A Mixed-Level Factorial Inference Approach for Ensemble Long-Term Hydrological Projections over the Jing River Basin

Xiong Zhou,[a] Guohe Huang,[a] Yurui Fan,[b] Xiuquan Wang,[c] and Yongping Li[a]

[a] *State Key Joint Laboratory of Environmental Simulation and Pollution Control, China-Canada Center for Energy, Environment and Ecology Research, UR-BNU, School of Environment, Beijing Normal University, Beijing, China*
[b] *Department of Civil and Environmental Engineering, Brunel University London, Uxbridge, Middlesex, United Kingdom*
[c] *School of Climate Change and Adaptation, University of Prince Edward Island, Charlottetown, Prince Edward Island, Canada*

ABSTRACT: Long-term hydrological projections can vary substantially depending on the combination of meteorological forcing dataset, hydrologic model (HM), emissions scenario, and natural climate variability. Identifying dominant sources of model spread in an ensemble of hydrologic projections is critically important for developing reliable hydrological projections in support of flooding risk assessment and water resources management; however, it is not well understood due to the multifactor and multiscale complexities involved in the long-term hydrological projections. Therefore, a stepwise clustered Bayesian (SCB) ensemble method will be first developed to improve the performance of long-term hydrological projections. Meanwhile, a mixed-level factorial inference (MLFI) approach is employed to estimate multiple uncertainties in hydrological projections over the Jing River basin (JRB). MLFI is able to reveal the main and interactive effects of the anthropogenic emission and model choices on the SCB ensemble projections. The results suggest that the daily maximum temperature under RCP8.5 in the 2050s and 2080s is expected to respectively increase by 3.2° and 5.2°C, which are much higher than the increases under RCP4.5. The maximum increase of the RegCM driven by CanESM2 (CARM)-projected changes in streamflow for the 2050s and 2080s under RCP4.5 is 0.30 and 0.59×10^3 m s^{-3} in November, respectively. In addition, in a multimodel GCM–RCM–HM ensemble, hydroclimate is found to be most sensitive to the choice of GCM. Moreover, it is revealed that the percentage of contribution of anthropogenic emissions to the changes in monthly precipitation is relatively smaller, but it makes a more significant contribution to the total variance of changes in potential evapotranspiration and streamflow.

SIGNIFICANCE STATEMENT: Increasing concerns have been paid to climate change due to its aggravating impacts on the hydrologic regime, leading to water-related disasters. Such impacts can be investigated through long-term hydrological projection under climate change. However, it is not well understood what factor plays a dominant role in inducing extensive uncertainties associated with the long-term hydrological projections due to plausible meteorological forcings, multiple hydrologic models, and internal variability. The stepwise cluster Bayesian ensemble method and mixed-level factorial inference approach are employed to quantify the contribution of multiple uncertainty sources. We find that the total variance of changes in monthly precipitation, potential evapotranspiration, and streamflow can be mainly explained by the model choices. The identified dominant factor accounting for projection uncertainties is critically important for developing reliable hydrological projections in support of flooding risk assessment and water resources management. It is suggested that more reliable models should be taken into consideration in order to improve the projection robustness from a perspective of the Loess Plateau.

1. Introduction

Global warming has the potential to significantly alter the hydrologic regime, leading to increasing concerns about the frequency and severity of future water-related disasters (Arnell and Gosling 2013; Marx et al. 2018). Regional climate models (RCMs) and hydrologic models (HMs) have been commonly used to examine climate change–related effects on the water cycle (Zhou et al. 2018b). Nevertheless, a single integration of RCM and HM tends to investigate some particular impacts of climate change on the hydrologic regime, such that decisions for planning water-related infrastructure may merely be effective in limited aspects without a comprehensive consideration. Meanwhile, long-term hydrological projections are subject to extensive uncertainties such as multiple meteorological forcings, modeling structures, and scheme parameterizations (Jones 2000; Knutti and Sedláček 2013; Murphy et al. 2009; Webster et al. 2002; Yip et al. 2011). Therefore, it is imperative to quantify uncertainty based on an ensemble approach in order to improve projection robustness.

Over the past decades, there have been numerous ensemble methods to improve forecasting performance under uncertainty based on multiple simulation models such as generalized linear modeling (GLM; Gregersen et al. 2013; Hundecha et al. 2009; Steinschneider et al. 2016; Thorarinsdottir et al. 2018),
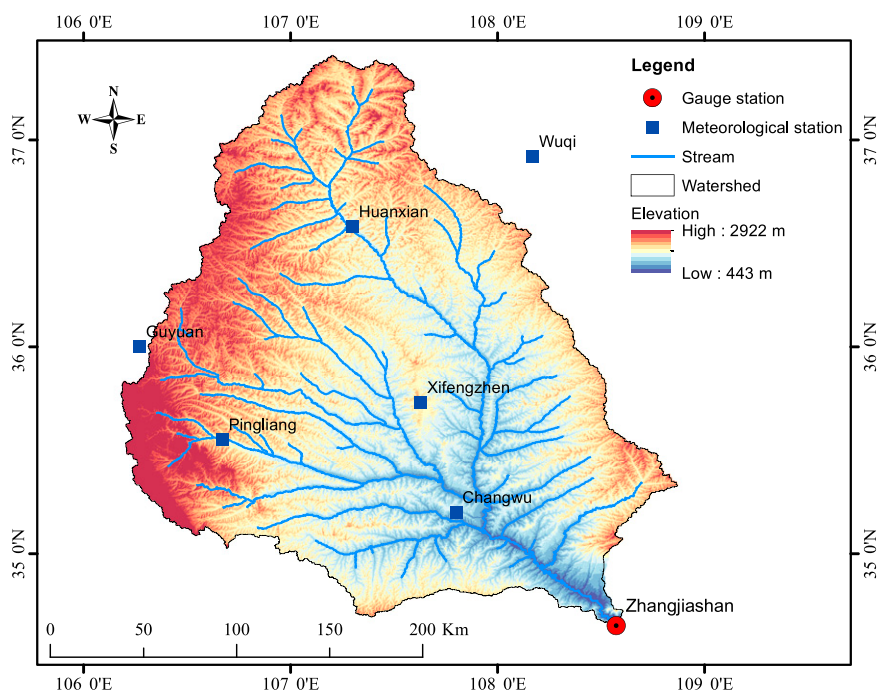
FIG. 1. Elevation of the JRB, which is derived from the hydrological data and maps based on the HydroSHEDS dataset. The red dots show the location of the Zhangjiashan gauge station (34.638°N, 108.608°E). The blue squares display the meteorological stations.

Bayesian model averaging (BMA; Baran et al. 2019; Duan et al. 2007; Madadgar and Moradkhani 2014; Rings et al. 2012; Tsai and Elshall 2013), and Bayesian hierarchical modeling (BHM; Huard and Mailhot 2006; Yan and Moradkhani 2016). Among them, BMA was widely applied to improve the meteorological and hydrological projection robustness in recent years (Fang and Li 2016; Strauch et al. 2012; Wang et al. 2012; Zhang et al. 2016). The approach supports probabilistic prediction through weighted predictive distributions based on the posterior probability. To improve BMA's performance, multiple sets of posterior coefficients may be used to capitalize on the merits of multiple climatic and hydrologic models. This is predominantly because various models perform differently in capturing particular ranges of observations. However, such multiple sets were often obtained by splitting the analysis period into several time intervals according to subjective measures.

Any merging of multimodel outputs should account for the uncertainty sources in modeling, including climate forcing, model structure and parameterizations, and numerous statistical inferences based on ensemble simulation (Madadgar and Moradkhani 2014). Manifold methods have been developed to partition such uncertainties (Evin et al. 2019; Hanel and Buishand 2015; Hingray and Saïd 2014; Yip et al. 2011; Zhang and Chen 2021). The factorial design based on analysis of variance (ANOVA) is one of the primary approaches to examine the main effects of multiple factors and their interactions (Keppel 1991). For example, Northrop and Chandler (2014) used a random-effect ANOVA model to partition uncertainties from three sources (i.e., climate models, emissions scenarios, and internal variability) in future climate projections.

In this study, we extend methods of previous studies to quantify the independent and combined effects of multiple sources of uncertainty in long-term hydrological projections over the Jing River basin (JRB). Our specific approach, a stepwise clustered Bayesian (SCB) ensemble method is shown to improve the performance of long-term hydrological projections by combining BMA and stepwise cluster analysis (SCA; Li et al. 2015; Qin et al. 2008; Wang et al. 2013; Zhuang et al. 2016). Specifically, model outputs from three global climate models (GCMs) under two representative concentration pathways (RCPs) have been dynamically downscaled using either Providing Regional Climates for Impacts Studies (PRECIS) or RegCM. The dynamically downscaled climate forcings are then used as inputs to three hydrologic models, which amounts to an ensemble size of three. To this ensemble, a factorial approach is applied to explore the impact of uncertainties (i.e., climate models, emissions scenarios, and internal variability) on the prediction of monthly precipitation, potential evapotranspiration, and streamflow in the JRB.

2. Study area and data collection

This study focused on the JRB located in the middle of the Loess Plateau, situated between 34.64° and 37.41°N and between 106.14° and 108.89°E. The Loess Plateau is known for having one of the largest rates of soil erosion on Earth (Sun et al. 2013; Zhao et al. 2013). The 455-km-long Jing River drains a total area of over 45 000 km^2 (Peng et al. 2015) (Fig. 1). As shown in Fig. 2a, the land cover in JRB is mainly composed of grassland (43.5%), open shrublands (27.5%), wooded grassland
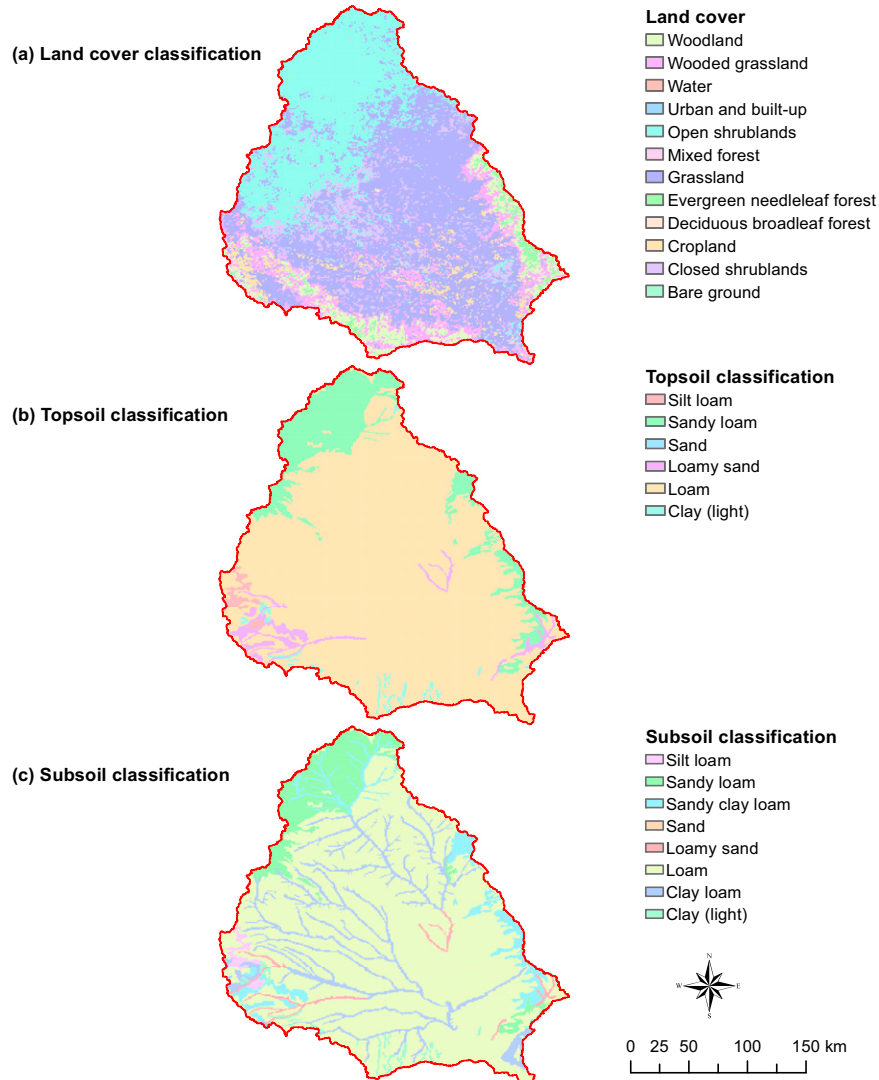
FIG. 2. Land- and soil-cover classifications in JRB based on the AVHRR Global Land Cover Classification and the Harmonized World Soil Database. The corresponding layer depths for topsoil and subsoil classification are 0–30 cm and 30–100 cm, respectively. (a) Land-cover classification, (b) topsoil classification, and (c) subsoil classification.

(9.4%), closed shrublands (8.3%), woodland (5.6%), and cropland (3.2%). Figure 2b shows that topsoil cover in the watershed includes loam (81.6%), sandy loam (13.4%), loamy sand (3.0%), clay (1.1%), and silt loam (0.9%), while Fig. 2c depicts the subsoil classification which contains loam (71.6%), clay loam (10.1%), sandy loam (9.0%), sandy clay loam (5.9%), loamy sand (1.5%), clay (1.1%), and silt loam (0.8%).

The study basin has a typical continental climate and is significantly affected by monsoons (Dong et al. 2021; Shi et al. 2021). There is a northwest–southwest precipitation gradient from semiarid to semihumid climates (Fig. 2). Significant impacts of climate change on flooding frequencies and drought durations have been investigated in a number of previous studies (Arnell and Gosling 2016; Blöschl et al. 2019; Das et al. 2020; Lehner et al. 2017; Pokhrel et al. 2021; Ukkola et al. 2020; Zhou et al. 2018b). The monthly average daily maximum temperature during the period of 1960–2005 ranges from 0.4°C in January to 26.6°C in July, while the daily minimum temperature is varied between −10.1°C in January and 16.1°C in July. Over the same period, the monthly average daily total precipitation varies from 0.09 mm day^{-1} in January to 3.22 mm day^{-1} in August. Based on the CRU datasets, the mean value and standard deviation of annual-averaged daily total precipitation for the period of 1985–2004 are 1.28 and 0.19 mm day^{-1}, respectively. As seen from a relatively low value of standard deviation, small spatial variability of precipitation is found over JRB. In addition, the average potential evapotranspiration ranges from 0.90 mm day^{-1} in December to 4.07 mm day^{-1} in June. Such values are all averaged over the JRB domain.

TABLE 1. The selected GCMs, RCMs, and HMs in each experiment.

Experiment	GCM	RCM	Hydrological model	Scenarios	Statistical model
1	HadGEM2-ES	PRECIS	VIC, SWAT, and SAC	RF	SCB
2	HadGEM2-ES	PRECIS	VIC, SWAT, and SAC	RCP45	SCB
3	HadGEM2-ES	PRECIS	VIC, SWAT, and SAC	RCP85	SCB
4	GFDL	RegCM	VIC, SWAT, and SAC	RF	SCB
5	GFDL	RegCM	VIC, SWAT, and SAC	RCP45	SCB
6	GFDL	RegCM	VIC, SWAT, and SAC	RCP85	SCB
7	CanESM2	RegCM	VIC, SWAT, and SAC	RF	SCB
8	CanESM2	RegCM	VIC, SWAT, and SAC	RCP45	SCB
9	CanESM2	RegCM	VIC, SWAT, and SAC	RCP85	SCB

To validate the performance of RCMs in reproducing spatial and temporal patterns of historical climate over JRB, simulation results are validated against the gridded monthly datasets of daily maximum temperature, minimum temperature, and total precipitation from the Climate Research Units Time series (CRU TS v4.05, herein CRU) (Harris et al. 2020). The CRU dataset has a spatial resolution of $0.5° \times 0.5°$ from 1901 to 2020 based on reanalyses of over four thousand individual meteorological station records (Harris et al. 2020). Furthermore, gridded observations of wind speed (WS) at a height of 10 m above ground level were obtained from the ERA5 dataset, which is the fifth-generation reanalysis from the European Centre for Medium-Range Weather Forecasts (Hersbach et al. 2020).

