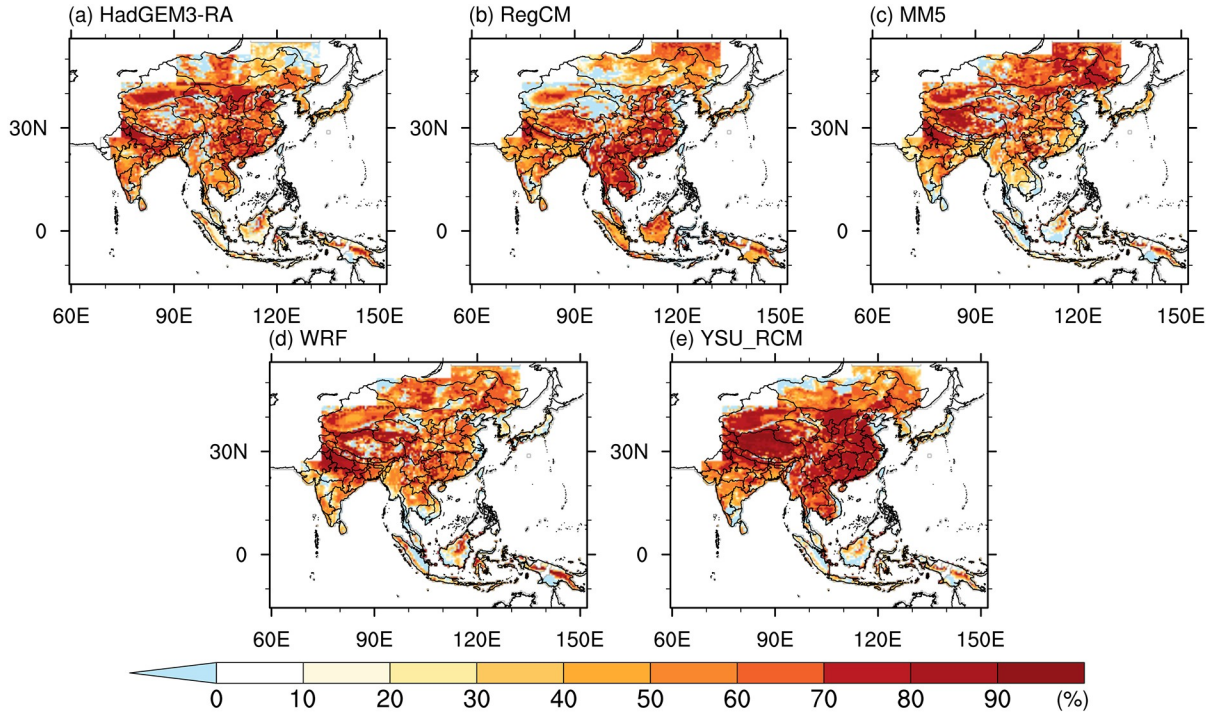


**Fig. 5**. Spatial distribution of relative decrease in MAE (%) for annual temperatures using the QM method. Panels (a–e) show results from each of the five RCMs. See Fig. S4 in the ESM for spatial distribution of relative decrease in MAE (%) for seasons.

specifically, the RCMs performed well for the regions at high latitudes (e.g., the NE, KJ and MG regions) and badly for regions at low latitudes (e.g., the SEA, TP, NW and InC regions). Several previous studies have analyzed the possible causes of the biases in these regions and suggested that careful configuration of the RCM parameterization schemes could help. For example, according to Hui et al. (2019), the radiation parameterization schemes in RCMs underestimate the monthly longwave upward and downward fluxes throughout the year, especially in cold months over subtropical regions, which leads to significant cold biases in subtropical regions. For the cold biases in the TP region, they may be attributable to the overestimation of upward shortwave radiation and the corresponding overestimation in albedo (Tangang et al., 2015; Chen et al., 2017; Hui et al., 2019; Yin et al., 2020). When the bias corrections of SMA and BMA were applied, the seasonal cycles of the raw models were adjusted and fitted the CRU data well. The bias corrections of SMA and BMA significantly improved the performance of the RCMs.

The model's ability to capture the real interannual variability is another important performance measure. Here, we used the variance among 26 historical years as the indicator of interannual variability. Figure 7 gives the results of interannual variability of the SMA (BC), BMA (BC), SMA (raw) and CRU temperature values. The results show that the RCMs' ability to capture the real interannual variability varied among subregions. For the NW, TP, NE, MG and SEA regions, the interannual variability of bias-corrected data was closer to the real interannual variability. But for other regions, the bias correction narrowed the variability. The results may be due to the fact that the bias correction methods mainly focus on the mean and trend of the data, with less focus on the variance (Ayugi et al., 2020). Although, for most subregions, the performances of SMA and SMA (BC) were similar, but for the MG region, where the interannual variability was greater, the variabilities of SMA (BC) were closer to the variability of CRU. Thus, SMA (BC) was considered as a better method. In addition, the results based on SMA (BC) were better than those based on BMA (BC) when compared to the variability of CRU. This is due to the fact that the objective function of BMA only considers the minimum bias without adjusting for variance (Raftery et al., 2005; Fragoso et al., 2018). In future studies, we will pay more attention to the variance and consider multi-objective optimization.

The spatial variability statistics of the models in reproducing annual average temperature are shown using Taylor diagrams in Fig. 8. The Taylor diagrams show that the bias correction improved the performance of the RCMs, contributing to a higher spatial correlation and lower normalized standard deviation. Furthermore, the BMA and SMA ensemble results both reduced the uncertainties in simulation with a closer distance to the observation. For some subregions, the performances of BMA and SMA were similar (e.g., in NE, NC, TP and SC). However, for other regions, such as NW, MG and SEA, the BMA method performed better. This is reasonable because the BMA weights are estimated according to the RCMs' performance in the training period (Duan et al., 2007). A previous study also showed that the BMA

