LETTER • OPEN ACCESS

Historical and future runoff changes in the Yangtze River Basin from CMIP6 models constrained by a weighting strategy

To cite this article: Jiazhen Zhao *et al* 2022 *Environ. Res. Lett.* **17** 024015

View the article online for updates and enhancements.

You may also like

ENVIRONMENTAL RESEARCH
LETTERS

LETTER

Historical and future runoff changes in the Yangtze River Basin from CMIP6 models constrained by a weighting strategy

Jiazhen Zhao[1], Shengping He[2,3] and Huijun Wang[1,3,4,*]

[1] Key Laboratory of Meteorological Disaster, Ministry of Education/Joint International Research Laboratory of Climate and Environment Change (ILCEC)/Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science & Technology, Nanjing 210044, People's Republic of China
[2] Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research, Bergen, Norway
[3] Nansen-Zhu International Research Center, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, People's Republic of China
[4] Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, People's Republic of China
* Author to whom any correspondence should be addressed.

E-mail: hjwang@nuist.edu.cn

Abstract

Based on the ERA5-Land datasets from 1981–2020, a decadal oscillation has been found in the variation of summer runoff in the middle and lower reaches of the Yangtze River Basin (MLYRB). The oscillation suggests that the MLYRB will experience increased runoff in the next few decades after 2020, which saw a record high runoff in the MLYRB. The decadal changes in summer runoff over the MLYRB under various climate change scenarios are then analyzed with direct runoff outputs from 28 general circulation models participating in the sixth phase of the Coupled Model Intercomparison Project. Given that the equal-weighted multi-model ensemble mean could not well represent the historical runoff changes in the MLYRB, in this paper we introduce a model weighting scheme that considers both the model skill and independence. It turns out that this scheme well constrains the models to represent the observed decadal changes of summer runoff. The weighted mean projections suggest that the summer runoff in the MLYRB during 2015–2100 under all warming scenarios will be higher than the present day; and 2021–2040 is likely to be a period with significantly increased summer runoff. Results of the present study have great implications for flood control and effective water resources management over the MLYRB in the future, and the weighting approach used in this paper can be applied to a wide range of projections at both regional and global scales.

1. Introduction

Annual-mean globally averaged surface air temperature has increased at an average rate of 0.08 °C per decade since 1880 (0.18 °C per decade since 1981) (NOAA National Centers for Environmental Information 2021), which has enormous impacts on the global water cycle (Sohoulande *et al* 2016). As one of the key elements of the hydrologic cycle, runoff has changed significantly over the past years (Nam *et al* 2018). Since runoff impacts people, agriculture, industries and ecosystems, an enhanced understanding of the spatiotemporal variations in runoff at a river basin and their response to climate change is of

great significance for effective flood prevention, ecological environment protection and policy formulation (Xiao *et al* 2018, Xing *et al* 2018).

The Yangtze River Basin (YRB), the largest river basin in China and the 12th largest river basin in the world, plays a crucial role in regional water cycle, economic and social development (Tao *et al* 2021). The middle and lower reaches of the YRB (MLYRB) have the most abundant water source in China, with the highest nationwide density and number of lakes (Li *et al* 2019). There has been considerable research on spatiotemporal variation of runoff in the YRB/MLYRB based on short-term observations and simulations by various land surface models (LSMs)

or global hydrological models (GHMs) (Chen *et al* 2014, Tang *et al* 2018, Xiao *et al* 2018, Gao *et al* 2020). However, previous studies mostly focused on the seasonal, interannual variability or long-term trend of the runoff in the MLYRB, while less attention has been paid to its interdecadal variability. 2020 summer saw the worst flooding disasters in the MLYRB since 1998 (another recorded severe flood occurred in 1983), which severely impacted the lives and activities of the local people (Wang *et al* 2021). This implies an interdecadal variation in the runoff at the MLYRB. To develop adaption strategies for climate risk reduction, it is necessary to analyze the historical decadal changes of summer runoff in the MLYRB, to reveal its linkage to atmospheric circulation and to project future runoff changes in this region.

