



# An evaluation of two statistical downscaling models for downscaling monthly precipitation in the Heihe River basin of China

Haifeng Su<sup>1,2</sup> · Zhe Xiong<sup>1</sup> · Xiaodong Yan<sup>3</sup> · Xingang Dai<sup>1</sup>

Received: 28 October 2016 / Accepted: 17 June 2019 / Published online: 16 July 2019  
© Springer-Verlag GmbH Austria, part of Springer Nature 2019

## Abstract

The stepwise regression model (SRM) is a widely used statistical downscaling method that could be used to establish a statistical relationship between observed precipitation and predictors. However, the SRM cannot reflect the contributions of predictors to precipitation reasonably, which may not be the best model based on several possible competing predictors. Bayesian model averaging (BMA) is a standard inferencing approach that considers multiple competing statistical models. The BMA infers precipitation predictions by weighing individual predictors based on their probabilistic likelihood measures over the training period, with the better-performing predictions receiving higher weights than the worse-performing ones. Furthermore, the BMA provides a more reliable description of all the predictors than the SRM, leading to a sharper and better calibrated probability density function (PDF) for the probabilistic predictions. In this study, monthly precipitation at fifteen meteorological stations over the period of 1971–2012 in the Heihe River basin (HRB), which is located in an arid area of Northwest China, was simulated using the Bayesian model averaging (BMA) and the stepwise regression model (SRM), which was then compared with the observed datasets (OBS). The results showed that the BMA produced more accurate results than the SRM when used to statistically downscale large-scale variables. The multiyear mean precipitation results for twelve of the fifteen meteorological stations that were simulated by the BMA were better than those simulated by the SRM. The RMSE and MAE of the BMA for each station were lower than those of the SRM. The BMA had a lower mean RMSE (– 13.93%) and mean MAE (– 14.37%) compared with the SRM. The BMA could reduce the RMSEs and MAEs of precipitation and improve the correlation coefficient effectively. This indicates that the monthly precipitation simulated by the BMA has better consistency with the observed values.

## 1 Introduction

Statistical downscaling techniques are important to simulate future climate scenarios at the regional scale, as it establishes statistical relationships between the outputs of large-scale GCMs or reanalysis data (predictors) and local-scale meteorological variables (predictands) to obtain future predictands from predictors based on these relationships (Spak 2007). Statistical downscaling can

be classified into three approaches: regression methods (Kim et al. 1984; von Storch et al. 1993; Maraun et al. 2011), weather-type approaches (Hay et al. 1991; Vrac and Naveau 2007; Cheng et al. 2011; Osca et al. 2013), and stochastic weather generators (Wilby and Wigley 1997; Murphy 1999; Fatichi et al. 2011, 2013). Several statistical downscaling methods have been applied to simulate precipitation at the basin scale in many foreign countries. Campozano et al. (2016) used the statistical downscaling model (SDSM), artificial neural network (ANN), and the least squares support vector machine (LS-SVM) approaches to simulate monthly precipitation in the Paute River basin in southern Ecuador. The classification and regression tree (CART) method was used to simulate precipitation in the Mahanadi River basin in India (Kannan and Ghosh 2010). Singh et al. (2015) simulated precipitation in the Tapi basin in India using the kernel-regression (KR) method. In China, there are also a large number of studies focused on simulating precipitation at the basin scale. Liu et al. (2016) compared the simulated precipitation of the Hanjiang River

✉ Zhe Xiong  
xzh@tea.ac.cn

<sup>1</sup> Key Laboratory of Regional Climate–Environment for Temperate East Asia, Institute of Atmospheric Physics, Chinese Academy of Sciences, No. 40 HuaYanLi, ChaoYang District, Beijing 100029, China

<sup>2</sup> University of Chinese Academy of Sciences, Beijing 100049, China

<sup>3</sup> State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100087, China

basin using the support vector machine (SVM), weather generators (WGs), and the statistical downscaling model (SDSM) and assembled these models based on BMA. Wang et al. (2015) used empirical statistical downscaling methods to simulate the daily precipitation of the Huaihe River basin. Huang et al. (2012) applied the statistical downscaling model (SDSM) to simulate extreme precipitation in the Yangtze River basin. However, the statistical downscaling technique has scarcely been applied to simulate precipitation at the basin scale in arid areas. Liu et al. (2016) evaluated the nonhomogeneous hidden Markov model (NHMM) and the statistical downscaling model (SDSM) for daily precipitation in the Tarim River basin, which is located in an arid area of Northwest China. Research on statistically downscaled precipitation in the Heihe River basin (HRB) is also being conducted.

The HRB, which is located in an arid area of Northwest China, is the second largest inland river region in China. A shortage of water resources leads to ecological deterioration and restricts the economic development of the HRB; thus, water resources have become the core of research on the HRB (Lan et al. 2005). As the main source of water resources for this area, future precipitation in the HRB has become a problem of concern. There are only 19 meteorological stations in the HRB, and they are unevenly distributed. Therefore, a statistical downscaling technique is a necessary tool to obtain precipitation data with high spatial and temporal resolutions. Su et al. (2016) used the stepwise regression method (SRM) as a statistical downscaling method and a regional climate model based on the regional integrated environmental model system (RIEMS 2.0) as a dynamical downscaling model to simulate rainy season precipitation over the period of 2003–2012 in the HRB, and the results showed that the SRM could reasonably simulate monthly precipitation. For statistical downscaling methods, the stepwise regression model (SRM) is a widely used statistical downscaling method (Huth 1999; Wilby et al. 1999) that is powerful in simulating precipitation at the regional scale, in which some predictors (e.g., sea level pressure, geopotential height, and specific humidity) that highly influence local precipitation are selected to establish linear relationships with precipitation data from meteorological observation stations. However, information from other predictors that influence local precipitation are ignored in this statistical model. Even more crucially, we do not know the contributions of each predictor to precipitation, which may make the SRM unstable. In a previous study, biases in the SRM for some stations were larger than 50%, and the correlation between the simulations and observed datasets for some stations is not significant, which illustrates that the linear regression model applied to this region should be improved.

To achieve this goal, Bayesian model averaging (BMA), a standard inference approach using multiple competing statistical models, was proposed for postprocessing of the SRM.

The BMA has been widely applied to research in social and health sciences. Viallefont et al. (2001), Raftery and Zheng (2003), Raftery et al. (2005) and Duan et al. (2007) extended the BMA to the study of multimodel ensembles. Recently, Zhang and Yan (2015) applied the BMA to the study of statistical downscaling to simulate monthly precipitation in China, and the results showed that the BMA obtained better results than the linear regression method. The BMA assigns a weight to each predictor, which reflects the degree of influence of that predictor on precipitation during the training period. The SRM was calibrated by assigning higher weights to the better-performing predictors instead of the worse-performing predictors. Thus, the precipitation simulated by the BMA is the weighted average of predictors.

In this study, monthly precipitation over the period of 1971–2012 at 15 meteorological stations around the HRB was simulated using the SRM and BMA and then compared with the observed datasets (OBS). The main goal was to systematically compare the skill of the BMA with that of the SRM to simulate monthly precipitation at the basin scale in arid areas of Northwest China.