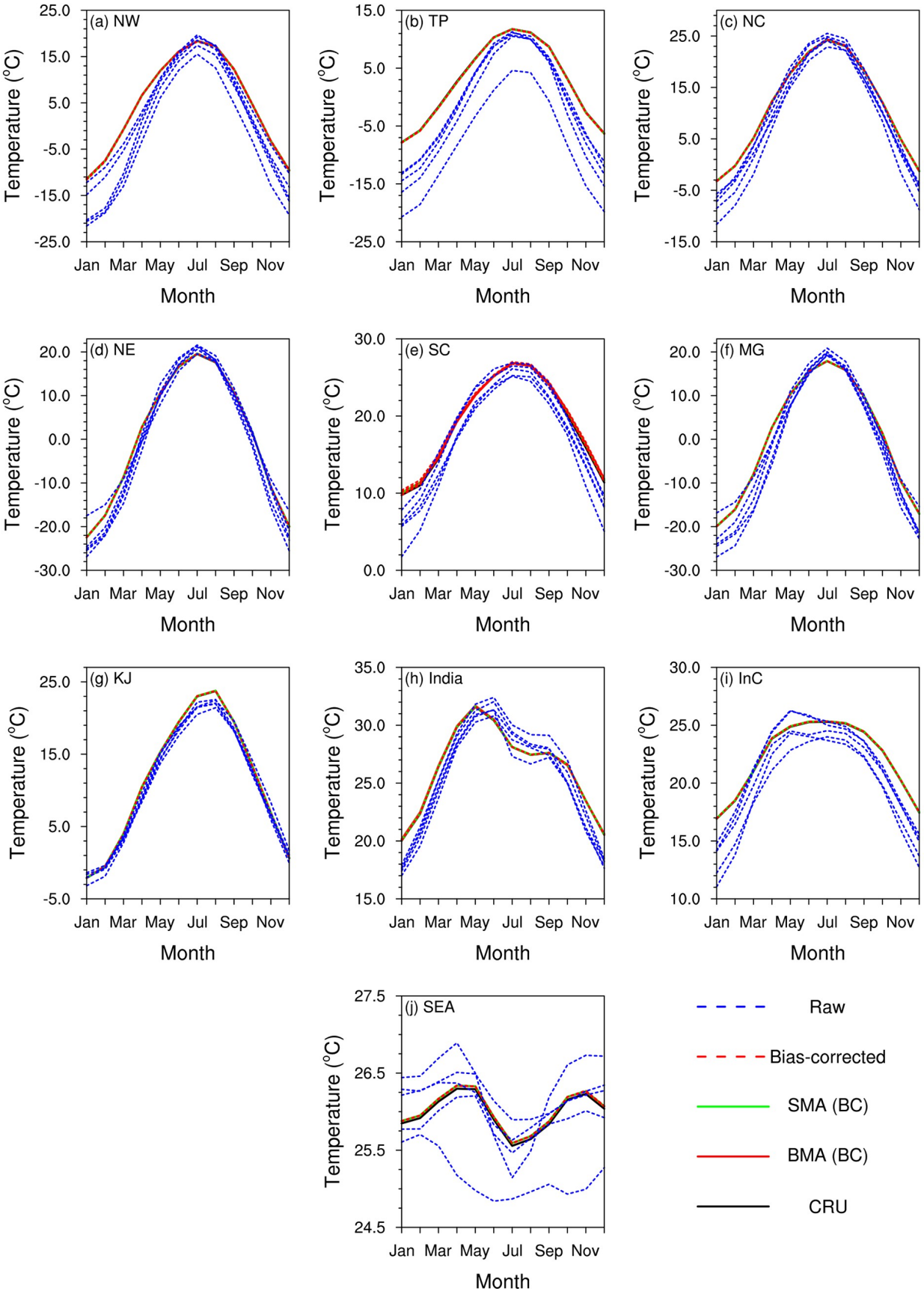


**Fig. 6**. Observed (CRU), raw simulated, bias-corrected, SMA [bias-corrected (BC)] and BMA [bias-corrected (BC)] monthly temperatures of 10 subregions in the validation period.

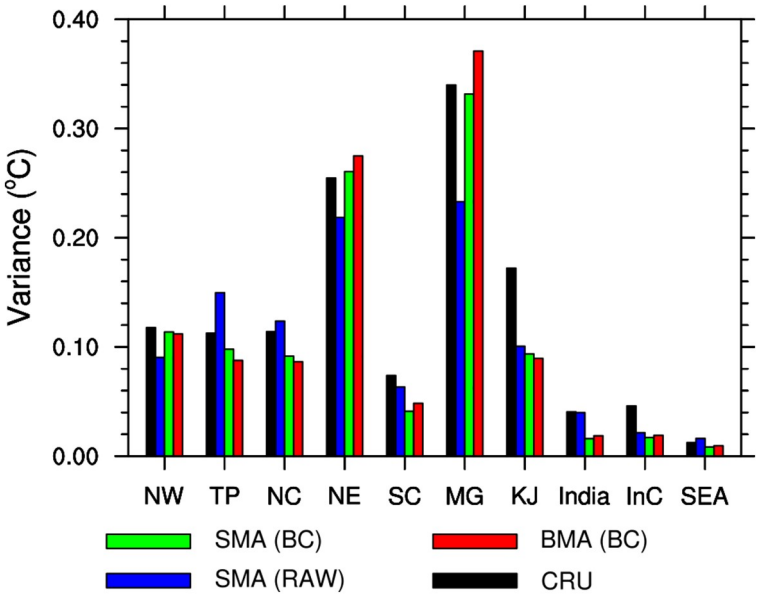


**Fig. 7**. Interannual variability for SMA [bias-corrected (BC)], BMA [bias-corrected (BC)], SMA (raw model), and CRU temperature values among the 26 years of the historical period.

method outperformed the SMA method when applied to the CORDEX-EA data (Kim and Suh, 2013). For the correlation coefficient, the improvement rate of the BMA method was between 2% and 31% when compared with individual RCMs. Although the SMA performed better with respect to interannual variability, we focus mainly on the mean and trend in the projection. Thus, we chose the BMA method for the projection to narrow the uncertainties. Although the BMA-related improvement was not substantial, the amount of improvement was reasonable considering the temperature had already been corrected by bias correction methods.

### 3.3. *Temperature projections*

Based on the most effective bias correction method (QM) and BMA weights derived from the training period (1980–2005), the bias-corrected and BMA-weighted temperature projections for the 10 subregions under the two scenarios (RCP4.5 and RCP8.5) were generated. Figure 9 shows the average temperature projections for the BMA ensemble under the RCP4.5 and RCP8.5 scenarios (results of five RCMs and the driving GCM are shown in Figs. S6 and S7 under the RCP4.5 and RCP8.5 scenarios, respectively). Similar warming trends were detected over the 10 subregions for the 2030–2049 period under both scenarios but with a more obvious warming trend under the RCP8.5 scenario. The warming trend was more remarkable in the northern part of CORDEX-EA than in the southern part, especially for the TP, NW, MG and NE regions, where the warming was over 3°C in the 2030–49 period. Moreover, analysis of seasonal warming results indicated that the warming was more remarkable in winter than in summer (not shown). Similar warming patterns have also been detected in previous studies (Ham et al., 2016; Gu et al., 2018), although these studies focused mainly on the China region. The BMA results indicate a clear increase in average temperature under the RCP4.5 and RCP8.5 scenarios for all subregions (Table 2). However, for a given subregion, the warming varied among RCMs. For example, the annual temperature increase over NE ranged from 0.5°C to 3.5°C under the RCP4.5 scenario. Figure 10 also illustrates the obvious warming trend over the 10 subregions from 2006 to 2049 under both scenarios. Note that the warming trends in the TP, NW, NE and MG regions [reaching 0.6°C (10 yr)$^{-1}$–0.7°C (10 yr)$^{-1}$] were more remarkable than in the other regions [0.3°C (10 yr)$^{-1}$–0.5°C (10 yr)$^{-1}$] under the RCP8.5 scenario, which is consistent with results in Fig. 9.

The future temperature changes varied among subregions and months. Figure 11 illustrates the changes of monthly temperature in the 10 subregions under the RCP8.5 scenario (changes under RCP4.5 were similar but with smaller amplitude; not shown). The BMA projection also indicates that monthly temperature increased more notably in the northerly regions of CORDEX-EA, especially the NE and MG regions, where the most rapid increases were in November (more than 4.5°C). Furthermore, we also found that the increases in monthly temperature varied by latitude. For example, the MG and NE regions exhibited similar increasing patterns; the same was true for the KJ, NC, NW and TP regions, as well as for the SEA, India and InC regions. The SC region was dissimilar to the other regions. According to Hui et al. (2019), this may be due to the cloud cover, but determining detailed reasons for how the pattern of increases in temperature varies with latitude needs further study. Finally, the monthly warming pattern in the subregions of China under the RCP4.5 scenario ranged from 0.8°C to 4.2°C. These values are similar to but larger than findings in a previous study (0.3°C–2.2°C) (Gu et al.,

