



Improving China's summer precipitation prediction in 2020 by observational constrained bias correction

Yun Wei¹ · Haipeng Yu² · Jianping Huang¹ · Xiaoyue Liu¹ · Jie Zhou¹

Received: 11 December 2020 / Accepted: 10 June 2021 / Published online: 29 June 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2021

Abstract

In summer 2020, precipitation in many parts of China reached its maximum in the past 60 years, and southern China suffered from severe flooding, which was mainly due to the strong westward subtropical high, the enhanced east Asian summer monsoon, the development of meridional circulation, and the active cold vortex in the middle and high latitudes. We tried to apply a statistical bias correction method to improve the prediction of summer precipitation at 160 stations in China in 2020. In summer 2020, the precipitation anomaly percentages were above 40% in the Yellow River Basin and the Yangtze River Basin. Using the bias correction method, the main rain belt of the corrected summer precipitation was located in the Yangtze River's middle reaches and the Yellow River Basin, which is in line with the observations, but it does not catch the trend of more summer precipitation in the Huaihe River Basin. In addition, the corrected predicted results can basically catch the trend of dryness in South and Southwest China. Relative to the original model output, the abnormal correlation coefficient (ACC) between the corrected predictions, and the observations increased from -0.02 to 0.30 , and the prediction score (PS) increased from 67.8 to 77.4 for summer rainfall in 2020. Moreover, the bias correction method showed stable and considerable improvement for the prediction of summer precipitation in China from 2014 to 2019. Overall, this study demonstrates that the bias correction method can effectively improve the accuracy of the summer precipitation predicted results in China.

Keywords Bias correction · Summer precipitation prediction · China

1 Introduction

According to multiple reports (reported by National Climate Center of China), in the summer of 2020, southern China suffered from severe flooding. In 2020, the summer precipitation in China was 14.7% more than that during the same period in normal years and the second largest since 1961, second only to 1998. The precipitation in the Yangtze River and Yellow River basins was 35% more than that in the same period during normal years and the highest in the past 60 years. The summer rainfall in the provinces of Gansu,

Sichuan, Hubei, and Anhui experienced once in the last 60 years. In addition, the accumulated numbers of rainstorms in China were also the largest in the last 60 years. Affected by heavy rainfall, the floods in southern China have affected more than 45 million people in the first two months of summer in 2020, of which more than 142 people died or were missing, and the economic losses were estimated to be approximately 116 billion RMB (approximately 16.5 billion US dollars) (Wei et al. 2020). A more accurate prediction of the precipitation distribution in the flood season in advance can provide a reliable reference for countries to better respond to drought and flood disasters, and can also effectively reduce casualties and economic losses.

According to the assessment report of the World Meteorological Organization (World Meteorological Organization (WMO) 2020), the past 10 years have been the hottest since the observational record existed. Under global warming, the global atmospheric circulation system has undergone significant changes, leading to the increased frequency of regional extreme precipitation and drought events (Hartmann et al. 2013). China's climate is also changing,

✉ Haipeng Yu
yuhp@lzb.ac.cn

¹ Key Laboratory for Semi-Arid Climate Change of the Ministry of Education, College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China

² Key Laboratory of Land Surface Process and Climate Change in Cold and Arid Regions, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China

and the frequency of droughts and floods has increased, which greatly restricts China's economic construction and social development (Ding et al. 2006; Huang et al. 2006). Moreover, China is located in southeastern Eurasia. In the southeast is the world's largest ocean, the Pacific Ocean, and the southwest has the world's highest elevation, the Qinghai-Tibet Plateau. The special geographical location leads to many climatic factors affecting summer precipitation in this area, including sea surface temperature (Feng et al. 2011; Newman et al. 2012; Sun et al. 2009), Eurasian winter snow cover (Yang and Xu 1994), Arctic sea ice (Wu et al. 2009), East Asian summer monsoon (EASW) (Wei et al. 2017), western Pacific subtropical high (WPSH) (Hu 1997), etc.. Affected by many of the above factors, seasonal rainfall prediction in China has been facing various difficulties and challenges.

Since the 1990s, general circulation models (GCMs) have become an important tool for predicting seasonal climate. Many researchers (Kang et al. 2002; Min et al. 2014; Wang et al. 2009) have evaluated the ability of several dynamic models to predict seasonal precipitation. It has been shown that dynamic models can provide skillful large-scale features of atmospheric variability (Rowell 1996), but they also exhibit a limited capability in predicting summer rainfall (Alessandri et al. 2010; Gong et al. 2018), which can be improved using dynamic-statistic techniques to correct the raw model output (Chu et al. 2008; Feng et al. 2013; Gong et al. 2016; Kang et al. 2004; Kug et al. 2008; Liu and Fan 2014; Wang and Fan 2009; Zheng et al. 2009; Zheng et al. 2013). Statistical downscaling has become a widely used tool (Chu et al. 2008; Chen et al. 2012; Sun and Chen 2012) because of its relatively lower computational consumption relative to dynamic downscaling (Benestad and Haugen 2007). For example, Feddersen et al. (Feddersen et al. 1999) and Uvo et al. (Uvo et al. 2001) used the leading modes of a singular value decomposition analysis (SVDA) to improve the skill of seasonal precipitation simulations. The results in Min et al. (Min et al. 2014) also demonstrated that model correction using the stepwise pattern projection method (SPM) has a positive effect for improving the multimodel predictions of precipitation. Wei and Huang (Wei and Huang 2010) used 500 hPa geopotential height as a predictor and found increased skill in predicting monthly rainfall in East China based on a canonical correlation analysis (CCA). Similar results using related statistical correction methods applied to a number of different atmospheric model simulations have been reported for many studies (Kang et al. 2004; Fan et al. 2008; Tippett et al. 2010; Wang et al. 2000a; Wang et al. 2010; Widmann et al. 2003). Although the predictability of GCMs was effectively increased by these proposed schemes, it still cannot meet the operational demand. Most current statistical techniques that are used in China are only available at the East China (Guo et al. 2011) and regional-mean scales but not at the site scale. Thus, in this study, based on the results of the BCC-CSM 1.1 model

provided in March of each year by the National Climate Center, we use an empirical orthogonal function (EOF) to decompose and establish a linear regression relationship between principal components (PCs) in historical observations and model simulations to correct model predictions and reduce predicted errors in station-based rainfall in China as a whole.

The other parts of this study are described as follows, Section 2 introduces the data used and the bias correction methodology. The basic characteristics of summer precipitation and atmospheric circulation in 2020 are analyzed in Section 3. Section 4 introduces the specific process of the summer precipitation correction in 2020 and analyzes the effect of this bias correction method on the prediction correction in 2020. Section 5 provides an assessment of this bias correction method's ability to improve rainfall during the past 6 years in China. The conclusions and discussions are presented in Section 6.

2 Data and methods

2.1 Data

In this study, the monthly precipitation data from 1951 to 2020 at 160 stations provided by the National Climate Center of the China Meteorological Administration are used as observational precipitation data. The model data include the monthly precipitation field predicted by BCC_CSM 1.1 in March from 1991 to 2020. BCC_CSM1.1 is the version 1.1 of the Beijing Climate Center-Climate System Model (BCC_CSM) developed at the Beijing Climate Center (BCC), China Meteorological Administration, based on NCAR CCSM2.0.1. Since the model is the grid data with a $2.5^{\circ} \times 2.5^{\circ}$ spatial resolution and the observation is site data, we used the '*dspnt2*' function of NCAR Command Language to interpolate the model data (gridded data) into the site (at specified points) before correction. The more information about '*dspnt2*' function can be found in <http://www.ncl.ucar.edu/Document/Functions/Built-in/dspnt2.shtml>.

To analyze the atmospheric circulation and water vapor transport in summer 2020, we also used the monthly geopotential heights, air temperature, wind velocity, surface pressure, and specific humidity since 1948 from the US National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) Reanalysis (Kalney et al. 1996).

2.2 Bias correction methodology

First, the model and observed data sets can be decomposed into eigenvectors and corresponding principal components (PCs) by applying an EOF analysis, as follows:

$$Z_p(x, t) = \sum_{i=1}^m EOF_i^p(x) PC_i^p(t) \tag{1}$$

$$Z_o(x, t) = \sum_{j=1}^k EOF_j^o(x) PC_j^o(t) \tag{2}$$

Among them, x is the spatial position of 160 stations and t is time. For prediction in 2020, t represents 1991–2019. Z_p is the model projection, and Z_o is the observation. Then, each PC of the observed data set is regressed by the leading PCs of model simulation, and a matrix of regression coefficients is obtained, which shows in Equation 3:

$$PC_j^o(t) = \sum_{i=1}^m \alpha_{i,j} PC_i^p(t) + \varepsilon_j \tag{3}$$

where $\alpha_{i,j}$ is the regression coefficient. According to the relationship in Equation 3, the observed PCs can be predicted by the simulated PCs from future model simulations in Equation 4:

$$Z_c(x, t) = \sum_{j=1}^K (\sum_{i=1}^M \alpha_{i,j} Z_p(x, t) + \varepsilon_j) \cdot EOF_j^o(x) \tag{4}$$

Among them, Z_c is the corrected result, K is the number of selected observed EOF modes, and M is the number of selected modeled EOF decomposition modes. For prediction in 2020, the t in Equation 4 represents 2020.