## 2 Datasets and methods

### 2.1 Predictor selection and observed datasets

The HRB is located at 98–101.5° E, 38–42° N, with a drainage area of approximately 290,000 km<sup>2</sup> and a total length of approximately 810 km (Fig. 1). Monthly precipitation from 15 meteorological observation stations in the HRB (Table 1) was used to establish statistical models for the period 1971–2012, and data were obtained from the Chinese Meteorological Data Sharing Service System (<http://cdc.cma.gov.cn>).

There are obvious differences in surface types and topography in the HRB; the factors affecting regional precipitation are complex. Therefore, the selection of potential predictors should be adequate and appropriate. The National Centers for Environmental Prediction and the National Center for Atmospheric Research (NCEP/NCAR) distribute a reanalysis data set, where the time resolution varies from hours to months and the spatial resolution is 2.5° that presents atmospheric conditions at different levels of the atmosphere (Kalnay et al. 1996). The NCEP/NCAR reanalysis data have been used to simulate precipitation in different regions worldwide (e.g., Sachindra et al. 2014; Su et al. 2016; Li and Smith 2009). After a series of tests, the predicted variables extracted from the NCEP/NCAR reanalysis data set include sea level pressure (SLP); wind speed and direction at 850, 700, and 500 hPa (U/V850, U/V700, and U/V500, respectively); geopotential height at 1,000, 850, 700, and 500 hPa (H1000, H850, H700, and H500, respectively); and specific humidity at 850, 700, and 500 hPa (S850, S700, and S500,

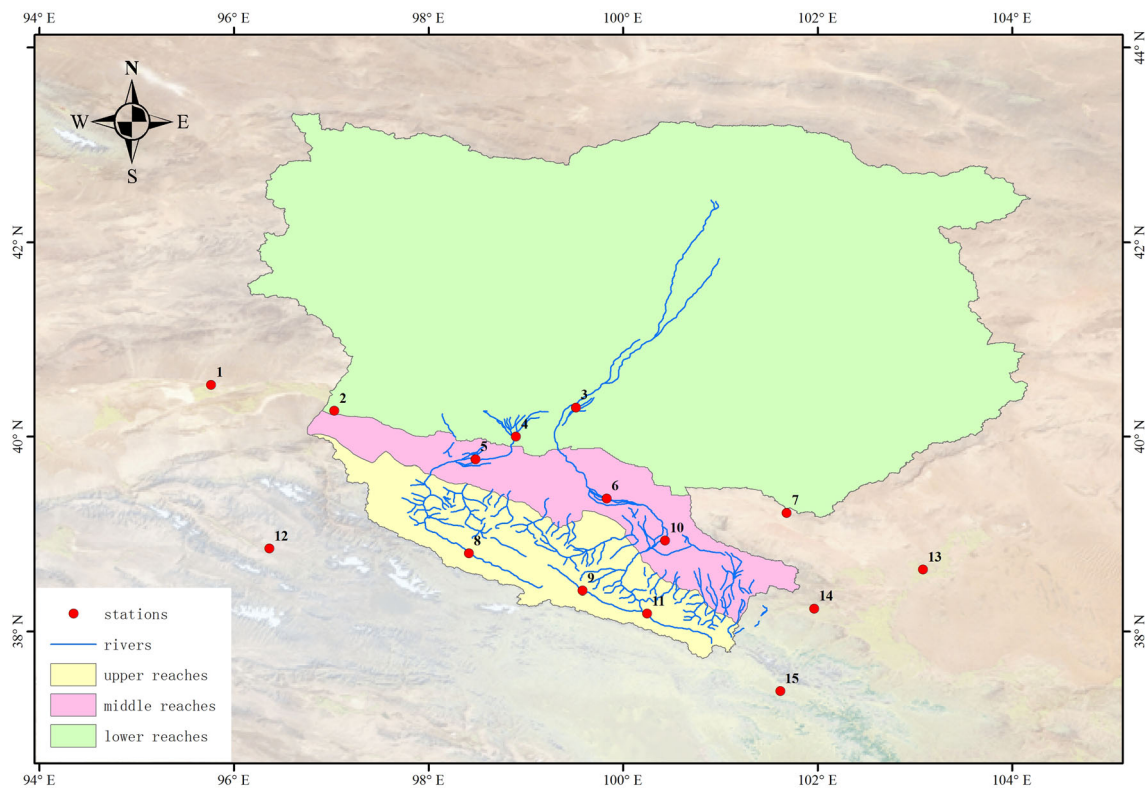


Fig. 1 Spatial distribution of meteorological stations in the HRB

respectively), which were obtained from the website <http://www.cdc.noaa.gov/>. Spatial domains of predicted variables influence the establishment of statistical models, which depends on subjective selection. To be physically reasonable, several grids around station were selected as the predictor domain, which may be adjusted appropriately to make all the models to pass the significance test.

Table 1 Coordinates and altitudes of stations

Station	No.	Longitude (°E)	Latitude (°N)	Altitude (m)
Anxi	52424	95.77	40.53	1170.9
Yumenzhen	52436	97.03	40.27	1526.0
Dingxin	52446	99.52	40.30	1177.4
Jinta	52447	98.90	40.00	1270.5
Jiuquan	52533	98.48	39.77	1477.2
Gaotai	52546	99.83	39.37	1332.2
Alashanyouqi	52576	101.68	39.22	1510.1
Tuole	52633	98.42	38.80	3367.0
Yeniugou	52645	99.58	38.42	3320.0
Zhangye	52652	100.43	38.93	1482.7
Qilian	52657	100.25	38.18	2787.4
Yongchang	52674	101.97	37.85	3173.2
Minqin	52681	103.08	38.63	1367.5
Dachaidan	52713	95.37	38.23	1976.9
Menyuan	52765	101.62	37.38	2850.0

### 2.2 Stepwise regression model

The fourteen predicted variables and the observed precipitation data over the period 1971–2002 were used as training samples to fit the SRM and are standardized as follows:

$$Y = \frac{X - \bar{X}}{\sigma} \tag{1}$$

where  $Y$  is the standard value,  $X$  is the monthly mean value of the predicted variables or the predictand (monthly rainfall),  $\bar{X}$  is the mean monthly value of  $X$ , and  $\sigma$  is the standard deviation of  $X$  (Zhang and Yan 2015).

Principal component analysis (PCA) is used to decrease the dimensionality of these fourteen predicted variables, which can be achieved by including only the first few PCs (Ruping and Straus 2002). In the study, the first four PCs of every potential predictor are selected to establish the SRM.

Fifty-six variables are used as potential predictors to develop models by the stepwise regression method as follows:

$$Y(t) = \sum_{n=1}^N \alpha_n X_n(t) + \varepsilon_t \tag{2}$$

where  $Y$  is the monthly mean observed precipitation,  $X$  is the potentially predicted variable,  $\alpha$  is the regression coefficient, and  $\varepsilon_t$  is the residual not described by the statistical model.