Due to the short observed runoff data, model outputs are indispensable for runoff projection. To obtain long-term runoff data, previous studies have commonly adopted the LSMs/GHMs driven by bias-corrected precipitation data, which is estimated by general circulation models (GCMs) (Koirala *et al* 2014, Schewe *et al* 2014, Nam *et al* 2019). However, obtaining simulated runoff data in accordance with the above procedure will face great challenges. For example, due to our limited knowledge of the highly complicated climate system, large uncertainties exist in the precipitation predictions from GCMs (Kim *et al* 2008, Hawkins and Sutton 2011), which contribute to the major source of uncertainties in runoff simulations (Wilby *et al* 2008). And the runoff simulated by LSMs/GHMs might be heavily dependent on the model used, the bias correction and the downscaling method applied to the model input data (Hagemann *et al* 2013, Schewe *et al* 2014). Therefore, it is meaningful to use as many models as possible to explore the future runoff changes in the study region. In recent years, with a rapid increase in computational power and the great development of model structures and spatial resolutions, GCMs have been extensively applied to climate projection (Kitoh *et al* 2009, Mizuta *et al* 2012). Currently, scientists commonly utilize multi-model ensembles to provide a range of plausible outcomes that a region may experience (Merrifield *et al* 2020). The ensemble projection allows researchers to analyze the consistencies and discrepancies of multiple models in the ensemble to evaluate the overall level or the individual model's ability in projecting long-term climate change.

As the most comprehensive collection of climate simulations produced to date, the Coupled Model Intercomparison Project Phase 6 (CMIP6) archive (Eyring *et al* 2016) contains the latest outputs of more than 30 GCMs developed at different institutes around the world (Jiang *et al* 2020). Compared with the CMIP5, the CMIP6 has designed new scenarios, i.e., the shared socioeconomic pathway (SSP) 1–2.6,

SSP2–4.5, SSP3–7.0, and SSP5–8.5, whose new framework filled in the critical gaps of intermediate forcing levels in CMIP5 (Eyring *et al* 2016, O'Neill *et al* 2016). It is therefore necessary to use CMIP6 GCMs to project runoff changes in the MLYRB under these new scenarios. Given the existing large model systematic errors, biases need to be constrained within the ensemble runoff projection. However, in most of the previous studies, each simulation/projection is taken to be equally important and the equal-weighted ensemble mean becomes a common way to reduce the uncertainties caused by internal variability and models' biases. However, the equal-weighted method has a fundamental drawback because an individual model has distinct performance and skill (Chen *et al* 2020, Jiang *et al* 2020, Zhou *et al* 2020). Some studies specifically selected those models that have better performance to calculate the equal-weighted mean. However, such a method has not considered the model independence (Lee and Wang 2014, Wang *et al* 2014). Additionally, some studies have suggested that both the model skill and the inter-dependence of multi-model archives are essential to constrain the uncertainties in climate simulations and projections (Knutti *et al* 2013, Sanderson *et al* 2015, 2017). Specifically, given a set of models, it is important to evaluate the value of adding another model to this set for projection. Taking an extreme case, for example, adding a duplicated model to the original set would not provide any new information; rather, it would even bias the combination of existing models' results toward the results of the duplicated model (Caldwell *et al* 2014). In the CMIP6 archive, some models are highly similar to each other because they are from the same institution or they share a large number of codes. In that case, these simulations and projections cannot be seen as independent samples of runoff changes. Therefore, it is essential to consider both the model skill and the model independence to effectively constrain the model biases and provide a more reliable projection of runoff in the MLYRB.

Therefore, in this study, we will first use the observational/reanalysis data to investigate the interdecadal variation of the summer (June, July and August) runoff in the MLYRB and reveal its linkage to large-scale atmospheric circulation. The results of this part are described briefly in section 3.1 and more details can be found in supplementary information (available online at stacks.iop.org/ERL/17/024015/mmedia). Subsequently, the CMIP6 archive will be processed by a model weighting scheme utilized by Sanderson *et al* (2017) to constrain the uncertainties in the historical simulation and projection of summer runoff in the MLYRB. The model weighting scheme considers both the model skill and the model independence, which can reduce respectively the influence of relatively poor-performing models and the influence of

duplicated models as much as possible. It is developed based on the assumption that if a model has skills as good as other models but has higher independence, the model should have greater potential to provide new and more useful information. Finally, based on the multi-model weighted mean projections, we provided a comprehensive picture of future runoff changes under different climate change scenarios. The results of this study could provide insights into the effects of climate change on hydrological processes and provide a valuable reference for future flood prevention in the MLYRB.