The detailed corrected calculation process can be seen in Huang et al. (Huang et al. 2016). It is worth noting that the climatology in the model simulations needed to be adjusted to be the same as the observations before correction. Since the model simulation started in 1991, for predicting the 2020 summer precipitation, we choose 1991–2019 to be the climatology period that needs to be adjusted.

In general, we first report the past, select the optimal principal component number by calculating the abnormal correlation coefficient (ACC), prediction score (PS), etc., and then, we use the same method (according to the selected principal component number) to correct the model results for the predicted year. Finally, we obtain the corrected predicted results.

ACC and PS are calculated as follows:

$$ACC = \frac{\sum_{i=1}^n (F_i - C_i - M_i^{F,C})(O_i - C_i - M_i^{O,C}) \cos \phi_i}{\sqrt{\sum_{i=1}^n (F_i - C_i - M_i^{F,C})^2 \cos \phi_i} \sqrt{\sum_{i=1}^n (O_i - C_i - M_i^{O,C})^2 \cos \phi_i}} \tag{5}$$

where n is the number of spatial grids, ϕ_i is the latitude of the grid point, F is the predicted value, O is the observed value, and C is the average climate, $M_i^{F,C} = \frac{1}{n} \sum_{i=1}^n (F_i - C_i)$, $M_i^{O,C} = \frac{1}{n} \sum_{i=1}^n (O_i - C_i)$. Here, all anomalies for observation are calculated based on 1981–2010 climatology. Because the climatology in the model simulations had been adjusted to be the same as the observations before correction, the climatology of the corrected model results when calculating ACC was also the observed 1981–2010 climatology. While before the climatology adjustment, the climatology of the original model results when calculating ACC was the modeled 1991–2010 climatology.

$$PS = \frac{a \times N0 + b \times N1 + c \times N2}{(N - N0) + a \times N0 + b \times N1 + c \times N2 + M} * 100 \tag{6}$$

Among them, $N0$ is the total number of stations where the precipitation trend was correctly predicted, N_1 is the total number of stations where the precipitation anomaly percentages between 20% and 50% were correctly predicted, N_2 is the total number of stations where the precipitation anomaly percentages above 50% were correctly predicted, M is the total number of underreported stations (the predicted precipitation anomaly percentage is less than 50%, while the observed is 100% or more), and N is the actual number of stations in the evaluation. a , b , and c are the weight coefficients of $N0$, $N1$, and $N2$, respectively, which are taken as $a = 2$, $b = 2$, and $c = 4$, respectively.

The PS is based on the assessment of the ability to predict a wide range of anomalies. It is expressed in a percentage system, which is relatively intuitive. In addition, the weighted score of the anomaly has a clear guiding effect on improving the predicted ability of an anomalous climate. The PS is 100 when the prediction and the actual situation are completely consistent.

3 2020 summer precipitation and circulation anomalies in China

3.1 2020 summer precipitation in China

In the summer of 2020, the cumulative precipitation during June and July in some areas of the Yangtze River's lower reaches was the severest flood year in the past 60 years, which caused severe flooding and economic losses in southern China (Wei et al. 2020). The precipitation in China showed a gradual increase from the Yangtze River Valley (YRV) to North China (Fig. 1a). Among them, the summer precipitation in the southeast area was only 500 mm, and most of the precipitation area was located in the YRV, up to above 1000 mm. In addition, the precipitation in the southwest areas reached more than 800 mm.

From the percentage of precipitation anomalies (Fig. 1b), the summer precipitation in China in 2020 was in a state of less in the south and more in the central and northern parts. The precipitation positive anomaly was located in the Yangtze River Basin and the Yellow River Basin, where the summer rainfall anomaly percentage was 35% more than the average climatic precipitation during 1981–2010. In addition, the summer rainfall in the Yangtze River's middle and lower reaches was 40% and 60% higher, respectively. Continuous heavy rainfall caused frequent floods and mudslides in the middle and lower reaches of the Yangtze River in 2020 and brought significant economic losses to these areas (Wei et al. 2020).

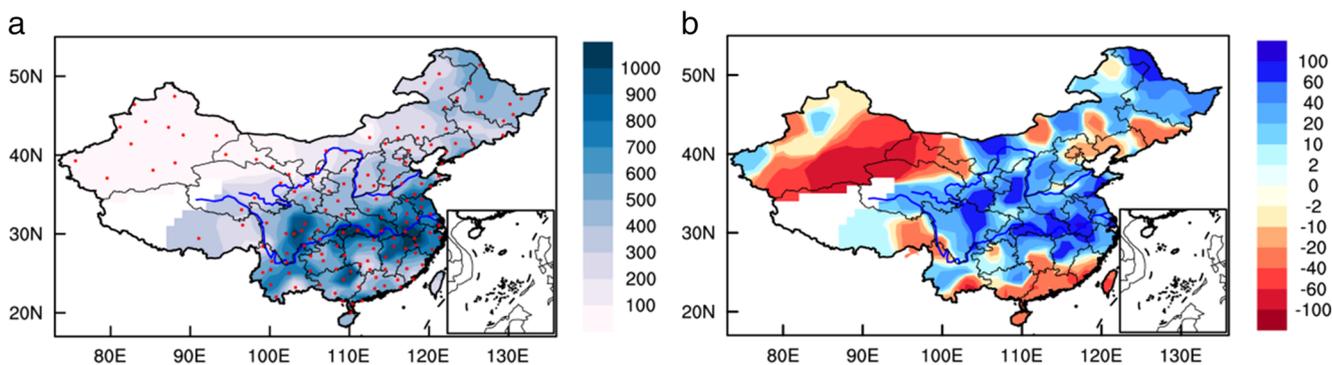


Fig. 1 Spatial distributions of observed summer rainfall in 2020. **a** Summer rainfall (units: mm) spatial pattern in China in 2020 with 160 station locations. The dots show the locations of the 160 stations in China. **b** Distributions of observed summer rainfall anomaly percentages (units: %) in 2020

Especially in some areas of southern Anhui Province, the summer rainfall was 100% more than the average from 1981–2010, which brought great pressure on local flood resistance. The areas with less precipitation were mainly located in the southeast region and Liaoning Province. The summer rainfall in the western part of Northwest China was 60% less than in the same period in history and even more than -100% in some areas. The percentage of precipitation anomalies was larger in the northwestern region due to the lower summer precipitation.

3.2 Atmospheric circulation and water vapor transportation in summer 2020

In the summer of 2020, the WPSH was located more southward and significantly stronger (Fig. 2c–f). The EASW was significantly stronger than normal (Fig. 2b), which was conducive to the warm and humid airflow in the western Pacific. It was transported to the middle and lower reaches of the Yangtze River to the north, causing water vapor convergence in these areas (Fig. 3a). The strong WPSH and the EASW made the warm and humid air flow less stable in South China, causing divergence of water vapor in South China (Fig. 3a). However, the formation of precipitation requires not only sufficient water vapor transport but also lower temperatures. From the perspective of the mid-to-high latitude circulation, in the summer of 2020, the mid-to-high latitude circulation was dominated by the meridian circulation in Europe and Asia (Fig. 2c–f). The polar vortex located near the new island transported the cold polar air along the northwest wind to northwest China (Fig. 2c), causing negative temperature anomalies in this area (Fig. 2a). The strong westerly wind was transported to the middle and lower reaches of the Yangtze River in China and its northern regions (Fig. 2c). In the middle and lower reaches of the Yangtze River and to the north, cold air from the polar vortex meets warm air from the western Pacific, causing more precipitation in these areas in the summer of 2020.

From June to July, the position of the western ridge point of the subtropical high reached approximately 110° E, its ridge line swinging between 10° N and 20° N, and its position was more westward and southerly than usual in the same period of the year (according to Zhao et al. (Zhao et al. 2015); the usual location of the western ridge was at approximately 135° E, and the ridge line was approximately 25° N) (Fig. 2d, e). It is conducive to the transportation of water vapor to the middle and lower reaches of the Yangtze River, especially the lower reaches, resulting in a strong water vapor convergence center in this area (Fig. 3b, c). In addition, from June to July, in the mid-to-high latitude regions over Eurasia, strong high-pressure ridges occurred near the west of the Ural Mountains and the Okhotsk Sea, and the low-pressure trough occurred near the Balkhash Lake (Fig. 2d, e). The cold air through the northwestern and/or northeastern paths could reach the middle and lower reaches of the Yangtze River in China. As a result, the rainfall in the middle and lower reaches of the Yangtze River in the summer of 2020 was abnormally high (Fig. 1b), which are consistent with the results in Liu and Ding (Liu and Ding 2020) and Chen et al. (Chen et al. 2020). In August, the atmospheric circulation in the mid-high latitudes over Eurasia was adjusted to “two troughs and one ridge” pattern. The position of the western ridge of the subtropical high returned to approximately 115° E, and its ridge reached approximately 28° N (Fig. 2f). The strong subtropical high pressure caused the warm and humid airflow to be transported to northern China, which met with the active cold air through the middle path from the polar regions on the northwest side of the WPSH, causing large-scale abnormal water vapor convergence in the Yellow River Basin and its northern regions (Fig. 3d), resulting in abnormally heavy precipitation in these areas than in the same period historically. Some studies (Liu et al. 2021; Wang 2020) pointed out that the abnormal atmospheric circulation from June to July 2020 may be mainly attributed to the continuous warmer of the tropical Indian Ocean, while Liu et al. (Liu et al. 2020) pointed out the sequential warm and cold Meiyu front regulated by the North Atlantic Oscillation (NAO) was responsible for the extreme