Precipitation over the periods 1971–1982, 1983–1992, and 1993–2002 is simulated using the method we mentioned above. All of the models reached a significance level of 0.05%.

### 2.3 Bayesian model averaging

The lower reaches of the HRB is located in a desert area, with a typical continental arid climate and little precipitation, especially in winter. It is difficult to establish a statistical relationship between the observed data and the predicted variables that strongly influences precipitation in other regions. We selected the PCs of many variables from the NCEP/NCAR reanalysis data set and used the SRM to select appropriate predictors. However, this may lead to overfitting by the SRM. Raftery et al. (2005) noted that the linear regression method may not obtain the best model from among several possible competing models. Other plausible models could give different answers to the scientific question at hand, which is a source of uncertainty in drawing conclusions. Bayesian model averaging overcomes this problem by conditioning the entire ensemble of statistical models first considered. Therefore, we used the BMA to simulate precipitation in the HRB and compare it with the SRM.

The BMA scheme, which is extended to statistical down-scaling, is briefly described as follows: consider that a quantity  $y$  is the precipitation to be simulated, where  $y^T$  represents training data with data length  $T$  and  $x_1 \dots x_n$  represents the  $N$  predictors. The probability density function (PDF)  $p(y|x_1 \dots x_n)$  simulated by the BMA can be represented as:

$$p(y|x_1 \dots x_n) = \sum_{n=1}^N w_n p_n(y|x_n) \tag{3}$$

where  $p_n(y|x_n)$  is the simulated PDF based on the predictor  $x_n$  and training data  $y^T$ .  $x_n$  represents the PCs of the monthly mean value of the predicted variables; thus,  $p(y|x_1 \dots x_n)$  often seems reasonable to approximate the conditional PDF by a normal distribution centered at a linear function of the predictor  $a_n x_n + b_n$  with variance  $\sigma_n^2$ , where  $a_n$  and  $b_n$  can be obtained by the regression between the predictor and the predictand.  $w_n$  is the posterior probability of predictor  $x_n$ , which add up to

one, namely,  $\sum_{n=1}^N w_n = 1$ ; the results can be considered the weights for the predictors influencing precipitation.

The deterministic Bayesian precipitation prediction is the conditional expectation of  $y$  given the simulation, which is calculated as:

$$E(y|x_1 \dots x_n) = \sum_{n=1}^N w_n (a_n x_n + b_n) \tag{4}$$

where the parameters  $w_n$  and  $\sigma_n^2$  could be estimated by the expectation-maximization (EM) algorithm (Zhang and Yan 2015; Duan et al. 2007; Raftery and Zheng 2003).

Similar to the SRM, first, the observed precipitation and fourteen predicted variables were standardized; then, the PCA was applied to decrease the dimensionality of these predicted variables. The first four PCs of every predicted variable were selected as potential predictors. Finally, the correlation coefficient between the observed data and the PCs reached a significance level of 0.05 and was selected as a predictor in the BMA. The precipitation values over the periods 1971–1982, 1983–1992, 1993–2002, and 2003–2012 were simulated by the BMA.

## 3 Results

### 3.1 Comparison of the spatial distributions of precipitation

The landscape patterns in the upper, middle, and lower reaches of the HRB comprise glaciers and permafrost, alpine meadows and forests, and deserts, respectively. Therefore, the HRB is divided into three subregions based on the spatial distribution of precipitation: upper reaches (Tuole, Yeniugou, Qilian, Yongchang, and Menyuan stations), middle reaches (Jiuquan, Gaotai, Alashanyouqi, Zhangye, Minqin, and Dachaidan stations), and lower reaches (Anxi, Yumenzhen, Dingxin, and Jinta stations) (Cheng et al. 2006).

Table 2 shows the regional mean precipitation. From Table 2, it can be observed that precipitation values in the three subregions and the HRB as a whole are underestimated

**Table 2** The OBS and regional mean precipitation (mm) simulated by the BMA and the SRM. Biases, RMSEs (mm), and MAEs (mm) of the two models

	OBS	BMA	SRM	Bias		RMSE		MAE	
				BMA (%)	SRM (%)	BMA	SRM	BMA	SRM
Upper	373.68	365.37	360.75	- 2.22	- 3.46	36.11	37.04	27.52	29.51
Middle	110.00	106.62	105.90	- 3.07	- 3.72	19.88	19.87	15.23	16.27
Lower	60.16	59.66	57.69	- 0.83	- 4.11	15.74	16.52	11.55	11.24
HRB	184.60	180.35	178.00	- 2.31	- 3.58	18.23	17.89	13.79	14.31



by the BMA and the SRM, and the two models can simulate the spatial distribution of observed precipitation, with a high-value center being concentrated predominantly over the upper reaches of the Qilian Mountain and a low-value center appearing predominantly over the lower reaches of the desert area. In the upper reaches, the observed precipitation is 373.68 mm, while precipitation simulated by the BMA is 365.37 mm; the bias of the BMA is  $-2.22\%$ , with an RMSE and MAE of 36.11 mm and 27.52 mm, respectively, whereas precipitation simulated by the SRM is 360.75 mm, and the bias of the SRM is  $-3.46\%$ , with an RMSE and MAE of 37.04 mm and 29.51 mm, respectively. In the middle reaches, the observed precipitation is 110.00 mm, while precipitation simulated by the BMA is 106.62 mm; the bias of the BMA is  $-3.07\%$ , with an RMSE and MAE of 19.88 mm and 15.23 mm, respectively, whereas the precipitation simulated by the SRM is 105.90 mm, and the bias of the SRM is  $-3.72\%$ , with an RMSE and MAE of 19.87 mm and 16.27 mm, respectively. In the lower reaches, the observed precipitation is 60.16 mm, and the precipitation simulated by the BMA is 59.66 mm; the bias of

the BMA is  $-0.83\%$ , with an RMSE and MAE of 15.74 mm and 11.55 mm, respectively, whereas the precipitation simulated by the SRM is 57.69 mm, and the bias of the SRM is  $-4.11\%$ , with an RMSE and MAE of 16.52 mm and 11.24 mm, respectively. Over the entire HRB, the observed precipitation is 184.60 mm, and the precipitation simulated by the BMA is 180.35 mm; the bias of the BMA is  $-2.31\%$ , with an RMSE and MAE of 18.23 mm and 13.79 mm, respectively, whereas precipitation simulated by the SRM is 178.00 mm, and the bias of the SRM is  $-3.58\%$ , with an RMSE and MAE of 17.89 mm and 14.31 mm, respectively.

Figure 2 shows the mean monthly precipitation in the three subregions and the entire HRB region. In the upper reaches, the biases of the BMA are in the range of  $-8.70\%$  (Sep)– $2.99\%$  (Jun), whereas biases of the SRM are in the range of  $-12.38\%$  (Nov)– $13.51\%$  (Jan). Precipitation simulated by the BMA is better than that by the SRM, except for in Mar, Jun, and Oct. In the middle reaches, biases of the BMA are in the range of  $-31.39\%$  (Nov)– $5.15\%$  (Feb), whereas biases of the SRM are in the range of  $-45.99\%$  (Nov)– $8.21\%$  (Feb).