2. Data and methods

2.1. Observations and CMIP6 archive

In this study, we use 1981–2014 as the baseline period to compare the observations with CMIP6 historical simulations. To investigate the atmospheric factors leading to the decadal changes of summertime runoff in the MLYRB, we have used the monthly gridded data from the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5) including the atmospheric fields, runoff, precipitation, evaporation with a spatial resolution of $1° \times 1°$ (Hersbach *et al* 2020; https://cds.climate.copernicus.eu/). Hourly and daily gridded precipitation from the ERA5-Land dataset and outputs of 28 GCMs from the CMIP6 (https://esgf-node.llnl.gov/search/cmip6/) are used to calculate the extreme precipitation index following the ETCCDI protocols (Sillmann *et al* 2013; http://etccdi.pacificclimate.org/) which is one of the essential objects of the evaluation to confirm model weighting scheme (supplementary information). The confirmed weighting scheme will be used to obtain the bias-constrained summer-mean runoff from the GCMs of CMIP6 in historical simulations and future projections. Supplementary table 1 lists the information on the models used in this study. To be consistent with the reanalysis data, the model outputs are re-gridded to $1° \times 1°$ through the bilinear interpolation method. Five experiments are investigated, including the historical simulation and the projections under different scenarios of SSPs (i.e. SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5) (Eyring *et al* 2016, O'Neill *et al* 2016). Only the first available realization of each model is selected for analysis.

2.2. Statistical methods

In addition to composite and correlation analyses (Zhu *et al* 2013, Huang *et al* 2019, Zhang *et al* 2020), empirical orthogonal function (EOF) analysis (Chen *et al* 2015), the Lanczos low-pass filter method (Duchon 1979) and the developed moving *t*-test technique (Xiao and Li 2007) are used in this study. The linear trends in all the reanalysis datasets are removed to eliminate the global warming effects and better analyze the interdecadal variability. The correlation

significance is evaluated using the two-tailed Student's *t*-test.