The CRU gridded datasets include monthly time series of variables, which could not perform hydrologic simulations at a daily time step. To calibrate and validate the models, the daily time series of observations (e.g., daily maximum temperature, minimum temperature, total precipitation, wind speed, relative humidity, and sunshine hours) for 6 meteorological stations across the study watershed from 1960 to 2005 were acquired from the China National Meteorological Information Center (http://data.cma.cn/). Fig. 1 presents the six meteorological stations across the study watershed. The streamflow observations at the Zhangjiashan (ZJS) gauge station (Fig. 1) from 1960 to 2005 were retrieved from the National Earth System Science Data Center (http://loess.geodata.cn/). The void-filled elevation datasets were retrieved from the Hydrological data and maps based on the Shuttle Elevation Derivatives at Multiple Scales (Hydro-SHEDS) dataset, which was developed by the Conservation Science Program of the World Wildlife Fund (Lehner et al. 2006; Zhou et al. 2018b). More detailed information could be referred to the product page (https://www.hydrosheds.org/hydrosheds-core-downloads). We also used leaf area index values and vegetation parameters at 1-km spatial resolution from the AVHRR Global Land Cover Classification, produced by the Department of Geography, University of Maryland (Hansen et al. 1998; Xiao et al. 2014; Zhou et al. 2018b). Moreover, the soil parameters were acquired from the Harmonized World Soil Database with 1-km spatial resolution, which is developed by the United Nations Food and Agriculture Organization of the United Nations (FAO/IIASA/IS-RIC/ISSCAS/JRC 2012; Zhou et al. 2018b).

3. Methodology

a. Climate modeling and hydrological projection

Table 1 presents how the modeling experiments are set up in this study. As shown in Table 1, RCMs, including PRECIS2.0 (herein PRECIS) and RegCM4.6.0 (herein RegCM), are employed to develop high-resolution climate projections using GCM historical (1985–2004, herein RF), RCP4.5 (2006–2100) and RCP 8.5 (2006–2100) output.

A lumped conceptual model [Sacramento Soil Moisture Accounting (SAC-SMA)] and two fully distributed physically based models [variable infiltration capacity (VIC) and soil and water assessment tool (SWAT)] are employed to simulate streamflow to climate inputs from RCMs over JRB at a daily time step. SAC-SMA (Duan et al. 2007; Fan et al. 2017) has 16 parameters, which has been extensively employed in National Weather Service for streamflow forecasting purposes. SWAT (Jayakrishnan et al. 2005), developed by the U.S. Department of Agriculture, is a comprehensive semidistributed hydrologic/water quality model. There are 587 hydrologic response units (HRUs) for SWAT in JRB. The spatial area of these units ranges from 0.024 to 2066.757 km^2, with an average of 77.38 km^2. VIC (Liang et al. 1996, 1994; Zhou et al. 2018b) is a spatially distributed macroscale hydrologic model, which is set up at a horizontal resolution of $0.22° \times 0.22°$ to couple with the regional climate models to assess the impacts of climate change on hydrology regimes. The streamflow of JRB simulated by VIC is derived through an offline routing model (https://vic.readthedocs.io/en/vic.4.2.d/Documentation/Routing/). It routes within the grids and to the outlet by using the unit hydrograph approach and the linearized Saint-Venant equation (Lohmann et al. 1996; Zhou et al. 2018b). The climate inputs for VIC are daily total precipitation, minimum temperature, maximum temperature, and wind speed. Additional inputs of solar radiation and relative humidity are required for SWAT and SAC-SMA. The calibrated HMs will be driven by each of GCM–RCM simulation to develop three members of hydrological projections. Each of the 27 model output streamflow time series is then inputted into the SCB method to infer a probabilistic projection of streamflow at the outlet.

The daily time series (i.e., 1976–88) of both the forcings (e.g., daily maximum temperature, minimum temperature, total precipitation, wind speed, relative humidity, and sunshine hours) from the six meteorological stations and the

TABLE 2. Selected parameters for calibration of the three hydrologic models (Fan et al. 2017; Gupta et al. 1999; Her and Chaubey 2015; Singh et al. 2005; Zhou et al. 2018b).

Parameters	Range	Unit	Calibrated value
VIC			
Variable infiltration curve parameter (B_i)	0–1		0.1065
Fraction of maximum soil moisture (W_s)	0–1		0.1290
Maximum velocity of baseflow (D_{smax})	0–30	mm day^{-1}	24.5324
Fraction of the maximum velocity of baseflow (D_s)	0–1		0.5547
Second soil layer depths (d_2)	0.1–1.5	m	1.1971
Third soil layer depths (d_3)	0.1–2	m	0.1779
SWAT			
Runoff curve number (CN2)	40–70		45.7817
Groundwater delay (GW_DELA)	0–500	days	120.5924
Baseflow alpha factor (ALPHA_BF)	0–1	days	0.4713
Threshold depth of water in the shallow aquifer (GWQMN)	0–5000	mm	4600.3874
Groundwater re-evaporation coefficient (GW_REVAP)	0.02–0.2		0.0632
Soil evaporation compensation factor (ESCO)	0–1		0.2311
Plant uptake compensation factor (EPCO)	0–1		0.9496
Manning's n for main channel (CH_N2)	−0.01 to 0.3		−0.0567
Effective hydraulic conductivity (CH_K2)	0–500	mm h^{-1}	168.8398
Baseflow alpha factor for bank storage (ALPHA_BNK)	0–1	days	0.5358
Moist bulk density for two soil layers (SOL_BD)	0.9–2.5	g cm^{-3}	0.7681 and 0.0127
Available water capacity for two soil layers (SOL_AWC)	0–1	mm mm^{-1}	0.2025 and 0.4628
Saturated hydraulic conductivity for two soil layers (SOL_K)	0–2000	mm h^{-1}	5.6161 and 36.9450
Snowfall temperature (SFTM)	−20 to 20	°C	4.9460
SAC-SMA			
Upper-zone tension water maximum storage (UZTWM)	5–300	mm	59.3312
Upper-zone free water maximum storage (UZFWM)	5–150	mm	46.9599
Lower-zone tension water maximum storage (LZTWM)	40–600	mm	234.7610
Lower-zone free water primary maximum storage (LZFPM)	40–600	mm	451.4140
Lower-zone free water supplemental maximum storage (LZFSM)	5–500	mm	207.4350
Additional impervious area (ADIMP)	0–0.4		0.2523
Upper-zone free water lateral drainage rate (UZK)	0.1–0.75	day^{-1}	0.5096
Lower-zone primary free water lateral drainage rate (LZPK)	0.001–0.05	day^{-1}	0.0201
Lower-zone supplemental free water lateral drainage rate (LZSK)	0.01–0.6	day^{-1}	0.4880
Maximum percolation rate (ZPERC)	0–350		12.3252
Exponent of the percolation equation (REXP)	0–5		2.3111
Impervious fraction of the watershed area (PCTIM)	0–0.1		0.0761
Fraction of water percolating from upper zone directly to lower zone free water storage (PFREE)	0–0.9		0.6330
Riparian vegetation area (RIVA)	0		0
Ratio of deep recharge to channel base flow (SIDE)	0		0
Fraction of lower zone free water not transferrable to lower zone tension water (RSERV)	0.3		0.3

streamflow from the ZJS gauge station are used to calibrate and validate the three hydrologic models. The calibration and validation periods for JRB are 1979–85 and 1986–88, respectively. Table 2 presents the detailed descriptions, possible ranges, and calibrated values of the selected parameters from the three hydrologic models, which are considered for calibration purposes in this study (Fan et al. 2017; Gupta et al. 1999; Her and Chaubey 2015; Singh et al. 2005; Zhou et al. 2018b). Notably, only 13 parameters of SAC-SMA are calibrated, since the other three parameters ratio of deep recharge to channel base flow (SIDE), riparian vegetation area (RIVA), and fraction of lower zone free water not transferrable to lower zone tension

water (RSERV) were fixed at prespecified values in accordance with previous studies (Gupta et al. 1999). The Nash–Sutcliffe coefficient of efficiency (NSE) at a monthly time step is selected as the objective function (Zhou et al. 2018b). The shuffle complex evolution algorithm, University of Arizona (SCE-UA) (Duan et al. 1994), is then employed to calibrate these parameters of the three hydrologic models in simulating the daily time series of historical streamflow at the ZJS station.

PRECIS2.0 is the latest version of the regional climate modeling system, which is developed by the Met Office Hadley Centre for Climate Science and Services. PRECIS can only be implemented at horizontal resolution of 0.44° × 0.44° and

$0.22° \times 0.22°$ at the rotated grid (Zhou et al. 2018a), which approximately provides 50 km $\times$ 50 km and 25 km $\times$ 25 km, respectively. Specific parameterizations of the PRECIS model are detailed by Jones et al. (2004) and Zhou et al. (2018c). Therefore, we ran our model extending out from JRB with the highest horizontal resolution of $0.22° \times 0.22°$ and time resolution of 1 h. The RCM simulations and observation/reanalysis datasets are regridded to the PRECIS grid cell for the purpose of performance evaluation.

PRECIS model can solely be driven by the boundary data from the Met Office Hadley Center (e.g., HadGEM2-ES). This is due to that the input and output data of PRECIS are written in the post processing (PP) binary data format, which is a Met Office proprietary format. To identify the uncertainty of boundary conditions, the RegCM4.6.0 modeling system from the International Center for Theoretical Physics (ICTP) is also used to dynamically downscale additional GCMs. As shown in Table 1, the first ensemble member (r1i1p1) of HadGEM2-ES, CanESM2, and GFDL-ESM2M from CMIP5 is employed to address the effects of model choices on future climate projections. The horizontal resolutions (i.e., latitude $\times$ longitude) of each model are $1.25° \times 1.875°$, $2.8125° \times 2.8125°$, and $2° \times 2.5°$, respectively. The temporal frequency of each lateral boundary forcing is 6-hourly. The boundary data of HadGEM2-ES for PRECIS experiments are obtained from the Met Office Hadley Centre, which is not available publicly (Met Office 2020). The boundary data of CanESM2 and GFDL-ESM2M are retrieved from their websites (http://clima-dods.ictp.it/regcm4/CanESM2/ and http://clima-dods.ictp.it/regcm4/GFDL-ESM2M/, respectively).

Two different future climate change scenarios, namely, RCP4.5 and RCP8.5, are chosen to account for the effects of greenhouse gas (GHG) emissions due to anthropogenic activities. More detailed descriptions of the two RCPs are presented in the study of Zhou et al. (2018b). Dynamical downscaling multiple GCMs under two RCPs is able to develop the range of possible climatic changes that can be expected for JRB. The RCM simulations and observation/reanalysis datasets are regridded to the PRECIS grid cell for the purpose of performance evaluation.

The future projection of climate changes for the daily maximum and minimum temperature, total precipitation, and wind speed are developed to force hydrologic simulations under RCP4.5 and RCP8.5 by using the validated RCMs driven by multiple GCMs. However, in order to explore the impacts of climate change on the hydrologic regime, we first analyze changes derived from the dynamically downscaled climate projections for the two 30-yr periods: 2050s (2036–65) and 2080s (2066–95) in relative to the reference period (i.e., 1985–2004, herein RF). The averaged values and standard deviations of annual and seasonal changes, as well as seasonal cycles of changes in the climate variables, are derived and examined.

b. SCB ensemble model

Based on multivariate analysis of variation, the SCA method is capable of cutting or merging sample sets of dependent variables given by whether there is a significant difference among them (Wang et al. 2015). Thus, the original sample set of observed streamflow at the ZJS station can be split into multiple subsets through numerous cutting and merging operations, indicating the complex relationships between the observed streamflow and the multiple simulated streamflow. The detailed information is described by Huang (1992) and Wang et al. (2015). Therefore, SCA is integrated into the SCB framework to generate multiple subsets of observed streamflow without any subjective assumptions, and then the projection of each subset based on BMA is combined into a probabilistic projection for the original dataset.

The proposed SCB thus can take its advantages to tackle nonlinear relationships among multiple continuous and discrete variables, as well as to generate more reliable probabilistic projections from multiple RCMs and HMs. However, the key prerequisite of SCB is to select independent hydrologic models in order to improve its reliable performance. It is practically difficult to acquire a mutually exclusive and collectively exhausted (MECE) ensemble of hydrologic models, which is expected to be a similar challenge to BMA (Madadgar and Moradkhani 2014; Refsgaard et al. 2012). The uncertainty of probabilistic projections may be overestimated without independent hydrologic models (Madadgar and Moradkhani 2014). It also should be noted that a limitation of the SCB method is that the model independency should be stationary under various climatic conditions. Figure 3 conveys the detailed processes from initial GCM–RCM simulations to final SCB projections.

Specifically, consider two sets of original datasets: $\mathbf{c}_1 = (\mathbf{o}, \mathbf{m})$ and $\mathbf{c}_2 = (\mathbf{p}, \mathbf{f})$, where $\mathbf{o}$ and $\mathbf{p}$ are column vectors of dependent variables with n_1 and n_2 samples, respectively; $\mathbf{m}$ and $\mathbf{f}$ are vectors of independent k models with a dimension of $n_1 \times k$ and $n_2 \times k$, respectively. We are interested in comparing possible differences in the mean responses for the two datasets. Thus, the appropriate null and alternative hypothesis can be formulated as follows:

$$H_0: \mu = \nu, \quad \text{and} \tag{1}$$

$$H_1: \mu \neq \nu, \tag{2}$$

where μ and ν respectively represents the average of $\mathbf{o}$ and $\mathbf{p}$:

$$\mu = \frac{1}{n_1} \sum_{i=1}^{n_1} o_i, \quad \text{and} \tag{3}$$

$$\nu = \frac{1}{n_2} \sum_{j=1}^{n_2} p_j. \tag{4}$$

According to the Wilks' statistic (Huang 1992; Kennedy and Gentle 1981; Wang et al. 2015), the sums of squares and cross-products (SSCP) for the within group (w) and between group (b) can be respectively obtained:

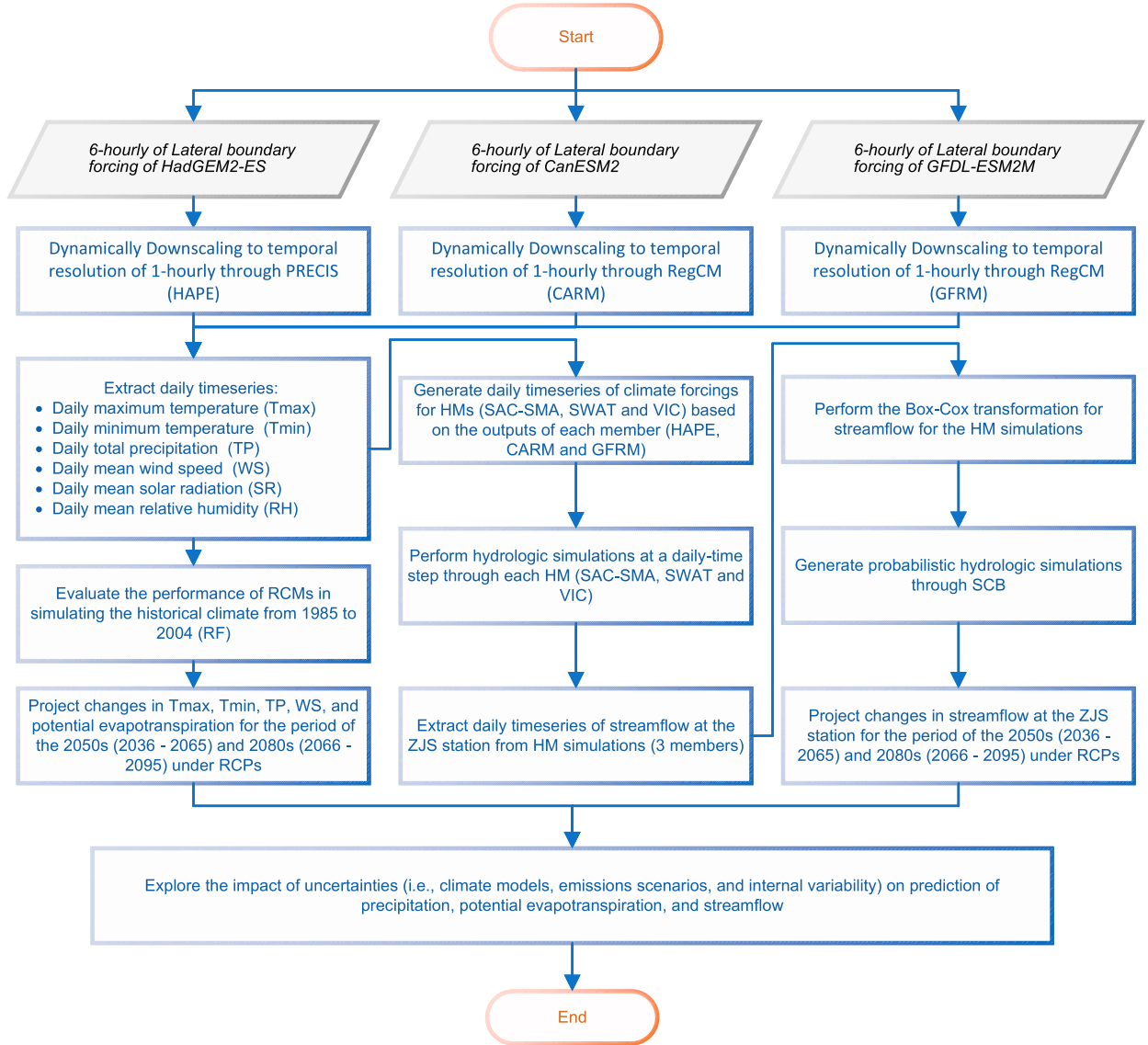$$w = \sum_{i=1}^{n_1} (o_i - \mu) + \sum_{j=1}^{n_2} (p_i - \nu), \quad \text{and} \tag{5}$$

FIG. 3. Flowchart for developing the climatic and hydrologic projections.

$$b = \frac{n_1 n_2}{n_1 + n_2} (\mu - \nu)^2. \tag{6}$$

Thus, the value of the Wilks' statistic λ can be calculated based on the w and b SSCP:

$$\lambda = \frac{w}{w + b}. \tag{7}$$

Based on Rao's F approximation (Huang 1992; Rao 1952; Wang et al. 2015), λ can be transformed to distribute as an exact F statistic with 1 and $n_1 + n_2 - 1$ degree of freedom:

$$F_{1, n_1 + n_2 - 1} = \frac{(n_1 + n_2 - 1)(1 - \lambda)}{\lambda}. \tag{8}$$

Therefore, Eq. (8) is the test statistic to determine whether to reject the null hypothesis H_0 given a significance level of α. We would conclude that there is no significant difference in the mean responses (i.e., μ and ν) of the two datasets at a significance level of α if $F < F_{\alpha, 1, n_1 + n_2 - 1}$, implying that SCB will merge two datasets. Otherwise, if $F \geq F_{\alpha, 1, n_1 + n_2 - 1}$, SCB will cut them into two different subsets, as the mean responses of the two datasets differ at the given significance level.

The SCB method will exhaustively cut and merge datasets based on the hypothesis test. The finalized multiple subsets for the entire training period will then be obtained until no hypotheses of further cutting or merging operations are accepted (Huang 1992; Li et al. 2016; Wang et al. 2015). Assume that a total of S subsets is generated through the above procedures,
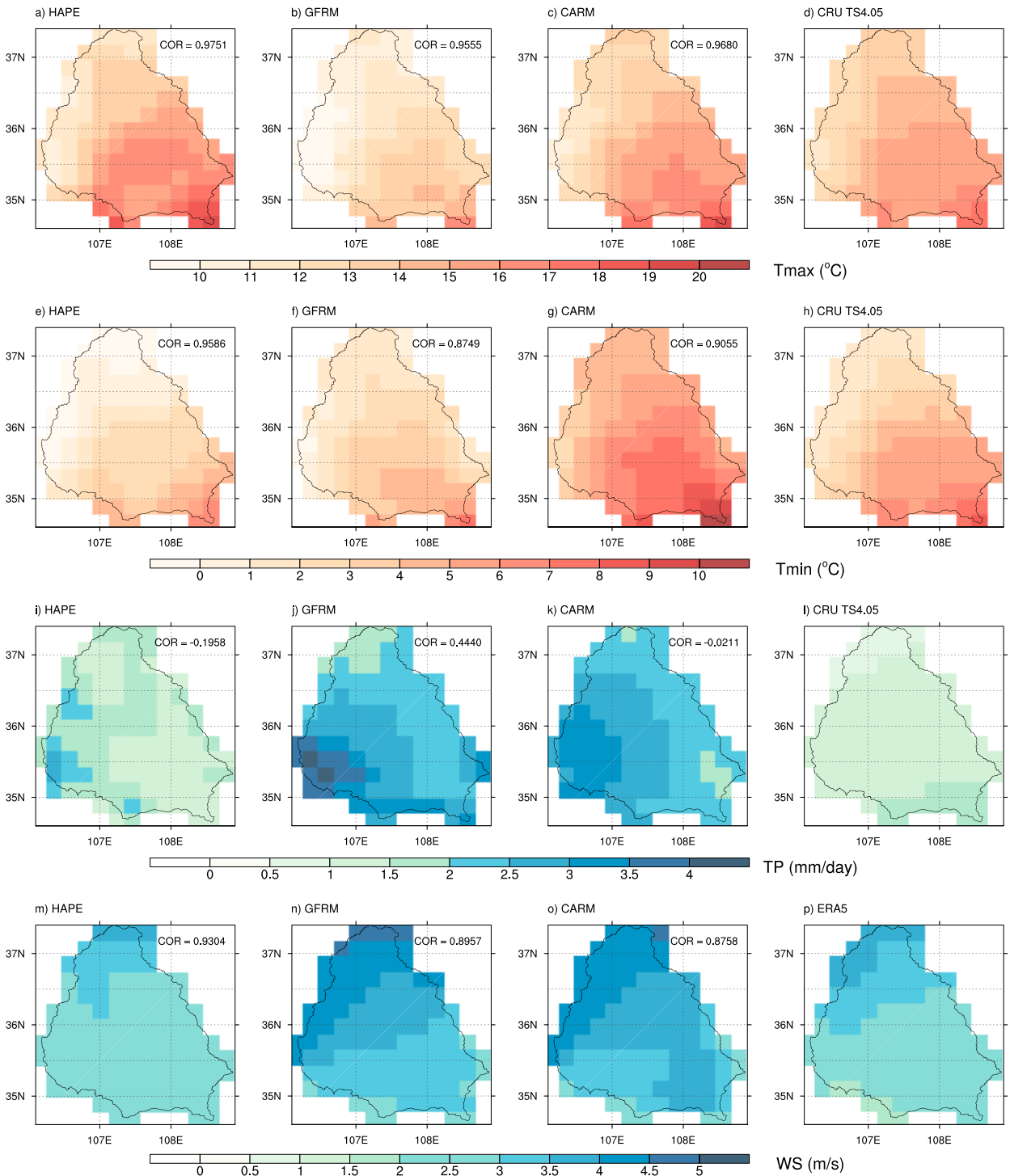
FIG. 4. Spatial distributions of annual averages of historical simulations for the period of 1985–2004 derived from HAPE, CARM, GFRM, and CRU/ERA5. COR is the spatial correlation between simulations and CRU/ERA5. (a)–(d) Tmax, (e)–(h) Tmin, (i)–(l) TP, and (m)–(p) WS.

and then we could have the dependent variable and independent variables for each subset s as follows:

$$\mathbf{o}'_s = (o'_{s1}, o'_{s2}, \ldots, o'_{st_s})', \quad \text{and} \tag{9}$$

$$\mathbf{m}'_s = (\mathbf{m}'_{sr}) = (\mathbf{m}'_{s1}, \mathbf{m}'_{s2}, \ldots, \mathbf{m}'_{sk}), \, r = 1, 2, \ldots, k, \tag{10}$$

where $\mathbf{m}'_{sr}$ is the column vector of model r with t_s samples. According to BMA (Duan et al. 2007), the probability density

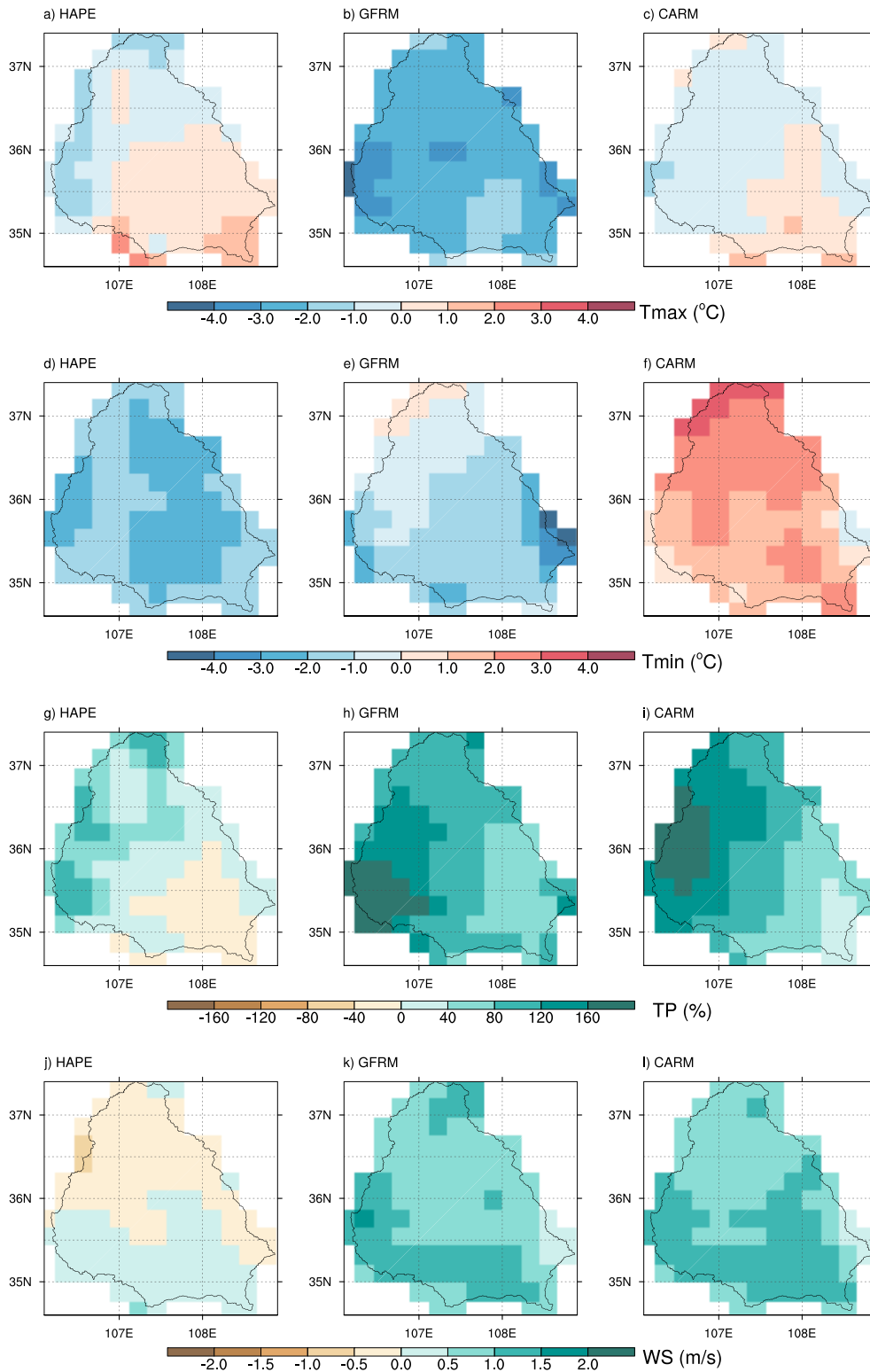FIG. 5. Differences in annual averages of (a)–(c) Tmax, (d)–(f) Tmin, (g)–(i) TP, and (j)–(l) WS between the simulations and the CRU/ERA5 for the period of 1985–2004.

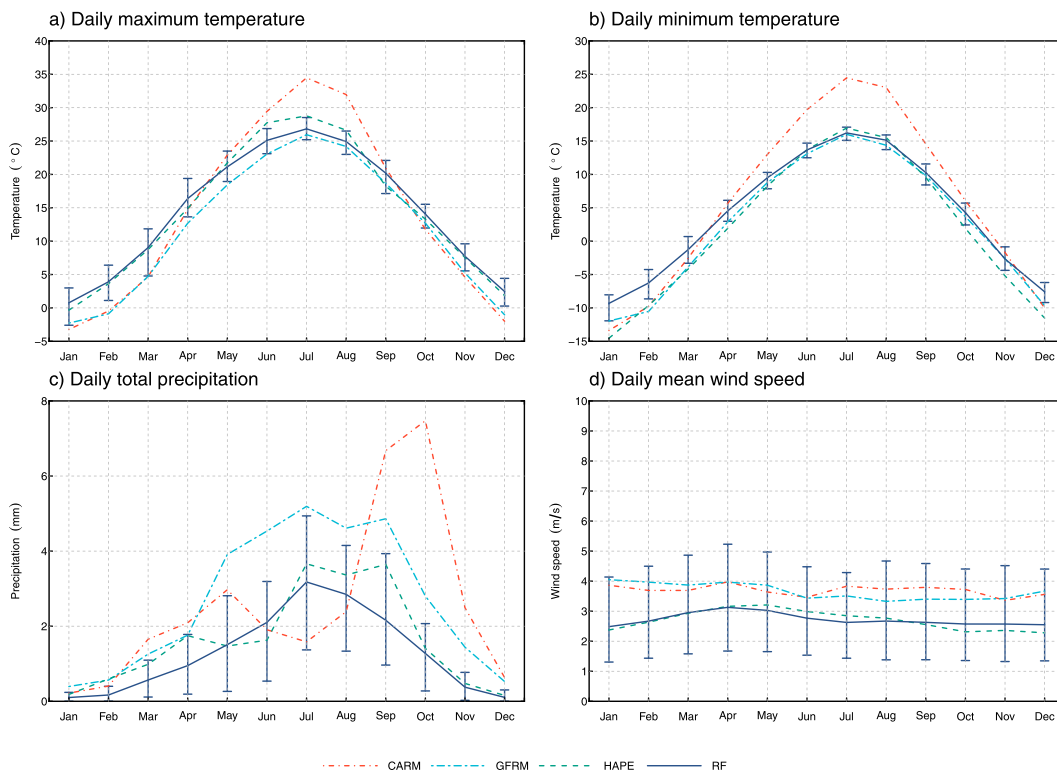FIG. 6. The domain-averaged seasonal cycles of (a) Tmax, (b) Tmin, (c) TP, and (d) WS over JRB obtained from HAPE, CARM, and GFRM for the RF. The blue line with the error bar shows the mean value bounded by a 90% confidence interval of CRU/ERA5.

function (PDF) of the probabilistic predictand y'_s in subset s can be expressed as follows:

$$p(y'_s|\mathbf{o}'_s) = \sum_{r=1}^{k} p(\mathbf{m}'_{sr}|\mathbf{o}'_s) \cdot p(y'_s|\mathbf{m}'_{sr}, \mathbf{o}'_s). \tag{11}$$

The posterior mean and variance of the predictand y'_s can be formulated as follows (Duan et al. 2007):

$$E[y'_s|\mathbf{o}'_s] = \sum_{r=1}^{k} p(\mathbf{m}'_{sr}|\mathbf{o}'_s) \cdot E[p(y'_s|\mathbf{m}'_{sr}, \mathbf{o}'_s)] = \sum_{r=1}^{k} \omega_{sr}\eta_{sr}, \quad \text{and} \tag{12}$$

$$V(y'_s|\mathbf{o}'_s) = \sum_{r=1}^{k} \omega_{sk}\left(\eta_{sr} - \sum_{r=1}^{k} \omega_{sr}\eta_{sr}\right)^2 + \sum_{r=1}^{k} \omega_{sr}\sigma_{sr}^2, \tag{13}$$

where $\omega_{sr} = p(\mathbf{m}'_{sr}|\mathbf{o}'_s)$ and the sum of ω_{sr} for all models is equal to 1; expectation η_{sr} and variance σ_{sr}^2 represent the average and uncertainty of model r given by the observation $\mathbf{o}'_s$, respectively. The expectation-maximization (EM) algorithm is employed to loop all the subsets in order to estimate the parameters ω_{sr} and σ_{sr}^2. The probabilistic projection y'_s for each subset s can be obtained individually using Eqs. (12) and (13). The SCB projection y' for the entire training datasets thus can be developed through the combination of each subset.

