# A weighted ensemble of regional climate projections for exploring the spatiotemporal evolution of multidimensional drought risks in a changing climate

**B. Zhang[1], S. Wang[1,2,*], J. Zhu[3]**

[1]Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China

[2]The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen, China

[3]School of Geography and Planning, Sun Yat-Sen University, Guangzhou, China

Corresponding author: Phone: (852) 3400-3896; Email: shuo.s.wang@polyu.edu.hk

**Abstract**

Understanding future drought risks plays a crucial role in developing climate change adaptation strategies and in enhancing disaster resilience. However, previous studies may lead to biased conclusions due to the neglect of two factors, including the relative performance of climate simulations and the uncertainty in drought characterization. In this study, Bayesian model averaging is used to merge five regional climate model simulations and to project future changes in hydroclimatic regimes over China under two representative emission scenarios (RCP4.5 and RCP8.5). Drought characteristics, including drought severity and duration, are extracted using the Standardized Precipitation Evapotranspiration Index (SPEI). A Bayesian copula approach is used to uncover underlying interactions of drought characteristics and associated uncertainties across 10 climate divisions of China. The regional return periods of drought characteristics are used to assess future changes in multidimensional drought risks and the probability of extreme droughts. Our findings reveal that the variations in drought characteristics are generally underestimated by the ensemble mean (AEM) simulation. The Bayesian framework improves the reliability and accuracy of hydroclimate simulations and better reproduces the drought regimes compared to the AEM simulation. The drought duration and severity are projected to substantially increase for most areas of China based on the Bayesian framework, but the AEM simulation may lead to multiple opposite behaviors, especially under RCP4.5. The estimated joint risk from drought duration and drought severity is expected to increase under both emission scenarios. The likelihood of extreme droughts is also projected to increase as the radiative forcing increases.

**Keywords:** Climate projection; Drought risk; Bayesian model averaging; Copula; China

## 1. Introduction

Droughts, one of the costliest and most widespread natural hazards, have caused massive economic losses, environmental degradation, and even loss of human life around the world (Dai 2013; Samaniego et al. 2018; Su et al. 2018). For example, a severe and prolonged drought episode during 2009 and 2010 affected millions of people and livestock in northern and southwestern China with billions of dollars in economic losses (Barriopedro et al. 2012). Considering the substantial impacts of droughts and the indisputable fact of global warming, assessing the evolution of drought hazards in a changing climate has received considerable attention in recent decades (Prudhomme et al. 2014; Cook et al. 2016; Chen et al. 2020).

Global climate models (GCMs) and regional climate models (RCMs) have been widely used to assess the implications of climate change for future drought hazards (Russo et al. 2013; Van Huijgevoort et al. 2014; Asadi Zarch et al. 2015; Zhu et al. 2019; Qing et al. 2020). The Coordinated Regional Downscaling Experiment (CORDEX) archive provides quite a few RCMs and has played a crucial role in the multi-model ensemble simulations of regional drought events in recent years (Samouly et al. 2018; Zhai et al. 2019; Li et al. 2020; Spinoni et al. 2020). Since each climate model has strengths and weaknesses in characterizing the hydroclimatic regimes, a multi-model ensemble simulation is commonly used to improve the reliability of drought projections. The arithmetic ensemble mean (AEM) of drought variables (e.g., precipitation) and the inter-model spread derived from multiple RCMs are widely used to assess climate change impacts on regional droughts (Parajka et al. 2016; Vidal et al. 2016; Rajsekhar and Gorelick 2017; Lee et al. 2019). Although the AEM simulation reduces the model bias compared to a single climate model, the systematic bias cannot be neglected and would hinder reliable projections of future droughts. An alternative approach of the AEM approach is Bayesian model averaging

(BMA), which has been proven to be a promising tool for improving multi-model hydroclimate simulations (Duan and Phillips 2010; Yang et al. 2011; Olson et al. 2016, 2018; Zhang et al. 2016; Ahmadalipour et al. 2018; Shin et al. 2019; Basher et al. 2020). However, little effort has been directed towards applying BMA to project future drought characteristics (Ahmadalipour et al. 2018; Chen et al. 2020; Miao et al. 2020). It is unclear whether the BMA approach can improve the reliability of climate-induced drought simulations. In addition, it is also unclear whether the AEM and BMA approaches would lead to different drought projections. It is necessary to elucidate these issues for better understanding future drought regimes and thus improving the resilience of water management system.

In addition to climate simulations, drought frequency analysis is also required to assess climate change impacts on drought hazards (Hao and AghaKouchak 2013; Borgomeo et al. 2015; Seager et al. 2015; Williams et al. 2015; Liu et al. 2016b). Since drought characteristics (i.e., drought severity, spatial extent, and duration, etc.) are commonly interdependent, the multivariate frequency analysis has been widely performed to quantify drought hazards and the potential risks (Maity et al. 2013; Kam et al. 2014; Ayantobo et al. 2018). Copula has gained remarkable success in multivariate drought analysis owing to its flexibility in capturing the complicated dependencies between drought characteristics regardless of their marginal distributions (Salvadori and De Michele 2004; AghaKouchak et al. 2014; Ganguli and Reddy 2014; Xu et al. 2015; Liu et al. 2016a; Salvadori et al. 2016; Masud et al. 2017). However, previous studies fail to explicitly address the underlying uncertainties of copula parameters, thus leading to a potential bias in drought risk assessment (Yan 2007). Such uncertainty is considerably large since the samples of drought episodes are typically limited, and ignoring the uncertainty diminishes the scientific credibility in drought assessments (De Michele et al. 2013; Sadegh et al. 2017). Therefore, it is necessary to explicitly address the uncertainty in copula-based multivariate drought assessments for advancing our understanding of complex mechanisms and potential impacts of droughts.

The aforementioned limitations of the AEM climate simulation and the copula-based drought characterization may lead to unreliable projections of future drought hazards. Therefore, in this study, we will develop a probabilistic projection of multidimensional drought hazards through BMA and Bayesian copula. We hypothesize that the reliability of climate-induced drought hazard projections can be improved by taking into account the relative performance of climate models and the uncertainty in drought characterization. Specifically, an ensemble of five regional climate simulations, including four from the CORDEX East Asia experiment and one from the Providing REgional Climate Impacts for Studies (PRECIS) simulation will be used to improve the performance of climate simulations in China based on BMA techniques. Drought episodes will be detected using the Standardized Precipitation-Evapotranspiration Index (SPEI) in 10 climate divisions of China (Vicente-Serrano et al. 2010). Drought hazards will be quantified using the joint return period of duration and severity calculated by a Bayesian copula approach. The hydroclimate regimes and drought characteristics generated from the BMA simulation will be also compared with those generated from the AEM simulation.

This paper is divided into four sections. Section 2 will describe models, algorithms, and datasets used to perform Bayesian multi-model climate simulations and multivariate drought hazard projections. Section 3 will systematically evaluate the BMA-based hydroclimate simulations and assess climate change impacts on multidimensional drought hazards. Finally, Section 4 will provide a summary and conclusions of this study.

## 2. Models, algorithms, and data sources

### 2.1 Bayesian multi-model climate projection

The PRECIS model developed by the UK Hadley Centre, together with four regional climate simulations from CORDEX available for the East Asia domain, were used to assess the changes in hydroclimatic regimes over China. Specifically, the COnsortium for Small-scale MOdelling in CLimate Mode (CCLM) RCM was used to dynamically downscale four Coupled Model Intercomparison Project Phase 5 (CMIP5) GCMs (CNRM-CM5, EC-EARTH, HadGEM2-ES, and MPI-ESM-LR) in the CORDEX East Asia experiment, while the PRECIS model was driven by the HadGEM2-ES (Rockel et al. 2008; Huang et al. 2018; Shrestha and Wang 2020; Zhu et al. 2021). All the five simulations have the same horizontal resolution of about 0.44° × 0.44° (~ 50 km) but differ in the model domain. The computational domain of the PRECIS simulation is configured to extend from about 64.68°E–139.04°E and 13.44°N–56.12°N with 109 × 88 50-km grid points and a lateral buffer zone of 8 grid points (see Fig. 1a). Such a choice of domain size is made by following relevant studies to capture the large-scale circulation and boundary forcing which play important roles in China's regional climatology, such as East Asian winter, summer and tropical oceanic monsoons (Centella-Artola et al. 2015; Guo et al. 2019; Wu et al. 2021). In comparison, the CCLM model domain is slightly different with 203×167 horizontal grid points (see Fig. 1b). The PRECIS climate simulation covers the historical period (1969–2005) and a future period (2006–2099), while the CCLM climate simulation covers the historical period (1951–2005) and a future period (2006–2100). Future simulations for both PRECIS in this study and CCLM in the CORDEX East Asia experiment are forced with two emission scenarios, including RCP4.5 and RCP8.5. The 30-year monthly hydroclimatic variables including precipitation and potential evapotranspiration (PET) for the historical (1975−2004) and future (2069−2098) periods are collected from the five climate projections to assess the impact of climate change on hydrological regimes. The FAO-56 Penman-Monteith Equation was applied to the calculation of PET, which was suggested to yield more realistic estimates than the temperature-only-based Thornthwaite method (Allen et al. 1998; Dai 2013).

Bayesian model averaging (BMA), as an effective tool of correcting under dispersion in ensemble climate projections, was used to improve the accuracy of monthly precipitation and PET simulations. Assume that $x = x_1,\ldots, x_K$ signify the ensemble of all considered climate simulations, and $y$ denotes the climate observations. $p_k(y|x_k)$ represents the conditional probability density function (pdf) of $y$ given $x_k$. The probabilistic forecast pdf of $y$ for the multi-model ensemble can be expressed as

$$p(y \mid x_1 \ldots x_k) = \sum_{k=1}^{K} w_k p_k(y \mid x_k) \tag{1}$$

Where $w_k$ is the BMA weight of model $k$ in the ensemble. The sum of all $w_k$ values is equal to 1 and they are nonnegative, which reflect how well an individual climate simulation matches the observations in the training period. Since a certain distribution cannot be appropriate for all climate variables, the conditional pdf, $p_k(y|x_k)$, is defined as the copula-based conditional probability distribution that has a wide range of parametric distribution as

$$p_k(y \mid x_k) = c_k(u_y, u_{x_k})\, p(y) \tag{2}$$

where $c_k(u_y, u_{x_k})$ represents the joint pdf of $y$ and $x_k$; $u$ represents the cumulative distribution function; $p(y)$ represents the pdf of $y$. Details of copulas are described in Section 2.2. The posterior mean of the BMA simulation can be expressed as

$$E(y \mid x_1 \ldots x_K) = \sum_{k=1}^{K} w_k x_k \tag{3}$$

BMA has been demonstrated to be a powerful approach to combine an ensemble of climate simulations since it is essentially an "intelligent" weighted average forecast based on the model performance (Raftery et al. 2005; Madadgar and Moradkhani 2014; Vrugt 2016; Zhang et al. 2016). Therefore, BMA was applied to monthly precipitation and PET for each grid cell with CRU's (Climatic Research Unit) gridded monthly precipitation and PET dataset as reference. The CRU dataset is a global gauge-based climate variable product with a 0.5° × 0.5° grid resolution based on thousands of weather stations (Harris et al. 2014). The CRU data is also consistent with the in-situ meteorological observations in terms of capturing drought durations and severities in China (see Figs. S7 and S8 of the supplementary material).

The BMA weights were estimated using the MCMC simulation instead of the EM algorithm. The MCMC simulation has been demonstrated to outperform the EM algorithm, which explicitly samples the posterior distribution of the BMA parameters for uncovering the uncertainty associated with model weights and thus improving the reliability of climate projections (Duan and Phillips 2010; Vrugt 2016; Wang et al. 2018a; Wang and Wang 2019). The MCMC simulation is implemented using the Differential Evolution Adaptive Metropolis (DREAM) algorithm (Vrugt 2016). According to the Bayes' theorem, the posterior distribution $p(w|x, y)$ of the BMA weights $w = (w_1,\dots,w_K)$ given the ensemble simulations $x$ and the observational variable $y$ can be expressed as

$$p(\boldsymbol{w} \mid \boldsymbol{x}, y) = \frac{p(\boldsymbol{w}) \times p(\boldsymbol{x}, y \mid \boldsymbol{w})}{p(\boldsymbol{x}, y)} \tag{4}$$

where $p(\boldsymbol{w})$ and $p(\boldsymbol{w}|\boldsymbol{x}, y)$ denote the prior and posterior distributions of BMA weights, respectively. $p(\boldsymbol{x}, y|\boldsymbol{w}) \cong L(\boldsymbol{w}|\boldsymbol{x}, y)$ denotes the likelihood function; $p(\boldsymbol{x}, y)$ denotes the evidence that acts as a normalization constant, which can be excluded from the Bayesian analysis in practice. Thus, the formulation of equation 4 can be simplified as

$$p(\boldsymbol{w} \mid \boldsymbol{x}, y) \propto p(\boldsymbol{w}) \times L(\boldsymbol{w} \mid \boldsymbol{x}, y) \tag{5}$$

The likelihood function $L(\cdot|\cdot)$ in the MCMC-based BMA projection is commonly logarithmically transformed to equation 6 for numerical stability and simplicity, where $n$ represents the number of observations in the training period.

$$\ell(w_1, ..., w_K \mid x_1, ..., x_K, y) = \sum_{t=1}^{n} \log\left( \sum_{k=1}^{K} w_k p_k (y^t \mid x_k^t) \right) \tag{6}$$

The prior distribution is set as a uniform prior distribution of $\boldsymbol{w} \in [0, 1]^K$. The MCMC simulation proceeds by running multiple Markov chains simultaneously and proposing a candidate point $z_p$ at each step (Vrugt 2016; Wang and Wang 2019). The acceptance or rejection of the candidate depends on the Metropolis acceptance probability:

$$p_{\text{accept}}(z_{\text{c}} \rightarrow z_{\text{p}}) = \min\left[1, \frac{p(z_{\text{p}})}{p(z_{\text{c}})}\right] \tag{7}$$

where $z_c$ represents the current point, and $p(\cdot)$ represents the probability density. The Markov chain moves to $z_p$ or not, depending on whether the candidate point is accepted. The convergence of Markov chains indicates that the MCMC evolution can stop, which is commonly monitored through the multi-chain $\hat{R}$ diagnostic of Gelman and Rubin (1992). Typically, a $\hat{R}$-statistic value below 1.2 indicates that the posterior distribution converges to the stationary distribution. A more detailed description of the MCMC simulation, together with the DREAM algorithm, is available in Vrugt et al. (2008) and Vrugt (2016).