2.3. Model weighting scheme

Instead of regarding each model as equally important in future projections like the previous studies noted in section 1, we introduce a model weighting scheme (Sanderson *et al* 2017) to reduce uncertainties in projections of future runoff. In this scheme, model weights are determined by evaluating models' performances in simulating the historical changes of two seasonal-mean variables (summer-mean runoff (mrro) and precipitation (pr)) and one extreme precipitation index (summer maximum consecutive 5 day precipitation (rx5day)) (Sillmann *et al* 2013; http://etccdi.pacificclimate.org/) during the baseline period. The scheme will provide each model with three sets of weights: skill weight, independence weight and overall weight. The skill weight and independence weight reflect model skill and model independence, respectively. With comprehensive consideration of the above two model characteristics, we calculated the model's overall weight as a combination of the skill weight and independence weight (equations (7) and (8) in the supplementary information). The sum of all models' overall weights is equal to one and only the model's overall weight will be eventually used to multi-model weighted mean runoff projection. Taking a model for an example, its skill weight can be obtained by considering its capability of reproducing of the observed variation of the above evaluation objects (mrro, pr and rx5day). It is determined through evaluating the root mean square error (RMSE) distance between this model and the observation (equation (6) in the supplementary information). A smaller RMSE represents better model performance and corresponds to a larger model skill weight. For a given model, its independence weight can be obtained by considering its uniqueness, which reflects whether the model's simulation/projection is representative and its ability to provide new and useful information. It is determined by the RMSEs between this model and all the other models in the CMIP6 archive (equations (3)–(5) in the supplementary information). Larger RMSEs correspond to larger model independence weight. RMSE distance is calculated using the following procedure (equation (1) in the supplementary information): given a model and an evaluation object, for each year we computed the area mean and summertime value of this object simulated by this model, so for *n* years, we can obtain a time series containing *n* values and the RMSE distance is calculated based on this series. Here we take the MLYRB region as a whole and just focus on its temporal variation, but not the internal spatial variation considered by Sanderson *et al* (2017). More details on the weighting scheme are described in the supplementary information.
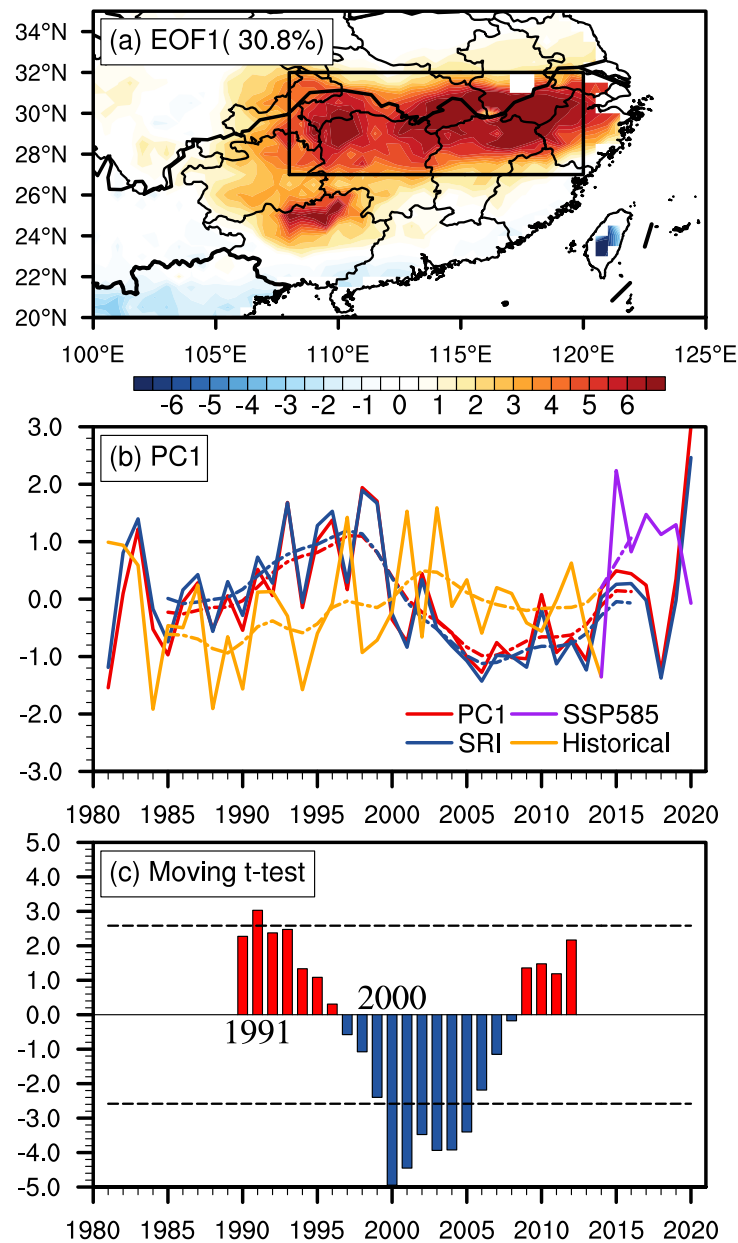
Figure 1. The first mode of EOF (a) and its corresponding standardized PC1 (red solid curve) (b) of the summertime runoff in southern China from 1981–2020. The black box in panel (a) indicates the domain of the MLYRB. The standardized area-averaged runoff over the MLYRB in the observation (blue solid curve) and the ensemble mean of 28 GCMs (orange and purple solid curve) are also presented in panel (b). The simulation consists of a combination of historical simulations in 1981–2014 and projection under SSP5-8.5 in 2015–2020. The dashed curves in panel (b) represent the nine year low-pass filtered series. Panel (c) represents the results calculated by the developed moving *t*-test technique (filled bars). The black dotted lines in (c) represents significance at the 0.05 level. 1991 and 2000 are the shift years.