To infer probabilistic projection through the SCA, BMA, and SCB, the one-parameter Box–Cox transformation is

performed on the observed and simulated streamflow of each hydrologic model. The detailed equation for the Box–Cox transformation is presented in the supporting information (Text S1). The performance of the three hydrologic models and the three ensemble methods are then evaluated by using four hydrological metrics, including NSE, the normalized root-mean-square error (NRMSE), the RMSE-observations standard deviation ratio (RSR), and the continuous ranked probability score (CRPS). CRPS is a measurement of the integrated squared difference between cumulative distribution functions (CDFs) of simulations and observations (Candille et al. 2007; Hersbach 2000). The detailed equations of these metrics (Eum et al. 2017; Fan et al. 2017; Moriasi et al. 2007; Wagener et al. 2009; Zhou et al. 2018b) are presented in the supporting information (Text S2). As for the probabilistic projections, the daily expected values are determined first in order to calculate these metrics.

The three validated hydrologic models along with SCB are then employed to develop the historical and future streamflow over JRB, driven by high-resolution simulations of GCMs through PRECIS and RegCM. In detail, dynamically downscaled climate variables are first employed to drive three hydrologic models to generate different deterministic simulations for the historical and future periods. SCB is then used to develop probabilistic simulations based on the ensemble hydrologic outputs under multiple GCMs downscaled by RCMs.
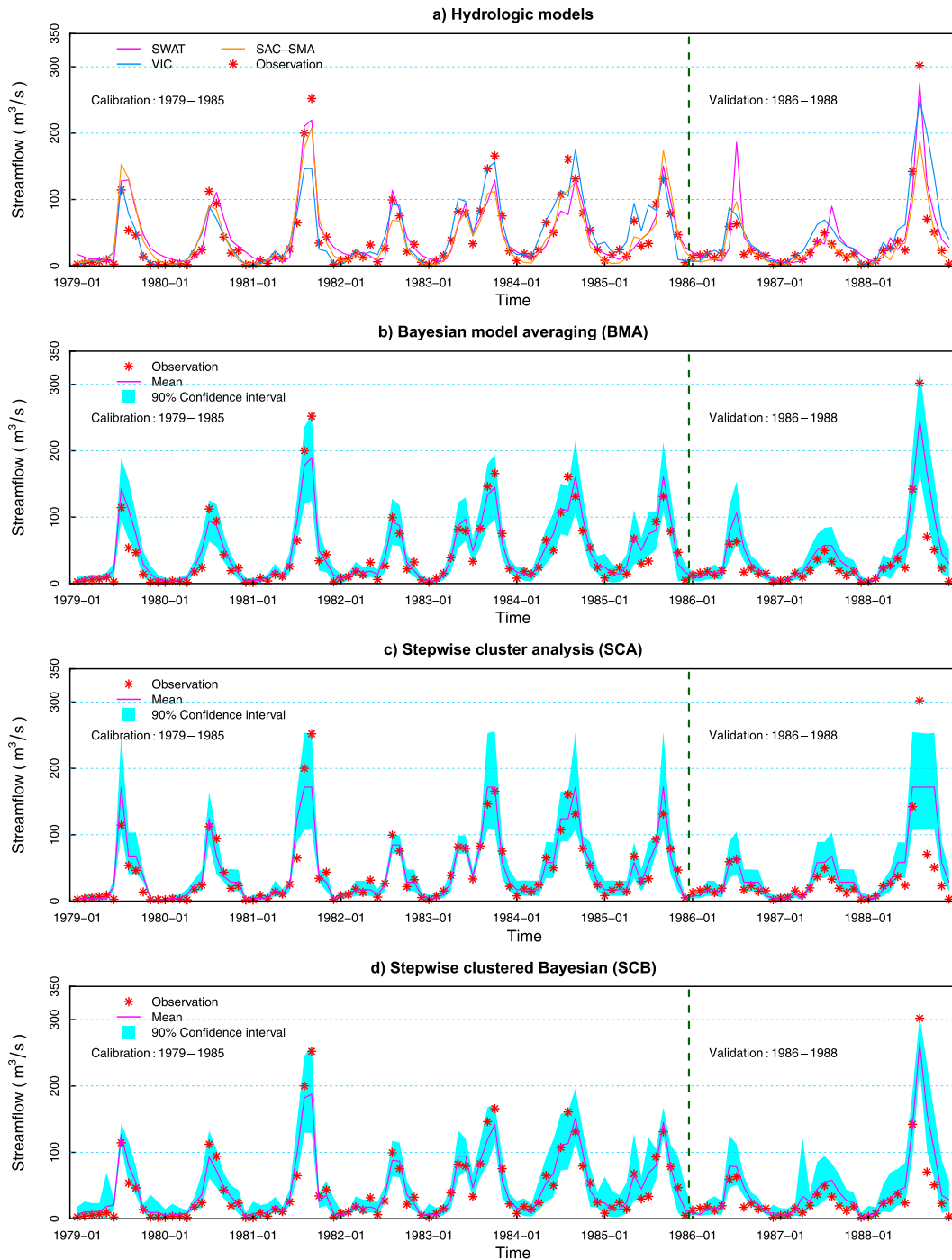
FIG. 7. Simulated and observed monthly streamflow at the Zhangjiashan station (34.63°N, 108.60°E). (a) HM, (b) BMA, (c) SCA, and (d) SCB.

c. Mixed-level factorial design

In this study, impacts of factors, including the anthropogenic emissions and model choices on the changes in precipitation, potential evapotranspiration, and simulated streamflow are investigated. Particularly, the anthropogenic

emissions mainly include two levels, namely, the RCP4.5 and RCP8.5 scenarios. The ensemble projections of PRECIS driven by HadGEM2-ES (herein HAPE), RegCM driven by CanESM2 (herein CARM), and RegCM driven by GFDL-ESM2M (herein GFRM) are composed the three levels of model choices. Changes for the two 30-yr periods,

TABLE 3. Evaluations of streamflow at the ZJS station from the hydrologic models for the calibration and verification period.

Model	Calibration period of 1979–85				Verification period of 1986–88			
	NSE	NRMSE	RSR	CRPS	NSE	NRMSE	RSR	CRPS
SWAT	0.821	0.484	0.423		0.703	0.845	0.545	
VIC	0.829	0.473	0.413		0.617	0.960	0.619	
SAC	0.829	0.472	0.413		0.792	0.708	0.456	
SCA	0.877	0.401	0.350	6.952	0.500	1.097	0.707	14.495
BMA	0.882	0.393	0.344	8.013	0.772	0.741	0.478	9.413
SCB	0.912	0.339	0.297	6.885	0.825	0.649	0.418	9.307

including the 2050s (2036–65) and the 2080s (2066–95) relative to the baseline period (i.e., 1985–2004), are taken into consideration. The mixed-level factorial experiments are then performed at all the possible combinations of GCM–RCMs (i.e., HAPE, CARM, and GFRM), RCPs (i.e., RCP4.5 and RCP8.5), and periods (i.e., the 2050s and 2080s).

The ANOVA approach (Zhu et al. 2018) is employed to investigate the main and interaction effects of different contributors to the changes in precipitation, potential evapotranspiration, and simulated streamflow. ANOVA is useful in estimating the effects of anthropogenic emissions and model choices, by decomposing the total variance into a sum of squares due to anthropogenic emissions, a sum of squares due to model choices, a sum of squares due to the interaction between these two factors, and an error sum of squares. Two-way ANOVA is used to decompose the contributions to the total uncertainty in the responses into not only the main effects of anthropogenic emissions and model choices but also their interactions. The analysis of variance model for projected changes driven by multiple climate simulations can be formulated as

$$y_{ijl} = \tau + A_i + U_j + (AU)_{ij} + \varepsilon_{ijl}, \qquad (14)$$

where τ is the overall mean effects of the projected changes in precipitation, potential evapotranspiration, or simulated streamflow; particularly, these changes can be calculated by subtracting the annual averages in the historical period from their averages in the future periods; A_i is the main effect of anthropogenic emission scenario i; U_j is the main effect driven climate model j; $(AU)_{ij}$ is the interaction effects of anthropogenic emission scenario i and driven climate model j; y_{ijkl} is the projected changes in precipitation, potential evapotranspiration, or simulated streamflow driven by climate simulation j under the RCP scenario i in period k; ξ_{ijkl} is residual terms. To perform the analysis of variance, the equations of the corresponding total sum of squares and its individual and interaction contribution components are derived as follows:

$$SS_A = \frac{1}{un} \sum_{i=1}^{a} y_{i..}^2 - \frac{y_{...}^2}{aun}, \qquad (15)$$

$$SS_U = \frac{1}{an} \sum_{j=1}^{u} y_{.j.}^2 - \frac{y_{...}^2}{aun}, \qquad (16)$$

$$SS_{AU} = \frac{1}{n} \sum_{i=1}^{a} \sum_{j=1}^{u} y_{ij.}^2 - \frac{y_{...}^2}{aun} - SS_A - SS_U, \quad \text{and} \qquad (17)$$

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{u} \sum_{l=1}^{n} y_{ijl}^2 - \frac{y_{...}^2}{aun}, \qquad (18)$$

where SS_A and SS_U are the main effect of anthropogenic emissions and model choices, respectively; SS_{AU} is the interactive effect of these two factors; SS_T is the total sum of squares; $y_{...}$ represents the sum of change responses; $y_{i..}$ and $y_{.j.}$ are respectively defined as the total of ith level of factor A and jth level of factor U; $y_{ij.}$ is the corresponding sum at the combination of two factors; a and u respectively denotes the total levels of factor A and U; and n is the replicates of each combination of three factors. Moreover, internal variability arises from the internal dynamics of the climate system (Frankcombe et al. 2015). In this study, a trend model will be fitted to these variables in order to estimate the internal variability since there is only one realization for each experiment (Hawkins and Sutton 2011; Lafaysse et al. 2014). The daily time series of total precipitation (mm day^{-1}), potential evapotranspiration (mm day^{-1}), and streamflow (m^3 s^{-1}) for the period of 1985–2095 are estimated with smoothing splines (Evin et al. 2019). The smooth.spline function in the R software (https://www.r-project.org/) is employed to fit the trends.