### 2.2 Multidimensional drought risk projection

Copulas are multivariate cumulative distribution functions that enable us to link the marginal distributions of multiple random variables together to form the joint distribution (Genest and Favre 2007; Zhang et al. 2019). The dependence of drought duration and severity, detected by the 6-month SPEI (SPEI6) over each of the 10 climate divisions in China (see Fig. 1a), was thus described using copulas in this study, leading to a bivariate return period of drought episodes. The

SPEI6 is used since it has been demonstrated to be useful for well capturing both short- and long-term meteorological droughts (Masud et al. 2015, 2017; Huang et al. 2018; Lee et al. 2019) and the duration of most droughts is less than 6 months in China during the 1950–2006 period (Wang et al., 2011). Drought duration and severity are defined as the number of months and the sum of the integral area below −1, respectively, when SPEI6 is persistently below −1. And the SPEI6 values below −1 are often considered as suffering from droughts (Ayantobo et al. 2018; Huang et al. 2018). The 10 climate divisions are created based on the long-term mean temperature and precipitation as well as the topography in China. Assume that $X = X_1,\ldots, X_n$ denote $n$ random variables, and $F_1(x_1),\ldots, F_n(x_n)$ represent their marginal cumulative distribution functions (CDFs), the joint CDF $F(x_1,\ldots, x_n)$ can be expressed as equation 8 according to Sklar's theorem (Sklar 1959).

$$F(x_1,...,x_n) = C(F_1(x_1),...,F_n(x_n)) = C(u_1,...,u_n) \tag{8}$$

where $C$ is an $n$-dimensional copula, i.e., a joint CDF with uniform margins $(u_1,\ldots,u_n) \in [0,1]^n$. For the bivariate copula, the joint CDF $p$ of drought severity $X$ and duration $Y$ can be formulated as

$$P(X \le x, Y \le y) = C[F(x), G(y)] = p \tag{9}$$

where $F(x) = P(X \le x)$ and $G(y) = P(Y \le y)$ are the marginal CDFs of drought severity and duration, respectively. To identify the marginal CDF of drought characteristics, several types of probability distributions, including Nakagami, exponential, Rayleigh, gamma, inverse Gaussian, t location scale, generalized Pareto, Birnbaum-Saunders, extreme value, logistic, lognormal, Weibull, log-logistic, Rician, generalized extreme value, and normal distributions were included as the CDF candidates (Results are shown in Table S1 of the supplementary material). The optimal copula families were chosen from a total of 10 widely used candidates, including Gaussian, Clayton, Frank, Gumbel, Joe, Nelson, Marshal-Olkin, BB1, BB5, and Tawn. Formulas of the copula families are provided in Table 1. Both the marginal CDF and copula families were selected using the Akaike information criterion (AIC). In addition, a randomization strategy (also known as "Jittering") was used to avoid the potentially adverse impact of repeated drought durations on the bivariate analysis (De Michele et al. 2013; Chambers et al. 2018).

The copula parameters were estimated through the MCMC simulation in a Bayesian framework similar to the BMA parameters, leading to the posterior parameter distribution instead of the deterministic maximum likelihood (ML) estimates. Here, the Multivariate Copula Analysis Toolbox (MvCAT) was adopted to infer the MCMC-based copula parameters (Sadegh et al. 2017). The log-likelihood function for copula parameter inference in the MvCAT is expressed as

$$\ell(\theta \mid \tilde{y}) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2}\sigma^{-2}\sum_{i=1}^{n}\left[\tilde{y}_i - y_i(\theta)\right]^2 \tag{10}$$

where $\theta$ is the copula parameter set; $n$ denotes the total number of observations; $\sigma$ denotes the standard deviation of measurement error; $\tilde{y}_i$ denotes the empirical joint probability of observation $i$ calculated using Gringorten plotting position (Gringorten 1963); $y_i(\theta)$ is the joint probability of observation $i$ calculated by the parametric copula with the given parameter $\theta$. Different from the BMA parameters, the prior distributions of copula parameters are drawn using Latin Hypercube Sampling (LHS) which is an efficient sampler and has been widely used for implementing robust MCMC simulations (Stein 1987; Vrugt 2016; Huang et al. 2018). The Bayesian inference of copula parameter values requires specifying the initial uncertainty ranges, which are provided in Table 1. More details about the MCMC-based inference of copula parameters can be found in Sadegh et al. (2017). The MCMC simulations showed that the Marshall-Olkin copula was optimal for describing the dependence between drought severity and duration in Divisions 1−3 and 8 according to the AIC values, while the Clayton and Gumbel copulas were chosen for Divisions 4−7 and Divisions 9−10, respectively. Detailed results on the selection of copula families are provided in Table S2 of the supplementary material. To better assess the performance of the

MCMC-based copula simulation, the MCMC-based posterior distribution will be compared against the ML estimates derived by the frequentist approach.

To project the future drought hazards, the joint return period of all the episodes in which drought severity (S) and duration (D) exceed their respective threshold is computed using inclusive probability ("OR" and "AND" case) (Salvadori and De Michele 2004). The drought return period is commonly proportional to the rarity of drought episodes and the relevant losses, and thus climate-induced drought hazards can be evaluated by comparing the return periods under past and future climates. The two cases of bivariate return period can be computed using the copula-based approach as

$$T_{DS}^{\vee} = \frac{\mu}{1 - F_{DS}(D \le d, S \le s)} = \frac{\mu}{1 - C_{DS}(D \le d, S \le s, \hat{\theta})} \tag{11}$$

$$T_{DS}^{\wedge} = \frac{\mu}{1 - F_D(D \le d) - F_S(S \le s) + C_{DS}(D \le d, S \le s, \hat{\theta})} \tag{12}$$

where $\mu$ denotes the average inter-arrival time between the occurrences of drought episodes (Zhang et al. 2017). It should be noted that the return period is not deterministic but probabilistic with uncertainty ranges due to the posterior distribution of BMA weights and copula parameters derived from the MCMC simulation.

### 2.3 Performance metrics

In this study, we used several verification measures to evaluate the performance of climate simulations, including Kling-Gupta efficiency (KGE) and the supportive quantitative scores of predictive quantile-quantile (Q-Q) plot. KGE is a comprehensive verification measure introduced by Gupta et al. (2009), which combines correlation ($r$), bias ($\beta$), and variability ($\gamma$). It is defined as follows:

$$\text{KGE} = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \tag{13}$$

where the correlation component $r$ represents Pearson's correlation coefficient. The bias component $\beta$ represents the ratio of simulated and observed means, while the variability component $\gamma$ represents the ratio of the simulated and observed coefficients of variation:

$$\beta = \frac{\mu_s}{\mu_o} \text{ and } \gamma = \frac{\sigma_s / \mu_s}{\sigma_o / \mu_o} \tag{14}$$

where $\mu_s$ and $\mu_o$ represent the mean of simulated and observed variable, respectively; $\sigma_s$ and $\sigma_o$ represent the standard deviation of simulation and observation, respectively. KGE = $r$ = $\beta$ = $\gamma$ = 1 for a perfect simulation.

The predictive Q-Q plot presents a visual comparison between the quantiles in which the observations fall within the predictive distribution and the cumulative uniform distribution, U[0, 1] (Laio and Tamea 2007; Thyer et al. 2009). Detailed interpretation of the predictive Q-Q plot can be found in Thyer et al. (2009). Two reliability indices, $\alpha$ and $\varepsilon$, as well as a sharpness index, $\pi$, derived from the Q-Q plot were used to quantitatively assess the reliability and sharpness of climate simulations. These quantitative scores are defined as follows:

$$\alpha = 1 - 2\left[\frac{1}{T}\sum_{t=1}^{T}\left|P_t(y^t) - U(y^t)\right|\right], \varepsilon = 1 - \frac{1}{T}\sum_{t=1}^{T} I\left[P_t(y^t) = 1 \text{ or } P_t(y^t) = 0\right], \pi = \frac{1}{T}\sum_{t=1}^{T}\frac{E\left[y^t \mid x_1^t \ldots x_k^t\right]}{\sigma\left[y^t \mid x_1^t \ldots x_k^t\right]} \tag{15}$$

where $P_t(y^t)$ represents the nonexceedance probability of observation $y^t$ using the prediction CDF; $U(y^t)$ represents the nonexceedance uniform probability of observation $y^t$; $I$ represents the indicator function. $E[y^t \mid x_1^t \cdots x_k^t]$ and $\sigma[y^t \mid x_1^t \cdots x_k^t]$ represent the expectation and standard deviation, respectively, of the predictive distribution. The $\alpha$-index and $\varepsilon$-index vary between 0 (worst reliability) and 1 (perfect reliability). The simulation with a larger $\pi$-index indicates greater sharpness and is preferred for similarly reliable simulations.

## 3. Results

### 3.1. Reproduction of historical hydroclimate regimes and drought characteristics

Fig. 2 displays the spatial distributions of the 30-year annual mean precipitation and PET, respectively. These spatial distributions are derived from the CRU observations, the AEM simulations, and the BMA ensemble simulations as well as the absolute model bias generated by the AEM and BMA approaches. In general, there are considerable discrepancies between the AEM simulations and the CRU observations in reproducing the spatial pattern of annual mean precipitation and PET. Compared to the AEM simulations, the BMA ensemble simulations better reproduce the spatial pattern and have significantly lower absolute model biases. For example, the CRU observation and the BMA simulation generate a similar spatial gradient of precipitation in Northwest China (Figs. 2a and 2c), but such a gradient is not captured by the AEM simulation. The AEM simulation tends to underestimate the annual precipitation over Southeast China but overestimate over the Tibetan Plateau (Fig. 2d), which is congruent with previous studies (Gu et al. 2018; Zhu et al. 2018). Such biases can be caused by the cumulus convective parameterization scheme of Tiedtke (1989) used in the COSMO-CLM (CCLM) regional climate model (Giorgi et al. 2012; Niu et al. 2015; Zhang et al. 2015; Gu et al. 2020). The Tiedtke scheme activates the convection process less efficiently, leading to the negative bias of summer monsoon precipitation in Eastern China (Bao 2013). The complex orography is also a major reason for the precipitation overestimation in the Tibetan Plateau, since the resolution of 50 km is not fine enough to well describe the topographical effects of complex terrains (Wang et al. 2018b). The bias in the AEM-simulated precipitation would hinder realistic characterization of drought hazards since precipitation is one of the most important driving factors of droughts. Such model bias has been largely reduced by the BMA simulation although dry biases remain over Southeast China (Fig. 2e).

The improvement of the BMA simulation upon the AEM simulation is more significant for PET than precipitation. The AEM-simulated annual mean PET generally has a positive bias of over 0.8 mm/day over Northwest and Southeast China, as well as a negative bias of more than 1 mm/day over the Tibetan Plateau. The bias in the AEM-simulated temperature can be a major reason for the PET bias since temperature is one of the most important input variables for calculating PET, and previous studies also show a similar spatial pattern of the temperature bias in China (Yu et al. 2020). This indicates that the AEM-based projection of drought hazards can be largely overestimated over Southeast China based on the climate simulations currently available in the CORDEX East Asia experiment due to the overestimated evapotranspiration and the underestimated precipitation.

To evaluate the accuracy of the AEM- and BMA-based climate simulations, Fig. 3 presents the bar plots of the KGE score and its components $r$, $\beta$, and $\gamma$ for the AEM- and BMA-based simulations of precipitation and PET. Results show that the BMA simulation leads to a higher KGE score than the AEM simulation for most climate divisions. The AEM and BMA simulations lead to a quite similar and high correlation with observations. The correlation of the BMA-based precipitation in Division 8 is relatively low, but it is higher than that for the AEM-based precipitation. Regarding the bias score and the variability score, the BMA approach is more effective in matching simulations to observations (i.e., $\beta = 1$) and in capturing the variability of observations (i.e., $\gamma = 1$). For example, the AEM-based precipitation in Divisions 3 and 5 (i.e., the Tibetan Plateau) and the AEM-based PET in Divisions 7−10 (i.e., Southeast China) have the bias scores and the variability scores higher than 1, but the corresponding BMA-based scores are closer to 1. This indicates that the BMA simulation improves upon the AEM simulation in terms of the accuracy of precipitation and PET.

Fig. 4 presents the predictive Q-Q plots for the precipitation and PET simulations. According to the guide presented in Thyer et al. (2009), the closer the predictive Q-Q plot is to the uniform line, the better the climate simulation. The Q-Q plot falls below/above the uniform line, indicating a positive/negative bias, respectively. Overall, the Q-Q plot indicates the higher reliability and smaller bias of the BMA simulations as compared with the AEM simulations. For example, there is a clear negative bias for precipitation (Fig. 4j) and a positive bias for PET (Fig. 4t) in Division 10 based on the AEM simulation. In comparison, the BMA simulation leads to an obviously smaller area between the Q-Q plot and the uniform line, indicating higher reliability of precipitation and PET simulations. However, a visual inspection of the Q-Q plot cannot quantify the relative reliability of climate simulations over all the climate divisions. For example, the AEM- and BMA-based precipitation simulations are both overconfident in Division 2 (Fig. 4b). Therefore, two reliability indices ($\alpha$ and $\varepsilon$) derived from the Q-Q plot and a sharpness index ($\pi$), were used to quantitatively evaluate the performance of the AEM and BMA simulations.