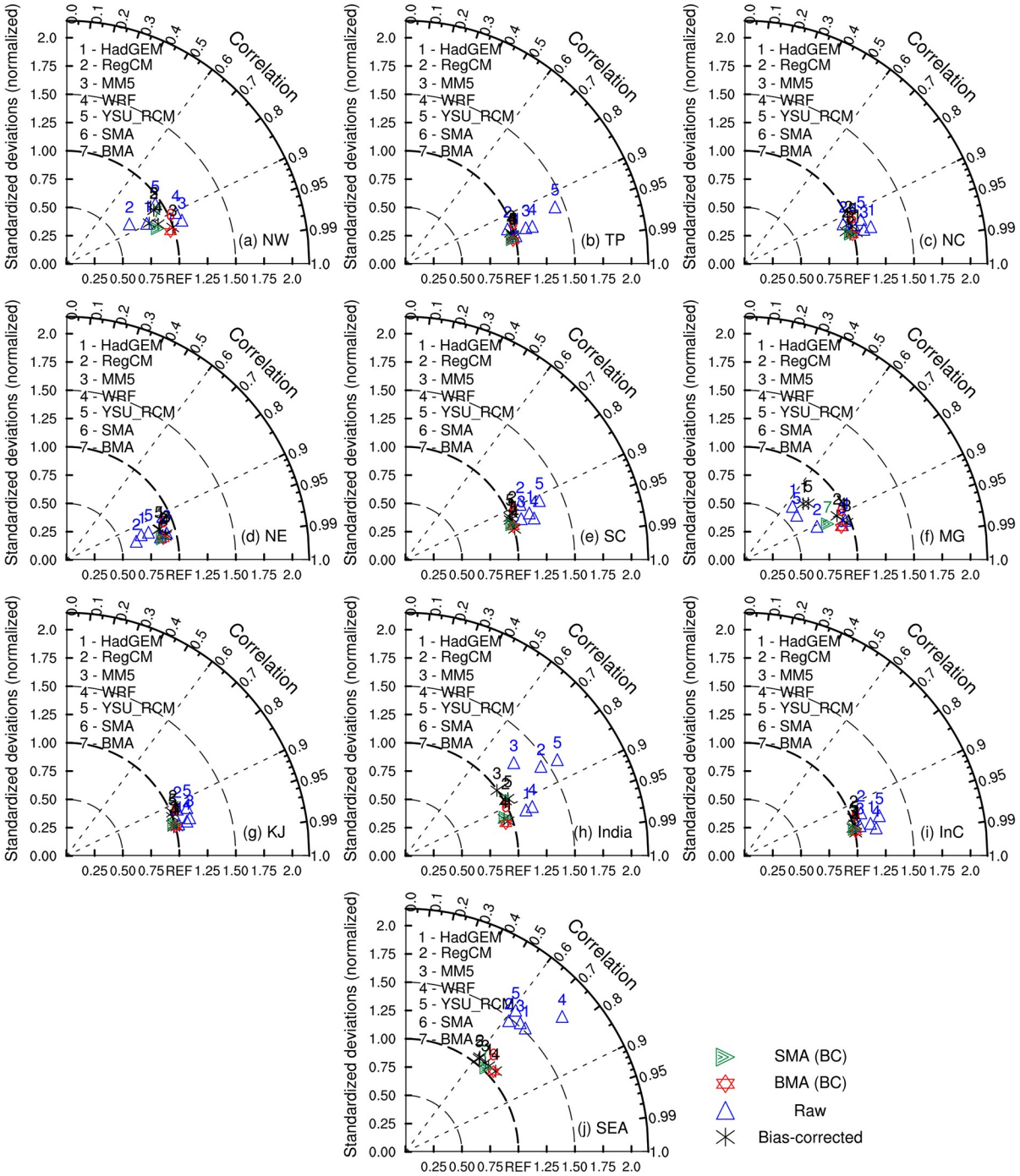


**Fig. 8**. Taylor diagrams evaluating the model skill in simulating the annual temperature and bias correction effects over 10 subregions. CRU observation data used as reference. The *x*- and *y*-axes refer to the standard deviations (normalized) and the azimuthal axis refers to the spatial pattern correlation between two fields.

2018), which was based on the raw temperature and simple multi-model ensemble averaging.

The uncertainty in projections needs to be considered when using them in applications (such as driving hydrological models). Here, we took the standard deviation among five RCMs as an uncertainty indicator (Nordhaus, 2018). Figure 12 gives the uncertainty of the raw model outputs and the bias-corrected results for both the RCP4.5 and RCP8.5 scenarios. The results show that the uncertainty had no apparent relationship with time or solar radiation forcing (RCP4.5 and RCP8.5). The results were different from uncertainties projected by GCMs in Asia, where the uncertainty increases with time (Miao et al., 2016). This may be due to the fact that the RCMs in CORDEX-EA use the same driving GCM, so there is only internal variability (Chen et al., 2019). The uncertainty was greatest for the TP and NW regions, exceeding 2.5°C both for the RCP4.5 and RCP8.5 scenarios throughout the projection period. For the SEA, KJ

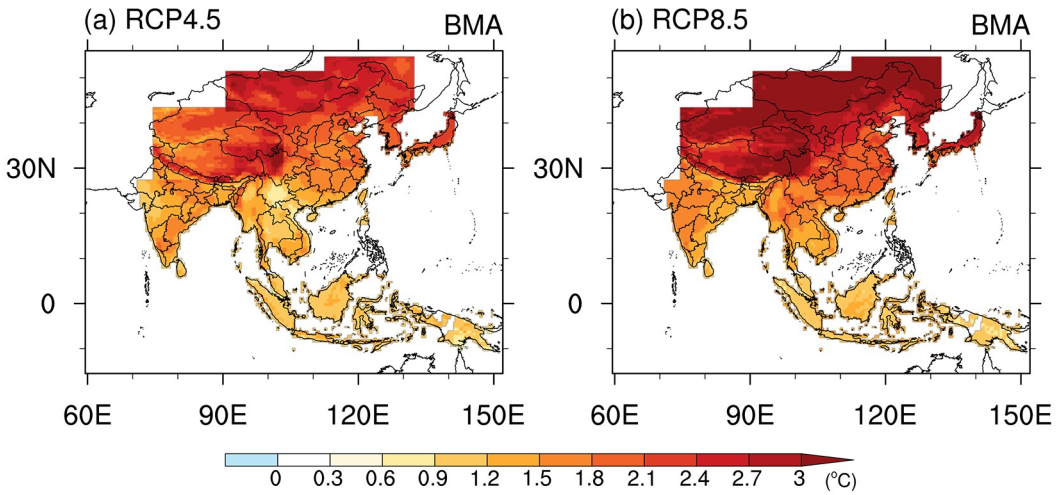


**Fig. 9**. Spatial distribution of annual temperature changes [RCP 8.5 (4.5) minus baseline] projected by five RCMs and the BMA method. See supplementary material for results of five RCMs and the driving GCM under RCP4.5 (Fig. S6) and RCP8.5 (Fig. S7).

**Table 2**. Annual temperature changes (future minus baseline) in °C projected by five RCMs and the BMA method for 10 subregions under the RCP4.5 and RCP8.5 scenarios.

| Subregion | Scenario | HadGEM3-RA | RegCM | MM5 | WRF | YSU-RCM | BMA |
|---|---|---|---|---|---|---|---|
| Northwestern China | RCP 4.5 | 2.6 | 3.1 | 1.8 | 3.4 | 2.7 | 2.1 |
| | RCP 8.5 | 2.9 | 3.8 | 3.2 | 3.2 | 3.3 | 3.1 |
| Tibetan Plateau | RCP 4.5 | 2.5 | 2.4 | 2.2 | 2.0 | 1.8 | 2.2 |
| | RCP 8.5 | 2.9 | 2.9 | 2.9 | 3.0 | 2.2 | 2.7 |
| Northern China | RCP 4.5 | 1.7 | 1.9 | 2.5 | 1.2 | 1.9 | 1.9 |
| | RCP 8.5 | 2.0 | 2.4 | 2.6 | 2.5 | 2.4 | 2.3 |
| Northeastern China | RCP 4.5 | 2.7 | 2.8 | 3.5 | 0.5 | 2.2 | 2.4 |
| | RCP 8.5 | 3.2 | 3.5 | 3.6 | 3.2 | 3.2 | 3.2 |
| Southern China | RCP 4.5 | 1.5 | 1.5 | 2.1 | 1.0 | 1.6 | 1.4 |
| | RCP 8.5 | 1.7 | 1.8 | 2.0 | 1.9 | 1.9 | 1.8 |
| Mongolia | RCP 4.5 | 2.9 | 3.1 | 3.0 | 3.0 | 3.0 | 2.6 |
| | RCP 8.5 | 3.4 | 3.9 | 3.9 | 3.3 | 4.0 | 3.5 |
| Korean Peninsula & Japan | RCP 4.5 | 2.0 | 2.0 | 2.4 | 1.8 | 2.1 | 2.0 |
| | RCP 8.5 | 2.2 | 2.4 | 2.6 | 2.6 | 2.6 | 2.4 |
| India | RCP 4.5 | 1.6 | 1.4 | 1.1 | 1.3 | 1.3 | 1.4 |
| | RCP 8.5 | 1.8 | 1.6 | 1.7 | 1.3 | 1.4 | 1.5 |
| Indochina | RCP 4.5 | 1.4 | 1.2 | 1.6 | 1.1 | 1.3 | 1.3 |
| | RCP 8.5 | 1.6 | 1.5 | 1.8 | 1.6 | 1.5 | 1.6 |
| Southeast Asia | RCP 4.5 | 1.1 | 0.9 | 0.9 | 1.0 | 0.9 | 1.0 |
| | RCP 8.5 | 1.4 | 1.1 | 1.2 | 1.1 | 1.1 | 1.2 |