4. Results

a. Evaluation of regional climate models

Figure 4 presents spatial distributions of annual averages of daily maximum temperature, minimum temperature, total precipitation, and wind speed over JRB extracted from different RCMs driven by multiple GCMs for the historical (i.e., 1985–2004). In general, as seen from the spatial correlations with CRU or ERA5 in Fig. 4, the ensemble ranges of the HAPE, CARM, and GFRM simulations can reasonably capture the spatial patterns of these variables except for daily total precipitation (Fig. 4). For example, the spatial correlations of daily maximum temperature, minimum temperature, and wind speed range from 0.8749 to 0.9751, while the maximum spatial correlation for daily total precipitation is 0.4440. Figure 5 further shows the calculated differences of the selected variables between simulations and observations/reanalyses (i.e., ERA5 and CRU TS, v4.05). For example, there are differences from −2.0° to 3.0°C in daily
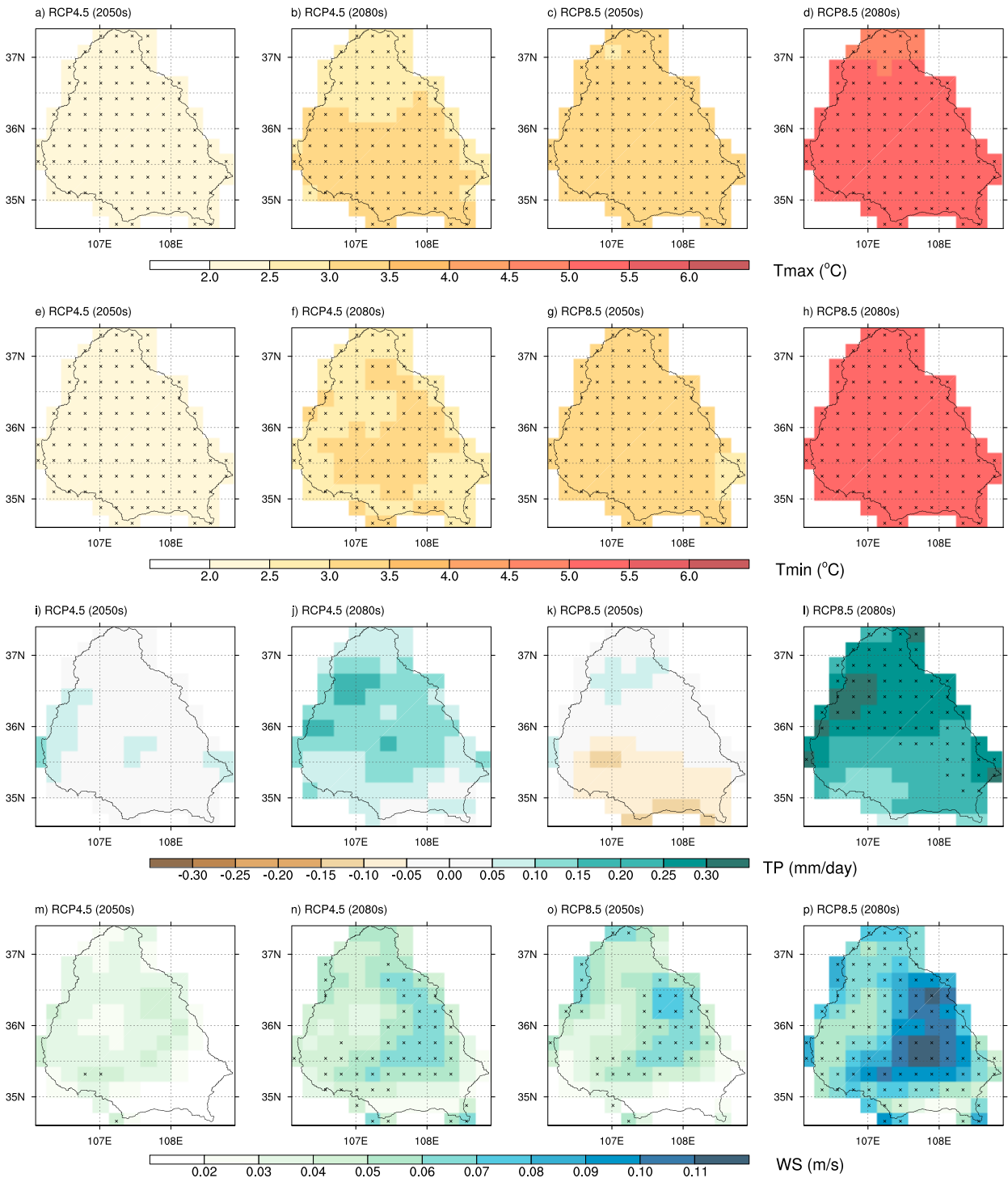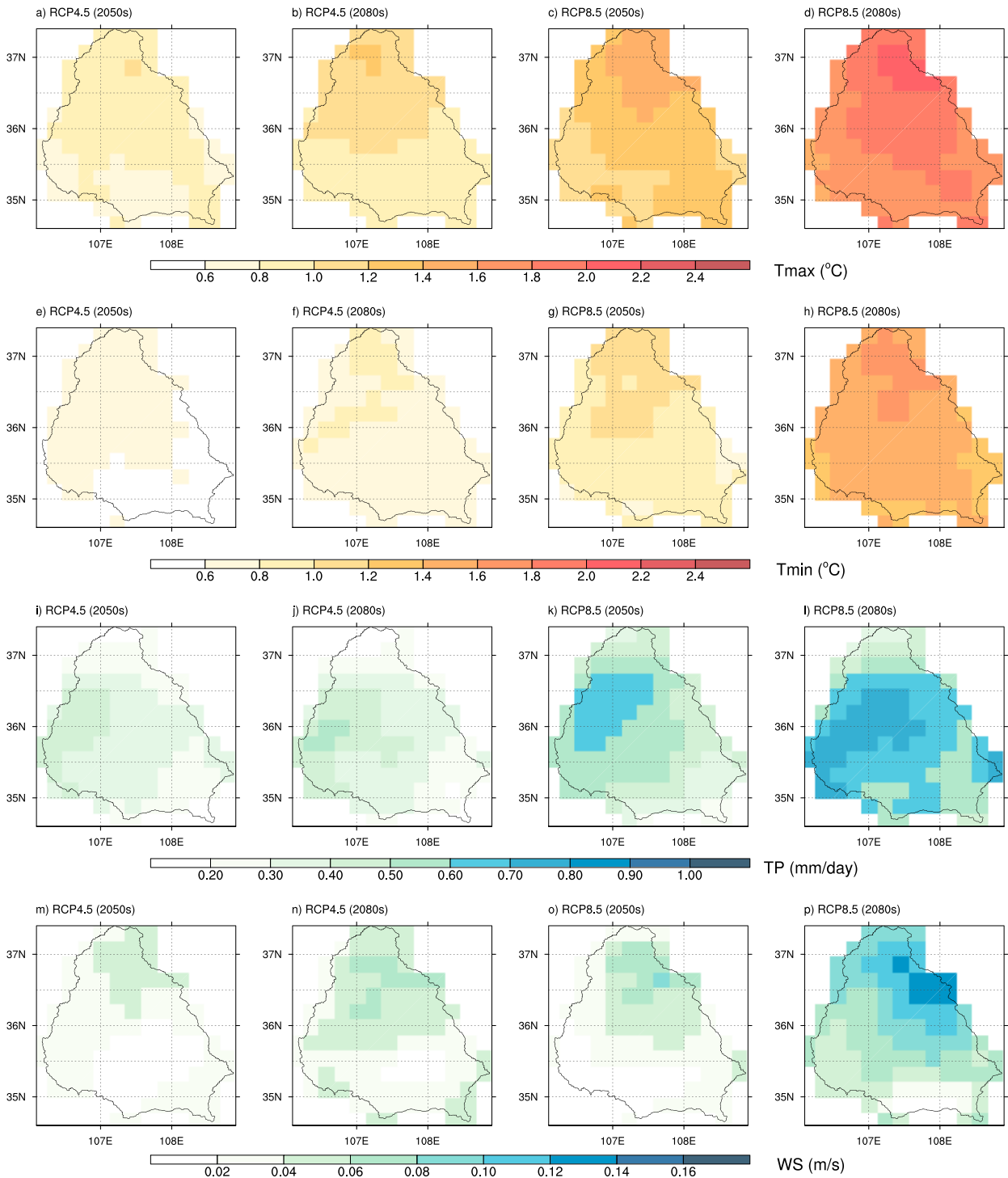
FIG. 8. Averaged values of the projected changes in (a)–(d) Tmax, (e)–(h) Tmin, (i)–(l) TP, and (m)–(p) WS under RCPs derived from the RCM simulations. Columns show the changes in the 2050s under RCP4.5, the 2080s under RCP4.5, the 2050s under RCP8.5, and the 2080s under RCP8.5, from left to right, respectively. Student's *t* test is employed to test the null hypothesis that the mean values in the baseline period (i.e., 1985–2004) and the future periods (i.e., 2036–65 and 2066–95) are from the same population. The grids with a cross mark are considered to be significant since *p* values are less than 0.05.

FIG. 9. Standard deviations over an ensemble of the projected changes in (a)–(d) Tmax, (e)–(h) Tmin, (i)–(l) TP, and (m)–(p) WS under RCPs derived from dynamically downscaled simulations. Columns show the changes in the 2050s under RCP4.5, the 2080s under RCP4.5, the 2050s under RCP8.5, and the 2080s under RCP8.5, from left to right, respectively.

maximum temperature (Tmax) as simulated by HAPE (Fig. 5a) and CARM (Fig. 5c). The GFRM simulation (Fig. 5b) tends to underestimate the Tmax value, with a cold bias from −4.0° to 0°C. However, relatively small differences of

total precipitation (TP; i.e., from −40% to 80%) are derived from HAPE (Fig. 5d), while there are large wet biases (i.e., up to 160%) in the CARM and GFRM simulations. This indicates that the performance in replicating daily total
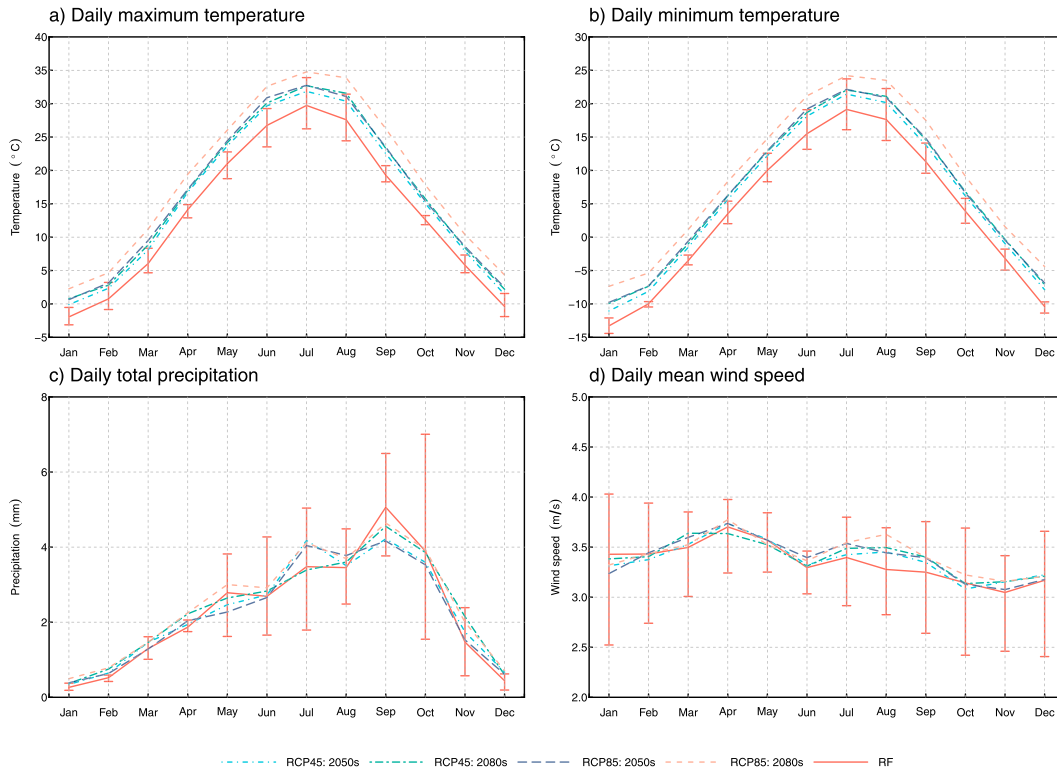
FIG. 10. The projected domain-averaged seasonal cycles of (a) Tmax, (b) Tmin, (c) TP, and (d) WS over JRB obtained from an ensemble mean, which is calculated based on the three members of RCM simulations (i.e., HAPE, CARM, and GFRM) for the RF and the future periods. The orange line with the error bar shows the mean value bounded by a 90% confidence interval of the three members.

precipitation is not as effective as that for temperature. As for WS (at a height of 10 m), the magnitude and pattern of wind speed obtained from ERA5 are not well replicated by RCMs except for HAPE (Figs. 4m and 5j). The mean differences between the RCM simulations and the ERA5 reanalyses vary from -0.56 m s^{-1} (i.e., from HAPE) to 1.58 m s^{-1} (i.e., from GFRM).

To further assess the RCM performance, we compared the averaged seasonal cycles of Tmax, daily minimum temperature (Tmin), TP, and WS over JRB against the observations/reanalyses. Figure 6 shows the comparison of the averaged seasonal cycles of the four variables for the period of 1985–2004 over JRB among HAPE, CARM, GFRM, and observations/reanalyses. The error bars indicate the 90% confidence interval of the observed datasets, implying that with 90% confidence the observation for any given month lies within such range. As can be seen in Fig. 6, HAPE and GFRM perform reasonably well in simulating Tmax and Tmin since they are close to or within the confidence interval of the averaged observations/reanalyses. However, it is interesting to note that CARM overestimates Tmax and Tmin during the summer months (i.e., June, July, and August), as well as TP in the autumn (i.e., September, October, and November). Furthermore, WS from all the simulations is within the confidence range of the observed data, meaning that the skill of

RCMs in reproducing the averaged seasonal cycles of this variable is even higher than the other three variables, and also that predictability may be low. Figures S1–S4 in the online supplemental material further evaluate the RCM performance in simulating seasonal climate through a comparative analysis against the observed Tmax, Tmin, TP, and WS over JRB.

b. Evaluation of hydrologic models

The three hydrologic models driven by the observed climate forcing are spun up from cold start for a 3-yr period (i.e., 1976–78), which is able to establish a stable moisture content of soil layer (Shrestha et al. 2014; Zhou et al. 2018b). Figure 7 compares the monthly mean observed and modeled streamflow at the Zhangjiashan station simulated by three hydrologic models (i.e., SAC-SMA, SWAT, and VIC). The results demonstrate that all three models have a good performance in simulating the historical monthly streamflow for the calibration and validation period. However, the results also indicate that the three models show different capabilities to capture various ranges of streamflow. Figures 7b, 7c, and 7d display the expected simulations of BMA, SCA, and SCB, respectively, along with the 90% confidence intervals of the three-member ensemble mean. As shown in the figures, the goodness of fit (i.e., NSE) with the observed streamflow for the multimodel
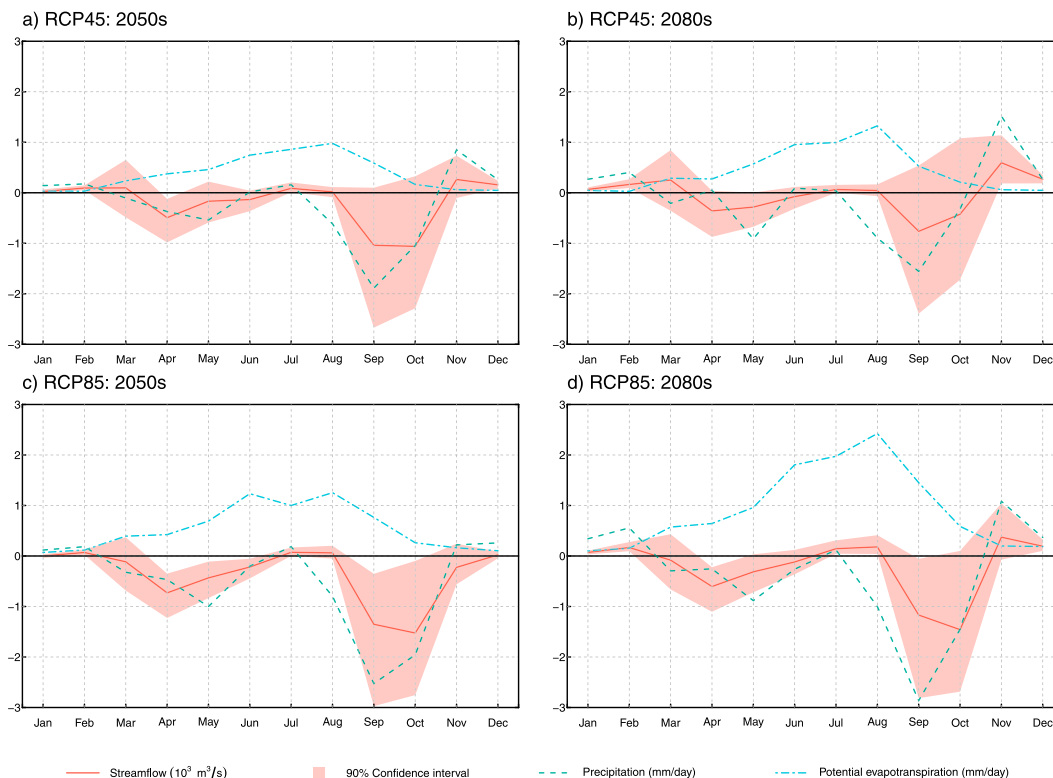
FIG. 11. Seasonal cycle changes in the CARM-projected precipitation and potential evapotranspiration, as well as the SCB-projected streamflow over JRB for (a) the 2050s under RCP4.5, (b) the 2080s under RCP4.5, (c) the 2050s under RCP8.5, and (d) the 2080s under RCP8.5 under RCPs relative to the baseline period (i.e., 1985–2004). The semitransparent shadow area shows the 90% confidence interval of changes in streamflow.

mean is improved as compared to that for the specific models (SAC-SMA, SWAT, and VIC).

The verification values of NSE, NRMSE, RSR, and CRPS for the SCB method achieve 0.825, 0.649, 0.418, and 9.307, respectively. During the calibration and validation period, the NSE values are larger than 0.8, while the RSR values are less than 0.5, which explains the good performance of the proposed SCB method (Moriasi et al. 2007). Compared with the three hydrologic models, the expected SCB simulation for the calibration and verification period have the highest NSE values and the lowest values in both NRMSE and RSR, which suggests that the developed method is much better than the three deterministic ones.