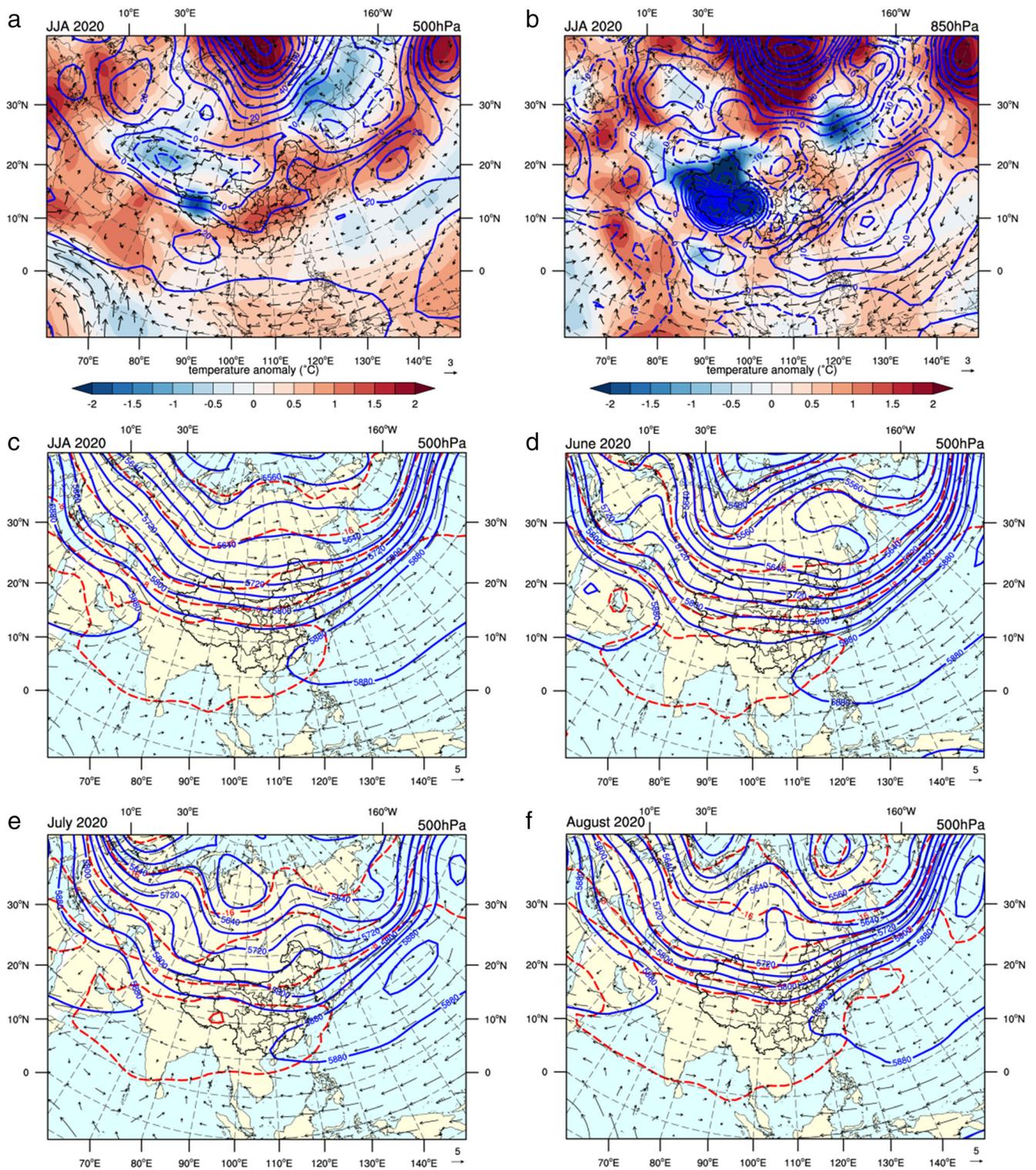


Fig. 2 Averaged atmospheric circulation anomalies in summer (JJA) 2020. Wind anomaly (arrows, units: m/s), temperature anomaly (color, units: °C), and geopotential height anomaly (blue line, units: gpm) field at

500h Pa (a) and 850h Pa (b) in summer 2020. c–f The distribution of 500 hPa geopotential height (blue line), temperature (red line), and wind (vector) in 2020 c summer, d June, e July, and f August

Meiyu precipitation. From the results of Liu et al. (Liu et al. 2021), the important reason for the significant adjustment of the

atmospheric circulation in mid-low latitude in August was the abnormal activity of the tropical Madden-Julian Oscillation.

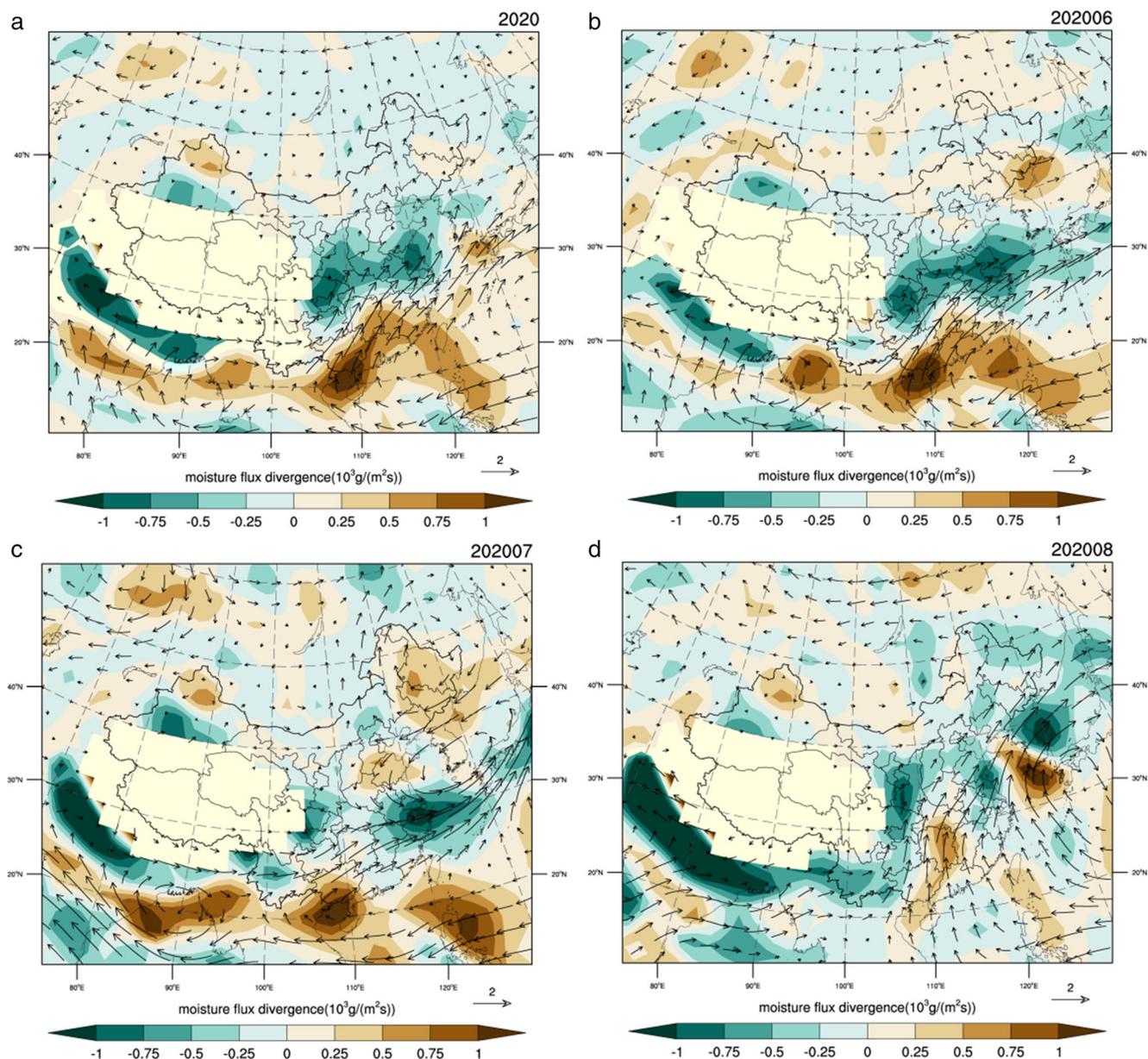


Fig. 3 The moisture flux integrated from the 1000 to 300 hPa anomaly (vectors, units: $\text{kg}/\text{m}/\text{s}$) and divergence anomaly (shading, units: $\text{kg}/\text{m}^2/\text{s}$) in summer (JJA) 2020. **a** The moisture flux anomaly and divergence

anomaly in summer (JJA) 2020. **b, c, d** are the same as **a** but in 2020 June (**b**), July (**c**), and August (**d**), respectively.

Based on the above analysis, a stronger WPSH with a westward shift from its normal position, abundant water vapor transported by the enhanced EASM to the inland areas, and frequent cold air activities in the middle and high latitudes provides favorable conditions for unusually increased precipitation in the Yangtze River region and Yellow River Basin of China.