and India subregions, the uncertainty was lowest–less than 0.7°C for both scenarios. This indicates that there was more uncertainty in the high-latitude subregions. Several previous studies have also shown that the uncertainty was greater at high latitude (Deser et al., 2012; Miao et al., 2016; Woldemeskel et al., 2016). Because the QM method narrowed the differences among RCMs, the uncertainty was reduced for most subregions and both scenarios (except for the NE region under the RCP4.5 scenario). The reductions were more remarkable for the RCP8.5 scenario, ranging from 66% to 94% across all subregions. The lower uncertainty in the RCP8.5 scenario indicates a consistent warming trend under the scenario for all subregions.

## 4. Summary and conclusions

In this study, five RCMs that participated in CORDEX-EA, with lateral and boundary forcing from the HadGEM2-AO model, were used to derive temperature projections over 10 subregions in CORDEX-EA under the RCP4.5 and RCP8.5 scenarios. To remove the biases in RCMs and narrow uncertainty in the projections, bias correction methods and ensemble calibration methods were used. At the same time, we analyzed the performance of bias correction and

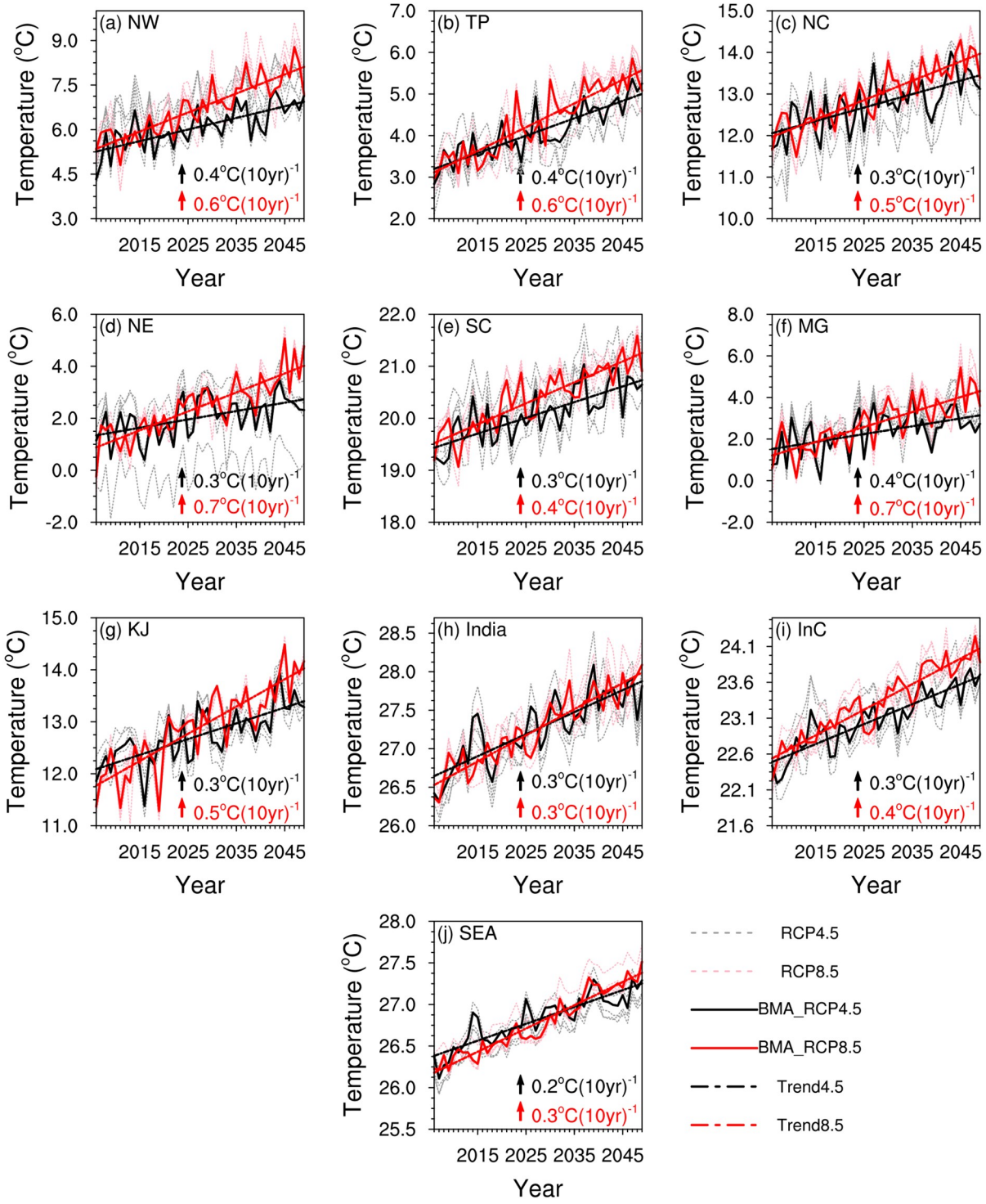


**Fig. 10**. Results from five RCMs and BMA for annual temperature change under two scenarios, RCP4.5 and 8.5. The straight lines, marked with uneven dashes, represent the trends of the temperature changes.

ensemble calibration methods over different subregions and RCMs. The major findings of the study can be summarized as follows:

(1) The RCMs all revealed reasonable representation of annual average temperature when compared with the reference data (CRU TS 4.03). All RCMs presented cold biases in most subregions, especially in regions with complex topography (e.g., the TP and NW regions). Moreover, the seasonal analysis shows that the cold bias mainly came from DJF, while the biases were relatively small in JJA. Therefore, there are obvious biases in the RCM dynamic downscaling results, and it is necessary to conduct bias correction.

(2) The validation results of the bias correction show that all bias correction methods could reduce the simulation biases, and the QM method was most effective in all subregions and RCMs. Compared to the additive scaling and vari-