3. Results

3.1. Interdecadal changes of runoff in the reanalysis data

Figures 1(a) and (b) show the observed first leading EOF mode and the responding principal component (PC1) for summer runoff over southern China. In this study, we refer the domain of (27°–32°N, 108°–120°E) as the MLYRB region, where summer runoff has the dominant interdecadal variability (figure 1(a)). The PC1 has well captured the abnormal hydrology events over the MLYRB, such as the extreme floods in 1983, 1998 and 2020 (solid

red curve in figure 1(b)). Additionally, the nine year low-pass filtered series of the PC1 (red dashed curve in figure 1(b)) in 1981–2020 exhibits an obvious interdecadal oscillation. To further interpret the results of the EOF analysis, we compare the PC1 with the standardized area-mean summertime runoff over the MLYRB region (SRI, blue curves in figure 1(b)). Obviously, the changes of SRI match the PC1 well at both the interannual and decadal time scales, with the correlation coefficient of 0.97 and 0.98, respectively. Given this consistency, we adopt the area-averaged summer runoff over the MLYRB to the following model assessment and projections. The result of the

moving *t*-test technique shown in figure 1(c), which is used for shift year detection, shows 1991 and 2000 are the shift years. Considering both the shift years and the nine year low-pass filtered series of the PC1 in figure 1(b), we identify the years from 1991–1999 as the high runoff period and the years from 2003–2013 as the low runoff period.

Previous studies have shown that the changes in annual runoff in the MLYRB are affected by many meteorological factors including precipitation and temperature (Xiao *et al* 2018). In summer, according to the composites during the two periods (supplementary figure 1), precipitation is the dominant factor that directly leads to the interdecadal oscillation of summertime runoff in the MLYRB. Since precipitation in the MLYRB can be significantly influenced by weather patterns (Xiao *et al* 2015, Su *et al* 2017, Li *et al* 2021), we have found a close linkage between runoff changes in this region and the atmospheric circulation (supplementary figure 2). We have confirmed the significant contributions of the anomalous anticyclone in the western North Pacific and the strengthening of zonal westerly in the upper troposphere to the observed runoff changes in the MLYRB through their impacts on the East Asian summer monsoon. It has been discussed in supplementary information in detail. The result agrees with that of Tang *et al* (2018) and it indicates that climate change plays an important role in the interdecadal variation of summer runoff in the MLYRB. On this basis, we attempt to further investigate the long-term decadal variation characteristics of summer runoff under climate change with 28 global climate models participating in CMIP6.