4 Bias correction based on BCC_CSM prediction

Figures 4 and 5 show the first three main spatial modes and their corresponding principal components obtained after the

EOF decomposition of the observations and models of summer precipitation from 1991 to 2019. From Fig. 4, the first main mode after EOF decomposition represents the precipitation trend during the flood season from 1991 to 2019. In general, the precipitation in the flood season in South and North China showed a decreasing trend, and the precipitation in central and northwest China showed an increasing trend. The second mode provides the characteristics of the interdecadal variation in precipitation during the flood season. South-Central-North China presents a negative-positive-negative tripole spatial mode, and the corresponding principal component also shows a 20-year cycle. The effect of the North

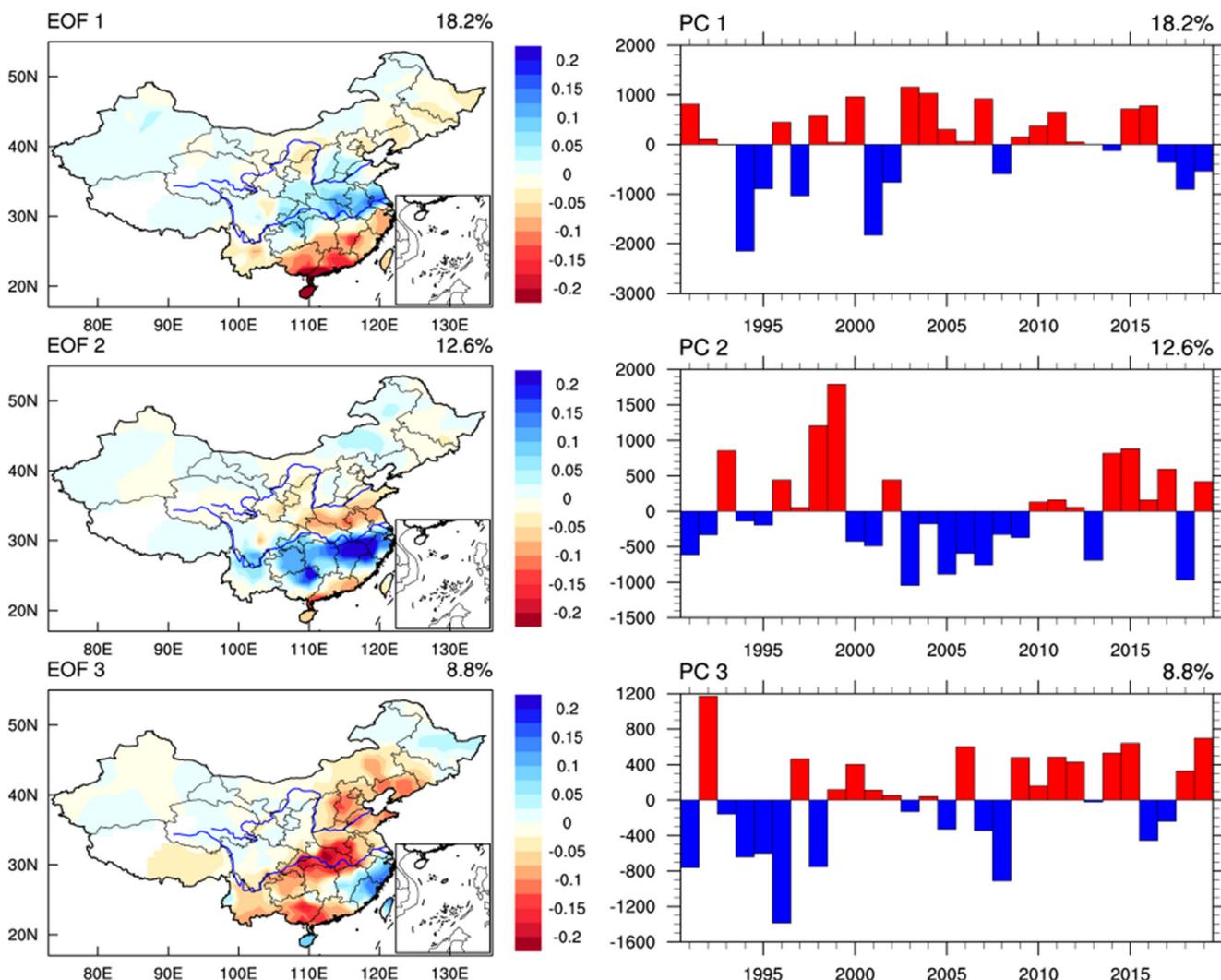


Fig. 4 The first 3 EOF modes of observed summer rainfall during 1991–2019 and its corresponding principal components (PCs)

Pacific SST is manifested in the fact that the correlation between the Pacific Decadal Oscillation (PDO) index and the interdecadal variation in precipitation during the flood season is consistent with the second mode. The third mode shows the interannual change in precipitation during the flood season, with a period of approximately 10 years. Its spatial mode shows the opposite dipole mode in South-Central China. Compared to the observation (Fig. 4), the spatial modes of the BCC_CSM model (Fig. 5) are basically similar to the observation, while their corresponding principal components are very different from the observations. Therefore, our bias correction method is mainly used to establish a multiple regression relationship between the PCs of the model and the observation to correct the predicted PC and then combine it with the observation spatial mode to give the corrected predicted result.

According to Equation 4, to know Z_c , we need to know $EOF_j^o(x)$, K , M , and $\alpha_{i,j}$. Among them, $EOF_j^o(x)$ and $\alpha_{i,j}$ can be obtained by Equations 2–3. K and M are the number

of selected observed and modeled EOFs and could be determined based on the average ACC and PS obtained by cross-validation over the past few decades (1991–2019), which are shown in Fig. 6. In the cross-validation, 1 year should be withheld from the predicted data set, and a prediction was made for the withheld year (Michaelsen 1987). To avoid overimprovement of the correction skills for the year before and after, 3 years were withheld from the prediction and data set. This procedure was repeated for all years, yielding 29 years (1991–2019) of predictions for validation. Figure 6 shows that when M are 5–7 and K are 12–13, the average PS and ACC of the cross-validation results from 1991 to 2019 are both higher. Figure 7 shows the ACC and PS scores for the cross-validation during 1991–2019 when K take 12–13 and M take 5–7. After cross-validation, the ACC increased from 0.01 to 0.08 for 1991–2019, and the mean PS improved from 66.7 to 69.9. Especially in 1993–1997, 2004, 2010–2011, and 2016, the bias correction method could significantly improve the ACC and PS scores of the predicted results. Therefore, for correction in