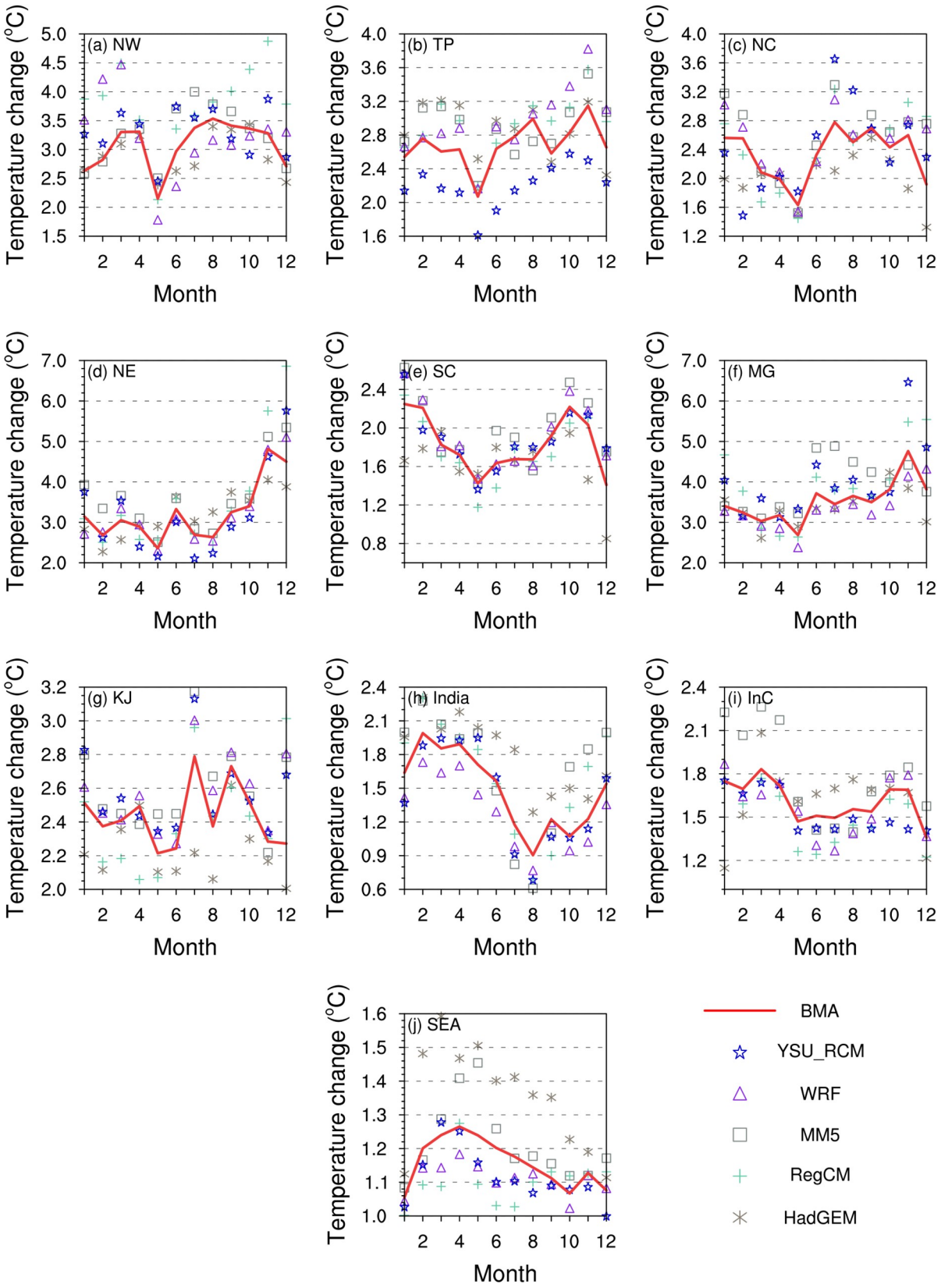


**Fig. 11**. Projected monthly temperature changes (RCP 8.5 minus baseline) for 10 subregions. The monthly temperature changes for the RCP4.5 scenario are similar to RCP8.5 and not shown.

ance scaling methods, the QM method corrected not only the mean but also the quantiles of the distribution. We found that the QM method could reduce the simulation bias by more than 50% when compared to the raw model output.

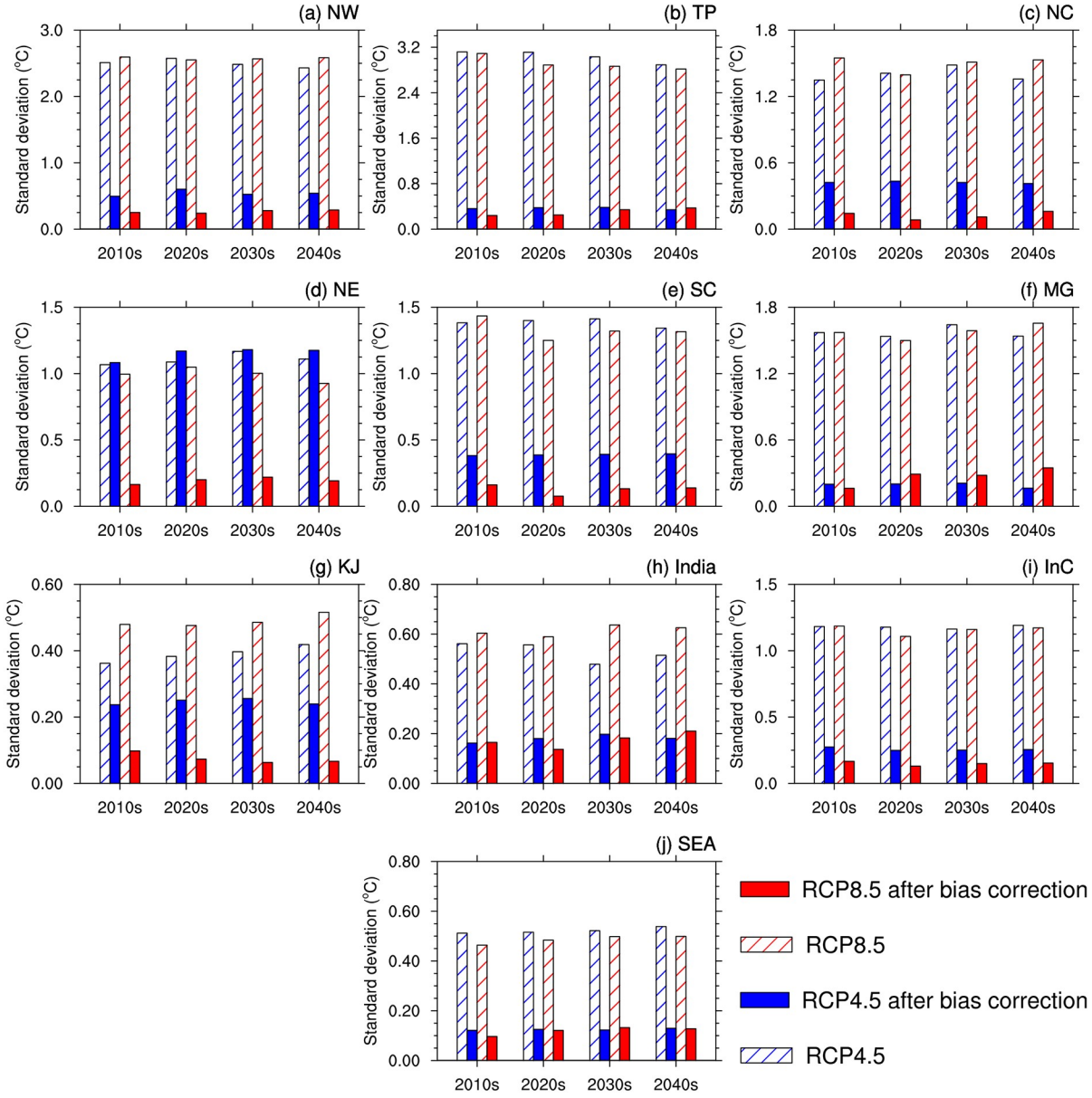


**Fig. 12**. Uncertainty of raw model data and bias-corrected data under two scenarios. We used the standard deviation as a measure of uncertainty.

For the TP and NW regions, the relative decreases in RMSE reached 61.5% and 80.7%, respectively. The spatial distribution of relative decreases in MAE further confirmed the results for RMSE. Finally, the QM method was used to correct the temperatures in projections of the future.

(3) The SMA and BMA methods could further narrow the uncertainty in the bias-corrected ensemble. Although the performances of the SMA and BMA methods were similar in some subregions, the BMA method performed better overall. Compared to the individual RCMs, the BMA method improved the correlation coefficient from 2% to 32% over the 10 subregions. Note that the BMA method may produce poorer results for interannual variability of the temperature, since it focuses only on the bias and trend of the data.

(4) For temperature projections (2030–49), both single RCMs and the BMA results showed consistent warming over all subregions under both scenarios (RCP4.5 and 8.5). BMA results indicated that the warming ranged from 1.2°C to 3.5°C over the 10 subregions under the RCP8.5 scenario (and from 1.0°C to 2.6°C for RCP4.5). The warming was more pronounced in the northern part of the CORDEX-EA domain. The monthly temperature changes seem to vary by latitude, and the detailed reasons need further study.