3.2. Ensemble-mean of historical simulations by CMIP6 models

In this section, utilizing runoff from 28 GCMs in the CMIP6 archive, we aim to objectively present the overall performance of these models in simulating historical summer runoff in the MLYRB during the baseline period. First of all, the multi-model mean (MME) has also shown decadal variation in the runoff of the MLYRB, however, the simulated evolution (orange and purple curves in figure 1(b)) is much different from the observation. Even though the climate models have well represented the human-induced effects on climate and the long-term increase of precipitation extremes in the MLYRB (figures not shown), the MME of the 28 CMIP6 GCMs have not well reproduced the observed climatology and trend of summer runoff in the MLYRB due to large model systematic errors (supplementary figure 3 and supplementary information). Note that the simulated runoff trend is important for future projections in worldwide climate impact studies, and the MME is obviously not appropriate for an impact study of future climate change.
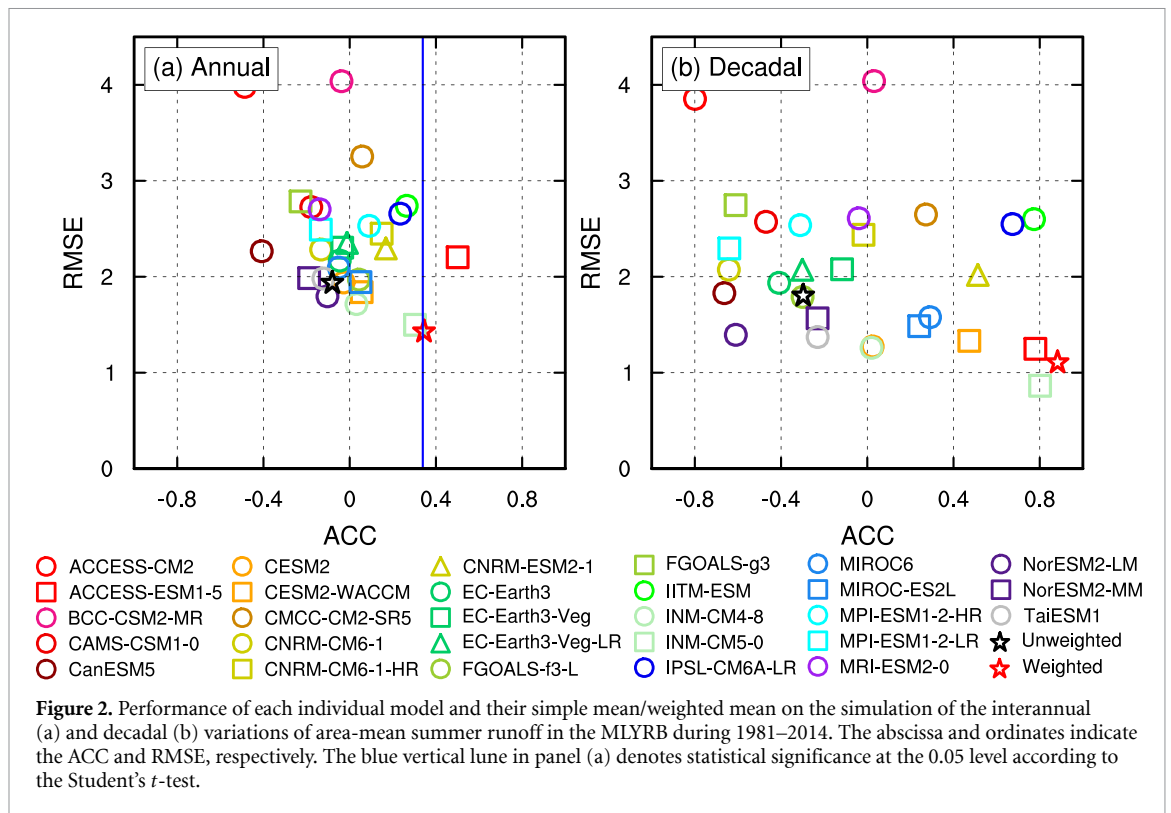
To quantitatively evaluate the models' performance in simulating the temporal variation of summer runoff in the MLYRB, we calculate anomaly correlation coefficient (ACC) and RMSE between the observed area-averaged runoff and simulations in individual GCMs and the MME at both interannual and decadal time scales (figure 2). It is clear that the CMIP6 models have diversities in reproducing temporal evolution of runoff changes. At the interannual time scale (figure 2(a)), (a) the ACCs in individual models range from −0.49 to 0.5 and most are centered around zero, only one model (ACCESS-ESM1-5) has well simulated the observed runoff changes (ACC passes the 95% significance level); and (b) the RMSEs range from 1.5–3.98. At decadal time scale (figure 2(b)), the simulations also have large biases from the observation with RMSEs ranging from 0.86–3.85 and ACCs ranging from −0.8 to 0.8. Because of the large model spread at both scales, the MME, a skill-weighted ensemble to some degree (Sanderson *et al* 2017), cannot well represent the interannual and decadal changes of summer runoff in the MLYRB (figures 1(b) and 2). The MME's ACC is −0.08 (−0.3) and the MME's RMSE is 1.95 (1.81) at the interannual (decadal) time scale over the MLYRB region. This means that the MME's projection may be potentially biased and cannot be used directly.

3.3. Constraining multi-model biases through a weighting strategy

In this section, we attempt to implement the weighting strategy (section 2.3) to constrain the model biases. Our goal is to obtain a weighted runoff projection that is more reliable and representative than that from the commonly used MME. The single set of weights provided in this section can be applied to a wide range of runoff projections.