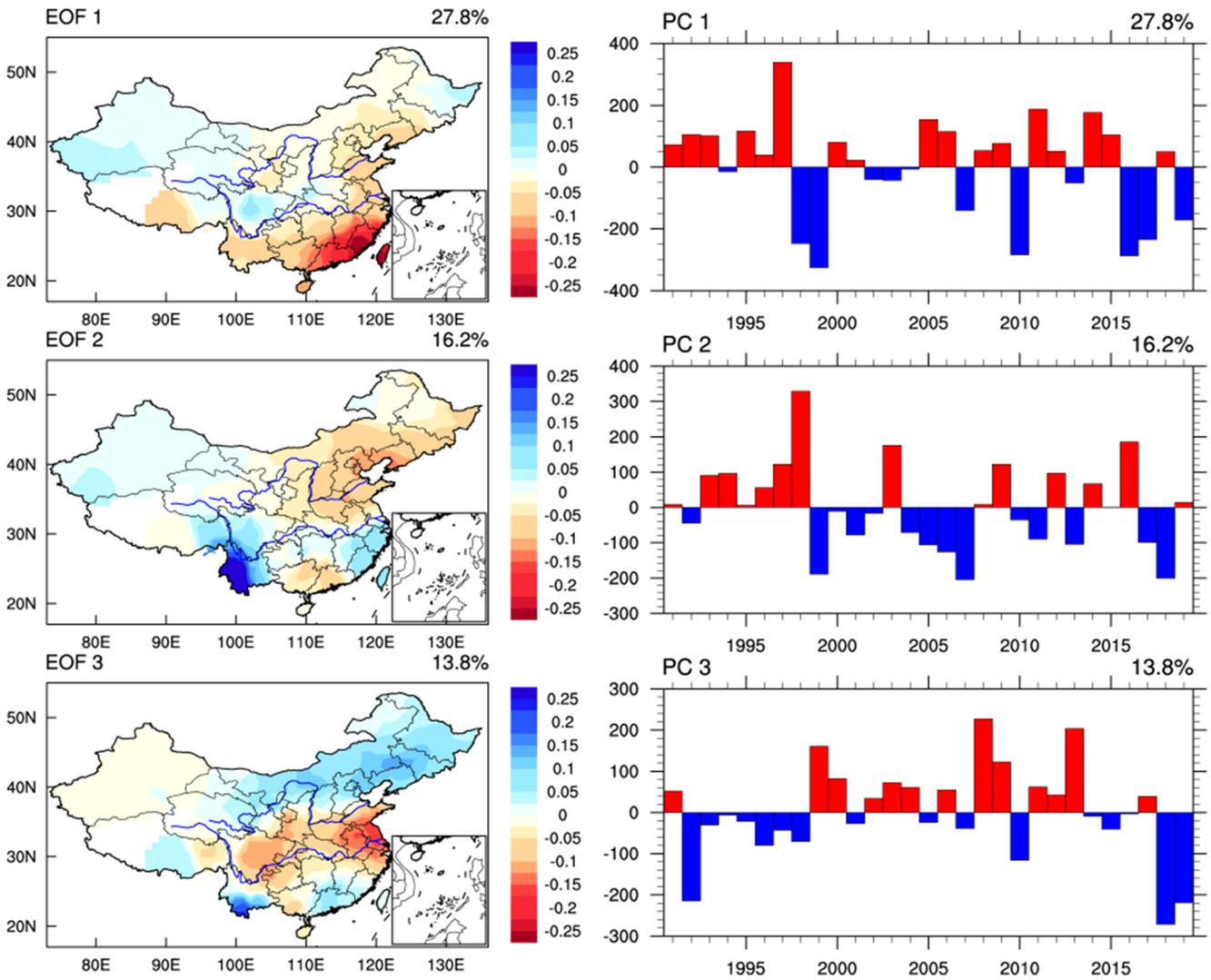


Fig. 5 Same as in Fig. 4 but for BCC_CSM1.1 modeled outputs

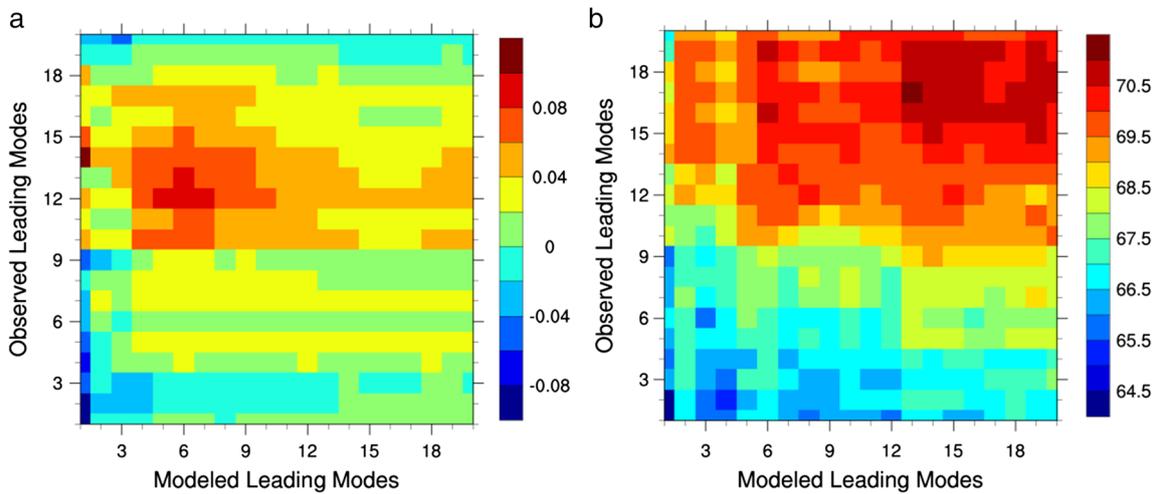
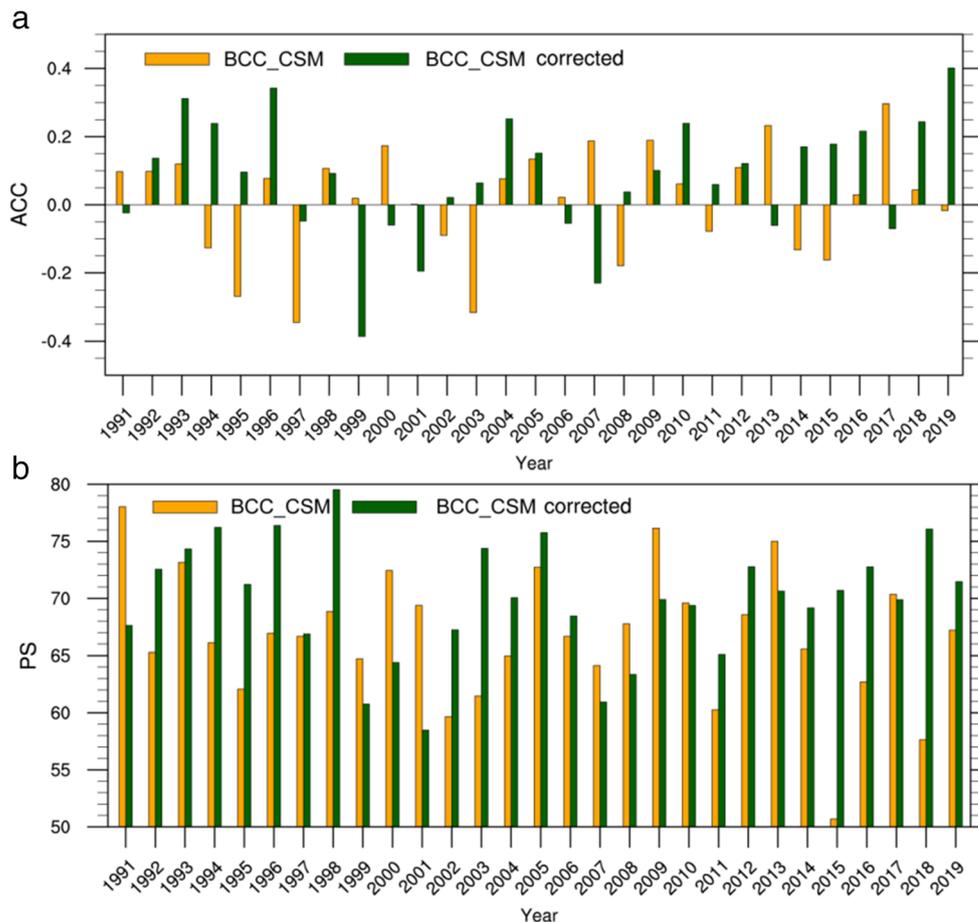


Fig. 6 The average ACC and PS scores between the cross validation results and observations in 1991–2019, where the abscissa is the number of retained modeled EOF modes and the ordinate is the number of retained observed EOF modes

Fig. 7 The corresponding ACC and PS scores between the observation and results before and after the cross validation when the observed leading modes take 12–13 and the modeled leading modes take 5–7 from 1991 to 2019. The average ACC between the original model projection and observation for the 1991–2019 period was 0.01, and the PS was 66.7. The average ACC after the cross validation was 0.08, and the PS was 69.9



2020, the first 12/13 EOF modes for the observed rainfall and the first 5–7 EOF modes for the modeled rainfall should be retained.

After the values of K and M were determined (K are 12–13, and M are 5–7), $EOF_j^o(x)$ and $\alpha_{i,j}$ can be obtained from Equations 1–3. By substituting these values into Equation 4, we can obtain six predictions (K = 12–13, M = 5–7). The final corrected prediction was the average of these six results (Fig. 8b). Figure 8 compares the spatial distribution of the summer precipitation anomaly percentage in 2020 before and after bias correction. The original prediction (Fig. 8a) showed that the positive summer precipitation anomalies were mainly located in northwest, east northeast, south China, and west southwest China, and dry areas were mainly located in central east southwest and northeast China. Compared with the observations, the negative-positive-positive patterns presented from South China to North China were not predicted. The dry anomalies in the southwest and northwest are also opposite to the observed wet anomalies. In general, the results from the original prediction by the BCC_CSM1.1 model differ greatly from the observations, with a PS score of 67.8 and an ACC score of -0.02.

As shown in Fig. 8b, the bias corrected predicted results showed that the areas with more precipitation during the flood season in 2020 were located in the Yangtze River’s middle reaches, the Yellow River Basin and in the north-east, and the areas with less precipitation were located in the southeast and Huaihe River region of China. Compared with the observation (Fig. 1b), although the corrected results did not capture the trend of more precipitation in the Huaihe River region of China in the summer of 2020, the results did capture the lower amounts of precipitation in south and northwest China and the greater amounts of precipitation in the Yangtze River’s middle reaches and the Yellow River Basin. The corrected result had a PS score of 77.4 and an ACC score of 0.30. Compared with the original model results (PS = 67.8, ACC = -0.02), the predicted accuracy significantly improved.