As seen from smaller CRPS in Table 3, the integrated squared difference in CDFs (i.e., between modeled and observed streamflow) is reduced in the ensemble simulations of SCB for both calibration and validation periods when compared to BMA. One reason for more uncertainties in BMA compared with SCB is that the BMA method is to infer a probabilistic projection based on one set of the posterior information, without accentuating different capabilities of the three hydrologic models in simulating different phases of the hydrograph. As seen from relatively larger NSE values and lower values of NRMSE, RSR, and CRPS compared to the SCA and BMA method, the probabilistic simulations of SCB

are more consistent with the observations, thus affirming better performance of the developed method. In general, the proposed SCB method has a good performance in reproducing the historical streamflow over JRB.

c. Projections of climatic changes

Figure 8 shows the averaged value of the projected changes in daily maximum and minimum temperature, total precipitation, and wind speed for the 2050s and 2080s under RCPs derived from the RCM simulations. As shown in Figs. 8a–h, it is found that the daily maximum and minimum temperature is projected to increase from the 2050s to the 2080s under RCPs. For example, the spatial range of the three-model mean values of daily maximum temperature for the 2050s and 2080s under RCP8.5 is 3.0°–3.4°C and 4.9°–5.5°C, respectively. However, the results further suggest that the spatial pattern between daily maximum and minimum temperature differs from each other. Much higher increases would occur in the southern JRB for the daily maximum temperature and in the middle area of JRB for the daily minimum temperature. Additionally, the daily maximum temperature under RCP8.5 in the 2050s and 2080s is expected to respectively increase by 3.2° and 5.2°C, which are much higher than the increases under RCP4.5. This implies the effects of GHG concentrations on future projections.
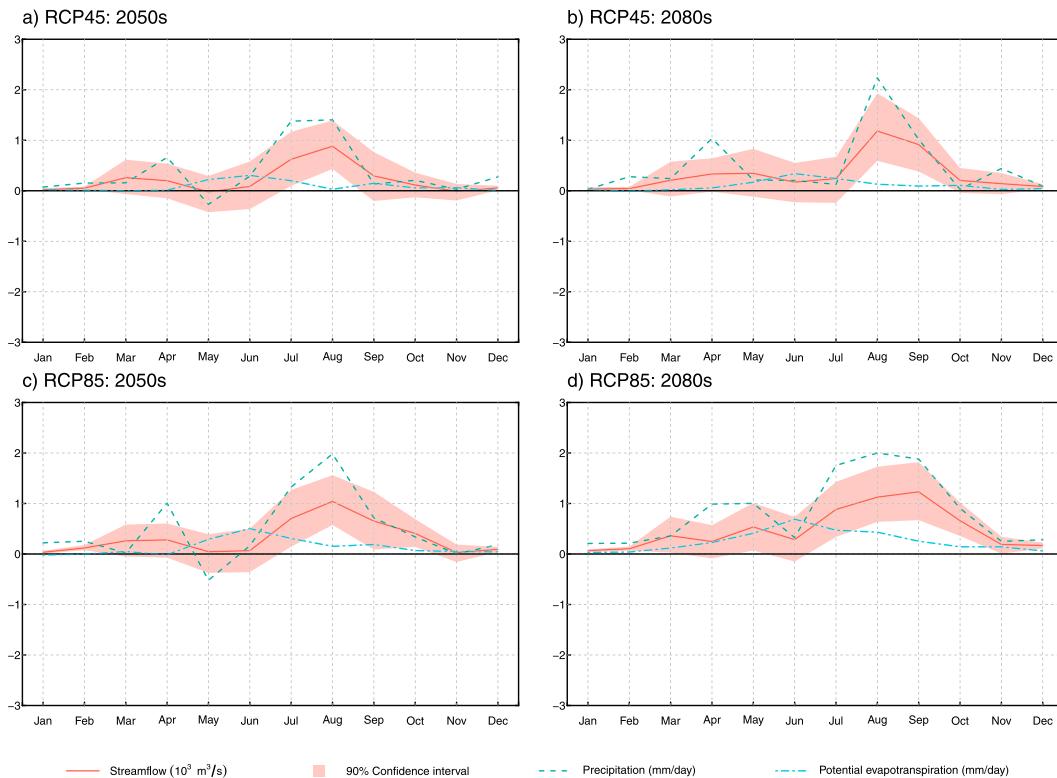
FIG. 12. As in Fig. 11, but for seasonal cycle changes in the GFRM-projected precipitation and potential evapotranspiration.

Similarly, the daily total precipitation derived from the ensemble climate simulations is projected to increase in the 2050s and 2080s under RCPs (Figs. 8i–8l). Student's t test is employed to test the null hypothesis that the mean values in the baseline period (i.e., 1985–2004) and the future periods (i.e., 2036–65 and 2066–95) are from the same population. The grids with a cross mark are considered to be significant since p values are less than 0.05. The projected changes for the daily total precipitation in the 2050s under RCP4.5, the 2080s under RCP4.5, the 2050s under RCP8.5, and the 2080s under RCP8.5 will increase as much as 0.11, 0.18, 0.08, and 0.33 mm day^{-1}, respectively. Moreover, the results also indicate that the northeastern JRB would experience a statistically significant increase in the 2080s under RCP8.5.

Nevertheless, the spatial distribution of RCMs-projected changes in the daily wind speed for the two future periods under RCPs (Figs. 8m–p) depicts a different pattern than other climate variables. The results indicate that there are no significant changes in the daily wind speed in the north part of JRB under RCP4.5. However, slight increases are found in the middle part of JRB, with a maximum increase of 0.05, 0.07, 0.08, and 0.11 m s^{-1} in the 2050s under RCP4.5, the 2080s under RCP4.5, the 2050s under RCP8.5, and the 2080s under RCP8.5, respectively.

Figure 9 presents the standard deviation of daily multimodel means of the projected changes in daily maximum and minimum temperature, total precipitation, and wind speed. As seen from much higher standard deviations, there is

more considerable variability in daily maximum and minimum temperature over the north than over the south. Moreover, under the scenario of RCP8.5, the uncertainties in Tmax and Tmin are further aggravated during the 2080s, indicating that climate change is associated with larger variability under RCP8.5 for the 2080s. However, it is indicated that greater variability is found in the western of JRB for the daily total precipitation, while the projected changes in daily wind speed have higher uncertainties in the north.

The projected seasonal cycles of daily maximum and minimum temperature, total precipitation, and wind speed are shown in Fig. 10. In general, the average daily maximum and minimum temperature in each month are projected to increase. For example, the maximum increase in average Tmax for the 2050s under RCP4.5, the 2080s under RCP4.5, the 2050s under RCP8.5, and the 2080s under RCP8.5 is 3.2°, 4.0°, 4.2°, and 7.0°C, respectively. In contrast, there would be a decrease in the daily total precipitation for a few months. Notably, the maximum decrease in the daily total precipitation is projected in September, which is 0.84, 0.13, 0.90, and 0.41 mm day^{-1} for the 2050s under RCP4.5, the 2080s under RCP4.5, the 2050s under RCP8.5, and the 2080s under RCP8.5, respectively. Moreover, the daily wind speed is projected to increase in the summer, while there are no significant changes in other months. For example, the maximum increase in the 2050s under RCP4.5 is 0.17 m s^{-1}. Moreover, spatial patterns of seasonal averages and
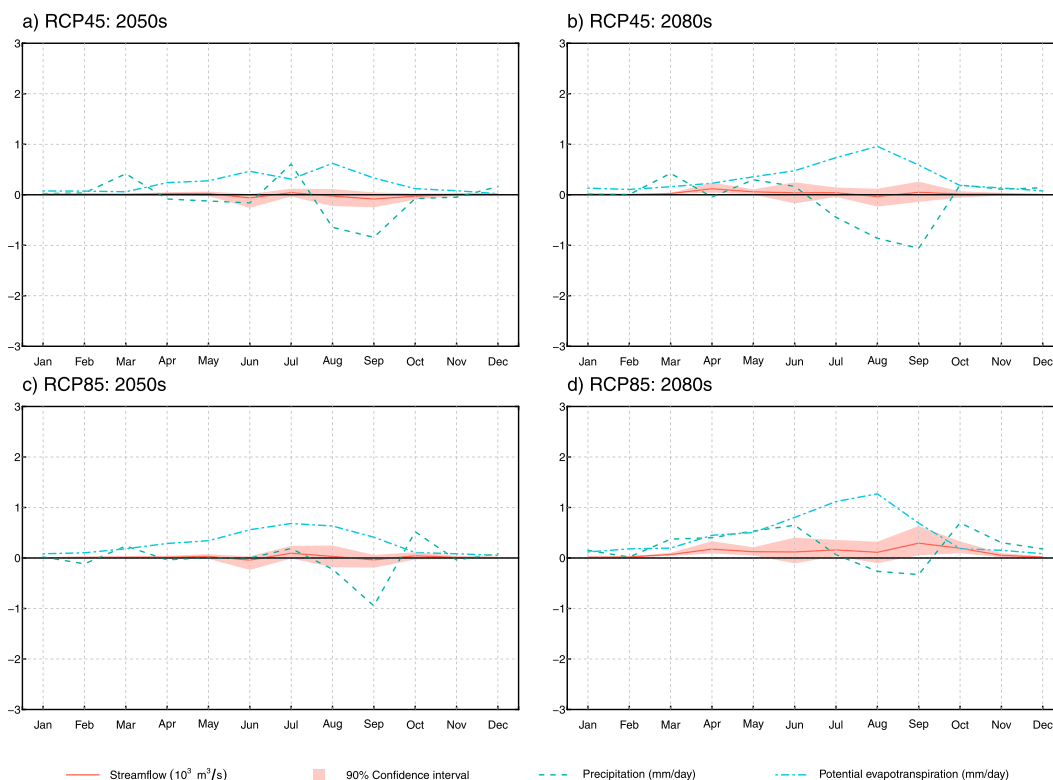
FIG. 13. As in Fig. 11, but for seasonal cycle changes in the HAPE-projected precipitation and potential evapotranspiration.

standard deviations of changes in the four variables are further shown in Figs. S5–S12.

d. Projections of hydrological changes

Subsequently, the RCM-projected climate projections driven by multiple GCMs are individually inputted to the three HMs to generate deterministic hydrologic simulations for the baseline and future period. The validated SCB method is then employed to develop more reliable probabilistic projections based on the Box–Cox transformed simulations of three hydrologic models over JRB. To explore the impacts of climate change on hydrology, the expected values and 90% confidence intervals of seasonal cycles of the SCB-projected changes in the streamflow for the 2050s and 2080s relative to the baseline period are derived. Moreover, the average changes in both precipitation and potential evapotranspiration are explored for an improved understanding of the hydrologic effects of climate change. The potential evapotranspiration for each hydrologic model is computed by using the Penman method with the RCM outputs.

The seasonal streamflow in JRB at the Zhangjiashan gauge to CARM under RCPs is shown in Fig. 11. It is indicated that decreased precipitation and streamflow are generally found across all months except for November–December. For example, the maximum increase of the expected changes in streamflow for the 2050s and 2080s under RCP4.5 is 0.30 and 0.59×10^3 m^3 s^{-1} in November, respectively. In addition, the maximum decreases in the streamflow are projected in October

for the 2050s under RCP4.5, the 2080s under RCP4.5, and the 2080s under RCP8.5, and in September for the 2080s under RCP4.5. One explanation for such decreased streamflow is due to a reduction of the daily total precipitation and in autumn without a significant intensification of daily potential evapotranspiration, resulting in less surface runoff. For example, as shown in the figure, the most substantial decrease in daily total precipitation with a value of 1.88 and 1.55 mm day^{-1} is also projected in September for the 2050s and 2080s under RCP4.5, respectively. Moreover, the semitransparent shadow area shown in Fig. 11 presents the changes in streamflow within the 90% confidence interval. The wide confidence interval in September–October streamflow indicates substantial uncertainty during that time of year. The large spread of streamflow from September to November might be attributed to significant changes in precipitation from CARM, which could generate a wide range of estimates by using different types of hydrologic models. However, it is expected to project large decreases in streamflow from September to November due to the CARM-projected reductions in precipitation.

The daily total precipitation and potential evapotranspiration, as well as the SCB-projected streamflow under the boundary condition of GFRM, are presented in Fig. 12, while the responses to HadGEM2-ES are depicted in Fig. 13. Compared with CARM, streamflow driven by GFRM is generally projected to increase for the 2050s and 2080s under RCPs due to higher projected precipitation. In contrast, there is no significant change in the streamflow based on HAPE except for a general

TABLE 4. Pearson correlation coefficients between the streamflow and the precipitation/potential evapotranspiration for the historical period.

	Daily total precipitation			Potential evapotranspiration		
	CARM	GFRM	HAPE	CARM	GFRM	HAPE
RF	0.81	0.90	0.95	−0.23	0.78	0.72

increase in the 2080s under RCP8.5. Tables 4 and 5 present the Pearson correlation coefficients between the monthly streamflow and the monthly precipitation/potential evapotranspiration for the historical period (Table 4), as well as between their changes for the future periods (Table 5). These results suggest that the change in streamflow over JRB is more positively related to monthly total precipitation than negatively related to monthly potential evapotranspiration. Specifically, the results indicate that the correlation between the streamflow and the potential evapotranspiration is generally weaker (except for HAPE in the 2080s) when compared to the monthly total precipitation. For example, GFRM has a Pearson's correlation value of 0.8801 and 0.9427 for monthly total precipitation in the 2050s and 2080s under RCP8.5, respectively. This confirms that there is a strong agreement between the projected monthly streamflow and precipitation. Nevertheless, smaller Pearson's correlation values suggest fewer agreements between the streamflow and potential evapotranspiration. Moreover, the correlations with monthly potential evapotranspiration/total precipitation are larger from GFRM than CARM.

e. Main and interaction effects of uncertain factors

The total variance of the changes in monthly variables (i.e., monthly precipitation, potential evapotranspiration, and streamflow) decomposed by uncertain factors and internal variability in the 2050s are shown in Fig. 14. The uncertain factors include anthropogenic emissions and model choices, as well as their interaction effects. The results indicate that the median contributions of each factor to the three variables are generally similar, except internal variability (IV). In general, the total variance of the long-term projected changes in three monthly variables over JRB can be primarily explained by the GCM model choices (Figs. 14a–f). The percentage of contribution of the RCP scenarios to the change in monthly precipitation has the least impact. However, it makes a more significant contribution to the total variance of changes in monthly potential evapotranspiration. The interaction effect is more significant than the main effects of anthropogenic emissions. The internal variability has the

greatest impact on the total uncertainty of monthly variables. Such influence is also found in the seasonal and annual variables. Figure 15 presents the results of the same analysis for the 2080s. It shows that the RCP scenario has a more significant contribution to the changes in these variables. Moreover, we found that the relative effect of the internal variability on changes in the three variables is relatively smaller in the 2080s than in the 2050s. The order of importance for these factors (i.e., U, IV, AU, and A) is still intact, but the RCP scenario has a larger contribution in the longer-term projection.

5. Discussion

In this study, the results suggest that the model choices can primarily explain the total variance of changes in monthly precipitation, potential evapotranspiration, and streamflow over JRB. Similarly, Evin et al. (2019) found that GCM choices have significant impacts on climate change of four French mountain massifs located in the Pyrenees and in the Alps. However, it is important to note that the main limitation of this study is that there are only three GCMs, which is not sufficient enough to conclusively determine relative contributions to uncertainty in future projections. The limited number of RCMs employed is another limitation of the current study. Moreover, as found by Evin et al. (2019), RCM choices have a smaller contribution to temperature, precipitation, and wind speed than GCM choices.

Driven by additional GCMs will dramatically increase the computational effort, which therein advanced methods such as fractional design are of necessity. With the assumption of insignificant in high order effects, the main and interactive effects can be quantitatively estimated through a fraction of the entire experiments at combinations of the levels of multiple factors. Therefore, quantifying the uncertainty associated with boundary conditions through factional factorial analysis will thus require future research efforts.

It should be noted that the choices of SCB, BMA, and SCA for developing the three-member ensemble mean (SAC, VIC, and SWAT) could also have impacts on the contribution study. Further calibration and validation of hydrological models with longer time series of observations could be performed, which will improve the suitability of calibrated model parameters.