(5) Furthermore, the QM method reduced uncertainty in temperature projections for all subregions. The reduction was more notable for the RCP8.5 scenario for all subregions. The QM method decreased the uncertainty among the 10 subregions by 66%–94% for the RCP8.5 scenario.

***Acknowledgements***. This work was supported by the Stra-

tegic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA20060401) and the National Natural Science Foundation of China (Grant Nos. 41622101 and 41877155). We acknowledge each of the CORDEX-EA modeling groups for making their simulations available for analysis, including the National Institute of Meteorological Research, three universities in the Republic of Korea (Seoul National University, Yonsei University and Kongju National University), and the World Climate Research Programme's Working Group on the Coordinated Regional Climate Downscaling Experiment, for making the CORDEX data set available (http://cordex-ea.climate.go.kr/cordex/treePage.do).

***Electronic supplementary material:*** Supplementary material is available in the online version of this article at https://doi.org/10.1007/s00376-020-0026-6.

## REFERENCES

Akhtar, M., N. Ahmad, and M. J. Booij, 2009: Use of regional climate model simulations as input for hydrological models for the Hindukush-Karakorum-Himalaya region. *Hydrology and Earth System Sciences*, **13**, 1075–1089, https://doi.org/10.5194/hess-13-1075-2009.

Arsenault, R., P. Gatien, B. Renaud, F. Brissette, and J.-L. Martel, 2015: A comparative analysis of 9 multi-model averaging approaches in hydrological continuous streamflow simulation. *J. Hydrol.*, **529**, 754–767, https://doi.org/10.1016/j.jhydrol.2015.09.001.

Ashfaq, M., D. Rastogi, R. Mei, D. Touma, and L. R. Leung, 2017: Sources of errors in the simulation of south Asian summer monsoon in the CMIP5 GCMs. *Climate Dyn.*, **49**, 193–223, https://doi.org/10.1007/s00382-016-3337-7.

Au-Yeung, A. Y. M., and J. C. L. Chan, 2012: Potential use of a regional climate model in seasonal tropical cyclone activity predictions in the western North Pacific. *Climate Dyn.*, **39**, 783–794, https://doi.org/10.1007/s00382-011-1268-x.

Ayugi, B., and Coauthors, 2020: Quantile mapping bias correction on Rossby Centre Regional Climate Models for precipitation analysis over Kenya, East Africa. *Water*, **12**, 801, https://doi.org/10.3390/w12030801.

Baek, H.-J., and Coauthors, 2013: Climate change in the 21st century simulated by HadGEM2-AO under representative concentration pathways. *Asia-Pacific Journal of Atmospheric Sciences*, **49**, 603–618, https://doi.org/10.1007/s13143-013-0053-7.

Bao, J. W., J. M. Feng, and Y. L. Wang, 2015: Dynamical downscaling simulation and future projection of precipitation over China. *J. Geophys. Res.*, **120**, 8227–8243, https://doi.org/10.1002/2015JD023275.

Bennett, J. C., M. R. Grose, S. P. Corney, C. J. White, G. K. Holz, J. J. Katzfey, D. A. Post, and N. L. Bindoff, 2014: Performance of an empirical bias-correction of a high-resolution climate dataset. *International Journal of Climatology*, **34**, 2189–2204, https://doi.org/10.1002/joc.3830.

Berg, P., H. Feldmann, and H.-J. Panitz, 2012: Bias correction of high resolution regional climate model data. *J. Hydrol.*, **448–449**, 80–92, https://doi.org/10.1016/j.jhydrol.2012.04.026.

Cha, D.-H., D.-K. Lee, and S.-Y. Hong, 2008: Impact of boundary layer processes on seasonal simulation of the East Asian summer monsoon using a regional climate model. *Meteorol. Atmos. Phys.*, **100**, 53–72, https://doi.org/10.1007/s00703-008-0295-6.

Chang, C.-P., Y. H. Lei, C.-H. Sui, X. H. Lin, and F. M. Ren, 2012: Tropical cyclone and extreme rainfall trends in East Asian summer monsoon since mid-20th century. *Geophys. Res. Lett.*, **39**, L18702, https://doi.org/10.1029/2012GL052945.

Chen, S. F., R. G. Wu, and W. Chen, 2019: Projections of climate changes over mid-high latitudes of Eurasia during boreal spring: Uncertainty due to internal variability. *Climate Dyn.*, **53**, 6309–6327, https://doi.org/10.1007/s00382-019-04929-4.

Chen, X. L., Y. M. Liu, and G. X. Wu, 2017: Understanding the surface temperature cold Bias in CMIP5 AGCMs over the Tibetan Plateau. *Adv. Atmos. Sci.*, **34**, 1447–1460, https://doi.org/10.1007/s00376-017-6326-9.

Deser, C., A. Phillips, V. Bourdette, and H. Y. Teng, 2012: Uncertainty in climate change projections: The role of internal variability. *Climate Dyn.*, **38**, 527–546, https://doi.org/10.1007/s00382-010-0977-x.

Ding, Y. H., and J. C. L. Chan, 2005: The East Asian summer monsoon: An overview. *Meteorol. Atmos. Phys.*, **89**, 117–142, https://doi.org/10.1007/s00703-005-0125-z.

Duan, Q. Y., N. K. Ajami, X. G. Gao, and S. Sorooshian, 2007: Multi-Model ensemble hydrologic prediction using Bayesian model averaging. *Advances in Water Resources*, **30**, 1371–1386, https://doi.org/10.1016/j.advwatres.2006.11.014.

Duan, Q. Y., and T. J. Phillips, 2010: Bayesian estimation of local signal and noise in multimodel simulations of climate change. *J. Geophys. Res.*, **115**, D18123, https://doi.org/10.1029/2009JD013654.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project Phase 6(CMIP6) experimental design and organization. *Geoscientific Model Development*, **9**, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016.

Ezéchiel, O., A. A. Eric, Z. E. Josué, B. I. Eliézer, C. Amédée, and A. Abel, 2016: Comparative study of seven bias correction methods applied to three regional climate models in Mekrou catchment (Benin, West Africa). *International Journal of Current Engineering and Technology*, **6**, 1831–1840.

Fang, G. H., J. Yang, Y. N. Chen, and C. Zammit, 2015: Comparing bias correction methods in downscaling meteorological variables for a hydrologic impact study in an arid area in China. *Hydrology and Earth System Sciences*, **19**, 2547–2559, https://doi.org/10.5194/hess-19-2547-2015.

Fragoso, T. M., W. Bertoli, and F. Louzada, 2018: Bayesian model averaging: A systematic review and conceptual classification. *International Statistical Review*, **86**, 1–28, https://doi.org/10.1111/insr.12243.

Fulakeza, M., L. M. Druyan, and T. N. Krishnamurti, 2002: A simple soil moisture scheme for regional climate simulations in the tropics. *Meteorol. Atmos. Phys.*, **79**, 105–126, https://doi.org/10.1007/s703-002-8231-7.

Gao, X. J., and F. Giorgi, 2017: Use of the RegCM System over East Asia: Review and perspectives. *Engineering*, **3**, 766–772, https://doi.org/10.1016/J.ENG.2017.05.019.

Giorgi, F., C. Jones, and G. R. Asrar, 2009: Addressing climate

information needs at the regional level: The CORDEX framework. *WMO Bulletin*, **58**, 175–183.

Giorgi, F., and Coauthors, 2012: RegCM4: Model description and preliminary tests over multiple CORDEX domains. *Climate Research*, **52**, 7–29, https://doi.org/10.3354/cr01018.

Grimm, N. B., and Coauthors, 2013: The impacts of climate change on ecosystem structure and function. *Frontiers in Ecology and the Environment*, **11**, 474–482, https://doi.org/10.1890/120282.

Gu, H. H., Z. B. Yu, C. G Yang, Q. Ju, T. Yang, and D. W. Zhang, 2018: High-resolution ensemble projections and uncertainty assessment of regional climate change over China in CORDEX East Asia. *Hydrology and Earth System Sciences*, **22**, 3087–3103, https://doi.org/10.5194/hess-22-3087-2018.