5 Evaluation of the bias correction method

To test the stability of this corrected method to improve the accuracy of predicted results, we use the above method to

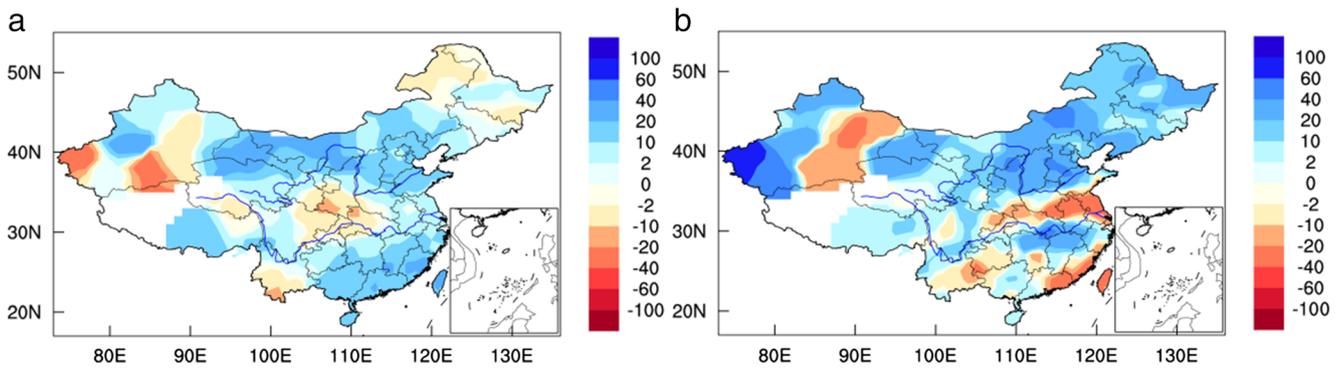


Fig. 8 The spatial distribution of the predicted summer precipitation anomaly percentage in 2020. **a** Prediction from the BCC_CSM1.1 original output. **b** Corrected prediction using the bias corrected method based on the outputs from BCC_CSM1.1. The ACC between the

corrected prediction and the observation is 0.30, and PS is 77.4, while the AAC and PS for the original model prediction from BCC_CSM1.1 are only -0.02 and 67.8, respectively

sequentially perform a correction on the 2014–2019 model predicted results. The results before the test year were used as historical years to establish multiple regression. For example, for the 2014 prediction correction, we used 1991–2013 as the historical period to establish the multiple regression, and so on in 2015–2019. Table 1 lists the number of the first N modes after the EOF decomposition in the observations and the model retained when the multivariate linear relationship is established for each year. Figure 9 provides a comparison of the predicted precipitation from the original output and after bias corrected output of the BCC_CSM1.1. The results after bias correction can well capture the wet and dry patterns of precipitation in the flood season in the eastern region. Specifically, the corrected results showed great skills for the observed less/more observed precipitation in Central/South China in 2014 and 2015. Similar improvements using the bias correction method applied to the BCC_CSM1.1 model projection can also be seen in 2018 and 2019. The results before the correction are basically opposite to the observations. In addition, the corrected results are also good for simulating wet and dry changes in precipitation during the flood season in the northwest, such as in 2015 and 2016. The bias correction method showed a

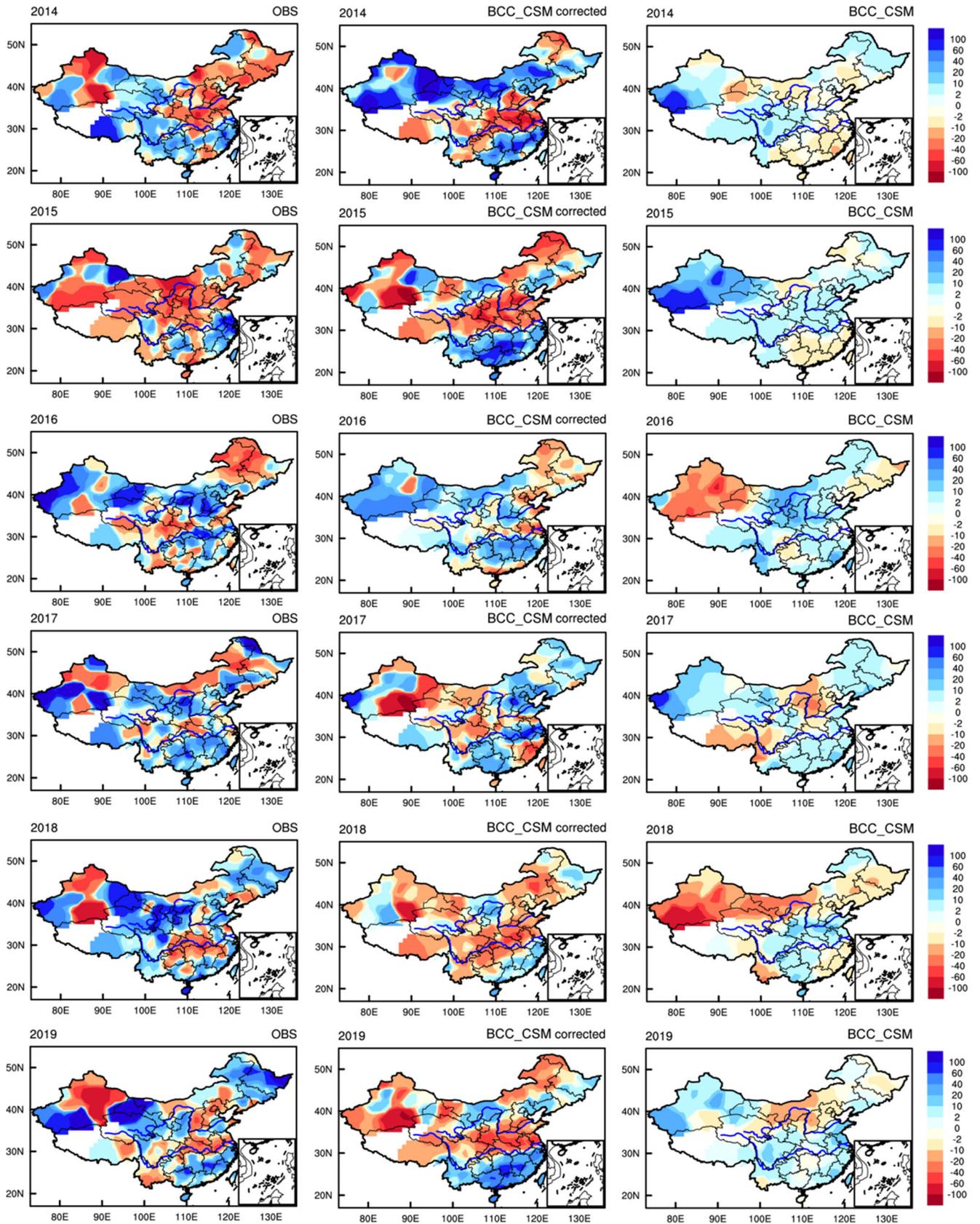
skillful and stable performance in reproducing summer rainfall anomalies over China.

The comparison of PS and ACC from the predicted results before and after correction during 2014–2019 is shown in Fig. 10. Except for 2017, the bias correction method can obviously improve the PS and ACC scores. For 2017, the ACC score after correction was 0.05 but was 0.29 before the correction, which was mainly because the corrected results are basically contrary to the observations in the southeast regions, the Yangtze River Basin and the Yellow River Basin (Fig. 9). The ACCs before and after the correction were 0.03 and 0.23 in 2016, respectively. For 2015 and 2016, the PS scores before the correction were 69.6 and 71.7, respectively, and all the PSs increased to more than 78 after the correction. For 2018 and 2019, the bias correction method shows great improvement of ACC scores, and the PS scores after the correction were increased to greater than 73. Overall, for 2014–2019, the bias correction method increased the ACC score of the predicted result from 0.01 to 0.25, and the PS score increased from 62.4 to 75.1. The bias correction method showed stable and considerable improvement for the prediction of summer rainfall in China.

Table 1 The number of selected observed leading modes and the modeled leading modes for bias correction in 2014–2019

Year	2014	2015	2016	2017	2018	2019
Observed	16-17	14	9-10	10-12	10-11	12-13
Modeled	15	9-11	6	6	6	6

Fig. 9 The summer rainfall anomaly percentage spatial patterns of the observations (left column), bias corrected results (middle column), and original model outputs (right column) in 2014–2019. The leftmost column is the spatial distribution of the observed anomalous percentage of summer precipitation. The middle column is the corrected BCC_CSM1.1 prediction result, and the rightmost column is the BCC_CSM1.1 original prediction result



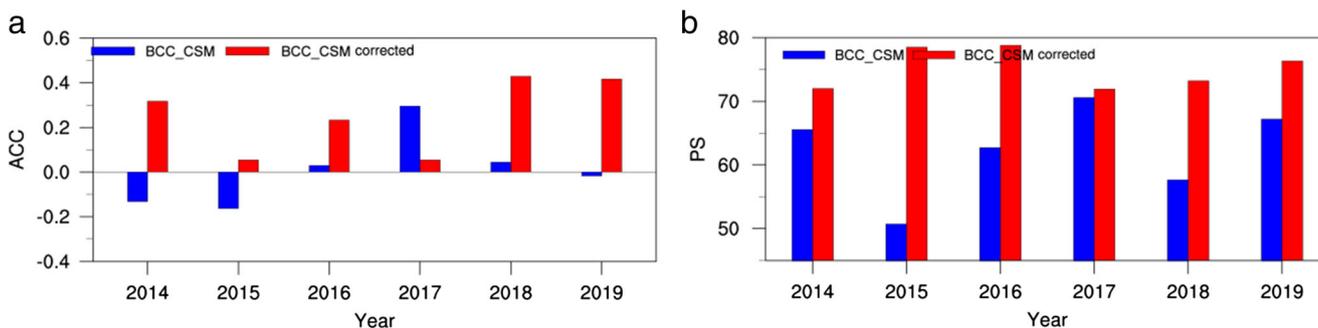


Fig. 10 ACC (a) and PS (b) scores between the observations and results from the BCC_CSM1.1 model and bias corrected model outputs in 2014–2019

6 Conclusions and discussions

In the summer of 2020, the precipitation distribution in China was in a state of less in the south and more in the central and northern areas. The precipitation anomaly percentages in the Yangtze River Basin and the Yellow River Basin were above 35%, and in the Huaihe Basin and the Yangtze River's lower reaches were 40% and 60% higher, respectively. Especially in some areas of southern Anhui Province, the summer rainfall was 100% more than the average during 1981–2010, which has put great pressure on local flood resistance. The stronger WPSH and the enhanced EASM promoted the water vapor convergence in the Yangtze River Basin and the Yellow River Basin, which caused more summer rainfall in these areas. In South China, the divergent water vapor led to less precipitation.