6. Conclusions

In this study, an SCB ensemble method has been developed to improve the performance of long-term hydroclimate

TABLE 5. Pearson correlation coefficients between the changes in streamflow and the changes precipitation/potential evapotranspiration for the future periods.

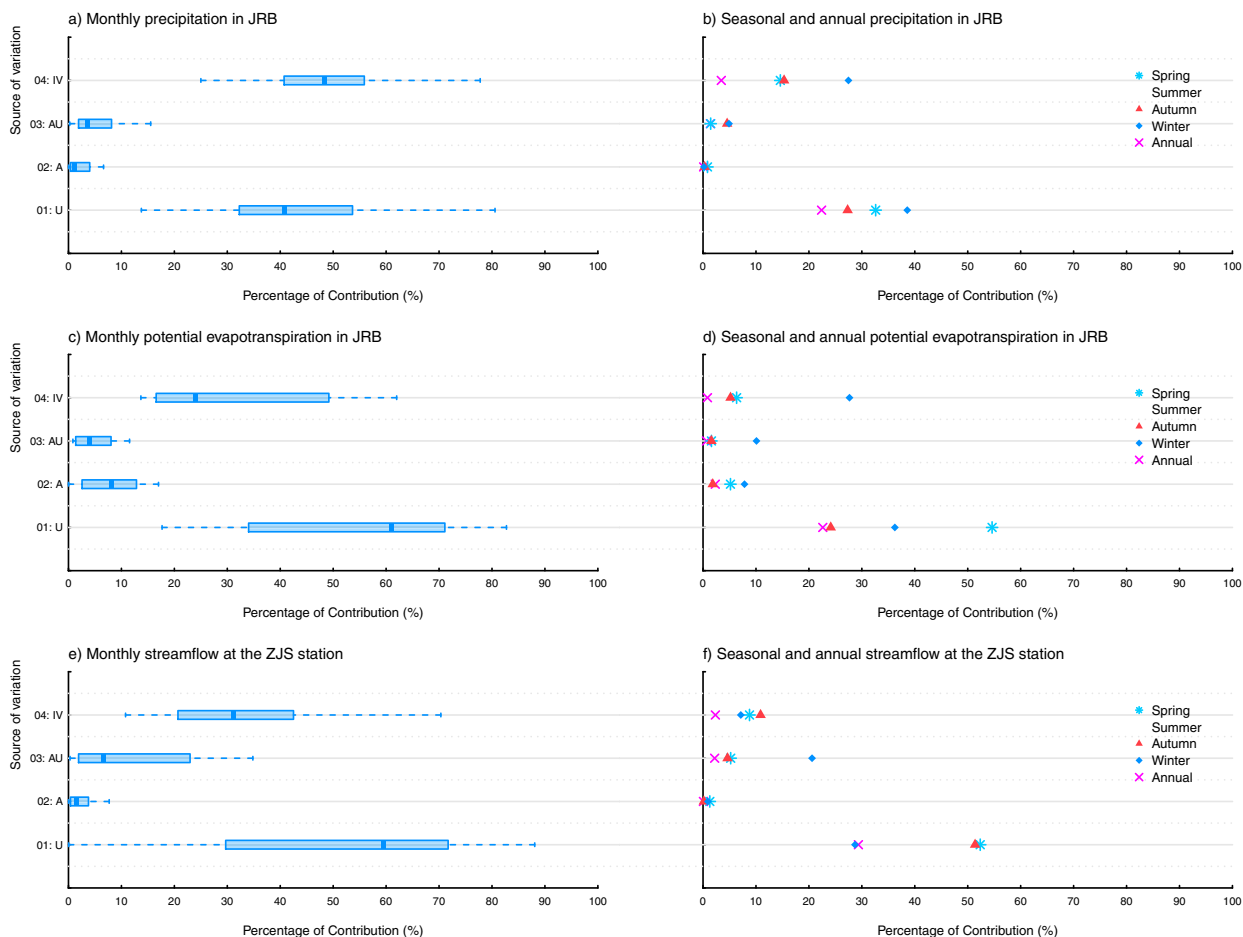|	Changes in total precipitation			Changes in potential evapotranspiration		
	CARM	GFRM	HAPE	CARM	GFRM	HAPE
RCP45 (2050s)	0.87	0.90	0.76	−0.11	−0.01	−0.46
RCP45 (2080s)	0.78	0.89	0.13	−0.21	0.15	−0.13
RCP85 (2050s)	0.89	0.88	0.56	0.05	0.14	0.17
RCP85 (2080s)	0.80	0.94	−0.11	−0.04	0.43	0.47

FIG. 14. Percentage of contribution from multiple uncertain factors in the 2050s. (left) The percentages of the contribution of each factor to the changes in monthly variables. (right) The percentage of contribution of each factor to the changes in seasonal and annual variables. (a),(b) Precipitation, (c),(d) potential evapotranspiration, and (e),(f) streamflow. Here, A and U denote the main effects of anthropogenic forcings (i.e., RCP4.5 and RCP8.5) and model choices (i.e., CARM, GFRM, and HAPE), respectively.

projections based on BMA and SCA. The developed ensemble method was then applied to investigate climate change impacts on the hydrologic regime within a general framework, which was developed through the integration of RCMs and HMs. Specifically, RCMs were first performed to downscale boundary conditions from multiple GCMs. Meanwhile, HMs were calibrated against the historical observations using SCE-UA. Afterward, the performance of climate and hydrologic simulations by using RCMs, HMs, and the ensemble methods (i.e., SCA, BMA, and SCB) was validated through comparison to the historical observations. The validated RCMs, HMs, and SCB were then applied to develop probabilistic long-term projections of hydrological changes. The mixed-level factorial inference approach is utilized to reveal the main and interaction effects of the anthropogenic emissions and model choices.

Based on the factorial analysis of multiple uncertain factors in long-term climate projections and hydrological modeling over JRB, the substantial findings of this study can be summarized as follows:

1) The range of ensemble RCMs simulations is able to reproduce the historical spatial and temporal patterns of annual and seasonal temperature. We also found that RCMs perform more effectively in capturing temperature variables than other precipitation and wind speed. However, uncertainty in the observations/reanalyses (i.e., CRU and ERA5) is another uncertainty source that future studies should consider.

2) The SCB streamflow driven by GFRM is generally projected to increase for the 2050s and 2080s under RCPs due to its larger precipitation. In contrast, there is no significant change in the SCB streamflow based on HAPE except for a general increase in the 2080s under RCP8.5. It is revealed that the change in streamflow in JRB is more positively related to daily total precipitation than negatively related to daily potential evapotranspiration.

3) Our findings suggest that the GCM model choices can primarily explain the total variance of changes in monthly precipitation, potential evapotranspiration, and
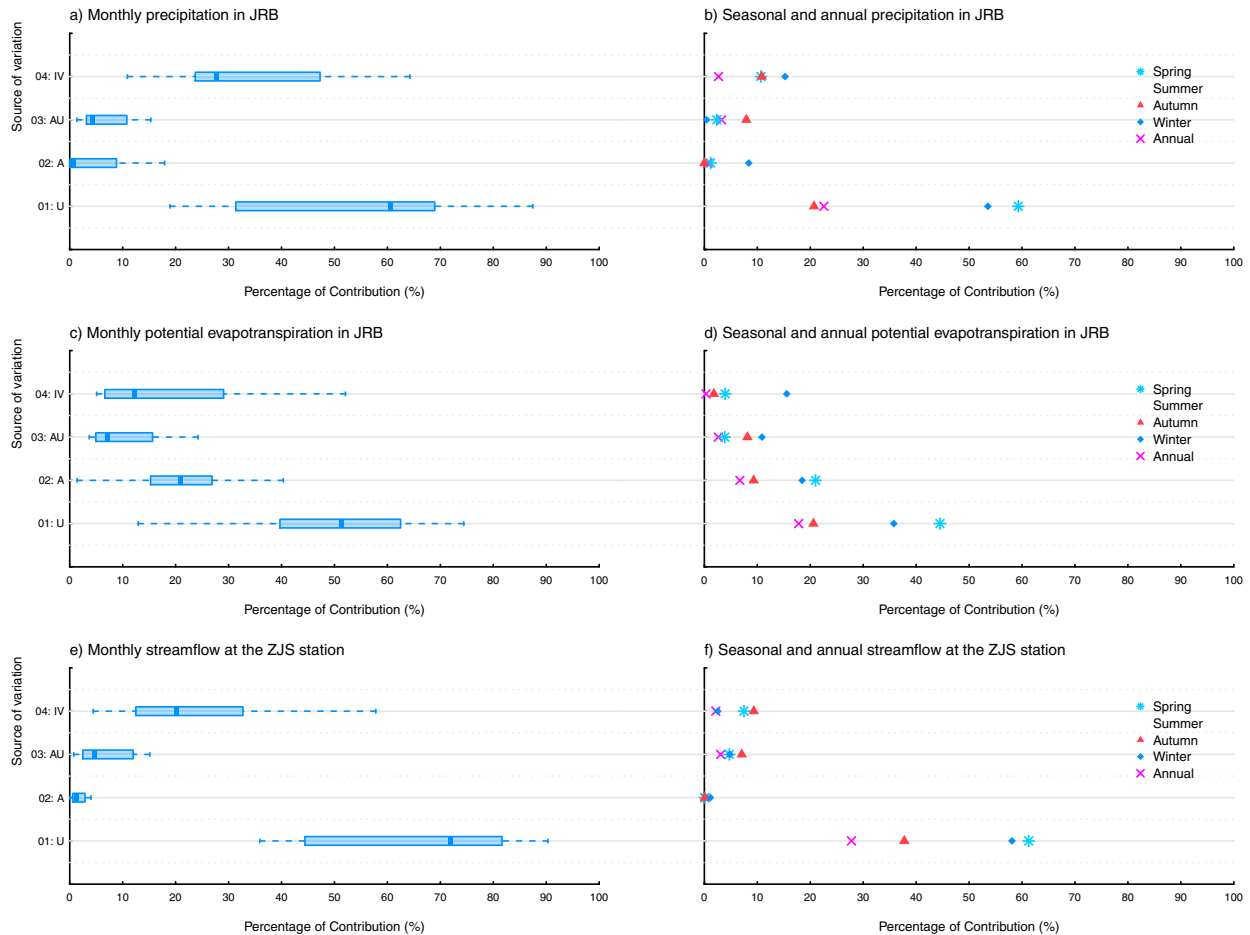
FIG. 15. As in Fig. 14, but for the 2080s.

streamflow over JRB. Moreover, it is revealed that the percentage of contribution of anthropogenic emissions to the changes in monthly precipitation is relatively smaller, but it makes a more significant contribution to the total variance of changes in potential evapotranspiration and streamflow. Meanwhile, the internal variability has a substantial impact on the total uncertainty of monthly variables. Nevertheless, such influence is decreased at the seasonal and annual time scales. In addition, the RCP scenario has a more significant contribution to the changes.

This study is the first attempt to develop the SCB method to improve the long-term projections based on BMA and SCA, without the requirements of any subjective assumption to split observations into multiple sets. The proposed mixed-level factorial inference approach was applied in order to reveal the main and interaction effects of multiple factors in long-term hydrological projections for the JRB. The results conclude that the developed framework based on SCB can thoroughly explore the potential impacts of climate change and the associated uncertainties. The results of this study are encouraging for further integration of RCMs and statistical methods to drive more HMs in order to identify model uncertainties within the context of a changing climate.

Data availability statement. The climate datasets presented in this research are available from the Climate Change Data Portal (http://ccdp.network/). The observations are acquired from the National Meteorological Information Center (http://data.cma.cn/). The elevation datasets are obtained from the hydrological data and maps website (https://www.hydrosheds.org/). The vegetation data are retrieved from the AVHRR Global Land Cover Classification (https://www.arcgis.com/home/

item.html?id=70c54b0b7b344c418dee4af9029fe6f2). The soil parameters are collected from the Harmonized World Soil Database (https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/).

REFERENCES

Arnell, N. W., and S. N. Gosling, 2013: The impacts of climate change on river flow regimes at the global scale. *J. Hydrol.*, **486**, 351–364, https://doi.org/10.1016/j.jhydrol.2013.02.010.

——, and ——, 2016: The impacts of climate change on river flood risk at the global scale. *Climatic Change*, **134**, 387–401, https://doi.org/10.1007/s10584-014-1084-5.

Baran, S., S. Hemri, and M. El Ayari, 2019: Statistical postprocessing of water level forecasts using Bayesian model averaging with doubly truncated normal components. *Water Resour. Res.*, **55**, 3997–4013, https://doi.org/10.1029/2018WR024028.

Blöschl, G., and Coauthors, 2019: Changing climate both increases and decreases European river floods. *Nature*, **573**, 108–111, https://doi.org/10.1038/s41586-019-1495-6.

Candille, G., C. Côté, P. L. Houtekamer, and G. Pellerin, 2007: Verification of an ensemble prediction system against observations. *Mon. Wea. Rev.*, **135**, 2688–2699, https://doi.org/10.1175/MWR3414.1.

Das, J., S. Jha, and M. K. Goyal, 2020: Non-stationary and copula-based approach to assess the drought characteristics encompassing climate indices over the Himalayan states in India. *J. Hydrol.*, **580**, 124356, https://doi.org/10.1016/j.jhydrol.2019.124356.

Dong, Y., Y. Zhao, J. Zhai, J. Zhao, J. Han, Q. Wang, G. He, and H. Chang, 2021: Changes in reference evapotranspiration over the non-monsoon region of China during 1961–2017: Relationships with atmospheric circulation and attributions. *Int. J. Climatol.*, **41**, E734–E751, https://doi.org/10.1002/joc.6722.

Duan, Q., Y., S. Sorooshian, and V. K. Gupta, 1994: Optimal use of the SCE-UA global optimization method for calibrating watershed models. *J. Hydrol.*, **158**, 265–284, https://doi.org/10.1016/0022-1694(94)90057-4.

——, N. K. Ajami, X. Gao, and S. Sorooshian, 2007: Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Adv. Water Resour.*, **30**, 1371–1386, https://doi.org/10.1016/j.advwatres.2006.11.014.

Eum, H.-I., Y. Dibike, and T. Prowse, 2017: Climate-induced alteration of hydrologic indicators in the Athabasca River Basin, Alberta, Canada. *J. Hydrol.*, **544**, 327–342, https://doi.org/10.1016/j.jhydrol.2016.11.034.

Evin, G., B. Hingray, J. Blanchet, N. Eckert, S. Morin, and D. Verfaillie, 2019: Partitioning uncertainty components of an incomplete ensemble of climate projections using data augmentation. *J. Climate*, **32**, 2423–2440, https://doi.org/10.1175/JCLI-D-18-0606.1.

Fan, Y. R., G. H. Huang, B. W. Baetz, Y. P. Li, and K. Huang, 2017: Development of a Copula-based particle filter (CopPF) approach for hydrologic data assimilation under consideration of parameter interdependence. *Water Resour. Res.*, **53**, 4850–4875, https://doi.org/10.1002/2016WR020144.

Fang, M., and X. Li, 2016: Application of Bayesian model averaging in the reconstruction of past climate change using PMIP3/CMIP5 multimodel ensemble simulations. *J. Climate*, **29**, 175–189, https://doi.org/10.1175/JCLI-D-14-00752.1.

FAO/IIASA/ISRIC/ISSCAS/JRC, 2012: Harmonized world soil database, version 1.2. FAO Soils Portal, accessed 15 May 2018, https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/.