Fig. 5 presents the reliability and sharpness of the AEM- and BMA-based simulations for precipitation and PET over each climate division. It can be seen that the BMA precipitation simulation is more reliable than the AEM simulation in respect to $\alpha$ for several climate divisions (i.e., Divisions 2 and 6−10), while the reliability of the AEM- and BMA-based precipitation simulations is similar for other climate divisions (i.e., Divisions 1 and 3−5). With respect to $\varepsilon$, BMA performs better than AEM for precipitation over most climate divisions, except for Divisions 1, 3, and 8 where BMA and AEM lead to similar $\varepsilon$. The BMA simulations also improve the sharpness ($\pi$) of the precipitation upon the AEM simulations for most divisions. Regarding PET, the BMA simulations achieve equal or higher reliability compared to the AEM simulations, especially for Divisions 3 and 5 where the BMA simulations show large improvements. Although there is no large improvement in the reliability of PET for the other divisions, the corresponding sharpness is largely improved through the BMA application. We can also observe that the BMA simulation leads to a lower sharpness for precipitation and PET than the AEM simulation for Divisions 3 and 5. This does not necessarily imply a poor performance of the BMA simulation since improving the forecast reliability and accuracy is the first priority in hydroclimate applications (Madadgar and Moradkhani 2014). Therefore, the BMA approach improves upon the AEM approach in terms of the reliability of precipitation and PET simulations.

Fig. 6 compares drought duration, severity, and frequency generated from the CRU observation and the AEM simulation for 10 climate divisions in China. Results show that the variations in drought characteristics are generally underestimated by the AEM simulation. For example, the interquartile range (IQR) of the drought duration in Division 10 generated from the AEM simulation is 1.5, while the IQRs generated from the BMA simulation and the CRU observation are both 4. The longest drought duration generated from the AEM simulation is much shorter than that generated from the CRU observation. Such a bias suggests that the AEM simulation fails to capture those megadroughts which are of very high severity and are long-lasting. In comparison, the BMA simulation greatly enhances the consistency between the observed and simulated drought characteristics, thereby providing the confidence that future drought projections are more credible.

### 3.2 Multidimensional drought risk assessment

To assess the climate-induced drought hazards, the dependence between the drought severity and duration detected by SPEI6 was simulated through the Bayesian copula. Note that the severity of a drought event is the sum of minus SPEI6 during a drought event, while the drought duration is the total number of months that a drought event lasts. Fig. 7 presents the marginal posterior distribution of parameters in copulas that describe the dependence between drought severity and duration for 10 climate divisions in China during 1975−2004. The red asterisk in each panel denotes the ML estimates derived by the frequentist copula approach. It can be seen that most of the posterior parameters are well constrained with normal distributions, but some are not, especially for the second parameter $\theta_2$ of the Marshall-Olkin copula (e.g., Figs. 7d and 7f), with a nearly uniform marginal distribution. Such unconstrained parameter distributions can be due to the limited samples of drought episodes. In addition, there is generally a plausible consistency between the posterior distribution of copula parameters inferred by the MCMC simulation and the ML estimates from the frequentist approach for most copula families, but divergent parameter estimates exist for several copulas (e.g., Figs. 7c and 7e). Such a divergence does not imply that the frequentist copula approach provides unreliable simulations, but it indicates that the frequentist approach gets trapped in local optima and provides only one plausible estimate, thereby leading to a biased representation of the dependence structure. In comparison, the MCMC-derived posterior parameter distribution provides multiple scenarios of copula simulations with equal or even higher likelihood. The uncertainty in copula parameters can lead to substantial uncertainty in drought risk assessments (see Figs. S3 and S4 of the supplementary material). This indicates that the frequentist and Bayesian copulas may lead to different drought assessments since the copula parameters determine the calculation of drought return period, which is commonly invoked in terms of quantifying and communicating risk (De Michele et al. 2013).

To examine the fit quality of copulas, the joint probability derived from the empirical copulas and the parametric copulas are compared against each other, as shown in Fig. 8. The comparisons between the MCMC-based "best" copula and the frequentist copula are distinguished by different colors. The closer the points are to the diagonal in the diagnostic plot, the better the copula fitting is. In general, both the MCMC-based and frequentist approaches provide plausible copula simulations, especially for Divisions 1 and 9. But the frequentist approach tends to underestimate the joint probability compared to the empirical joint probability. Such an underestimation does not necessarily lead to biased copula simulations but can be potentially risky since the frequentist approach fails to guarantee the global optimization for reproducing the joint distribution of observations.

### 3.3 Multi-model drought risk projection

Fig. 9 presents the comparison of drought severity, duration, and frequency detected by the SPEI6 and the run theory between the historical (1975−2004) and future (2069−2098) periods over 10 climate divisions in China. Both the drought severity and duration are projected to increase for most climate divisions. For example, the median drought durations are approximately 2 months over Division 5 for the historical period (1975−2004) and are projected to increase to 5 months for the future period (2069−2098). The increase of the radiative forcing leads to an obvious increase in the drought duration and severity for most climate divisions. For example, the median drought duration and severity in Division 2 (Northwest China) are projected to increase from 7 months to 20 months and from 13 to 75, respectively, from RCP4.5 to RCP8.5. On the other hand, the frequency of drought episodes is projected to increase for most climate divisions. For example, Division 5 experienced 20 drought episodes during 1975−2004, while the corresponding number of drought occurrences is expected to increase to 35 under RCP4.5. In addition, the increase in the radiative forcing shows no significant impacts on the frequency of drought occurrences for most climate divisions. For example, Division 7 is projected to experience 32 and 31 drought episodes under RCP4.5 and RCP8.5, respectively.

To further quantify the climate-induced change in drought risks, the return periods ("AND" and "OR" cases) of drought episodes based on drought duration and severity are assessed for the historical and future periods, as shown in Fig. 10. The historical drought duration and severity were used to construct the parametric copula, which was then used to calculate the return period for each drought episode under past and future climates, leading to the box-and-whisker plots of return period in Fig. 10. Results show that the median drought return period does not show a significant difference between past and future climates for several divisions. However, the likelihood of megadroughts with long return periods is projected to increase due to the increase in drought duration and severity over most climate divisions. For example, the percentage of droughts with the "AND" return period of at least 10 years is 24%, 62%, and 41% under the historical climate, RCP4.5, and RCP8.5, respectively, over Division 1. This may indicate an elevated probability of recurrence of the 2014 Northeast China drought which was the worst on record and led to decreased maize production by 3.93 million tons in the Liaoning province (Wang et al. 2020). We also observe that the increase in the radiative forcing leads to an obvious amplification of the likelihood of extreme droughts for most climate divisions. The increase in the likelihood of droughts with the "AND" return period of at least 10 years is from 24% in Division 4 to 345% in Division 2 under RCP4.5, while the corresponding increase under RCP8.5 is from 70% in Division 1 to 1,075% in Division 2. Such a great increase may suggest an increased risk of recurrence of record-breaking drought events, such as the severe drought of 2000 in northern China, which affected agricultural areas for more than 40 million hectares (Zou et al. 2005).

**3.4 Comparison of drought projections**

Although the Bayesian simulations better reproduce the historical drought regimes, it is desired to compare the drought projections generated from the AEM and BMA simulations. Fig. 11 presents the box-and-whisker plots of the AEM-based drought duration, severity, and frequency between past and future climates over the 10 climate divisions. Results show that there is an obvious difference between the AEM and BMA simulations in projecting future changes of drought regimes. For example, the AEM-based drought frequency is projected to decrease for most climate divisions under RCP4.5 (Fig. 11c), but the corresponding number generated from the BMA simulation is projected to increase for most climate divisions (e.g., Divisions 2 and 4−10 in Fig. 11c). The AEM and BMA simulations can also lead to differences in future changes of drought severity and duration. Overall, the future drought severity and duration are projected to increase based on the BMA simulation for most climate divisions, but they are projected to decrease based on the AEM simulation for several climate divisions, especially under RCP4.5 (e.g., Divisions 5−7). Such differences bewteen the AEM and BMA climate projections lead to different conclusions on drought risk assessments in China (see Fig. S2 of the supplementary material). Since the BMA approach improves the reliability of hydroclimate simulations and drought characterization, and shows an acceptable model transferability based on a split-sample test (see Figs. S9 and S10), the BMA-based conclusions should be preferred.

To explore the underlying reason for different drought projections based on the AEM and BMA simulations, Fig. 12 presents the spatial patterns of BMA weights for precipitation and PET. Results show that the CNRM-CM5 and PRECIS simulations make major contributions to reproducing the historical distribution of precipitation, while the other three simulations make little contribution since their BMA weights are close to zero for most divisions. Regarding PET, the PRECIS simulation makes the largest contribution in East China, while the other four simulations are relatively capable of reproducing historical PET in Northwest China. The MOHC-HadGEM2-ES and MPI-ESM-LR simulations make little contribution to reproducing historical precipitation and PET for most regions. This is inconsistent with the assumption of the AEM approach which treats each member of the ensemble as an equally likely outcome. Therefore, the AEM approach assigns equal weights to each member of the ensemble, thus leading to a large bias in precipitation and PET, but the BMA approach more heavily weights the simulations that perform relatively well in reproducing historical climate (e.g., the PRECIS simulation in this study). Such a weighted climate simulation leads to projections of future changes in precipitation and PET different from the AEM-based projections (see Figs. S5 and S6 of the supplementary material), which can be the main reason for different drought projections based on the AEM and BMA simulations.

## 4. Conclusions

In this study, a probabilistic projection of multidimensional drought risks was developed by integrating copula with BMA. An ensemble of five regional climate simulations was used to project future changes in hydroclimatic regimes over China. A Bayesian copula approach was also introduced to explicitly uncover potential interactions of the SPEI-detected drought characteristics and associated uncertainties, thereby improving the multidimensional drought risk assessment. We examined the performance of arithmetic ensemble mean (AEM) and BMA simulations in reproducing the historical climate and the drought regimes, as well as Bayesian and frequentist copula approaches used for multidimensional drought simulations. We also compared the AEM- and BMA-based future changes in drought regimes and discussed possible reasons for the resulting difference.

The AEM climate simulations show large biases in most areas of China. In comparison, the BMA climate simulation can largely improve the simulation of precipitation and PET, with a higher level of reliability and accuracy as well as a smaller bias than the AEM simulation. The variations in drought characteristics are generally underestimated by the AEM simulation, but they are better reproduced by the BMA simulation. The introduced Bayesian copula approach not only provides equally plausible estimates compared to the frequentist copula approach but also explicitly uncovers the equifinality in the copula simulation. Such an uncovered equifinality can improve the multidimensional drought assessment by providing multiple scenarios.

The drought duration and severity are projected to substantially increase for most areas of China based on the Bayesian framework, but the AEM simulation leads to multiple opposite behaviors, especially under RCP4.5. Such a discrepancy can be attributed to the systematic bias of the AEM simulation in reproducing historical hydroclimatic regimes, which propagates into future drought projections. The BMA-based drought projection should be more credible since it provides a more accurate simulation of present-day droughts. The estimated joint risk from drought duration and drought severity in China is expected to increase under both emission scenarios. The likelihood of extreme droughts (e.g., the 10-year drought) is also projected to increase as the radiative forcing increases. These findings reveal that China will experience more frequent extreme droughts, and the associated risks would be elevated due to the increase in the radiative forcing.

It should be noted that although the MCMC-based BMA approach significantly improves the ensemble mean climate simulation, the potential errors are not completely corrected. It is thus desired to further improve regional climate simulations using the high-resolution convection-permitting modeling systems in future studies. In addition, the time-invariant BMA weights determined by the historical data in multi-model climate projections may not well represent the nonstationary nature of climate dynamics. Although the underlying uncertainty in the BMA weights was explicitly addressed in this study and previous studies also yielded plausible results

(Terando et al. 2012; Olson et al. 2016, 2018; Shin et al. 2019), it is desired to develop nonstationary frameworks to further improve the credibility of climate projections.

**Acknowledgments**

This research was supported by the National Natural Science Foundation of China (Grant No. 51809223) and the Hong Kong Research Grants Council Early Career Scheme (Grant No. PP5Z). The CRU dataset was provided by the Climate Research Unit at the University of East Anglia and is publicly available at https://catalogue.ceda.ac.uk/uuid/3f8944800cc48e1cbc29a5ee12d8542d. We acknowledge the World Climate Research Program's Working Group on Coupled Modeling, which is responsible for CORDEX, and we thank the climate modeling groups for producing and making their model outputs available at https://esg-dn1.nsc.liu.se/projects/cordex/. The PRECIS model outputs and other related data used in this paper are available in Zhang (2019), doi: 10.17632/n8ckgdy2rr.1. The source code of the methodology is available from the corresponding author upon request. We would like to express our sincere gratitude to the editor and three anonymous reviewers for their constructive comments and suggestions.