Gulizia, C., and I. Camilloni, 2015: Comparative analysis of the ability of a set of CMIP3 and CMIP5 global climate models to represent precipitation in South America. *International Journal of Climatology*, **35**, 583–595, https://doi.org/10.1002/joc.4005.

Guo, D.-L., J.-Q. Sun, and E.-T. Yu, 2018: Evaluation of CORDEX regional climate models in simulating temperature and precipitation over the Tibetan Plateau. *Atmos. Ocean. Sci. Lett.*, **11**, 219–227, https://doi.org/10.1080/16742834.2018.1451725.

Gutowski, W. J., and Coauthors, 2016: WCRP COordinated Regional Downscaling EXperiment (CORDEX): A diagnostic MIP for CMIP6. *Geoscientific Model Development*, **9**, 4087–4095, https://doi.org/10.5194/gmd-9-4087-2016.

Halder, S., S. K. Saha, P. A. Dirmeyer, T. N. Chase, and B. N. Goswami, 2016: Investigating the impact of land-use land-cover change on Indian summer monsoon daily rainfall and temperature during 1951–2005 using a regional climate model. *Hydrology and Earth System Sciences*, **20**, 1765–1784, https://doi.org/10.5194/hess-20-1765-2016.

Ham, S., J.-W. Lee, and K. Yoshimura, 2016: Assessing future climate changes in the East Asian summer and winter monsoon using regional spectral model. *J. Meteor. Soc. Japan*, **94A**, 69–87, https://doi.org/10.2151/jmsj.2015-051.

Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister, 2014: Updated high-resolution grids of monthly climatic observations--the CRU TS3. *10 Dataset. International Journal of Climatology*, **34**, 623–642, https://doi.org/10.1002/joc.3711.

Hijioka, Y., and Coauthors, 2014: Asia. *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel of Climate Change*, Barros et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1327-1370.

Huang, J. P., H. P. Yu, X. D. Guan, G. Y. Wang, and R. X. Guo, 2016: Accelerated dryland expansion under climate change. *Nature Climate Change*, **6**, 166–171, https://doi.org/10.1038/NCLIMATE2837.

Hui, P. H., Y. Li, Y. Chen, L. L. Zhang, F. F. Wei, S. Y. Wang, and J. P. Tang, 2019: The impact of radiation parameterization schemes on the regional climate simulations over the CORDEX-EA domain. *Atmospheric Research*, **224**, 81–98, https://doi.org/10.1016/j.atmosres.2019.03.020.

IPCC, 2001: *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, Houghton et al., Eds., Cambridge University Press, 881 pp.

IPCC, 2013: AISM-Annex I: Atlas of global and regional climate projections supplementary material RCP6.0. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Stocker et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1311-1394.

IPCC, 2014: *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Field et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1332 pp.

Jones, A. R., and N. A. Brunsell, 2009: A scaling analysis of soil moisture-precipitation interactions in a regional climate model. *Theor. Appl. Climatol.*, **98**, 221–235, https://doi.org/10.1007/s00704-009-0109-x.

Kang, S., and E. A. B. Eltahir, 2018: North China Plain threatened by deadly heatwaves due to climate change and irrigation. *Nature Communications*, **9**, 2894, https://doi.org/10.1038/s41467-018-05252-y.

Kim, C., and M.-S. Suh, 2013: Prospects of using Bayesian model averaging for the calibration of one-month forecasts of surface air temperature over South Korea. *Asia-Pacific Journal of Atmospheric Sciences*, **49**, 301–311, https://doi.org/10.1007/s13143-013-0029-7.

Kim, Y., M. Jun, S.-K. Min, M.-S. Suh, and H.-S. Kang, 2016: Spatial analysis of future East Asian seasonal temperature using two regional climate model simulations. *Asia-Pacific Journal of Atmospheric Sciences*, **52**, 237–249, https://doi.org/10.1007/s13143-016-0022-z.

Lee, J.-Y., and Coauthors, 2017: The long-term variability of Changma in the East Asian summer monsoon system: A review and revisit. *Asia-Pacific Journal of Atmospheric Sciences*, **53**, 257–272, https://doi.org/10.1007/s13143-017-0032-5.

Li, D. L., B. S. Yin, J. L. Feng, A. Dosio, B. Geyer, J. F. Qi, H. Q. Shi, and Z. H. Xu, 2018a: Present climate evaluation and added value analysis of dynamically downscaled simulations of CORDEX–East Asia. *J. Appl. Meteorol. Climatol.*, **57**, 2317–2341, https://doi.org/10.1175/JAMC-D-18-0008.1.

Li, J. P., H.-H. Hsu, W.-C. Wang, K.-J. Ha, T. Li, and A. Kitoh, 2018b: East Asian climate under global warming: Understanding and projection. *Climate Dyn.*, **51**, 3969–3972, https://doi.org/10.1007/s00382-018-4523-6.

Luo, M., T. Liu, F. H. Meng, Y. C. Duan, A. Frankl, A. M. Bao, and P. De Maeyer, 2018: Comparing bias correction methods used in downscaling precipitation and temperature from regional climate models: A case study from the Kaidu River basin in Western China. *Water*, **10**, 1046, https://doi.org/10.3390/w10081046.

Mann, M. E., and P. H. Gleick, 2015: Climate change and California drought in the 21st century. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 3858–3859, https://doi.org/10.1073/pnas.1503667112.

Martin, G. M., and Coauthors, 2011: The HadGEM2 family of Met Office Unified Model climate configurations. *Geoscientific Model Development*, **4**, 723–757, https://doi.org/10.5194/gmd-4-723-2011.

McSweeney, C. F., R. G. Jones, R. W. Lee, and D. P. Rowell, 2015: Selecting CMIP5 GCMs for downscaling over mul-

tiple regions. *Climate Dyn.*, **44**, 3237–3260, https://doi.org/10.1007/s00382-014-2418-8.

Mearns, L. O., D. P. Lettenmaier, and S. McGinnis, 2015: Uses of results of regional climate model experiments for impacts and adaptation studies: The example of NARCCAP. *Current Climate Change Reports*, **1**, 1–9, https://doi.org/10.1007/s40641-015-0004-8.

Meng, X., and Coauthors, 2018: Simulated cold bias being improved by using MODIS time-varying albedo in the Tibetan Plateau in WRF model. *Environmental Research Letters*, **13**, 044028, https://doi.org/10.1088/1748-9326/aab44a.

Miao, C. Y., and Coauthors, 2014: Assessment of CMIP5 climate models and projected temperature changes over Northern Eurasia. *Environmental Research Letters*, **9**, 055007, https://doi.org/10.1088/1748-9326/9/5/055007.

Miao, C. Y., Q. Y. Duan, Q. H. Sun, and J. D. Li, 2013: Evaluation and application of Bayesian multi-model estimation in temperature simulations. *Progress in Physical Geography: Earth and Environment*, **37**, 727–744, https://doi.org/10.1177/0309133313494961.

Miao, C. Y., L. Su, Q. H. Sun, and Q. Y. Duan, 2016: A nonstationary bias-correction technique to remove bias in GCM simulations. *J. Geophys. Res.*, **121**, 5718–5735, https://doi.org/10.1002/2015JD024159.

Miao, C. Y., Q. Y. Duan, Q. H. Sun, X. H. Lei, and H. Li, 2019: Non-uniform changes in different categories of precipitation intensity across China and the associated large-scale circulations. *Environmental Research Letters*, **14**, 025004, https://doi.org/10.1088/1748-9326/aaf306.

Ngai, S. T., F. Tangang, and L. Juneng, 2017: Bias correction of global and regional simulated daily precipitation and surface mean temperature over Southeast Asia using quantile mapping method. *Global and Planetary Change*, **149**, 79–90, https://doi.org/10.1016/j.gloplacha.2016.12.009.

Nordhaus, W., 2018: Projections and uncertainties about climate change in an Era of minimal climate policies. *American Economic Journal: Economic Policy*, **10**, 333–360, https://doi.org/10.1257/pol.20170046.