However, the original prediction of summer precipitation in 2020 by BCC_CSM1.1 did not capture the spatial distribution characteristics of summer precipitation in China. Therefore, in this study, a statistical downscaling method based on EOF for a given season was applied to correct the original model outputs for summer precipitation in China to improve the accuracy of summer precipitation prediction. The results show that although this bias correction method is relatively simple compared with dynamic downscaling, it has a certain ability to improve the model predicted results. For the summer of 2020, the main rain belt of the corrected summer precipitation was located in the Yangtze River's middle reaches and Yellow River Basin, which is in line with the observations. Relative to the original model output, the ACC between the corrected predictions and the observations increased from -0.02 to 0.30 and the PS increased from 67.8 to 77.4. Moreover, for 2014–2019, the bias correction method increased the ACC score of the prediction from 0.01 to 0.25 and the PS score increased from 62.4 to 75.1, which demonstrated that the bias correction method showed stable and considerable improvement for the prediction of summer precipitation in China.

However, this method lacks the analysis and utilization of precipitation mechanisms in China during the flood season. Many existing studies point out that there are many factors affecting rainfall in China during the flood season, and some examples include Pacific Decadal Oscillation (PDO), snow cover, El Nino-Southern Oscillation (ENSO), North Atlantic Oscillation (NAO), EASW, and WPSH (Chan and Zhou 2005; Ding et al. 2008; Feng et al. 2018; Gong and Ho 2002; Ma 2007; Si and Ding 2016; Wang et al. 2000b; Zhou and Yu 2005; Zhou et al. 2009; Zhu et al. 2011; Zhu et al. 2015). Many studies (Liu and Fan 2014; Chu and Yu 2010) used these factors (such as the pressure field, wind field and geopotential height field) as predictors for statistical downscaling and showed stable improvement in reproducing historical seasonal rainfall anomalies.

In the future, we look forward to incorporating these mechanisms into the correction methods. In addition, many other good predicted methods have slowly emerged in recent years, such as the year-to-year increment approach (Fan et al. 2009; Fan et al. 2012), artificial intelligence neural networks, and deep learning (Ham et al. 2019; Le et al. 2017; Reichstein et al. 2019; Shen et al. 2020; Voyant et al. 2017), which can also be integrated with the prediction corrected methods. We can also attempt to correct the results month-by-month from June to August, which has important implications for the prediction of precipitation in South China and the Yangtze River's middle and lower reaches during June and July. Considering that this statistical downscaling correction may have the best skill for correcting the trend of precipitation, and the skill of the correction of interdecadal and interannual changes needs to be improved. The ensemble empirical mode decomposition (EEMD) method can be applied to first separate the observations and model results and use this method to directly correct the trend item. For the interdecadal and interannual variation items, factors such as PDO and ENSO, which have the greatest impact on the interdecadal and interannual changes of precipitation in China (Chan and Zhou 2005; Gu et al. 2017; Shen et al. 2006; Wang et al. 2012; Yang et al. 2017; Yu 2013), could be combined with the correction to

obtain better predicted results. All the methods mentioned above are worth trying in the future but are beyond the main content of this article. Generally, the bias correction method used in this paper shows good improvement for the summer precipitation prediction in China, and more research and development are necessary in the future.

Acknowledgements The authors acknowledge the China National Climate Center for providing BCC_CSM1.1 prediction results and precipitation observation data of 160 stations and making them available for our analysis. The authors also acknowledge the NOAA/OAR/ESRL PSL, Boulder, Colorado, USA, for providing the NCEP reanalysis derived data (which can be downloaded from Web site at <https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.derived.pressure.html>).

Code availability Not applicable.

Author contribution H. Y. designed the study and contributed to the ideas, interpretation, and manuscript writing. Y. W. contributed to the data analysis, interpretation, and manuscript writing. J. H. contributed to the ideas and interpretation. All of the authors discussed and reviewed the manuscript.

Funding This work was jointly supported by the National key research and development program (2017YFC1502305), National Natural Science Foundation of China (42075043), and the Youth Innovation Promotion Association CAS.

Data availability The authors declare that the data supporting the findings of this study are available within the article.

Declarations

Ethics approval We confirm that this article is an original research and has not been published or presented previously in any journal or conference in any language (in whole or in part).

Consent to participate and consent for publication All the authors consented to participate and for publication.

Conflicts of interest The authors declare no competing financial interests.

References

- Alessandri A, Borrelli A, Masina S, Cherchi A, Gualdi S, Navarra A, di Pietro P, Carril AF (2010) The INGV–CMCC Seasonal Prediction System: Improved Ocean Initial Conditions. *Mon Weather Rev* 138: 2930–2952
- Benestad RE, Haugen JE (2007) On complex extremes: flood hazards and combined high spring-time precipitation and temperature in Norway. *Clim Chang* 85:381–406
- Chan JCL, Zhou W (2005) PDO, ENSO and the early summer monsoon rainfall over south China. *Geophys Res Lett* 32:L08810. <https://doi.org/10.1029/2004GL022015>
- Chen H, Sun J, Wang H (2012) A statistical downscaling model for forecasting summer rainfall in China from DEMETER Hindcast Datasets. *Weather Forecast* 27:608–628
- Chen T, Zhang F, Yu C et al (2020) Synoptic analysis of extreme Meiyu precipitation over Yangtze River Basin during June–July 2020. *Meteorological Monthly* 46:1515–1426 (in Chinese)
- Chu JL, Yu PS (2010) A study of the impact of climate change on local precipitation using statistical downscaling. *J Geophys Res Atmos* 115
- Chu JL, Kang H, Tam CY, Park CK, Chen CT (2008) Seasonal forecast for local precipitation over northern Taiwan using statistical downscaling. *J Geophys Res* 113:D12118
- Ding Y, Ren GY, Shi GY (2006) National Assessment Report of Climate Change (I): Climate change in China and its future trend. *Adv Clim Chang Res* 03:1–5
- Ding Y, Wang Z, Sun Y (2008) Inter-decadal variation of the summer precipitation in East China and its association with decreasing Asian summer monsoon. Part I: observed evidences. *Int J Climatol* 28: 1139–1161
- Fan K, Wang H, Choi YJ (2008) A physically-based statistical forecast model for the middle-lower reaches of the Yangtze River Valley summer rainfall. *Chin Sci Bull* 53:602–609
- Fan K, Lin M, Gao Y (2009) Forecasting the summer rainfall in North China using the year-to-year increment approach. *Sci China Ser D Earth Sci* 52:532–539
- Fan K, Liu Y, Chen H (2012) Improving the Prediction of the East Asian summer monsoon: New Approaches. *Weather Forecast* 27:1017–1030
- Fedderson H, Navarra A, Ward MN (1999) Reduction of model systematic error by statistical correction for dynamical seasonal predictions. *J Clim* 12:1974–1989
- Feng J, Chen AW, Tam CY, Zhou W (2011) Different impacts of El Nio and El Nio Modoki on China rainfall in the decaying phases. *Int J Climatol* 31:2091–2101
- Feng G, Zhao J, Zhi R, Gong Z, Zheng Z, Yang J, Xiong K (2013) Recent progress on the objective and quantifiable forecast of summer precipitation based on dynamical-statistical method. *J Appl Meteorol Sci* 24:656–665
- Feng G, Zou M, Qiao S, Zhi R, Gong Z (2018) The changing relationship between the December North Atlantic Oscillation and the following February East Asian trough before and after the late 1980s. *Clim Dyn* 51:1–14
- Gong DY, Ho CH (2002) Shift in the summer rainfall over the Yangtze River valley in the late 1970s. *Geophys Res Lett* 29:78–71
- Gong Z, Hutin C, Feng G (2016) Methods for improving the prediction skill of summer precipitation over East Asia–West Pacific. *Weather Forecast* 31:1381–1392
- Gong Z, Dogar MM, Qiao S, Hu P, Feng G (2018) Assessment and correction of BCC_CSM's performance in capturing leading modes of summer precipitation over North Asia. *Int J Climatol* 38:2201–2214
- Gu B, Zheng Z, Feng G, Wang X (2017) Interdecadal transition in the relationship between the western Pacific subtropical high and sea surface temperature. *Int J Climatol* 37:2667–2678
- Guo Y, Li JP, Li Y (2011) Statistically downscaled summer rainfall over the middle-lower reaches of the yangtze River. *Atmospheric and Oceanic Science Letters* 4:191–198
- Ham YG, Kim JH, Luo J (2019) Deep learning for multi-year ENSO forecasts. *Nature* 573:568–572
- Hartmann C, Tank A, Alexander M, Brönnimann S, Viterbo P (2013) Climate Change 2013: The Physical Science Basis, Working Group I Contribution to the Fifth Assessment Report of the