Frankcombe, L. M., M. H. England, M. E. Mann, and B. A. Steinman, 2015: Separating internal variability from the externally forced climate response. *J. Climate*, **28**, 8184–8202, https://doi.org/10.1175/JCLI-D-15-0069.1.

Gregersen, I. B., H. Madsen, D. Rosbjerg, and K. Arnbjerg-Nielsen, 2013: A spatial and nonstationary model for the frequency of extreme rainfall events. *Water Resour. Res.*, **49**, 127–136, https://doi.org/10.1029/2012WR012570.

Gupta, H. V., S. Sorooshian, and P. O. Yapo, 1999: Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *J. Hydrol. Eng.*, **4**, 135–143, https://doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135).

Hanel, M., and T. A. Buishand, 2015: Assessment of the sources of variation in changes of precipitation characteristics over the Rhine basin using a linear mixed-effects model. *J. Climate*, **28**, 6903–6919, https://doi.org/10.1175/JCLI-D-14-00775.1.

Hansen, M., R. DeFries, J. R. G. Townshend, and R. Sohlberg, 1998: UMD Global Land Cover Classification, 1 Kilometer, 1.0. Department of Geography, University of Maryland, 1981–1994, accessed 15 May 2018, http://app.earth-observer.org/data/basemaps/images/global/LandCover_512/LandCoverUMD_512/LandCoverUMD_512.html.

Harris, I., T. J. Osborn, P. Jones, and D. Lister, 2020: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data*, **7**, 109, https://doi.org/10.1038/s41597-020-0453-3.

Hawkins, E., and R. Sutton, 2011: The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dyn.*, **37**, 407–418, https://doi.org/10.1007/s00382-010-0810-6.

Her, Y., and I. Chaubey, 2015: Impact of the numbers of observations and calibration parameters on equifinality, model performance, and output and parameter uncertainty. *Hydrol. Processes*, **29**, 4220–4237, https://doi.org/10.1002/hyp.10487.

Hersbach, H., 2000: Decomposition of the continuous ranked probability score for ensemble prediction systems. *Wea. Forecasting*, **15**, 559–570, https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2.

——, and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, https://doi.org/10.1002/qj.3803.

Hingray, B., and M. Saïd, 2014: Partitioning internal variability and model uncertainty components in a multimember multimodel ensemble of climate projections. *J. Climate*, **27**, 6779–6798, https://doi.org/10.1175/JCLI-D-13-00629.1.

Huang, G., 1992: A stepwise cluster analysis method for predicting air quality in an urban environment. *Atmos. Environ.*, **26B**, 349–357, https://doi.org/10.1016/0957-1272(92)90010-P.

Huard, D., and A. Mailhot, 2006: A Bayesian perspective on input uncertainty in model calibration: Application to hydrological model "abc". *Water Resour. Res.*, **42**, W07416, https://doi.org/10.1029/2005WR004661.

Hundecha, Y., M. Pahlow, and A. Schumann, 2009: Modeling of daily precipitation at multiple locations using a mixture of distributions to characterize the extremes. *Water Resour. Res.*, **45**, W12412, https://doi.org/10.1029/2008WR007453.

Jayakrishnan, R., R. Srinivasan, C. Santhi, and J. G. Arnold, 2005: Advances in the application of the SWAT model for water resources management. *Hydrol. Processes*, **19**, 749–762, https://doi.org/10.1002/hyp.5624.

Jones, R. G., M. Noguer, D. Hassell, D. Hudson, S. Wilson, G. Jenkins, and J. Mitchell, 2004: Generating high resolution

climate change scenarios using PRECIS. Met Office Hadley Centre, 35 pp.

Jones, R. N., 2000: Managing uncertainty in climate change projections—Issues for impact assessment. *Climatic Change*, **45**, 403–419, https://doi.org/10.1023/A:1005551626280.

Kennedy, W. J., and J. E. Gentle, 1981: *Statistical Computing.* Statistics: Textbooks and Monographs, Vol. 33, Marcel Dekker, 591 pp.

Keppel, G., 1991: *Design and Analysis: A Researcher's Handbook.* 3rd ed. Prentice-Hall, Inc., 594 pp.

Knutti, R., and J. Sedláček, 2013: Robustness and uncertainties in the new CMIP5 climate model projections. *Nat. Climate Change*, **3**, 369–373, https://doi.org/10.1038/nclimate1716.

Lafaysse, M., B. Hingray, A. Mezghani, J. Gailhard, and L. Terray, 2014: Internal variability and model uncertainty components in future hydrometeorological projections: The Alpine Durance basin. *Water Resour. Res.*, **50**, 3317–3341, https://doi.org/10.1002/2013WR014897.

Lehner, B., K. Verdin, and A. Jarvis, 2006: HydroSHEDS technical documentation, version 1.0. World Wildlife Fund US Tech. Doc., 29 pp.

Lehner, F., S. Coats, T. F. Stocker, A. G. Pendergrass, B. M. Sanderson, C. C. Raible, and J. E. Smerdon, 2017: Projected drought risk in 1.5°C and 2°C warmer climates. *Geophys. Res. Lett.*, **44**, 7419–7428, https://doi.org/10.1002/2017GL074117.

Li, Z., G. Huang, J. Han, X. Wang, Y. Fan, G. Cheng, H. Zhang, and W. Huang, 2015: Development of a stepwise-clustered hydrological inference model. *J. Hydrol. Eng.*, **20**, 04015008, https://doi.org/10.1061/(ASCE)HE.1943-5584.0001165.

——, ——, X. Wang, J. Han, and Y. Fan, 2016: Impacts of future climate change on river discharge based on hydrological inference: A case study of the Grand River watershed in Ontario, Canada. *Sci. Total Environ.*, **548–549**, 198–210, https://doi.org/10.1016/j.scitotenv.2016.01.002.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, 1994: A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res.*, **99**, 14 415–14 428, https://doi.org/10.1029/94JD00483.

——, E. F. Wood, and D. P. Lettenmaier, 1996: Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Global Planet. Change*, **13**, 195–206, https://doi.org/10.1016/0921-8181(95)00046-1.

Lohmann, D., R. Nolte-Holube, and E. Raschke, 1996: A large-scale horizontal routing model to be coupled to land surface parametrization schemes. *Tellus*, **48A**, 708–721, https://doi.org/10.3402/tellusa.v48i5.12200.

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

Marx, A., and Coauthors, 2018: Climate change alters low flows in Europe under global warming of 1.5, 2, and 3°C. *Hydrol. Earth Syst. Sci.*, **22**, 1017–1032, https://doi.org/10.5194/hess-22-1017-2018.

Met Office, 2020: PRECIS: A regional climate modelling system. Accessed 25 December 2021, https://www.metoffice.gov.uk/research/applied/international/precis.

Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith, 2007: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE*, **50**, 885–900, https://doi.org/10.13031/2013.23153.

Murphy, J. M., and Coauthors, 2009: UK Climate Projections Science Report: Climate change projections. Met Office Hadley Centre, 21–35, accessed 15 May 2019, http://ukclimateprojections.metoffice.gov.uk/22530.

Northrop, P. J., and R. E. Chandler, 2014: Quantifying sources of uncertainty in projections of future climate. *J. Climate*, **27**, 8793–8808, https://doi.org/10.1175/JCLI-D-14-00265.1.

Peng, H., Y. Jia, C. Tague, and P. Slaughter, 2015: An eco-hydrological model-based assessment of the impacts of soil and water conservation management in the Jinghe River basin, China. *Water*, **7**, 6301–6320, https://doi.org/10.3390/w7116301.

Pokhrel, Y., and Coauthors, 2021: Global terrestrial water storage and drought severity under climate change. *Nat. Climate Change*, **11**, 226–233, https://doi.org/10.1038/s41558-020-00972-w.

Qin, X. S., G. H. Huang, G. M. Zeng, and A. Chakma, 2008: Simulation-based optimization of dual-phase vacuum extraction to remove nonaqueous phase liquids in subsurface. *Water Resour. Res.*, **44**, W04422, https://doi.org/10.1029/2006WR005496.

Rao, C. R., 1952: *Advanced Statistical Methods in Biometric Research.* John Wiley and Sons, 390 pp.

Refsgaard, J. C., S. Christensen, T. O. Sonnenborg, D. Seifert, A. L. Højberg, and L. Troldborg, 2012: Review of strategies for handling geological uncertainty in groundwater flow and transport modeling. *Adv. Water Resour.*, **36**, 36–50, https://doi.org/10.1016/j.advwatres.2011.04.006.

Rings, J., J. A. Vrugt, G. Schoups, J. A. Huisman, and H. Vereecken, 2012: Bayesian model averaging using particle filtering and Gaussian mixture modeling: Theory, concepts, and simulation experiments. *Water Resour. Res.*, **48**, W05520, https://doi.org/10.1029/2011WR011607.

Shi, W., and Coauthors, 2021: Drought-flood abrupt alternation dynamics and their potential driving forces in a changing environment. *J. Hydrol.*, **597**, 126179, https://doi.org/10.1016/j.jhydrol.2021.126179.

Shrestha, R. R., D. L. Peters, and M. A. Schnorbus, 2014: Evaluating the ability of a hydrologic model to replicate hydro-ecologically relevant indicators. *Hydrol. Processes*, **28**, 4294–4310, https://doi.org/10.1002/hyp.9997.

Singh, J., H. V. Knapp, J. G. Arnold, and M. Demissie, 2005: Hydrological modeling of the Iroquois river watershed using HSPF and SWAT 1. *J. Amer. Water Resour. Assoc.*, **41**, 343–360, https://doi.org/10.1111/j.1752-1688.2005.tb03740.x.

Steinschneider, S., M. Ho, E. R. Cook, and U. Lall, 2016: Can PDSI inform extreme precipitation? An exploration with a 500 year long paleoclimate reconstruction over the US. *Water Resour. Res.*, **52**, 3866–3880, https://doi.org/10.1002/2016WR018712.

Strauch, M., C. Bernhofer, S. Koide, M. Volk, C. Lorz, and F. Makeschin, 2012: Using precipitation data ensemble for uncertainty analysis in SWAT streamflow simulation. *J. Hydrol.*, **414–415**, 413–424, https://doi.org/10.1016/j.jhydrol.2011.11.014.

Sun, W., Q. Shao, and J. Liu, 2013: Soil erosion and its response to the changes of precipitation and vegetation cover on the Loess Plateau. *J. Geogr. Sci.*, **23**, 1091–1106, https://doi.org/10.1007/s11442-013-1065-z.

Thorarinsdottir, T. L., K. H. Hellton, G. H. Steinbakk, L. Schlichting, and K. Engeland, 2018: Bayesian regional flood frequency analysis for large catchments. *Water Resour. Res.*, **54**, 6929–6947, https://doi.org/10.1029/2017WR022460.

Tsai, F. T.-C., and A. S. Elshall, 2013: Hierarchical Bayesian model averaging for hydrostratigraphic modeling:

Uncertainty segregation and comparative evaluation. *Water Resour. Res.*, **49**, 5520–5536, https://doi.org/10.1002/wrcr.20428.

Ukkola, A. M., M. G. De Kauwe, M. L. Roderick, G. Abramowitz, and A. J. Pitman, 2020: Robust future changes in meteorological drought in CMIP6 projections despite uncertainty in precipitation. *Geophys. Res. Lett.*, **47**, e2020GL087820 https://doi.org/10.1029/2020GL087820.

Wagener, T., K. van Werkhoven, P. Reed, and Y. Tang, 2009: Multiobjective sensitivity analysis to understand the information content in streamflow observations for distributed watershed modeling. *Water Resour. Res.*, **45**, W02501, https://doi.org/10.1029/2008WR007347.

Wang, Q. J., A. Schepen, and D. E. Robertson, 2012: Merging seasonal rainfall forecasts from multiple statistical models through Bayesian model averaging. *J. Climate*, **25**, 5524–5537, https://doi.org/10.1175/JCLI-D-11-00386.1.

Wang, X. Q., and Coauthors, 2013: A stepwise cluster analysis approach for downscaled climate projection—A Canadian case study. *Environ. Modell. Software*, **49**, 141–151, https://doi.org/10.1016/j.envsoft.2013.08.006.

——, G. Huang, S. Zhao, and J. Guo, 2015: An open-source software package for multivariate modeling and clustering: Applications to air quality management. *Environ. Sci. Pollut. Res. Int.*, **22**, 14 220–14 233, https://doi.org/10.1007/s11356-015-4664-7.

Webster, M. D., M. Babiker, M. Mayer, J. M. Reilly, J. Harnisch, R. Hyman, M. C. Sarofim, and C. Wang, 2002: Uncertainty in emissions projections for climate models. *Atmos. Environ.*, **36**, 3659–3670, https://doi.org/10.1016/S1352-2310(02)00245-5.

Xiao, Z., S. Liang, J. Wang, P. Chen, X. Yin, L. Zhang, and J. Song, 2014: Use of general regression neural networks for generating the GLASS leaf area index product from time-series MODIS surface reflectance. *IEEE Trans. Geosci. Remote Sens.*, **52**, 209–223, https://doi.org/10.1109/TGRS.2013.2237780.

Yan, H., and H. Moradkhani, 2016: Toward more robust extreme flood prediction by Bayesian hierarchical and multimodeling. *Nat. Hazards*, **81**, 203–225, https://doi.org/10.1007/s11069-015-2070-6.

Yip, S., C. A. T. Ferro, D. B. Stephenson, and E. Hawkins, 2011: A simple, coherent framework for partitioning uncertainty in climate predictions. *J. Climate*, **24**, 4634–4643, https://doi.org/10.1175/2011JCLI4085.1.

Zhang, S., and J. Chen, 2021: Uncertainty in projection of climate extremes: A comparison of CMIP5 and CMIP6. *J. Meteor. Res.*, **35**, 646–662, https://doi.org/10.1007/s13351-021-1012-3.

Zhang, X., X. Yan, and Z. Chen, 2016: Reconstructed regional mean climate with Bayesian model averaging: A case study for temperature reconstruction in the Yunnan-Guizhou Plateau, China. *J. Climate*, **29**, 5355–5361, https://doi.org/10.1175/JCLI-D-15-0603.1.

Zhao, G., X. Mu, Z. Wen, F. Wang, and P. Gao, 2013: Soil erosion, conservation, and eco-environment changes in the Loess Plateau of China. *Land Degrad. Dev.*, **24**, 499–510, https://doi.org/10.1002/ldr.2246.

Zhou, X., G. Huang, B. W. Baetz, X. Wang, and G. Cheng, 2018a: PRECIS-projected increases in temperature and precipitation over Canada. *Quart. J. Roy. Meteor. Soc.*, **144**, 588–603, https://doi.org/10.1002/qj.3231.

——, ——, J. Piwowar, Y. Fan, X. Wang, Z. Li, and G. Cheng, 2018b: Hydrologic impacts of ensemble-RCM-projected climate changes in the Athabasca River Basin, Canada. *J. Hydrometeorology*, **19**, 1953–1971, https://doi.org/10.1175/JHM-D-17-0232.1.

——, ——, X. Wang, and G. Cheng, 2018c: Dynamically-downscaled temperature and precipitation changes over Saskatchewan using the PRECIS model. *Climate Dyn.*, **50**, 1321–1334, https://doi.org/10.1007/s00382-017-3687-9.

Zhu, Z., D. B. Wright, and G. Yu, 2018: The impact of rainfall space-time structure in flood frequency analysis. *Water Resour. Res.*, **54**, 8983–8998, https://doi.org/10.1029/2018WR023550.

Zhuang, X. W., Y. P. Li, G. H. Huang, and X. Q. Wang, 2016: A hybrid factorial stepwise-cluster analysis method for streamflow simulation—A case study in northwestern China. *Hydrol. Sci. J.*, **61**, 2775–2788, https://doi.org/10.1080/02626667.2015.1125482.