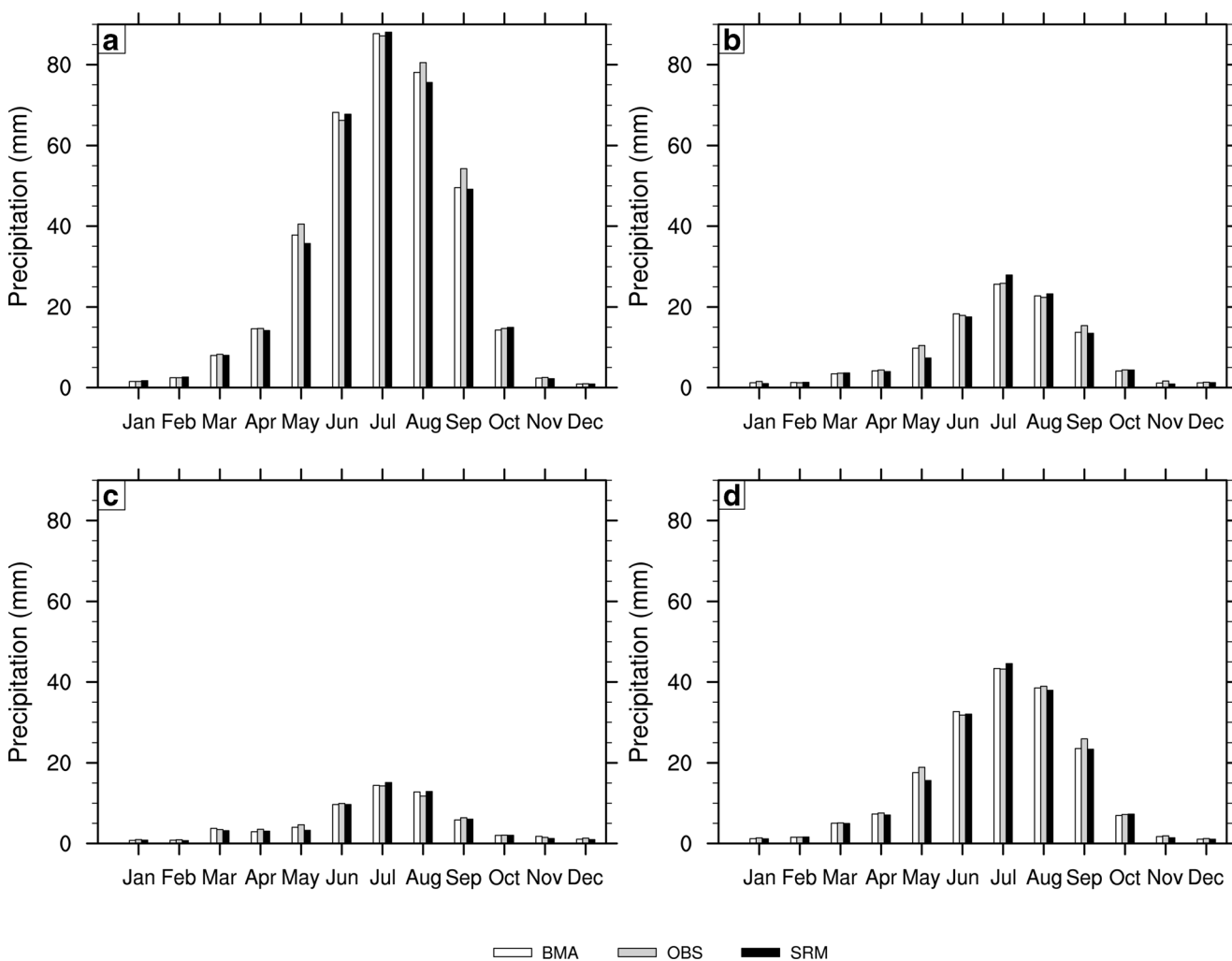
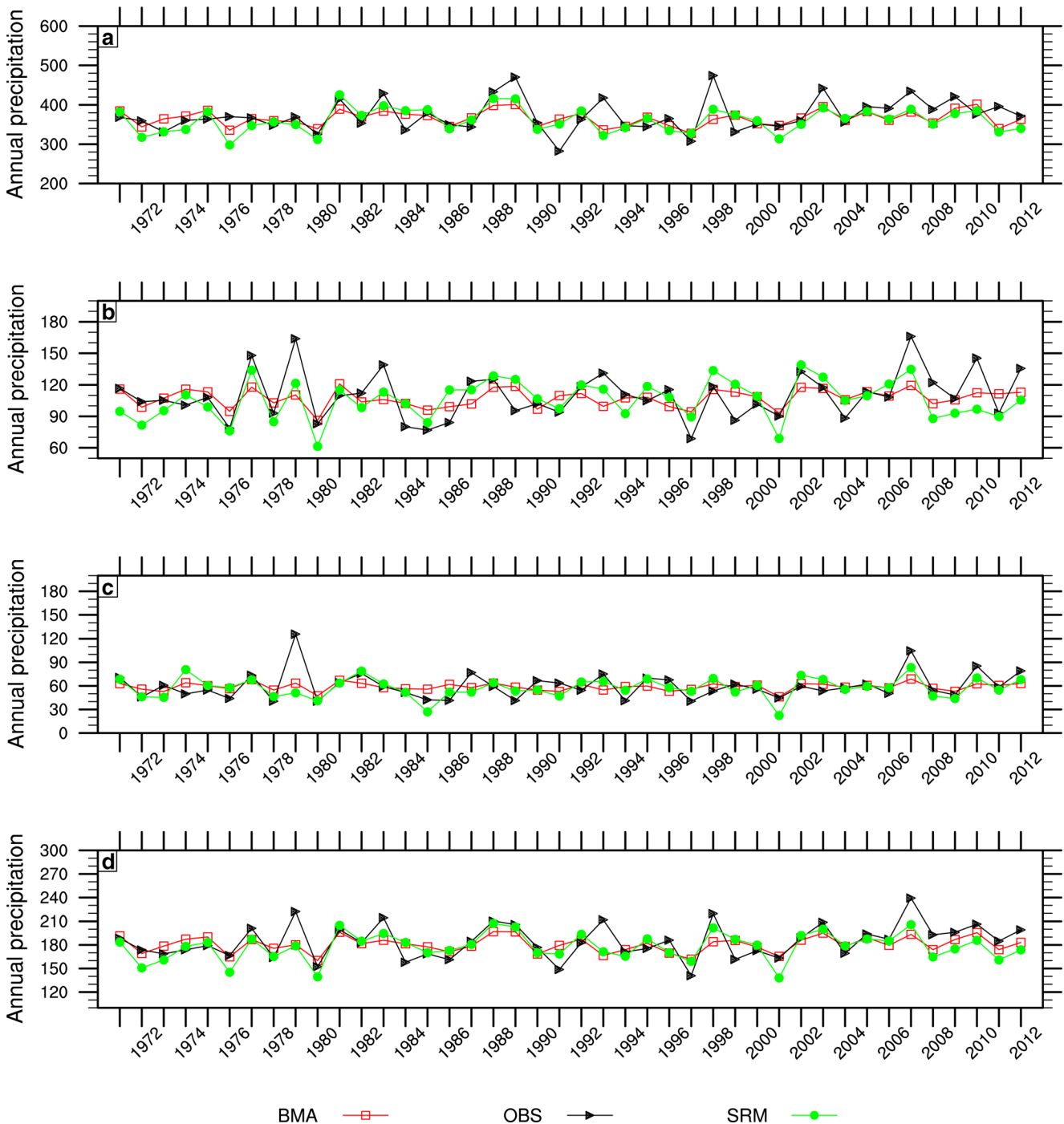