## References


AghaKouchak A, Cheng L, Mazdiyasni O, Farahmand A (2014) Global warming and changes in risk of concurrent climate extremes: Insights from the 2014 California drought. Geophys Res Lett 41:8847–8852. https://doi.org/10.1002/2014GL062308

Ahmadalipour A, Moradkhani H, Rana A (2018) Accounting for downscaling and model uncertainty in fine-resolution seasonal climate projections over the Columbia River Basin. Clim Dyn 50:717–733. https://doi.org/10.1007/s00382-017-3639-4

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Fao, Rome 300:D05109

Asadi Zarch MA, Sivakumar B, Sharma A (2015) Droughts in a warming climate: A global assessment of Standardized precipitation index (SPI) and Reconnaissance drought index (RDI). J Hydrol 526:183–195. https://doi.org/10.1016/j.jhydrol.2014.09.071

Ayantobo OO, Li Y, Song S, et al (2018) Probabilistic modelling of drought events in China via 2-dimensional joint copula. J Hydrol 559:373–391. https://doi.org/10.1016/j.jhydrol.2018.02.022

Bao Y (2013) Simulations of summer monsoon climate over East Asia with a Regional Climate Model (RegCM) using Tiedtke convective parameterization scheme (CPS). Atmos Res 134:35–44. https://doi.org/10.1016/j.atmosres.2013.06.009

Barriopedro D, Gouveia CM, Trigo RM, Wang L (2012) The 2009/10 drought in China: Possible causes and impacts on vegetation. J Hydrometeorol 13:1251–1267. https://doi.org/10.1175/JHM-D-11-074.1

Basher A, Islam AKMS, Stiller-Reeve MA, Chu PS (2020) Changes in future rainfall extremes over Northeast Bangladesh: A Bayesian model averaging approach. Int J Climatol 40:3232–3249. https://doi.org/10.1002/joc.6394

Borgomeo E, Pflug G, Hall JW, Hochrainer-Stigler S (2015) Assessing water resource system vulnerability to unprecedented hydrological drought using copulas to characterize drought duration and deficit. Water Resour Res 51:8927–8948. https://doi.org/10.1002/2015WR017324

Centella-Artola A, Taylor MA, Bezanilla-Morlot A, et al (2015) Assessing the effect of domain size over the Caribbean region using the PRECIS regional climate model. Clim Dyn. https://doi.org/10.1007/s00382-014-2272-8

Chambers JM, Cleveland WS, Kleiner B, Tukey PA (2018) Graphical methods for data analysis

Chen H, Wang S, Wang Y, Zhu J (2020) Probabilistic projections of hydrological droughts through convection-permitting climate simulations and multimodel hydrological predictions. J Geophys Res Atmos 125:e2020JD032914. https://doi.org/10.1029/2020JD032914

Chen Y, Yuan H, Yang Y, Sun R (2020) Sub-daily soil moisture estimate using dynamic Bayesian model averaging. J Hydrol 590:. https://doi.org/10.1016/j.jhydrol.2020.125445

Cook BI, Anchukaitis KJ, Touchan R, et al (2016) Spatiotemporal drought variability in the mediterranean over the last 900 years. J Geophys Res 121:2060–2074. https://doi.org/10.1002/2015JD023929

Dai A (2013) Increasing drought under global warming in observations and models. Nat Clim Chang 3:52–58. https://doi.org/10.1038/nclimate1633

De Michele C, Salvadori G, Vezzoli R, Pecora S (2013) Multivariate assessment of droughts: Frequency analysis and dynamic return period. Water Resour Res 49:6985–6994. https://doi.org/10.1002/wrcr.20551

Duan Q, Phillips TJ (2010) Bayesian estimation of local signal and noise in multimodel simulations of climate change. J Geophys Res Atmos 115:1–15. https://doi.org/10.1029/2009JD013654

Ganguli P, Reddy MJ (2014) Evaluation of trends and multivariate frequency analysis of droughts in three meteorological subdivisions of western India. Int J Climatol 34:911–928. https://doi.org/10.1002/joc.3742

Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Stat Sci 7:457–472. https://doi.org/10.1214/ss/1177011136

Genest C, Favre AC (2007) Everything you always wanted to know about copula modeling but were afraid to ask. J Hydrol Eng 12:347–368. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347)

Giorgi F, Coppola E, Solmon F, et al (2012) RegCM4: Model description and preliminary tests over multiple CORDEX domains. Clim Res 52:7–29. https://doi.org/10.3354/cr01018

Gringorten II (1963) A plotting rule for extreme probability paper. J Geophys Res 68:813–814. https://doi.org/10.1029/jz068i003p00813

Gu H, Yu Z, Peltier WR, Wang X (2020) Sensitivity studies and comprehensive evaluation of RegCM4.6.1 high-resolution climate simulations over the Tibetan Plateau. Clim Dyn 54:3781–3801. https://doi.org/10.1007/s00382-020-05205-6

Gu H, Yu Z, Yang C, et al (2018) High-resolution ensemble projections and uncertainty assessment of regional climate change over China in CORDEX East Asia. Hydrol Earth Syst Sci 22:3087–3103. https://doi.org/10.5194/hess-22-3087-2018

Guo J, Huang G, Wang X, Li Y (2019) Improved performance of a PRECIS ensemble in simulating near-surface air temperature over China. Clim Dyn 52:6691–6704. https://doi.org/10.1007/s00382-018-4540-5

Gupta H V., Kling H, Yilmaz KK, Martinez GF (2009) Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J Hydrol 377:80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hao Z, AghaKouchak A (2013) Multivariate Standardized Drought Index: A parametric multi-index model. Adv Water Resour 57:12–18. https://doi.org/10.1016/j.advwatres.2013.03.009

Harris I, Jones PD, Osborn TJ, Lister DH (2014) Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. Int J Climatol 34:623–642. https://doi.org/10.1002/joc.3711

Huang J, Zhai J, Jiang T, et al (2018) Analysis of future drought characteristics in China using the regional climate model CCLM. Clim Dyn 50:507–525. https://doi.org/10.1007/s00382-017-3623-z

Kam J, Sheffield J, Wood EF (2014) A multiscale analysis of drought and pluvial mechanisms for the southeastern United States. J Geophys Res 119:7348–7367. https://doi.org/10.1002/2014JD021453

Laio F, Tamea S (2007) Verification tools for probabilistic forecasts of continuous hydrological variables. Hydrol Earth Syst Sci 11:1267–1277. https://doi.org/10.5194/hess-11-1267-2007

Lee MH, Im ES, Bae DH (2019) A comparative assessment of climate change impacts on drought over Korea based on multiple climate projections and multiple drought indices. Clim Dyn 53:389–404. https://doi.org/10.1007/s00382-018-4588-2

Li X, Li Z, Huang W, Zhou P (2020) Performance of statistical and machine learning ensembles for daily temperature downscaling. Theor Appl Climatol 140:571–588. https://doi.org/10.1007/s00704-020-03098-3

Liu XF, Wang SX, Zhou Y, et al (2016a) Spatial analysis of meteorological drought return periods in China using Copulas. Nat Hazards 80:367–388. https://doi.org/10.1007/s11069-015-1972-7

Liu Z, Törnros T, Menzel L (2016b) A probabilistic prediction network for hydrological drought identification and environmental flow assessment. Water Resour Res 52:6243–6262. https://doi.org/10.1002/2016WR019106

Madadgar S, Moradkhani H (2014) Improved Bayesian multimodeling: Integration of copulas and Bayesian model averaging. Water Resour Res 50:9586–9603. https://doi.org/10.1002/2014WR015965

Maity R, Ramadas M, Govindaraju RS (2013) Identification of hydrologic drought triggers from hydroclimatic predictor variables. Water Resour Res 49:4476–4492. https://doi.org/10.1002/wrcr.20346

Masud MB, Khaliq MN, Wheater HS (2015) Analysis of meteorological droughts for the Saskatchewan River Basin using univariate and bivariate approaches. J Hydrol 522:452–466. https://doi.org/10.1016/j.jhydrol.2014.12.058

Masud MB, Khaliq MN, Wheater HS (2017) Future changes to drought characteristics over the Canadian Prairie Provinces based on NARCCAP multi-RCM ensemble. Clim Dyn 48:2685–2705. https://doi.org/10.1007/s00382-016-3232-2

Miao L, Li S, Zhang F, et al (2020) Future drought in the dry lands of Asia under the 1.5 and 2.0 °C warming scenarios. Earth's Futur 8:. https://doi.org/10.1029/2019EF001337

Niu X, Wang S, Tang J, et al (2015) Multimodel ensemble projection of precipitation in eastern China under A1B emission scenario. J Geophys Res 120:9965–9980. https://doi.org/10.1002/2015JD023853

Olson R, An SI, Fan Y, et al (2018) North Atlantic observations sharpen meridional overturning projections. Clim Dyn 50:4171–4188. https://doi.org/10.1007/s00382-017-3867-7

Olson R, Fan Y, Evans JP (2016) A simple method for Bayesian model averaging of regional climate model projections: Application to southeast Australian temperatures. Geophys Res Lett 43:7661–7669. https://doi.org/10.1002/2016GL069704

Parajka J, Paul Blaschke A, Blöschl G, et al (2016) Uncertainty contributions to low-flow projections in Austria. Hydrol Earth Syst Sci 20:2085–2101. https://doi.org/10.5194/hess-20-2085-2016

Prudhomme C, Giuntoli I, Robinson EL, et al (2014) Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. Proc Natl Acad Sci 111:3262–3267. https://doi.org/10.1073/pnas.1222473110

Qing Y, Wang S, Zhang B, Wang Y (2020) Ultra-high resolution regional climate projections for assessing changes in hydrological extremes and underlying uncertainties. Clim Dyn 55:2031–2051. https://doi.org/10.1007/s00382-020-05372-6

Raftery AE, Gneiting T, Balabdaoui F, Polakowski M (2005) Using Bayesian model averaging to calibrate forecast ensembles. Mon Weather Rev 133:1155–1174. https://doi.org/10.1175/MWR2906.1

Rajsekhar D, Gorelick SM (2017) Increasing drought in Jordan: Climate change and cascading Syrian land-use impacts on reducing transboundary flow. Sci Adv 3:1–16. https://doi.org/10.1126/sciadv.1700581

Rockel B, Will A, Hense A (2008) The regional climate model COSMO-CLM (CCLM). Meteorol. Zeitschrift 17:347–348

Russo S, Dosio A, Sterl A, et al (2013) Projection of occurrence of extreme dry-wet years and seasons in Europe with stationary and nonstationary Standardized Precipitation Indices. J Geophys Res Atmos 118:7628–7639. https://doi.org/10.1002/jgrd.50571

Sadegh M, Ragno E, AghaKouchak A (2017) Multivariate Copula Analysis Toolbox (MvCAT): Describing dependence and underlying uncertainty using a Bayesian framework. Water Resour Res 53:5166–5183. https://doi.org/10.1002/2016WR020242

Salvadori G, De Michele C (2004) Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. Water Resour Res 40:1–17. https://doi.org/10.1029/2004WR003133

Salvadori G, Durante F, De Michele C, et al (2016) A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities. Water Resour Res 52:3701–3721. https://doi.org/10.1002/2015WR017225

Samouly A Al, Luong CN, Li Z, et al (2018) Performance of multi-model ensembles for the simulation of temperature variability over Ontario, Canada. Environ Earth Sci 77:. https://doi.org/10.1007/s12665-018-7701-2

Samaniego L, Thober S, Kumar R, et al (2018) Anthropogenic warming exacerbates European soil moisture droughts. Nat Clim Chang 8:421–426. https://doi.org/10.1038/s41558-018-0138-5

Seager R, Hoerling M, Schubert S, et al (2015) Causes of the 2011-14 California drought. J Clim 28:6997–7024. https://doi.org/10.1175/JCLI-D-14-00860.1

Shin J, Olson R, An S Il (2019) Improved probabilistic twenty-first century projections of sea surface temperature over East Asian marginal seas by considering uncertainty owing to model error and internal variability. Clim Dyn 53:6075–6087. https://doi.org/10.1007/s00382-019-04911-0

Shrestha NK, Wang J (2020) Water quality management of a cold climate region watershed in changing climate. J Environ Inform 35:56–80. https://doi.org/10.3808/jei.201900407

Sklar A (1959) Fonctions de r{é}partition {à} {n} dimensions et leurs marges (Distribution functions of {n} dimensions and their marginals). Publ l'Institut Stat l'Universit{é} Paris 8:229–231

Spinoni J, Barbosa P, Bucchignani E, et al (2020) Future global meteorological drought hot spots: A study based on CORDEX data. J Clim 33:3635–3661. https://doi.org/10.1175/JCLI-D-19-0084.1

Stein M (1987) Large sample properties of simulations using latin hypercube sampling. Technometrics 29:143–151. https://doi.org/10.1080/00401706.1987.10488205

Su B, Huang J, Fischer T, et al (2018) Drought losses in China might double between the 1.5 °C and 2.0 °C warming. Proc Natl Acad Sci 115:201802129. https://doi.org/10.1073/pnas.1802129115

Terando A, Keller K, Easterling WE (2012) Probabilistic projections of agro-climate indices in North America. J Geophys Res Atmos 117:D08115. https://doi.org/10.1029/2012JD017436

Thyer M, Renard B, Kavetski D, et al (2009) Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. Water Resour Res 45:W00B14. https://doi.org/10.1029/2008WR006825

Van Huijgevoort MHJ, Van Lanen HAJ, Teuling AJ, Uijlenhoet R (2014) Identification of changes in hydrological drought characteristics from a multi-GCM driven ensemble constrained by observed discharge. J Hydrol 512:421–434. https://doi.org/10.1016/j.jhydrol.2014.02.060

Vicente-Serrano SM, Beguería S, López-Moreno JI (2010) A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. J Clim 23:1696–1718. https://doi.org/10.1175/2009JCLI2909.1