Panofsky, H. A. and G.W. Brier, 1968: Some applications of statistics to meteorology. Earth and Mineral Science Continuing Education, College of Earth and Mineral Sciences.

Park, C., and Coauthors, 2016: Evaluation of multiple regional climate models for summer climate extremes over East Asia. *Climate Dyn.*, **46**, 2469–2486, https://doi.org/10.1007/s00382-015-2713-z.

Prömmel, K., B. Geyer, J. M. Jones, and M. Widmann, 2010: Evaluation of the skill and added value of a reanalysis-driven regional simulation for Alpine temperature. *International Journal of Climatology*, **30**, 760–773, https://doi.org/10.1002/joc.1916.

Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski, 2005: Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Wea. Rev.*, **133**, 1155–1174, https://doi.org/10.1175/MWR2906.1.

Rocheta, E., J. P. Evans, and A. Sharma, 2017: Can bias correction of regional climate model lateral boundary conditions improve low-frequency rainfall variability? *J. Climate*, **30**, 9785–9806, https://doi.org/10.1175/JCLI-D-16-0654.1.

Ruan, Y. F., Z. F. Liu, R. Wang, and Z. J. Yao, 2019: Assessing the performance of CMIP5 GCMs for projection of future temperature change over the Lower Mekong Basin. *Atmosphere*, **10**, 93, https://doi.org/10.3390/atmos10020093.

Salzmann, N., J. Nötzli, C. Hauck, S. Gruber, M. Hoelzle, and W. Haeberli, 2007: Ground surface temperature scenarios in complex high-mountain topography based on regional climate model results. *J. Geophys. Res.*, **112**, F02S12, https://doi.org/10.1029/2006JF000527.

Schlaepfer, D. R., and Coauthors, 2017: Climate change reduces extent of temperate drylands and intensifies drought in deep soils. *Nature Communications*, **8**, 14196, https://doi.org/10.1038/ncomms14196.

Singh, A., R. K. Sahoo, A. Nair, U. C. Mohanty, and R. K. Rai, 2017: Assessing the performance of bias correction approaches for correcting monthly precipitation over India through coupled models. *Meteorological Applications*, **24**, 326–337, https://doi.org/10.1002/met.1627.

Soden, B. J., W. D. Collins, and D. R. Feldman, 2018: Reducing uncertainties in climate models. *Science*, **361**, 326–327, https://doi.org/10.1126/science.aau1864.

Sperber, K. R., H. Annamalai, I.-S. Kang, A. Kitoh, A. Moise, A. Turner, B. Wang, and T. Zhou, 2013: The Asian summer monsoon: An intercomparison of CMIP5 vs. *CMIP3 simulations of the late 20th century. Climate Dyn.*, **41**, 2711–2744, https://doi.org/10.1007/s00382-012-1607-6.

Sun, Q. H., C. Y. Miao, and Q. Y. Duan, 2015: Comparative analysis of CMIP3 and CMIP5 global climate models for simulating the daily mean, maximum, and minimum temperatures and daily precipitation over China. *J. Geophys. Res.*, **120**, 4806–4824, https://doi.org/10.1002/2014JD022994.

Sun, Q. H., C. Y. Miao, and Q. Y. Duan, 2016: Extreme climate events and agricultural climate indices in China: CMIP5 model evaluation and projections. *International Journal of Climatology*, **36**, 43–61, https://doi.org/10.1002/joc.4328.

Sun, Q. H., C. Y. Miao, A. AghaKouchak, I. Mallakpour, D. Y. Ji, and Q. Y. Duan, 2020: Possible increased frequency of ENSO-related dry and wet conditions over some major watersheds in a warming climate. *Bull. Amer. Meteor. Soc.*, **101**, E409–E426, https://doi.org/10.1175/BAMS-D-18-0258.1.

Tang, J. P., and Coauthors, 2016: Building Asian climate change scenario by multi-regional climate models ensemble. *Part I: Surface air temperature. International Journal of Climatology*, **36**, 4241–4252, https://doi.org/10.1002/joc.4628.

Tang, J. P., X. G. Sun, P. H. Hui, Y. Li, Q. Zhang, and J. Y. Liu, 2018: Effects of spectral nudging on precipitation extremes and diurnal cycle over CORDEX-East Asia domain. *International Journal of Climatology*, **38**, 4903–4923, https://doi.org/10.1002/joc.5706.

Tangang, F., and Coauthors, 2015: The Southeast Asia Regional Climate Downscaling (SEACLID) / CORDEX Southeast Asia project and the results of its sensitivity experiments of RegCM4 cumulus and ocean fluxes parameterization schemes on temperature and extremes. *Proc. EGU General Assembly Conference*, EGU, Vienna, Austria.

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

Terink, W., R. T. W. L. Hurkmans, P. J. J. F. Torfs, and R. Uijlenhoet, 2010: Evaluation of a bias correction method applied to downscaled precipitation and temperature reanalysis data for the Rhine basin. *Hydrology and Earth System Sciences*, **14**, 687–703, https://doi.org/10.5194/hess-14-687-2010.

Teutschbein, C., and J. Seibert, 2010: Regional climate models for hydrological impact studies at the catchment scale: A review of recent modeling strategies. *Geography Compass*,

**4**, 834–860, https://doi.org/10.1111/j.1749-8198.2010.00357.x.

Teutschbein, C., and J. Seibert, 2012: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *J. Hydrol.*, **456–457**, 12–29, https://doi.org/10.1016/j.jhydrol.2012.05.052.

von Storch, H., H. Langenberg, and F. Feser, 2000: A spectral nudging technique for dynamical downscaling purposes. *Mon. Wea. Rev.*, **128**, 3664–3673, https://doi.org/10.1175/1520-0493(2000)128<3664:ASNTFD>2.0.CO;2.

Wang, D. N., C. Menz, T. Simon, C. Simmer, and C. Ohlwein, 2013: Regional dynamical downscaling with CCLM over East Asia. *Meteorol. Atmos. Phys.*, **121**, 39–53, https://doi.org/10.1007/s00703-013-0250-z.

Wilcke, R. A. I., T. Mendlik, and A. Gobiet, 2013: Multi-variable error correction of regional climate models. *Climatic Change*, **120**, 871–887, https://doi.org/10.1007/s10584-013-0845-x.

Woldemeskel, F. M., A. Sharma, B. Sivakumar, and R. Mehrotra, 2016: Quantification of precipitation and temperature uncertainties simulated by CMIP3 and CMIP5 models. *J. Geophys. Res.*, **121**, 3–17, https://doi.org/10.1002/2015JD023719.

Yin, Z. L., Q. Feng, L. S. Yang, R. C. Deo, J. F. Adamowski, X. H. Wen, B. Jia, and J. H. Si, 2020: Projected spatial patterns in precipitation and air temperature for China's northwest region derived from high-resolution regional climate models. *International Journal of Climatology*, **40**, 3922–3941, https://doi.org/10.1002/joc.6435.

Zheng, H. Y., C. Y. Miao, J. W. Wu, X. H. Lei, W. H. Liao, and H. Li, 2019: Temporal and spatial variations in water discharge and sediment load on the Loess Plateau, China: A high-density study. *Science of the Total Environment*, **666**, 875–886, https://doi.org/10.1016/j.scitotenv.2019.02.246.

Zhou, W. D., J. P. Tang, X. Y. Wang, S. Y. Wang, X. R. Niu, and Y. Wang, 2016: Evaluation of regional climate simulations over the CORDEX-EA-II domain using the COSMO-CLM model. *Asia-Pacific Journal of Atmospheric Sciences*, **52**, 107–127, https://doi.org/10.1007/s13143-016-0013-0.

Zou, L. W., and T. J. Zhou, 2016: Future summer precipitation changes over CORDEX-East Asia domain downscaled by a regional ocean-atmosphere coupled model: A comparison to the stand-alone RCM. *J. Geophys. Res.*, **121**, 2691–2704, https://doi.org/10.1002/2015JD024519.