Fig. 2 The mean monthly precipitation (mm) in a upper reaches, b middle reaches, c lower reaches, and d the HRB

Precipitation simulated by the BMA is better than that by the SRM, except for in Mar, Jun, Oct, and Dec. In the lower reaches, biases of the BMA are in the range of  $-22.33\%$  (Jan)– $18.88\%$  (Nov), whereas biases of the SRM are in the range of  $-29.86\%$  (May)– $9.69\%$  (Aug). Precipitation simulated by the BMA is better than that by the SRM over 6 months. In the entire HRB, biases of the BMA are in the range of  $-14.76\%$  (Jan)– $2.74\%$  (Jun), whereas biases of the SRM

are in the range of  $-25.15\%$  (Nov)– $4.34\%$  (Feb). Precipitation simulated by the BMA is better than that by the SRM, except for Jun and Oct. Precipitation in the three sub-regions and the entire HRB region simulated by the BMA is better than that simulated by the SRM for most months.

Figure 3 shows the time series of annual precipitation. Two models could reasonably reproduce the time series of annual precipitation in different regions. The correlation coefficients



**Fig. 3** The time series of mean annual precipitation (mm) in **a** upper reaches, **b** middle reaches, **c** lower reaches, and **d** the HRB. The red line is the precipitation simulated by the BMA, the green line is the precipitation simulated using the SRM, and black line is the OBS

between the BMA and OBS in the three subregions and the entire HRB are 0.50, 0.49, 0.43, and 0.56, and the correlation coefficients are 0.55, 0.56, 0.43, and 0.65 between the SRM and OBS, respectively.

### 3.2 Comparison of stations

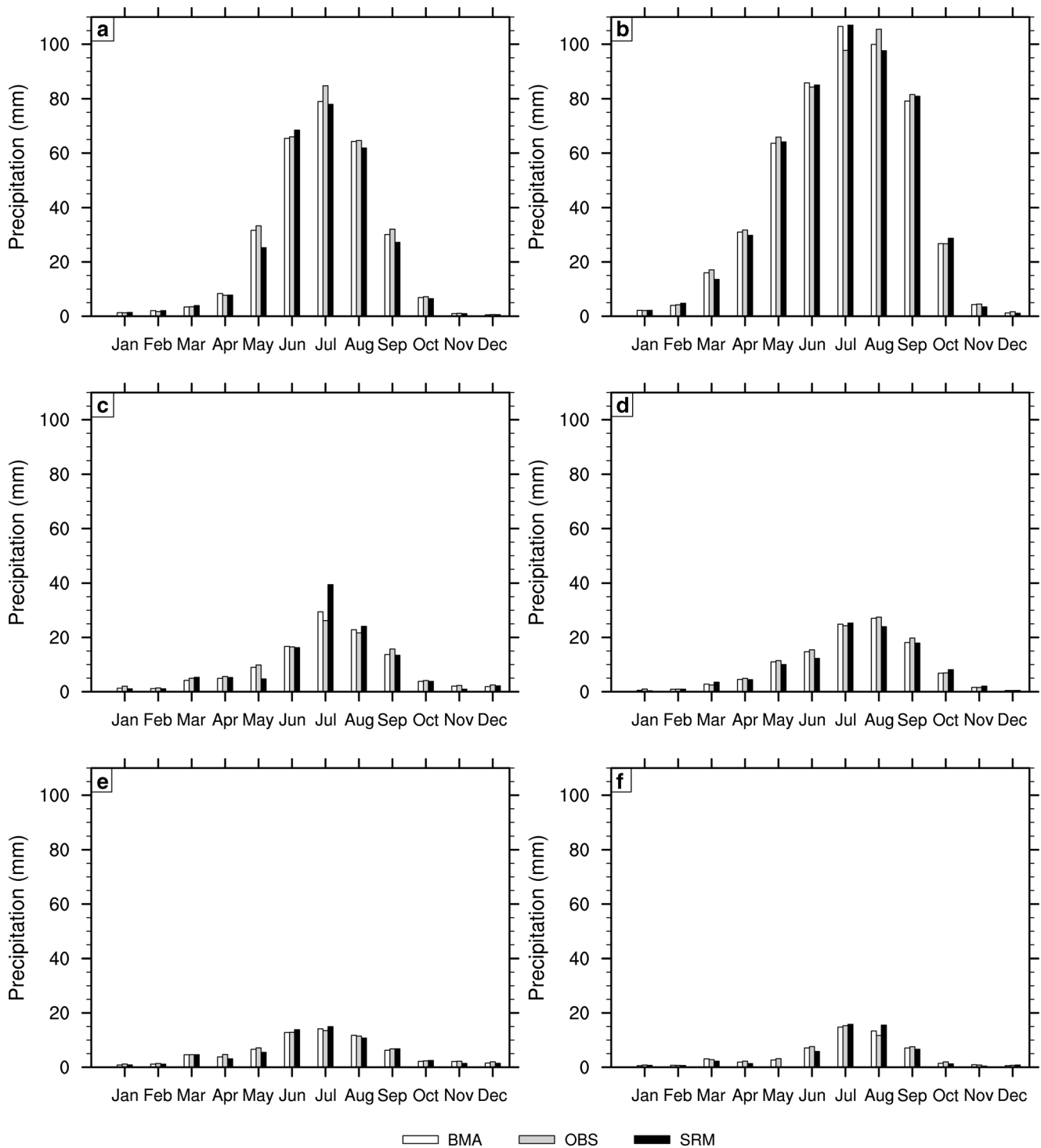
Monthly precipitation simulated by the BMA and the SRM is compared with the observed datasets (OBS). Table 3 shows the mean multiyear precipitation at fifteen stations. From Table 3, it can be observed that both models reasonably simulate the mean multiyear precipitation at each meteorological station, and precipitation simulated by the two models at most meteorological stations is underestimated. Precipitation in the HRB has a minimal value of 50.81 mm (Anxi station) and a maximal value of 522.95 mm (Menyuan station). Precipitation simulated by the BMA ranges from 53.98 mm (Anxi station) to 520.45 mm (Menyuan station), with biases in the range of - 4.79% (Jiuquan station) to 6.23% (Anxi station). Precipitation simulated using the SRM is in the range of 51.10 mm (Dingxin station) to 518.24 mm (Menyuan station), with biases in the range of - 7.92% (Alashanyouqi station) to 4.11% (Gaotai station). Precipitation simulated by the BMA is better than that by the SRM, except at the Anxi, Jiuquan, and Yeniugou stations. The RMSEs and MAEs of the BMA are in the range of 20.20–72.41 mm and 15.03–60.77 mm, while those of the SRM are in the range of 21.23–82.42 mm and 16.70–64.15 mm, respectively. The RMSE and MAE between the BMA and the OBS for each station were lower than those between the SRM and OBS. The mean RMSE and MAE of the BMA are 36.03 mm and 27.98 mm, respectively, and

those of the SRM are 41.86 mm and 32.68 mm, respectively. The mean RMSEs and MAEs for the BMA are 13.93% and 14.37% lower than those of the SRM, respectively.