Vidal JP, Hingray B, Magand C, et al (2016) Hierarchy of climate and hydrological uncertainties in transient low-flow projections. Hydrol Earth Syst Sci 20:3651–3672. https://doi.org/10.5194/hess-20-3651-2016

Vrugt JA (2016) Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. Environ Model Softw 75:273–316. https://doi.org/10.1016/j.envsoft.2015.08.013

Vrugt JA, Diks CGH, Clark MP (2008) Ensemble Bayesian model averaging using Markov Chain Monte Carlo sampling. Environ Fluid Mech 8:579–595. https://doi.org/10.1007/s10652-008-9106-3

Wang A, Lettenmaier DP, Sheffield J (2011) Soil moisture drought in China, 1950-2006. J Clim 24:3257–3271. https://doi.org/10.1175/2011JCLI3733.1

Wang C, Linderholm HW, Song Y, et al (2020) Impacts of drought on maize and soybean production in northeast China during the past five decades. Int J Environ Res Public Health 17:. https://doi.org/10.3390/ijerph17072459

Wang S, Ancell BC, Huang GH, Baetz BW (2018a) Improving robustness of hydrologic ensemble predictions through probabilistic pre- and post-processing in sequential data assimilation. Water Resour Res 54:2129–2151. https://doi.org/10.1002/2018WR022546

Wang S, Wang Y (2019) Improving probabilistic hydroclimatic projections through high-resolution convection-permitting climate modeling and Markov chain Monte Carlo simulations. Clim Dyn 53:1613–1636. https://doi.org/10.1007/s00382-019-04702-7

Wang X, Pang G, Yang M (2018b) Precipitation over the tibetan plateau during recent decades: A review based on observations and simulations. Int J Climatol 38:1116–1131

Williams AP, Seager R, Abatzoglou JT, et al (2015) Contribution of anthropogenic warming to California drought during 2012-2014. Geophys Res Lett 42:6819–6828. https://doi.org/10.1002/2015GL064924

Wu Y, Guo J, Lin H, et al (2021) Spatiotemporal patterns of future temperature and precipitation over China projected by PRECIS under RCPs. Atmos Res 249:. https://doi.org/10.1016/j.atmosres.2020.105303

Xu K, Yang D, Xu X, Lei H (2015) Copula based drought frequency analysis considering the spatio-temporal variability in Southwest China. J Hydrol 527:630–640. https://doi.org/10.1016/j.jhydrol.2015.05.030

Yan J (2007) Enjoy the joy of copulas: With a package copula. J Stat Softw 21:1–21. https://doi.org/10.18637/jss.v021.i04

Yang T, Wang X, Zhao C, et al (2011) Changes of climate extremes in a typical arid zone: Observations and multimodel ensemble projections. J Geophys Res Atmos 116:1–18. https://doi.org/10.1029/2010JD015192

Yu K, Hui P, Zhou W, Tang J (2020) Evaluation of multi-RCM high-resolution hindcast over the CORDEX East Asia Phase II region: Mean, annual cycle and interannual variations. Int J Climatol 40:2134–2152. https://doi.org/10.1002/joc.6323

Zhai Y, Huang G, Wang X, et al (2019) Future projections of temperature changes in Ottawa, Canada through stepwise clustered downscaling of multiple GCMs under RCPs. Clim Dyn 52:3455–3470. https://doi.org/10.1007/s00382-018-4340-y

Zhang B, Wang S, Wang Y (2019) Copula-based convection-permitting projections of future changes in multivariate drought characteristics. J Geophys Res Atmos 124:7460–7483. https://doi.org/10.1029/2019JD030686

Zhang H, Wu C, Chen W, Huang G (2017) Assessing the impact of climate change on the waterlogging risk in coastal cities: A case study of Guangzhou, South China. J Hydrometeorol 18:1549–1562. https://doi.org/10.1175/JHM-D-16-0157.1

Zhang S, Lü S, Bao Y, Ma D (2015) Sensitivity of precipitation over China to different cumulus parameterization schemes in RegCM4. J Meteorol Res 29:119–131. https://doi.org/10.1007/s13351-014-4042-2

Zhang X, Xiong Z, Zhang X, et al (2016) Using multi-model ensembles to improve the simulated effects of land use/cover change on temperature: a case study over northeast China. Clim Dyn 46:765–778. https://doi.org/10.1007/s00382-015-2611-4

Zhu J, Huang G, Wang X, et al (2018) High-resolution projections of mean and extreme precipitations over China through PRECIS under RCPs. Clim Dyn 50:4037–4060. https://doi.org/10.1007/s00382-017-3860-1

Zhu J, Wang S, Huang G (2019) Assessing climate change impacts on human-perceived temperature extremes and underlying uncertainties. J Geophys Res Atmos 124:3800–3821. https://doi.org/10.1029/2018JD029444

Zhu J, Wang S, Zhang B, Wang D (2021) Adapting to changing labor productivity as a result of intensified heat stress in a changing climate. GeoHealth 5:e2020GH000313. https://doi.org/10.1029/2020GH000313

Zou X, Zhai P, Zhang Q (2005) Variations in droughts over China: 1951-2003. Geophys Res Lett 32:1–4. https://doi.org/10.1029/2004GL021853

## List of Figure Captions

**Fig. 1 a** The PRECIS model domain with topography and 10 climate divisions including: 1. Cold-temperature and humid zone; 2. Warm-temperature and arid zone; 3. Plateau and semi-arid zone; 4. Warm-temperature and semi-arid zone; 5. Plateau and semi-humid zone; 6. Mid-temperature and humid zone; 7. Warm-temperature and humid zone; 8. North-subtropical and humid zone; 9. Mid-subtropical and humid zone; 10. South-subtropical and humid zone. The 10 climate divisions are generated based on the long-term mean temperature and precipitation as well as the topography in China. The buffer zone of 8 grids is between red and blue rectangle boxes. **b** The COnsortium for Small-scale MOdelling in CLimate Mode (CCLM) model domain with topography

**Fig. 2** Spatial patterns of 30-year (1975−2004) annual mean **a**−**e** precipitation and **f**−**j** PET (unit: mm/day) generated from the CRU observation, the ensemble mean (AEM) simulation, the BMA simulation, as well as the absolute model biases for the AEM and BMA simulations. BMA is the Bayesian model averaging, CRU is the Climatic Research Unit, and AEM is the arithmetic ensemble mean

**Fig. 3** The Kling-Gupta efficiency (KGE) score and its three components: correlation coefficient, bias, and variability ratio for the AEM- and BMA-based simulations of **a**−**d** precipitation and **e**−**h** PET during the historical 30-year period (1975−2004)

**Fig. 4** Comparison of predictive QQ plots produced by the AEM- and BMA-based simulations of **a**−**j** precipitation and **k**−**t** PET for 10 climate divisions in China during the historical 30-year period (1975−2004)

**Fig. 5** Comparison of the performance of the AEM- and BMA-based climate simulations indicated by reliability and sharpness for **a**−**c** precipitation and **d**−**f** PET during the historical 30-year period (1975−2004)

**Fig. 6** **a** Drought duration, **b** severity, and **c** frequency (i.e., number of drought episodes) generated from the CRU observation, the ensemble mean (AEM) simulation, and the BMA simulation for 10 climate divisions in China during the historical 30-year period (1975−2004). The thick black horizontal bars in **a** and **b** represent the median value, and the lower and upper edges of the box represent the 25$^{th}$ ($Q_1$) and 75$^{th}$ ($Q_3$) percentile values, respectively. The upper and lower whiskers represent the values of $Q_3 + 1.5 \times \mathrm{IQR}$ and $Q_1 - 1.5 \times \mathrm{IQR}$, respectively, where IQR denotes the interquartile range that is equal to $Q_3 - Q_1$. The values beyond the end of the whiskers are indicated by outlier points

**Fig. 7** The posterior distributions of parameters in copulas that describe the dependence between drought severity and duration for 10 climate divisions in China during 1975-2004. The red asterisk in each panel represents the maximum likelihood (ML) estimates derived by the frequentist approach

**Fig. 8** Comparison of the empirical and fitted copula-based joint probability between drought severity and duration for 10 climate divisions in China during 1975−2004. The fitted joint probability is separately calculated using copulas inferenced by Bayesian and frequentist approaches, as represented by the red and blue dots, respectively

**Fig. 9** Same as Fig. 6 but generated from the CRU observations (1975−2004) and the BMA-based projections (2069−2098) over the 10 climate divisions

**Fig. 10** The **a** AND- and **b** OR-case return periods of all drought episodes for the past (1975−2004) and future (2069−2098) climates over the 10 climate divisions. The setting of the box-and-whisker plot is the same as Fig. 6. The return periods are calculated by the parametric copula constructed for the historical drought duration and severity that are detected by the 6-month SPEI

**Fig. 11** Box-and-whisker plots of the **a** drought duration, **b** severity, and **c** frequency generated from the CRU observations (1975−2004) and the ensemble mean projections (2069−2098) over the 10 climate divisions

**Fig. 12** Spatial patterns of BMA weights for **a**−**e** precipitation and **f**−**j** PET

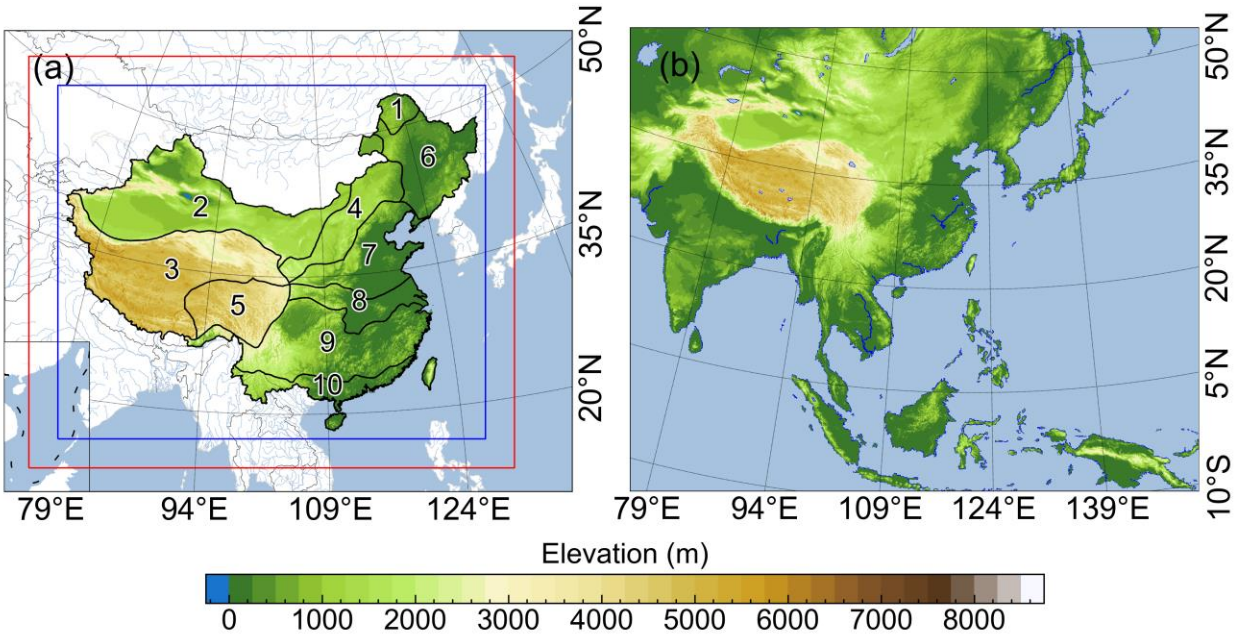


**Fig. 1 a** The PRECIS model domain with topography and 10 climate divisions including: 1. Cold-temperature and humid zone; 2. Warm-temperature and arid zone; 3. Plateau and semi-arid zone; 4. Warm-temperature and semi-arid zone; 5. Plateau and semi-humid zone; 6. Mid-temperature and humid zone; 7. Warm-temperature and humid zone; 8. North-subtropical and humid zone; 9. Mid-subtropical and humid zone; 10. South-subtropical and humid zone. The 10 climate divisions are generated based on the long-term mean temperature and precipitation as well as the topography in China. The buffer zone of 8 grids is between red and blue rectangle boxes. **b** The COnsortium for Small-scale MOdelling in CLimate Mode (CCLM) model domain with topography

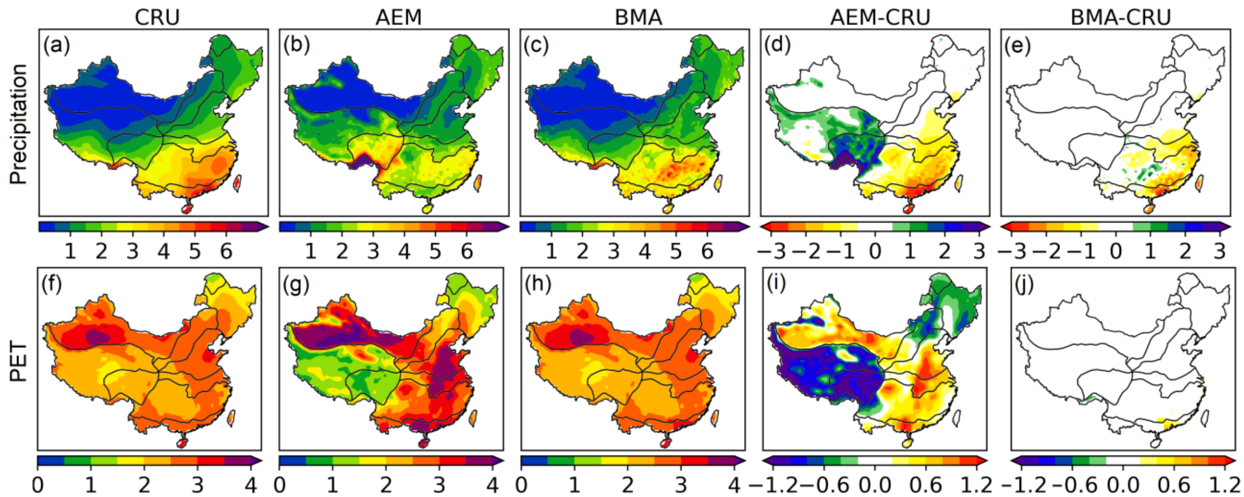


**Fig. 2** Spatial patterns of 30-year (1975−2004) annual mean **a−e** precipitation and **f−j** PET (unit: mm/day) generated from the CRU observation, the ensemble mean (AEM) simulation, the BMA simulation, as well as the absolute model biases for the AEM and BMA simulations. BMA is the Bayesian model averaging, CRU is the Climatic Research Unit, and AEM is the arithmetic ensemble mean