- Intergovernmental Panel on Climate Change, pp 159–254. IPCC, Cambridge Univ. Press, Cambridge
- Hu ZZ (1997) Interdecadal variability of summer climate over East Asia and its association with 500 hPa height and global sea surface temperature. *J Geophys Res Atmos* 102:19403–19412
- Huang RH, Cai RS, Chen JL, Zhou LT (2006) Interdecadal variations of drought and flooding disasters in China and their association with the East Asian climate system. *Chin J Atmos Sci* 30:730–743
- Huang J, Yu H, Guan X, Wang G, Guo R (2016) Accelerated dryland expansion under climate change. *Nat Clim Chang* 6:166–171
- Kalnay E et al (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 74:789–799
- Kang IS et al (2002) Intercomparison of the climatological variations of Asian summer monsoon precipitation simulated by 10 GCMs. *Clim Dyn* 19:383–395
- Kang IS, Lee JY, Park CK (2004) Potential predictability of summer mean precipitation in a dynamical seasonal prediction system with systematic error correction. *J Clim* 17:834–844
- Kug JS, Lee JY, Kang IS (2008) Systematic error correction of dynamical seasonal prediction of sea surface temperature using a stepwise pattern project method. *Mon Weather Rev* 136:3501–3512
- Le JA, El-Askary HM, Allali M, Struppa DC (2017) Application of recurrent neural networks for drought projections in California. *Atmos Res* 188:100–106
- Liu Y, Ding Y (2020) Characteristics and possible causes for the extreme Meiyu in 2020. *Meteorological Monthly* 46:1393–1404 (in Chinese)
- Liu Y, Fan K (2014) An application of hybrid downscaling model to forecast summer precipitation at stations in China. *Atmos Res* 143:17–30
- Liu B, Yan Y, Zhu C, Ma S, Li J (2020) Record-breaking Meiyu rainfall around the Yangtze River in 2020 regulated by the subseasonal phase transition of the North Atlantic oscillation. *Geophys Res Lett* 47:e2020GL090342. <https://doi.org/10.1029/2020GL090342>
- Liu Y, Wang Y, Ke Z (2021) Characteristics and possible causes for the climate anomalies over China in summer 2020. *Meteorological Monthly* 47:117–126 (in Chinese)
- Ma Z (2007) The interdecadal trend and shift of dry/wet over the central part of North China and their relationship to the Pacific Decadal Oscillation (PDO). *Chin Sci Bull* 15:116–125
- Michaelsen J (1987) Cross-validation in statistical climate forecast models. *J Appl Meteorol Climatol* 26:1589–1600
- Min YM, Kryjov VN, Oh SM (2014) Assessment of APCC multimodel ensemble prediction in seasonal climate forecasting: retrospective (1983–2003) and real-time forecasts (2008–2013). *J Geophys Res Atmos* 119:12,132–112,150
- Newman M et al (2012) The Pacific decadal oscillation, Revisited. *J Clim* 29:4399–4427
- Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais NP (2019) Deep learning and process understanding for data-driven Earth system science. *Nature* 566:195–204
- Rowell DP (1996) Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations. *J Clim* 11:109–120
- Shen C, Wang WC, Gong W, Hao Z (2006) A Pacific Decadal Oscillation record since 1470 AD reconstructed from proxy data of summer rainfall over eastern China. *Geophys Res Lett* 33:L03702
- Shen H, Luo Y, Zhao Z, Wang H (2020) Prediction of summer precipitation in China based on LSTM network. *Climate Change Research* 16:263–275 (in Chinese)
- Si D, Ding Y (2016) Oceanic forcings of the interdecadal variability in East Asian summer rainfall. *J Clim* 29:7633–7649
- Sun J, Chen H (2012) A statistical downscaling scheme to improve global precipitation forecasting. *Meteorol Atmos Phys* 117:87–102
- Sun J, Wang H, Yuan W (2009) A possible mechanism for the covariability of the boreal spring Antarctic Oscillation and the Yangtze River valley summer rainfall. *Int J Climatol* 29:1276–1284
- Tippett MK, Barlow M, Lyon B (2010) Statistical correction of central southwest Asia winter precipitation simulations central southwest Asia precipitation; Afghanistan; East Asia jet stream; CCA. *Int J Climatol* 23:1421–1433
- Uvo CB, Olsson J, Morita O, Jinno K, Nakashima T (2001) Statistical atmospheric downscaling for rainfall estimation in Kyushu Island, Japan. *Hydrol Earth Syst Sci* 5:259–271
- Voyant C, Notton G, Kalogirou S, Nivet ML, Paoli C, Motte F, Fouilloy A (2017) Machine learning methods for solar radiation forecasting: a review. *Renew Energy* 105:569–582
- Wang J (2020) Relationships between Jianghuai Meiyu Anomaly and the Collaborative Evolution of Wave Trains in the Upper and Lower Troposphere in Mid-July of 2020. *Front Earth Sci* 8:597930. <https://doi.org/10.3389/feart.2020.597930>
- Wang H, Fan K (2009) A new scheme for improving the seasonal prediction of summer precipitation anomalies. *Weather Forecast* 24:548–554
- Wang H, Zhou G, Zhao Y (2000a) An effective method for correcting the seasonal-interannual prediction of summer climate anomaly. *Adv Atmos Sci* 17:234–240
- Wang B, Wu R, Fu X (2000b) Pacific–East Asian teleconnection: how does ENSO affect East Asian climate? *J Clim* 13:1517–1536
- Wang B, Lee JY, Kang IS, Shukla J, Park CK, Kumar A, Schemm J, Cocke S, Kug JS, Luo JJ, Zhou T, Wang B, Fu X, Yun WT, Alves O, Jin EK, Kinter J, Kirtman B, Krishnamurti T, Lau NC, Lau W, Liu P, Pegion P, Rosati T, Schubert S, Stern W, Suarez M, Yamagata T (2009) Advance and prospectus of seasonal prediction: assessment of the APCC/CLIPAS 14-model ensemble retrospective seasonal prediction (1980–2004). *Clim Dyn* 33:93–117
- Wang H, Zhang Y, Lang X (2010) On the predictand of short-term climate prediction. *Climatic and Environmental Research* 15:225–228 (in Chinese)
- Wang H, Kumar A, Wang W, Xue Y (2012) Seasonality of the Pacific decadal oscillation. *J Clim* 25:25–38
- Wei FY, Huang JY (2010) A study of predictability for summer precipitation on east china using downscaling techniques. *Journal of Tropical Meteorology*.
- Wei Y, Yu H, Huang J, He Y, Yang B, Guan X, Liu X (2017) Comparison of the Pacific Decadal Oscillation in climate model simulations and observations. *Int J Climatol* 38:e99–e118
- Wei K, Ouyang C, Duan H, Li Y, Chen M, Ma J, An H, Zhou S (2020) Reflections on the Catastrophic 2020 Yangtze River Basin Flooding in Southern China. *The Innovation* 1:100038
- Widmann M, Bretherton CS, Salathe EP (2003) Statistical precipitation downscaling over the northwestern United States using numerically simulated precipitation as predictor. *J Clim* 16:799–816
- World Meteorological Organization (WMO) (2020) The global climate in 2015–2019.
- Wu B, Zhang R, Wang B, D'Arrigo R (2009) On the association between spring Arctic sea ice concentration and Chinese summer rainfall. *Geophys Res Lett* 36:L09501
- Yang S, Xu L (1994) Linkage between Eurasian winter snow cover and regional Chinese summer rainfall. *Int J Climatol* 14:739–750
- Yang Q, Ma Z, Fan X, Yang ZL, Xu Z, Wu P (2017) Decadal modulation of precipitation patterns over East China by sea surface temperature anomalies. *J Clim* 16 JCLI-D-16-0793.0791

- Yu L (2013) Potential correlation between the decadal East Asian summer monsoon variability and the Pacific decadal oscillation. *Atmospheric and Oceanic Science Letters* 05:394–397
- Zhao J, Yang J, Gong Z, Feng G (2015) Analysis of and discussion about dynamic-statistical climate prediction for summer rainfall of 2013 in China. *Advances in Meteorological Science Technology* 5:24–28 (in Chinese)
- Zheng Z, Ren H, Huang J (2009) Analogue correction of errors based on seasonal climatic predictable components and numerical experiments. *Acta Phys Sin* 58:7359–7367
- Zheng Z, Huang J, Feng G, Chou J (2013) Forecast scheme and strategy for extended-range predictable components. *Sci China Earth Sci* 56: 878–889
- Zhou T, Yu R (2005) Atmospheric water vapor transport associated with typical anomalous summer rainfall patterns in China. *J Geophys Res* 110:D08104
- Zhou T, Gong D, Li J, Li B (2009) Detecting and understanding the multi-decadal variability of the East Asian Summer Monsoon – recent progress and state of affairs. *Meteorologische Zeitschrift* 18: 455–467
- Zhu Y, Wang H, Zhou W, Ma J (2011) Recent changes in the summer precipitation pattern in East China and the background circulation. *Clim Dyn* 36:1463–1473
- Zhu Y, Wang H, Ma J, Wang T, Sun J (2015) Contribution of the phase transition of Pacific Decadal Oscillation to the late 1990s' shift in East China summer rainfall. *J Geophys Res Atmos* 120:8817–8827

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