Figure 4 shows the mean monthly precipitation obtained by the BMA and the SRM at six stations that was selected for comparison with the OBS. From Fig. 4, it can be observed that two models could simulate the mean monthly precipitation with a high level of skill, and precipitation simulated by the BMA is better than that simulated by the SRM in most months. Biases of the BMA at the Tuole station are in the range of - 17.86% (Dec) to 19.40% (Feb), whereas biases of SRM at the same station are in the range of - 24.22% (May) to 19.40% (Feb). Precipitation simulated by the BMA at the Tuole station is better than that simulated by the SRM, except for in Feb, Apr, and Dec. Biases of the BMA at the Qilian station are in the range of - 28.49% (Dec) to 9.01% (July), whereas biases of SRM at the same station are in the range of - 35.84% (Dec) to 13.50% (Feb). Precipitation simulated by the BMA at the Qilian station is better than that simulated by the SRM, except for in Jan, May, Jun, and Sep. Biases of the BMA at the Gaotai station are in the range of - 36.88% (Jan) to 12.35% (Jul), whereas biases of SRM are in the range of - 60.00% (Nov) to 50.62% (Jul). Precipitation simulated by the BMA at the Gaotai station is better than that by the SRM, except for in Mar, Apr, Oct, and Dec. Biases of the BMA at the Minqin station are in the range of - 46.23% (Jan) to 13.32% (Mar), whereas biases of the SRM at the same station are in the range of - 63.20% (Jan) to 2.75% (Mar). Precipitation simulated by the BMA at Minqin is better than that by the SRM, except for in Feb and Dec. Biases of the BMA at the Yumenzhen station are in the range of - 29.09%

**Table 3** OBS and mean multiple-annual precipitation (mm) simulated by the BMA and the SRM. Biases, RMSEs (mm), and MAEs (mm) of two models

Stations	OBS	BMA	SRM	Bias		RMSE		MAE	
				BMA (%)	SRM (%)	BMA	SRM	BMA	SRM
Anxi	50.81	53.98	52.25	6.23	2.83	20.20	21.23	15.03	16.70
Yumenzhen	69.98	67.61	66.80	- 3.38	- 4.54	24.67	27.56	17.53	20.71
Dingxin	55.20	54.08	51.10	- 2.04	- 7.42	20.23	22.78	16.10	17.69
Jinta	64.66	62.98	60.60	- 2.59	- 6.27	22.60	26.70	17.47	18.71
Jiuquan	89.83	85.53	92.00	- 4.79	2.41	29.26	35.53	22.47	27.81
Gaotai	112.89	110.94	117.53	- 1.73	4.11	28.99	30.02	21.39	24.14
Alashanyouqi	118.00	113.70	108.65	- 3.64	- 7.92	28.43	35.90	21.75	29.35
Tuole	303.84	293.96	283.89	- 3.25	- 6.56	49.58	59.20	38.58	44.89
Yeniugou	426.01	407.59	408.08	- 4.32	- 4.21	72.41	82.42	60.77	64.15
Zhangye	132.40	129.90	123.22	- 1.89	- 6.94	35.31	38.68	26.36	30.41
Qilian	408.28	405.44	397.14	- 0.69	- 2.73	54.69	59.47	41.24	45.14
Yongchang	207.31	199.38	196.40	- 3.82	- 5.26	40.84	49.32	32.99	40.17
Minqin	116.65	113.37	109.40	- 2.81	- 6.21	27.34	37.41	22.02	30.51
Dachaidan	90.21	86.27	84.63	- 4.38	- 6.19	29.18	33.45	23.17	25.86
Menyuan	522.95	520.45	518.24	- 0.48	- 0.90	56.69	68.16	42.82	53.88



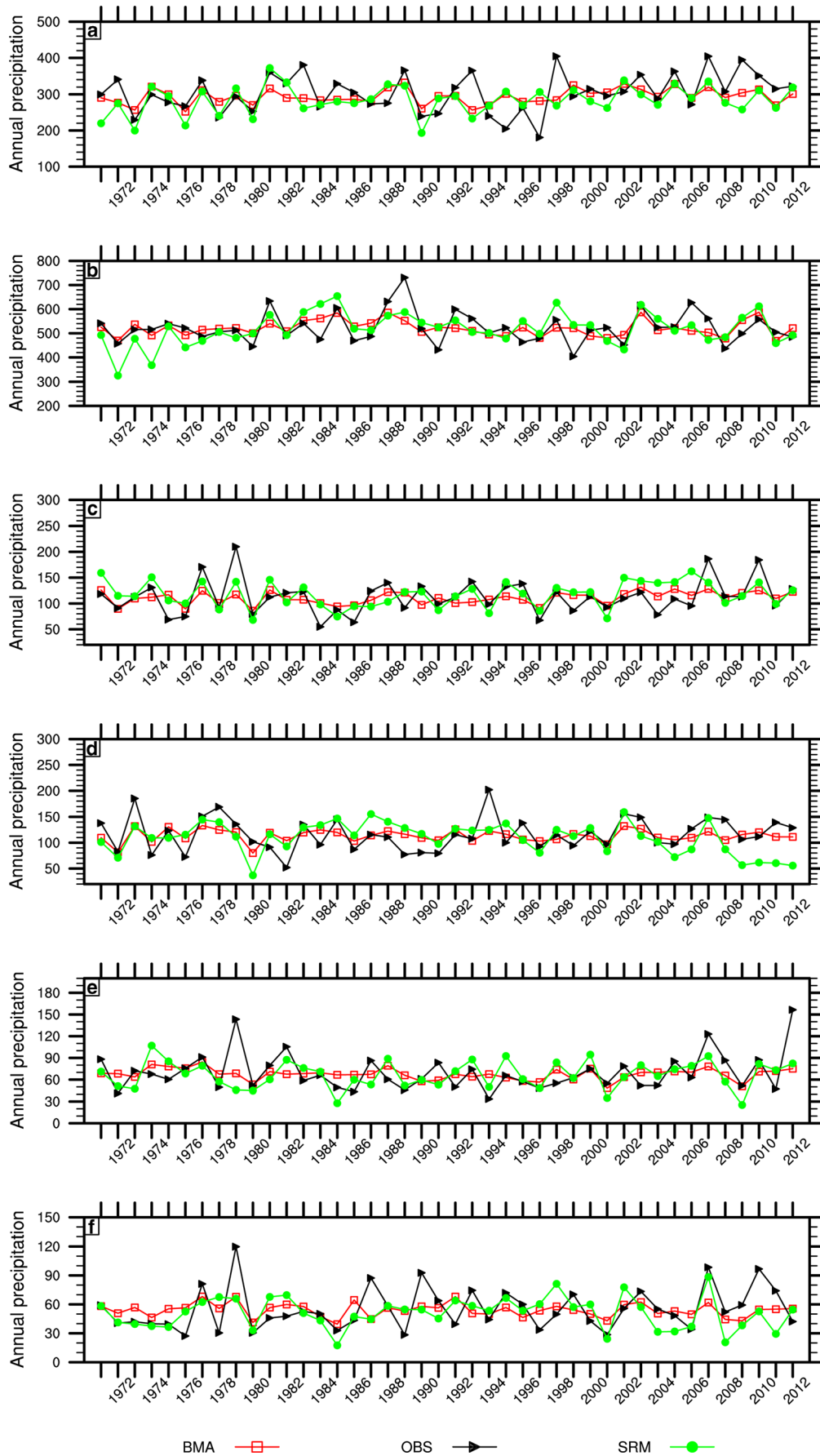
**Fig. 4** The mean monthly precipitation (mm) from six meteorological stations. **a** Tuole. **b** Menyuan. **c** Gaotai. **d** Minqin. **e** Yumenzhen. **f** Dingxin