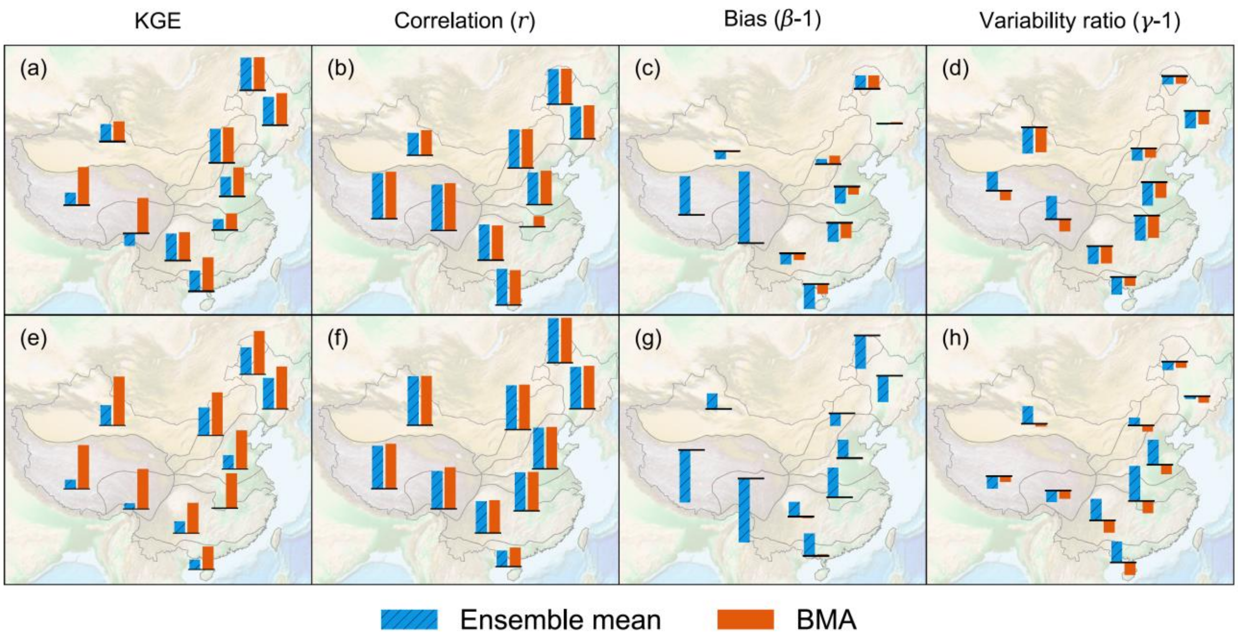


**Fig. 3** The Kling-Gupta efficiency (KGE) score and its three components: correlation coefficient, bias, and variability ratio for the AEM- and BMA-based simulations of **a**−**d** precipitation and **e**−**h** PET during the historical 30-year period (1975−2004)

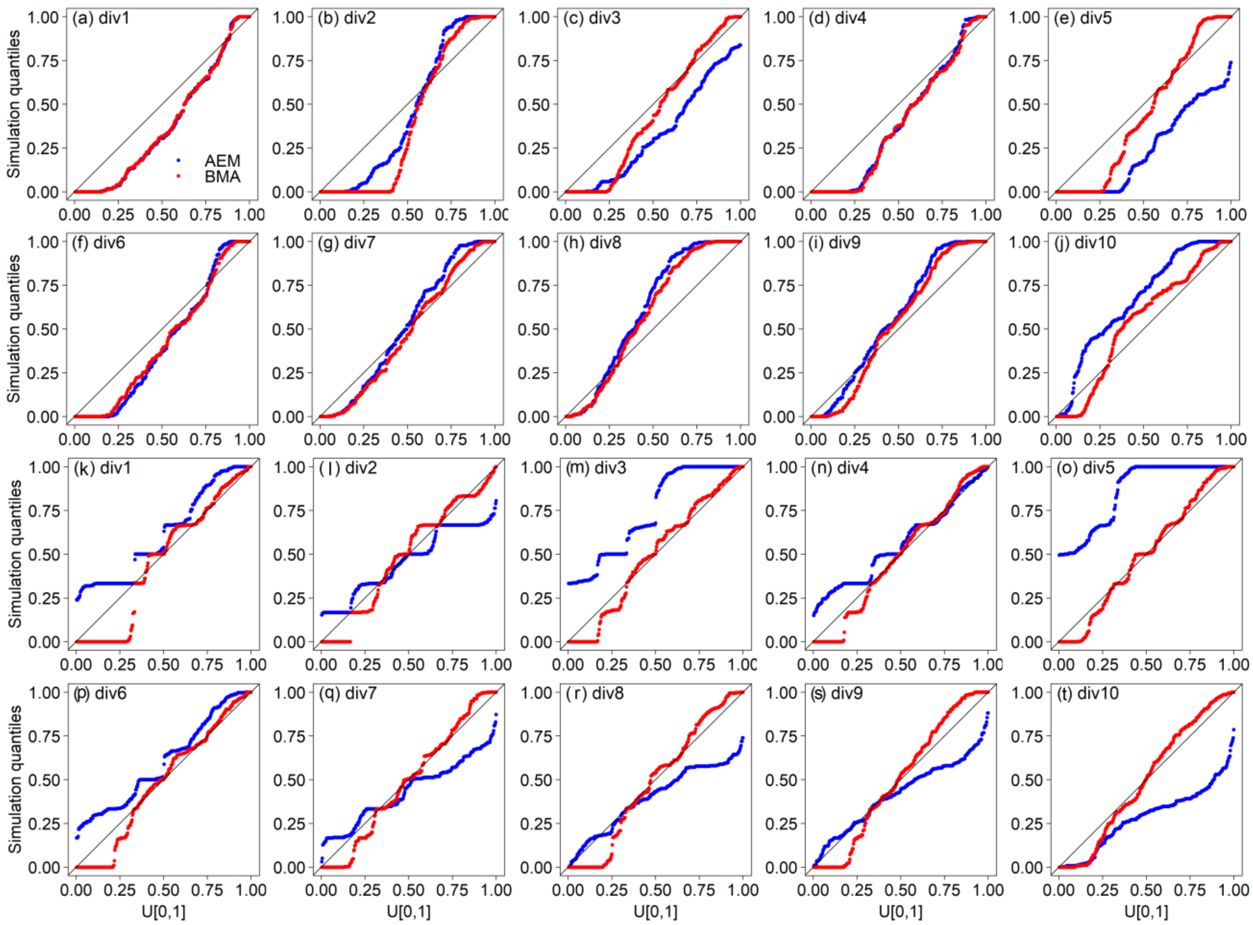


**Fig. 4** Comparison of predictive QQ plots produced by the AEM- and BMA-based simulations of **a**−**j** precipitation and **k**−**t** PET for 10 climate divisions in China during the historical 30-year period (1975−2004)

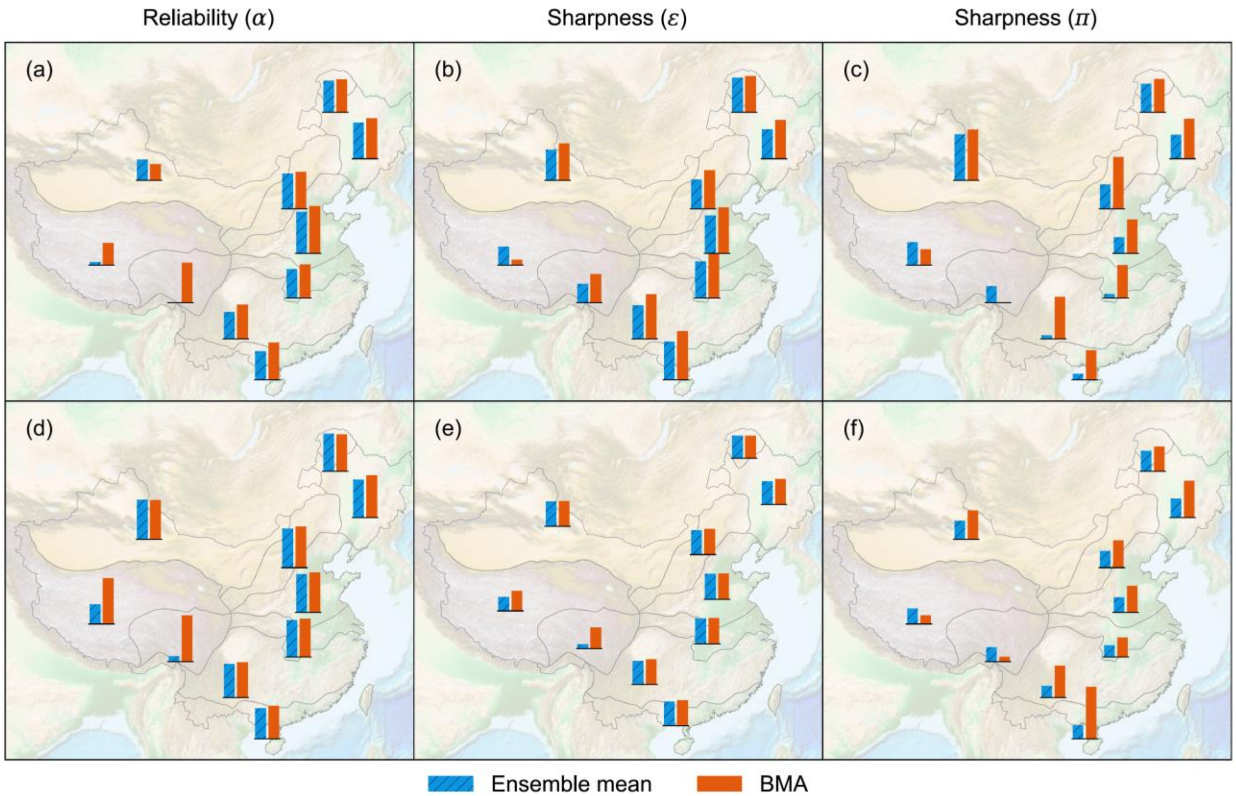


**Fig. 5** Comparison of the performance of the AEM- and BMA-based climate simulations indicated by reliability and sharpness for **a**−**c** precipitation and **d**−**f** PET during the historical 30-year period (1975−2004)

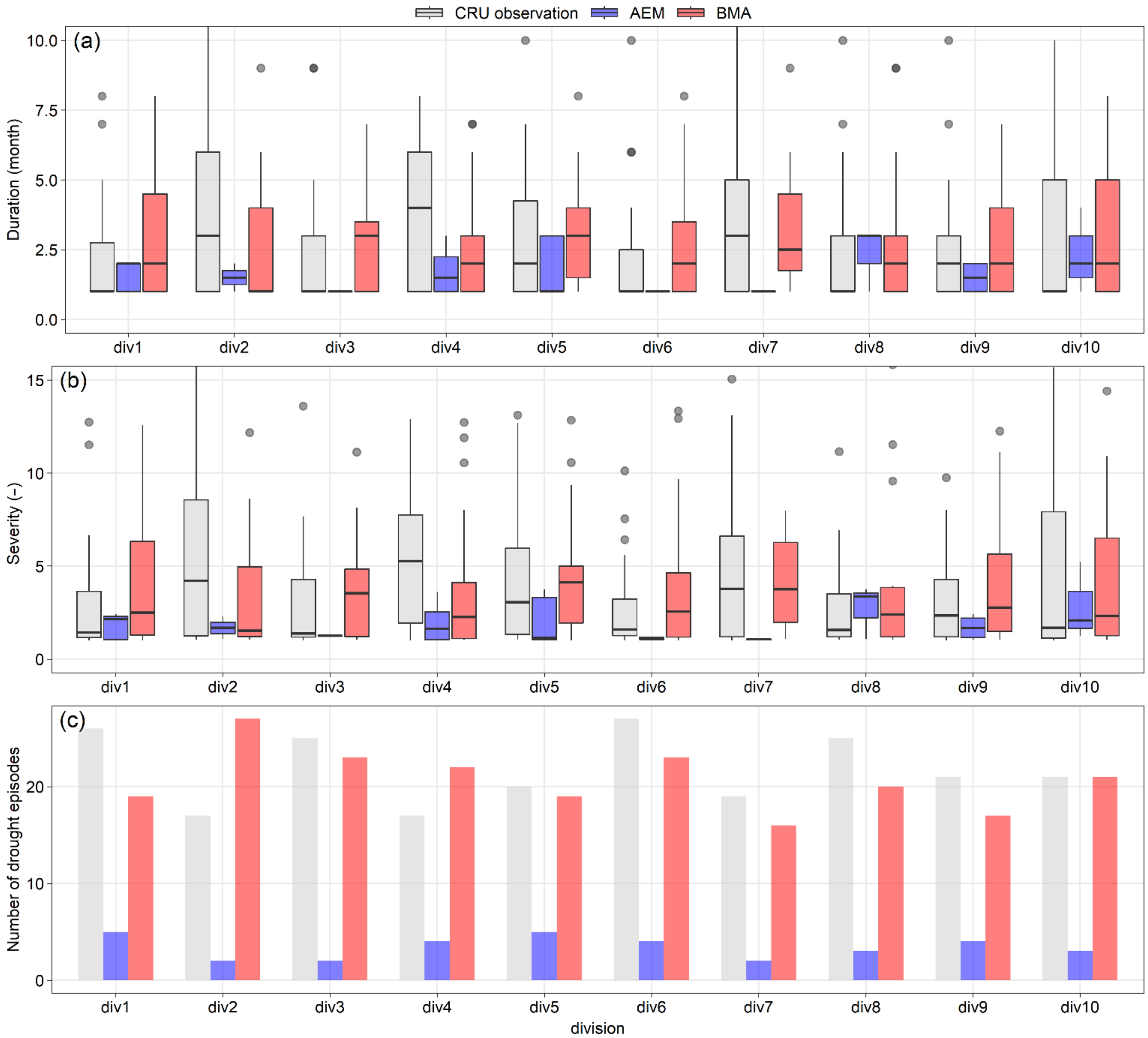


**Fig. 6** **a** Drought duration, **b** severity, and **c** frequency (i.e., number of drought episodes) generated from the CRU observation, the ensemble mean (AEM) simulation, and the BMA simulation for 10 climate divisions in China during the historical 30-year period (1975−2004). The thick black horizontal bars in **a** and **b** represent the median value, and the lower and upper edges of the box represent the 25th ($Q_1$) and 75th ($Q_3$) percentile values, respectively. The upper and lower whiskers represent the values of $Q_3 + 1.5 \times \text{IQR}$ and $Q_1 - 1.5 \times \text{IQR}$, respectively, where IQR denotes the

interquartile range that is equal to $Q_3 – Q_1$. The values beyond the end of the whiskers are indicated by outlier points