3.3.1. Inter-model distance matrix

As the basis for the model weighting strategy, the multivariate inter-model distance matrix δ is illustrated in figure 3(a). Since the magnitude of the RMSE distance may be sensitive to the object being evaluated, for each object each model's distances were divided by the mean distance across all models to ensure the mean inter-model distance is equal to one and each object gets about the same weight, like Sanderson *et al* (2017) or Knutti *et al* (2017). Compared with the mean pairwise inter-model distance, it can be found that some individual models (ACCESS-CM2, ACCESS-ESM1-5) have larger distances than any other model in the archive. This implies that these models are likely to be unique in future projections. However, in some previous studies, some of these unique models were abandoned only because their performance for the simulation of the present climate is poor. In addition, in several subsets of the

Figure 2. Performance of each individual model and their simple mean/weighted mean on the simulation of the interannual (a) and decadal (b) variations of area-mean summer runoff in the MLYRB during 1981–2014. The abscissa and ordinates indicate the ACC and RMSE, respectively. The blue vertical lune in panel (a) denotes statistical significance at the 0.05 level according to the Student's *t*-test.

archive (Institute of Numerical Mathematics Climate Model (INMCM) variants, Norwegian Earth System Model version 2 (NorESM2) variants), the models in a single subset may have smaller distances from each other, whereas they all have larger distances from other models outside the subset. In this situation, using the traditional method used in the previous studies (e.g. MME) to project runoff can lead to poor accuracy and poor estimation of uncertainty because well-performed models or poor-performed models together with their duplicates all are used, which would bias the ensemble projection toward to the results of the duplicated models. These two kinds of problems mentioned above are closely related to the model independence, which can be solved effectively in our weighting scheme through calculating the RMSE distances across the models in the archive to determine the independence weight of each model. The results of the independence weight scheme will be discussed in the following section.

3.3.2. Independence weight
The key point for estimating the independence weight is the determination of the appropriate value of the radius of similarity D_u (equation (3) in the supplementary information). It is a free parameter that determines the distance scale over which the model's independence weight should be reduced. D_u is sampled by considering the distribution of inter-model distances. Here we determine its value based on several typical individual models or subsets in the CMIP6 archive mentioned in section 3.3.1. Figure 3(b) shows the relationship between D_u

and their independence weights. As mentioned in section 3.3.1, we expect that those models that have no obvious duplicates (i.e. ACCESS-CM2 and ACCESS-ESM1-5), should not be down-weighted by the method. In addition, for some subsets in which there are several closely related variants submitted from the same institution (i.e. INMCM and NorESM2), it is appropriate to select a value of D_u that can produce a weight of about $1/n$ (n is the number of variants submitted) for each variant (Sanderson *et al* 2017). Finally, we set D_u to be 0.67 times the distance between the best-performing model and observation, and then, the independence weight of each model in the archive can be estimated (supplementary figure 6).