(Jan) to 4.90% (Jul), whereas biases of the SRM at the same station are in the range of  $-35.19\%$  (Nov) to  $10.91\%$  (Jul). Precipitation simulated by the BMA at the Yumenzhen station is better than that by the SRM, except for in Jan, Mar, Sep, and Oct. Biases of the BMA at the Dingxin station are in the range of  $-30.08\%$  (Jan) to  $14.81\%$  (Nov), whereas biases of the SRM at the same station are in the range of  $-99.22\%$  (May)

to  $32.53\%$  (Aug). Precipitation simulated by the BMA at the Dingxin station is better than that by the SRM, except for in Jan and Dec. The SRM is unstable for simulating precipitation

**Fig. 5** The time series of annual precipitation (mm) from six meteorological stations. **a** Tuole. **b** Menyuan. **c** Gaotai. **d** Minqin. **e** Yumenzhen. **f** Dingxin





BMA —□— OBS —▲— SRM —●—

in some months, with biases larger than 50% (e.g., May, Jul, and Oct at the Gaotai station; Jan at the Minqin station; and May at the Dingxin station). However, the BMA could effectively overcome this defect.

The time series of annual precipitation at the six stations mentioned above are shown in Fig. 5. The correlation coefficients between the BMA and the OBS for these stations are 0.40, 0.47, 0.51, 0.55, 0.31, and 0.37 and 0.30, 0.45, 0.51, 0.29, 0.28, and 0.33, respectively. In terms of the correlation coefficient, the two models have their own advantages, which depend on the meteorological station.

## 4 Discussion and conclusions

In this study, we extended the BMA to statistical downscaling to simulate precipitation in the HRB, which is located in an arid area of Northwest China. The observed monthly rainfall in the HRB and fourteen reanalysis variables were used to establish the SRM and BMA. Monthly precipitation in the HRB over the period 2003–2012 was simulated using the SRM and BMA to compare with the OBS. The results showed the following: (1) the BMA and the SRM reasonably reproduced the spatial pattern of precipitation in the HRB with a high level of skill. The biases of precipitation simulated by the BMA were in the range of  $-3.07$  to  $-0.83\%$ , with RMSEs in the range of 15.74 to 36.11 mm and MAEs in the range of 11.55 to 27.52 mm in the three subregions and across the entire HRB; the biases of the SRM were in the range of  $-4.11$  to  $-3.46\%$ , with RMSEs in the range of 16.52 to 37.04 mm and MAEs in the range of 11.24 to 29.51 mm. Both models could reasonably reproduce the time series of annual precipitation in the HRB. The correlation coefficients between the BMA and the OBS in the three subregions and across the entire HRB were 0.50, 0.49, 0.43, and 0.56, and they were 0.55, 0.56, 0.43, and 0.65 between the SRM and OBS. Two models could reasonably simulate the monthly precipitation at single stations. The biases of multiyear mean precipitation at the fifteen meteorological stations simulated by the BMA were in the range of  $-4.79$  to  $6.23\%$ , and the biases of the SRM were in the range of  $-7.92$  to  $4.11\%$ . (2) The BMA produced more accurate results than the SRM. The multiyear mean precipitation for twelve of the fifteen meteorological stations that were simulated by the BMA was better than that simulated by the SRM. The RMSE and MAE of the BMA for each station were lower than those of the SRM. The BMA had a mean RMSE and MAE that were  $-13.93\%$  and  $-14.37\%$  less than those of the SRM, respectively. (3) The BMA gave a weight to each predictor, which reflected the degree of predictor influence on precipitation during the training period. The SRM was calibrated by the better-performing predictors, which received higher weights compared with the worse-performing predictors. Both methods could simulate

the monthly mean precipitation powerfully, and the BMA performed slightly better than the SRM. However, the BMA could effectively reduce the RMSE and MAE of precipitation and improve the correlation coefficient. This indicates that the monthly precipitation simulated by the BMA has better consistency with the observed values. The BMA is more suitable for studies with high-precision requirements.

In the SRM, using too many predictors to establish a statistical model may lead to overfitting and thus decrease the predictive power, whereas using too few predictors would cause information of other predictors influencing local precipitation to be ignored. The BMA (Leamer 1978; Kass and Raftery 1995; Hoeting et al. 1999) overcomes this problem by integrating the overall linear models of predictors influencing precipitation. Furthermore, in the study, it is assumed that the predicted variables extracted from the NCEP/NCAR reanalysis data sets are normal distribution. However, the multiyear mean precipitation values for three of the fifteen meteorological stations that were simulated by the SRM were better than those simulated by the BMA, and some correlation coefficients of the BMA were lower than those of the SRM at some stations. The cause of these issues may be that some selected predictors for these stations did not follow a normal distribution. Converting non-normal distribution predictors into normal distribution predictors needs to be studied further.

**Acknowledgments** The authors thank the Environmental and Ecological Science Data Center of Western China, National Natural Science Foundation of China, for providing the meteorological and hydrological observation datasets.

**Funding information** This study was financially supported by the National Natural Science Foundation of China (Grant No. 91425304 and 51339004) and the Chinese Academy of Sciences Strategic Priority Program (Grant No. XDA05090206).