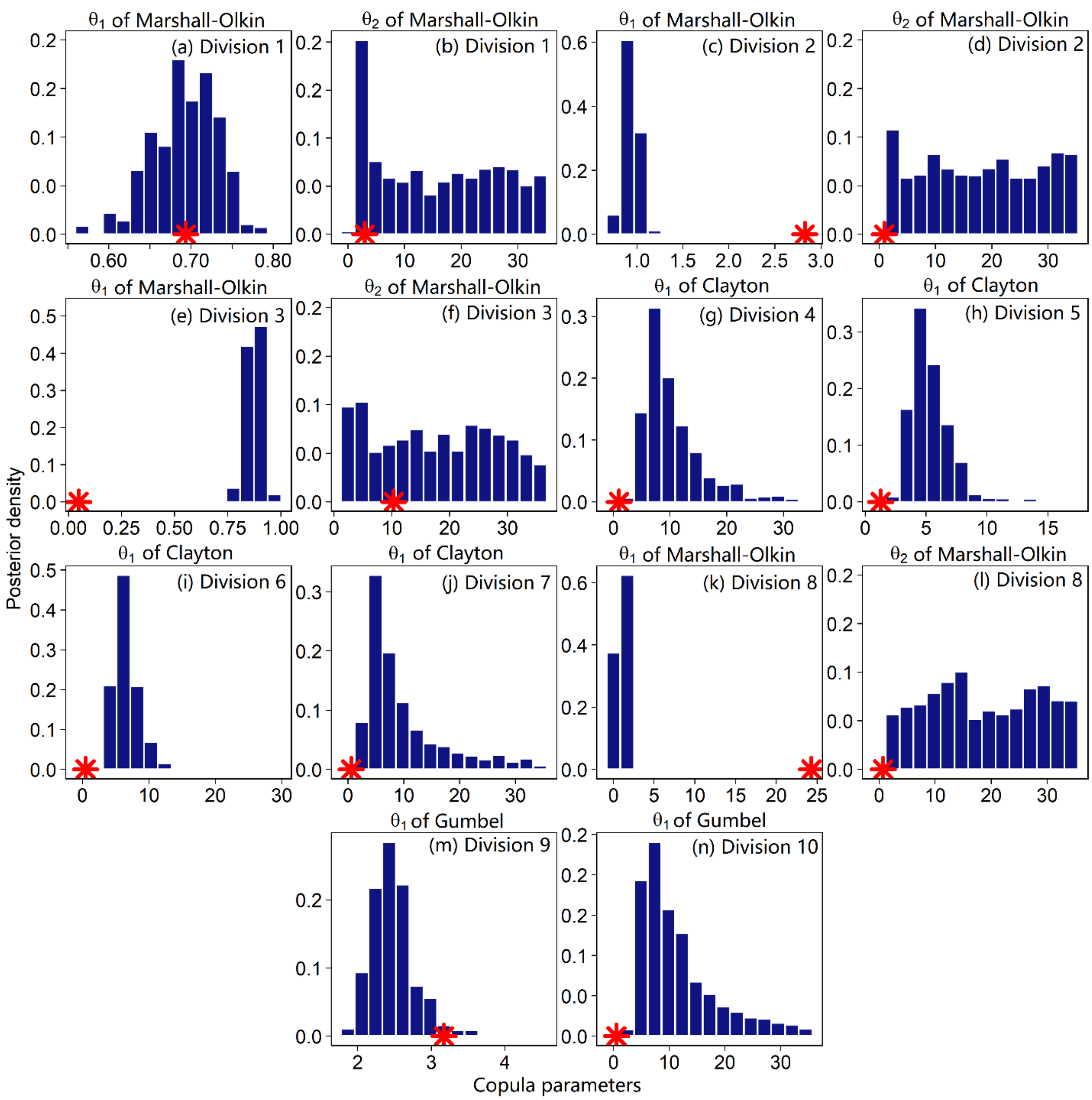


**Fig. 7** The posterior distributions of parameters in copulas that describe the dependence between drought severity and duration for 10 climate divisions in China during 1975−2004. The red asterisk in each panel represents the maximum likelihood (ML) estimates derived by the frequentist approach

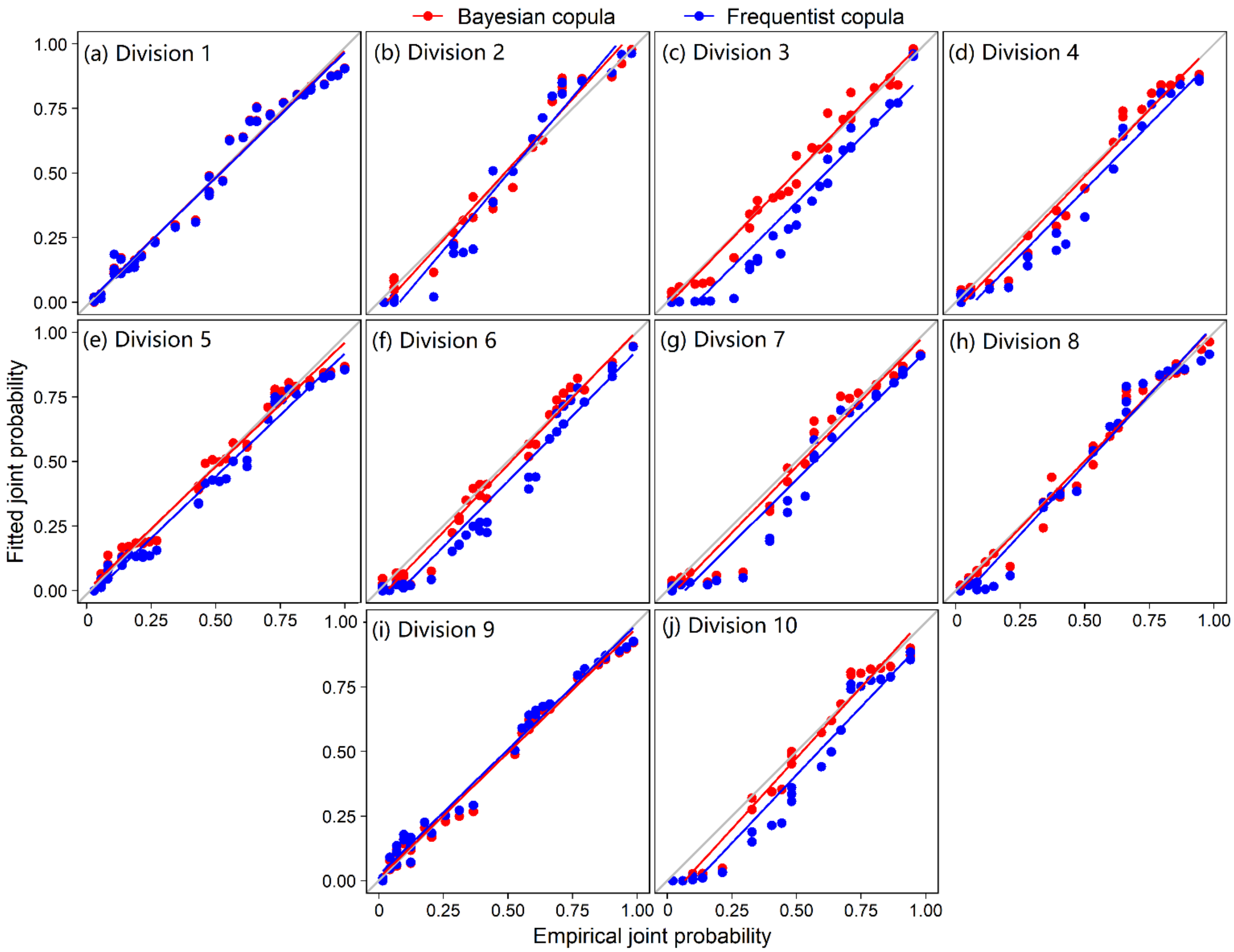


**Fig. 8** Comparison of the empirical and fitted copula-based joint probability between drought severity and duration for 10 climate divisions in China during 1975−2004. The fitted joint probability is separately calculated using copulas inferenced by Bayesian and frequentist approaches, as represented by the red and blue dots, respectively

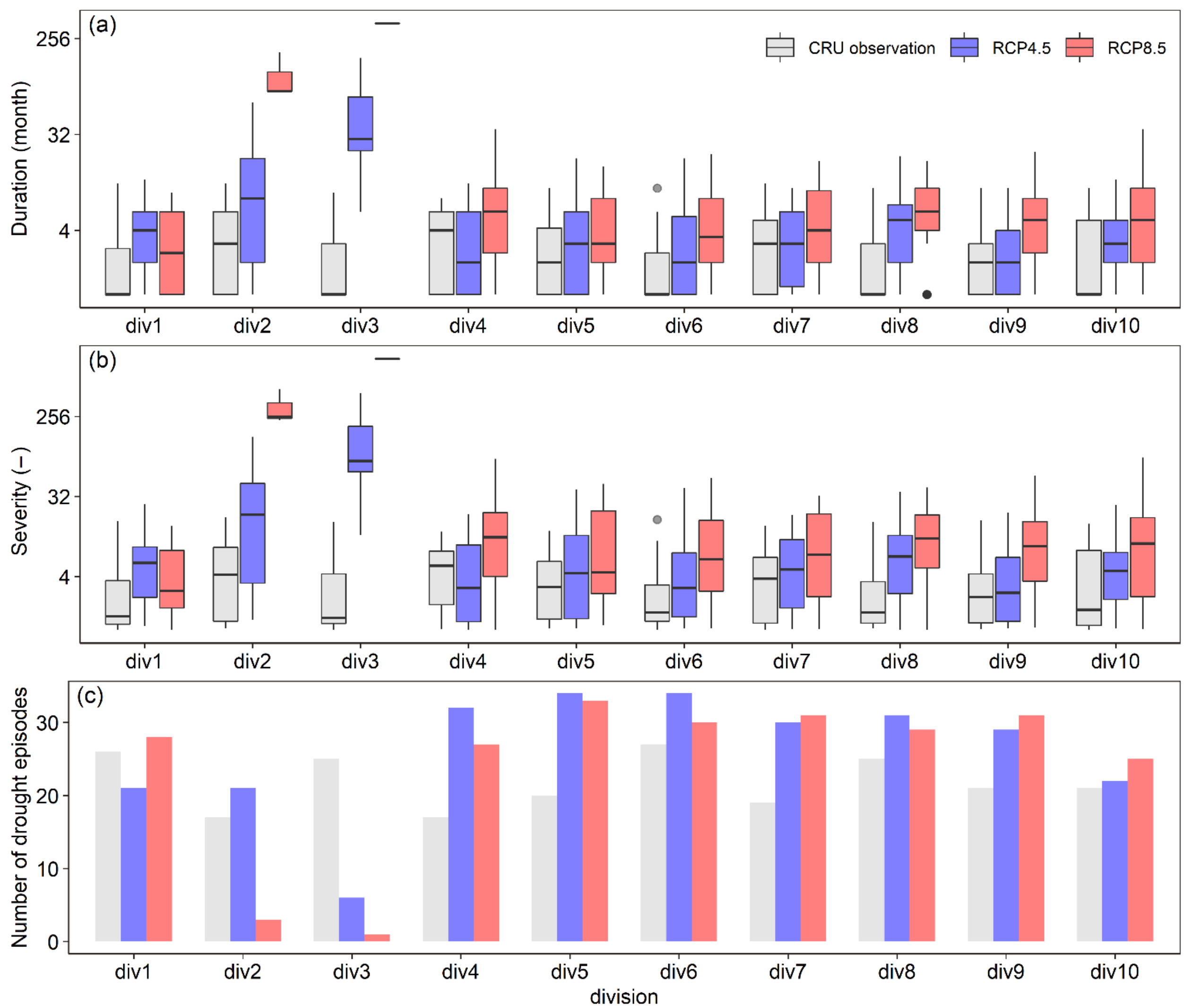


**Fig. 9** Same as Fig. 6 but generated from the CRU observations (1975−2004) and the BMA-based projections (2069−2098) over the 10 climate divisions