## References

- Campozano L, Tenelanda D, Sanchez E, Samaniego E, Feyen J (2016) Comparison of statistical downscaling methods for monthly total precipitation: case study for the Paute River basin in Southern Ecuador. *Adv Meteorol* 2016:6526341–6526313. <https://doi.org/10.1155/2016/6526341>
- Cheng GD, Xiao HL, Xu ZM et al (2006) Water issue and its countermeasure in the inland river basin of Northwest China (in Chinese). *J Glaciol Geocryol* 28:406–413
- Cheng CS, Li GL, Li Q, Auld H (2011) A synoptic weather-typing approach to project future daily rainfall and extremes at local scale in Ontario, Canada. *J Clim* 24:3667–3685. <https://doi.org/10.1175/2011JCLI3764.1>
- Duan QY, Ajami NK, Gao X, Sorooshian S (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>
- Fatichi S, Ivanov VY, Caporali E (2011) Simulation of future climate scenarios with a weather generator. *Adv Water Resour* 34:448–467. <https://doi.org/10.1016/j.advwatres.2010.12.013>

- Fatichi S, Ivanov VY (2013) Assessment of a stochastic downscaling methodology in generating an ensemble of hourly future climate time series. *Climate Dynamics* 40:1841–1861
- Hay LE, McCabe GJ, Wolock DM, Ayers MA (1991) Simulation of precipitation by weather type analysis. *Water Resour Res* 27:493–501. <https://doi.org/10.1029/90WR02650>
- Hoeting JA, Madigan D, Raftery AE, Volinsky CT (1999) Bayesian model averaging: A tutorial (with discussion). *Stat Sci* 14:382–401
- Huang J, Zhang JC, Zhang ZX (2012) Simulation of extreme precipitation indices in the Yangtze River basin by using statistical downscaling method (SDSM). *Theor Appl Climatol* 108:325–343. <https://doi.org/10.1007/s00704-011-0536-3>
- Huth R (1999) Statistical downscaling in central Europe: evaluation of methods and potential predictors. *Clim Res* 13:91–101. <https://doi.org/10.3354/cr013091>
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 77:437–471. [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2)
- Kannan S, Ghosh S (2010) Prediction of daily rainfall state in a river basin using statistical downscaling from GCM output. *Stoch Env Res Risk A* 25:457–474. <https://doi.org/10.1007/s00477-010-0415-y>
- Kass RT, Raftery AE (1995) Bayes Factors. *J Am Stat Assoc* 90:773–795
- Kim JW, Chang JT, Baker NL, Wilks DS, Gates WL (1984) The statistical problem of climate inversion-determination of the relationship between local and large-scale climate. *Mon Weather Rev* 112:2069–2077. [https://doi.org/10.1175/1520-0493\(1984\)112<2069:TSPOCI>2.0.CO;2](https://doi.org/10.1175/1520-0493(1984)112<2069:TSPOCI>2.0.CO;2)
- Lan YC et al (2005) Change of water resources in mountainous area of Heihe River under global-warming scene (in Chinese). *J Desert Res* 25:863–868
- Leamer EE (1978) *Specification Searches*. Wiley 370 pp
- Li Y, Smith I (2009) A statistical downscaling model for southern Australia winter rainfall. *J Clim* 22:1142–1158. <https://doi.org/10.1175/2008JCL2160.1>
- Liu JM, Yuan D, Zhang LP, Zou X, Song XY (2016) Comparison of three statistical downscaling methods and ensemble downscaling method based on Bayesian model averaging in Upper Hanjiang River basin, China. *Adv Meteorol* 2016:7463963. <https://doi.org/10.1155/2016/7463963>
- Maraun D, Osborn TJ, Rust HW (2011) The influence of synoptic airflow on UK daily precipitation extremes. Part I: observed spatio-temporal relationships. *Clim Dyn* 36:261–275. <https://doi.org/10.1007/s00382-009-0710-9>
- Murphy JM (1999) An evaluation of statistical and dynamical techniques for downscaling local climate. *J Clim* 12:2256–2284
- Osca J, Romero R, Alonso S (2013) Precipitation projections for Spain by means of a weather typing statistical method. *Glob Planet Chang* 109:46–63. <https://doi.org/10.1016/j.gloplacha.2013.08.001>
- Raftery AE, Zheng Y (2003) Discussion: performance of Bayesian model averaging. *J Am Stat Assoc* 98:931–938. <https://doi.org/10.1198/016214503000000891>
- 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>
- Ruping M, Straus DM (2002) Statistical–dynamical seasonal prediction based on principal component regression of GCM ensemble integrations. *Mon Weather Rev* 130:2167–2178. [https://doi.org/10.1175/1520-0493\(2002\)130<2167:SDSPBO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<2167:SDSPBO>2.0.CO;2)
- Sachindra DA, Huang F, Barton A, Perera BJC (2014) Statistical downscaling of general circulation model outputs to precipitation - part 2: bias-correction and future projections. *Int Journal Climatol* 34:3282–3303
- Singh D, Jain SK, Gupta RD (2015) Statistical downscaling and projection of future temperature and precipitation change in middle catchment of Sutlej River basin, India. *J Earth Syst Sci* 124:843–860. <https://doi.org/10.1007/s12040-015-0575-8>
- Spak S (2007) A comparison of statistical and dynamical downscaling for surface temperature in North America. *J Geophys Res* 112:D08101. <https://doi.org/10.1029/2005JD006712>
- Su HF, Xiong Z, Yan XD, Dai XD, Wei WG (2016) Comparison of monthly rainfall generated from dynamical and statistical downscaling methods: a case study of the Heihe River basin in China. *Theor Appl Climatol* 129:437–444. <https://doi.org/10.1007/s00704-016-1771-4>
- Viallefond V, Raftery AE, Richardson S (2001) Variable selection and Bayesian model averaging in case-control studies. *Stat Med* 20:3215–3230. <https://doi.org/10.1002/sim.976.abs>
- Von Storch H, Zorita E, Cubasch U (1993) Downscaling of global climate change estimate to regional scales: an application to Iberian rainfall in wintertime. *J Clim* 6:1161–1171. [https://doi.org/10.1175/1520-0442\(1993\)006<1161:DOGCCCE>2.0.CO;2](https://doi.org/10.1175/1520-0442(1993)006<1161:DOGCCCE>2.0.CO;2)
- Vrac M, Naveau P (2007) Stochastic downscaling of precipitation: from dry events to heavy rainfalls. *Water Resour Res* 43:W07402. <https://doi.org/10.1029/2006WR005308>
- Wang L, Ranasinghe R, Maskey S, van Gelder PHAJM, Vrijling JK (2015) Comparison of empirical statistical methods for downscaling daily climate projections from CMIP5 GCMs: a case study of the Huai River basin, China. *Int J Climatol* 36:145–164. <https://doi.org/10.1002/joc.4334>
- Wilby RL, Wigley TML (1997) Downscaling general circulation model output: a review of methods and limitations. *Prog Phys Geogr* 21:530–548. [https://doi.org/10.1016/S0022-1694\(99\)00136-5](https://doi.org/10.1016/S0022-1694(99)00136-5)
- Wilby R, Hay LE, Leavesley GH (1999) A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River basin, Colorado. *J Hydrol* 225:67–91. [https://doi.org/10.1016/S0022-1694\(99\)00136-5](https://doi.org/10.1016/S0022-1694(99)00136-5)
- Zhang XL, Yan XD (2015) A new statistical precipitation downscaling method with Bayesian model averaging: a case study in China. *Clim Dyn* 45:2541–2555. <https://doi.org/10.1007/s00382-015-2491-7>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.