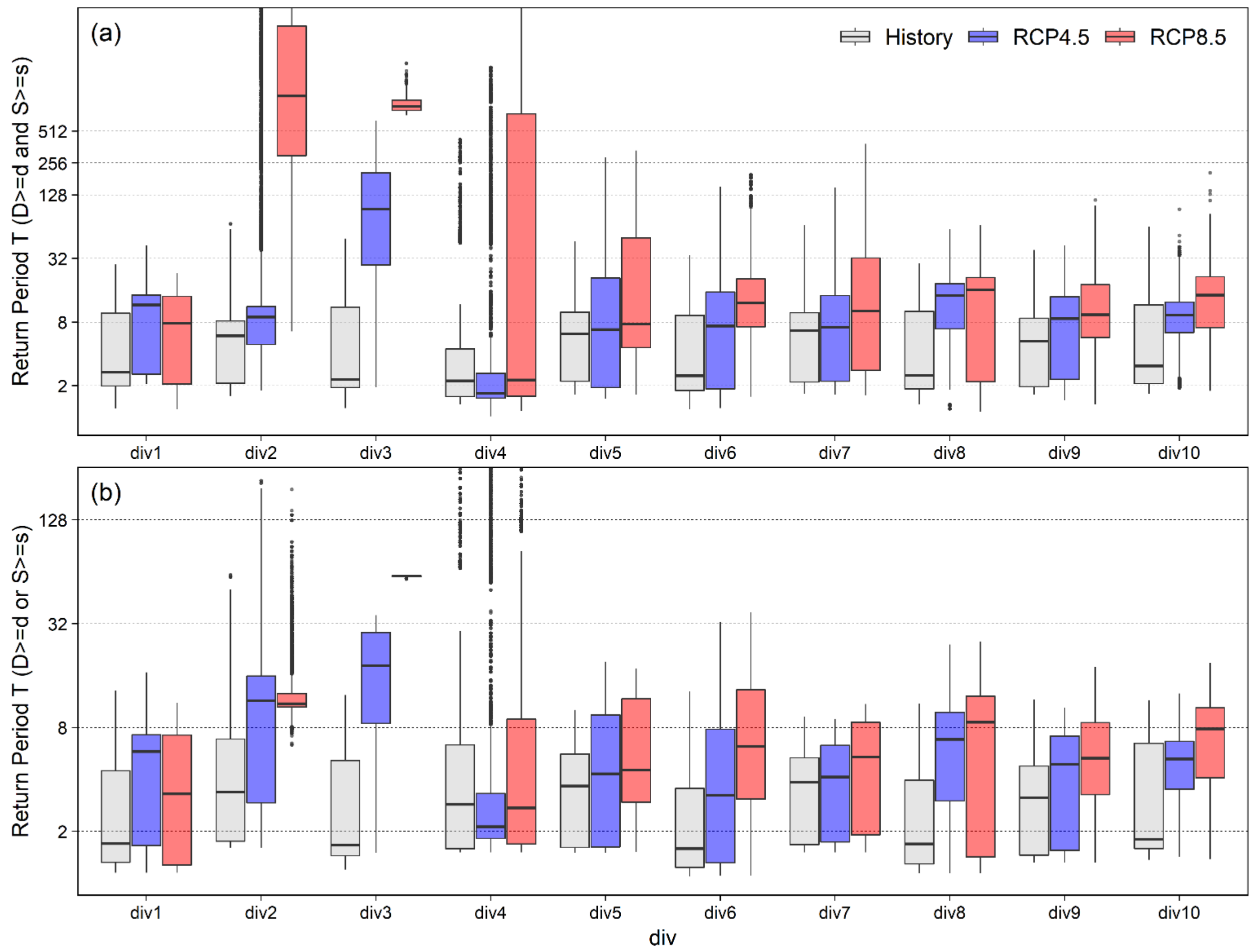


**Fig. 10** The **a** AND- and **b** OR-case return periods of all drought episodes for the past (1975−2004) and future (2069−2098) climates over the 10 climate divisions. The setting of the box-and-whisker plot is the same as Fig. 6. The return periods are calculated by the parametric copula constructed for the historical drought duration and severity that are detected by the 6-month SPEI

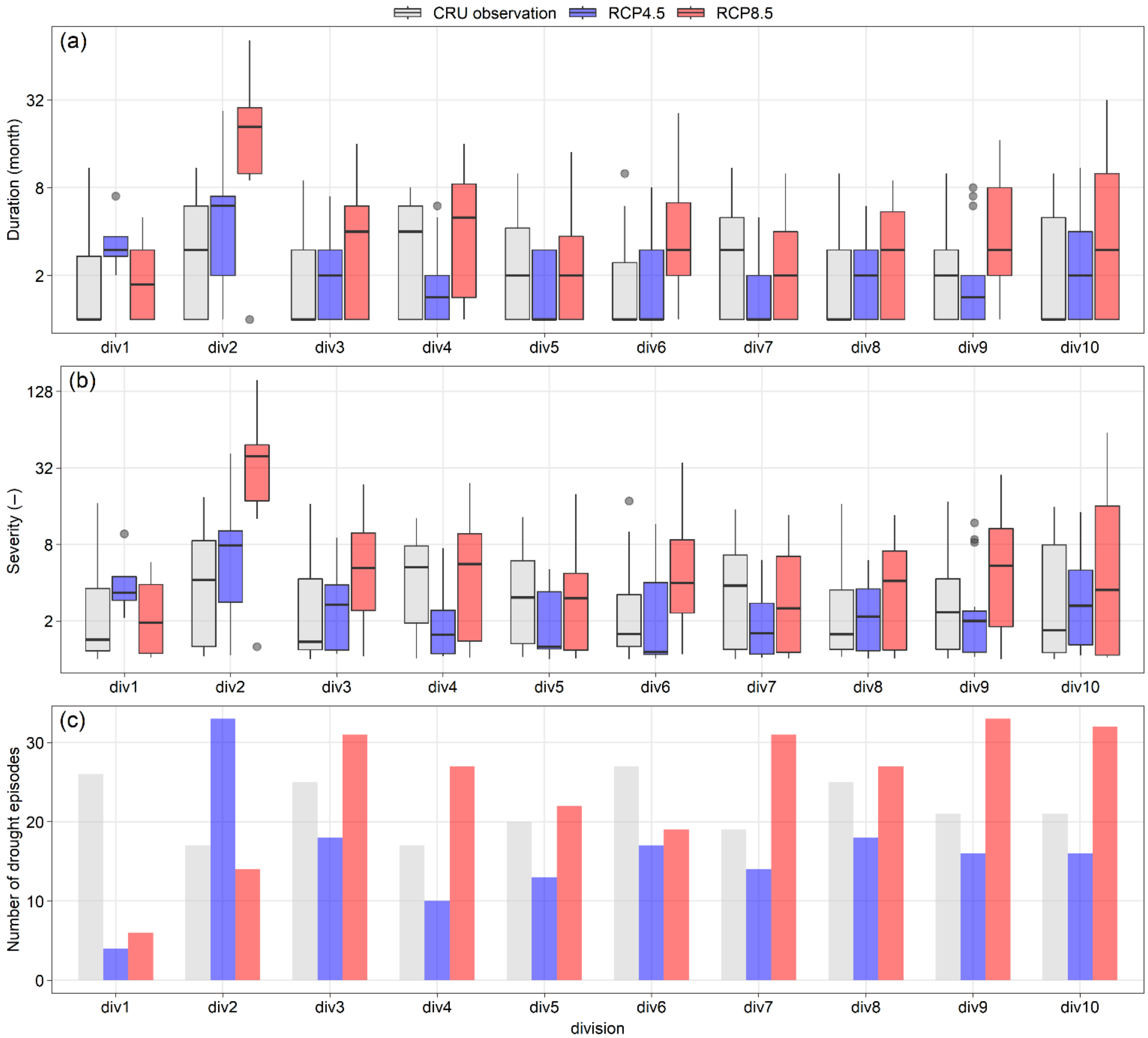


**Fig. 11** Box-and-whisker plots of the **a** drought duration, **b** severity, and **c** frequency generated from the CRU observations (1975−2004) and the ensemble mean projections (2069−2098) over the 10 climate divisions

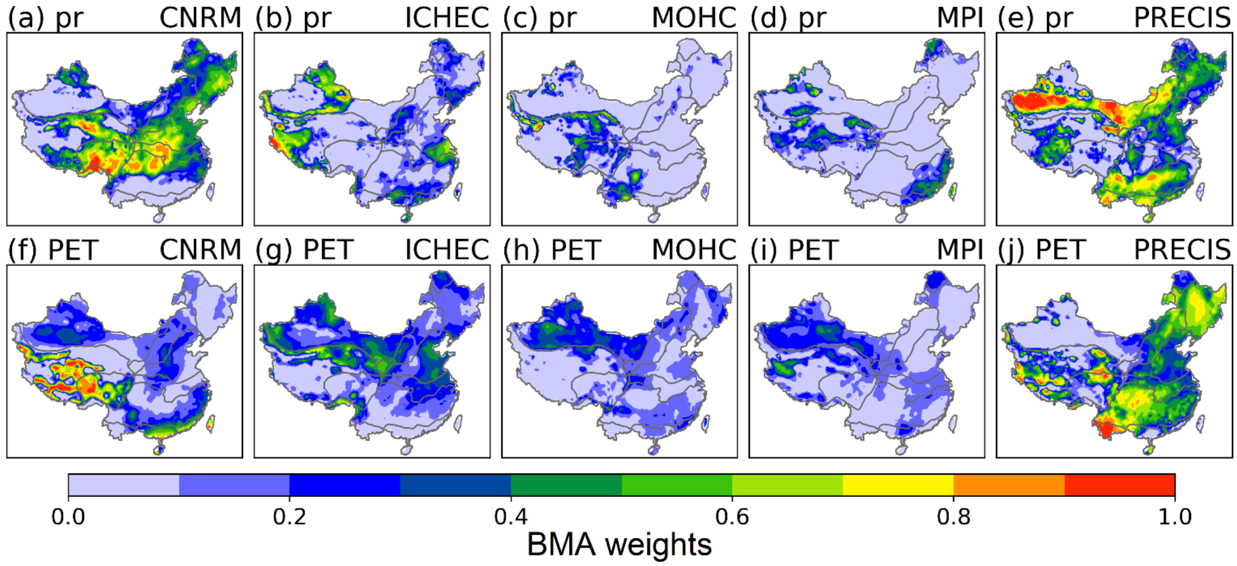


**Fig. 12** Spatial patterns of BMA weights for **a**−**e** precipitation and **f**−**j** PET

## List of Table Captions

**Table 1.** Summary of 10 copula families and the corresponding initial parameter uncertainty ranges for the MCMC-based inference

| Name | Mathematical Description for $C(u, v)$ | Parameter Range |
|---|---|---|
| Gaussian | $\int_{-\infty}^{\phi^{-1}(u)}\int_{-\infty}^{\phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\theta^2}} \exp\left(\frac{2\theta xy - x^2 - y^2}{2(1-\theta^2)}\right) \mathrm{d}x\mathrm{d}y$ | $\theta \in [-1,1]$ |
| Clayton | $C(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$ | $\theta \in (0,35]$ |
| Frank | $-\frac{1}{\theta}\ln\left[1+\frac{(\exp(-\theta u)-1)(\exp(-\theta v)-1)}{\exp(-\theta)-1}\right]$ | $\theta \in [-35,35]\setminus 0$ |
| Gumbel | $\exp\left\{-\left[(-\ln(u))^{\theta} + (-\ln(v))^{\theta}\right]^{1/\theta}\right\}$ | $\theta \in [1,35]$ |
| Joe | $1-\left[(1-u)^{\theta} + (1-v)^{\theta} - (1-u)^{\theta}(1-v)^{\theta}\right]^{1/\theta}$ | $\theta \in [1,35]$ |
| Nelson | $-\frac{1}{\theta}\log\left\{1+\frac{[\exp(-\theta u)-1][\exp(-\theta v)-1]}{\exp(-\theta)-1}\right\}$ | $\theta \in (0,35]$ |
| Marshal-Olkin | $\min[u^{(1-\theta_1)}v, uv^{(1-\theta_2)}]$ | $\theta_1, \theta_2 \in [0,35]$ |
| BB1 | $\left\{1+\left[(u^{-\theta_1}-1)^{\theta_2} + (v^{-\theta_1}-1)^{\theta_2}\right]^{1/\theta_2}\right\}^{-1/\theta_1}$ | $\theta_1 \in (0,35], \theta_2 \in (1,35]$ |
| BB5 | $\exp\left\{-\left[(-\ln(u))^{\theta_1} + (-\ln(v))^{\theta_1} - \left((-\ln(u))^{-\theta_1\theta_2} + (-\ln(v))^{-\theta_1\theta_2}\right)^{-1/\theta_2}\right]^{1/\theta_1}\right\}$ | $\theta_1 \in [1,35], \theta_2 \in (0,35]$ |
| Tawn | $\exp\left\{\ln(u^{1-\theta_1}) + \ln(v^{1-\theta_2}) - \left[(-\theta_1 \ln(u))^{\theta_3} + (-\theta_2 \ln(v))^{\theta_3}\right]^{1/\theta_3}\right\}$ | $\theta_1, \theta_2 \in [0,1], \theta_3 \in [1,35]$ |