3.3.3. Skill weight
According to the RMSEs between each model and observations, the model skill weight can be determined. A larger RMSE corresponds to a smaller skill weight, and vice versa. Similar to independence weight, the radius of model quality D_q needs to be determined appropriately (equation (6) in the supplementary information). Once D_q is determined, the model's skill weight and overall weight would finally be produced. In this paper, we provide D_q with a relatively appropriate value considering the errors both in present and future climate changes reproduced by a multi-model weighted average. Keep D_u invariant, which has been determined in section 3.3.2, and sample a range of D_q, figure 3(c) shows the corresponding changes in multivariate errors from a multi-model weighted mean relative to that from the
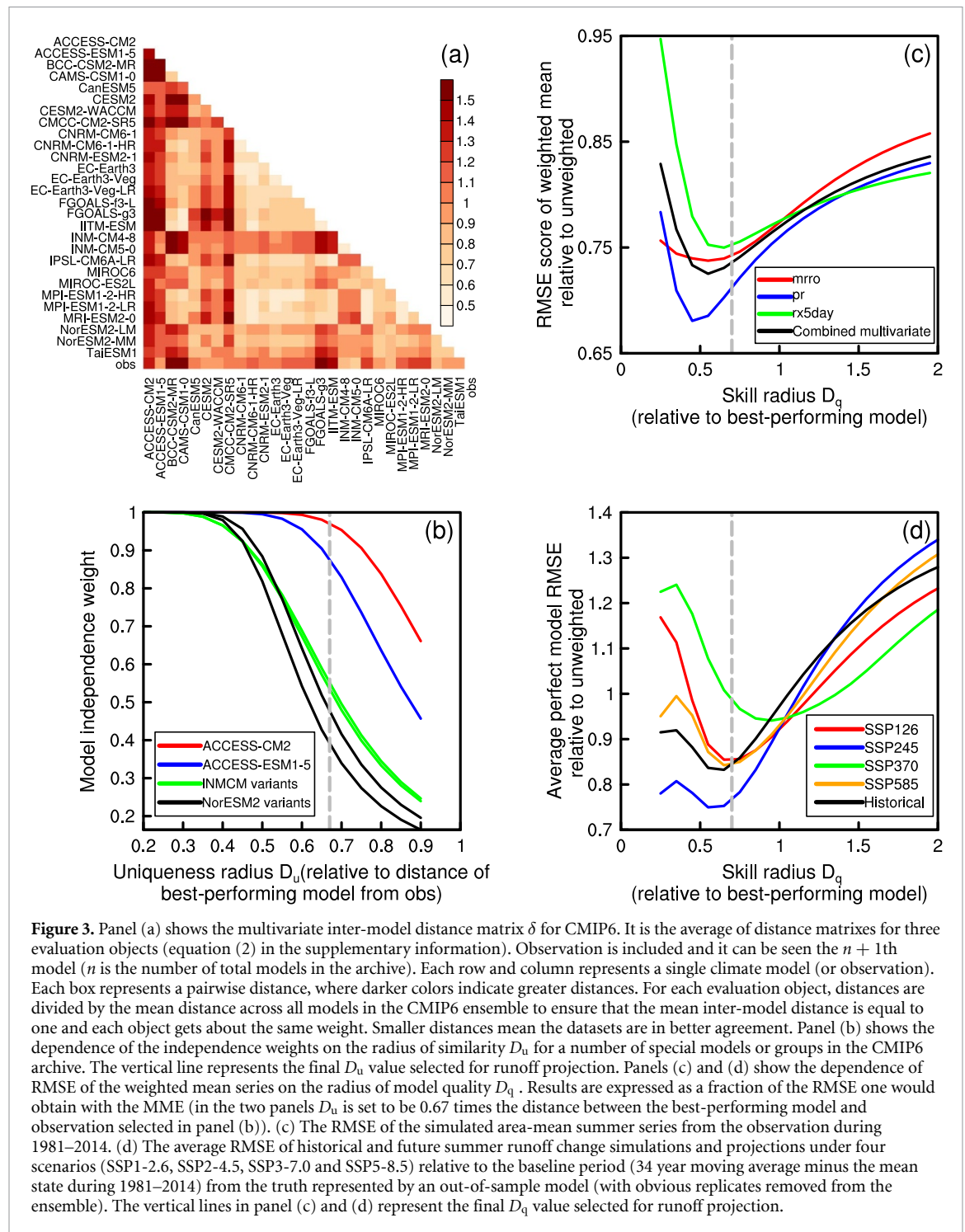
Figure 3. Panel (a) shows the multivariate inter-model distance matrix δ for CMIP6. It is the average of distance matrixes for three evaluation objects (equation (2) in the supplementary information). Observation is included and it can be seen the $n + 1$th model (n is the number of total models in the archive). Each row and column represents a single climate model (or observation). Each box represents a pairwise distance, where darker colors indicate greater distances. For each evaluation object, distances are divided by the mean distance across all models in the CMIP6 ensemble to ensure that the mean inter-model distance is equal to one and each object gets about the same weight. Smaller distances mean the datasets are in better agreement. Panel (b) shows the dependence of the independence weights on the radius of similarity D_u for a number of special models or groups in the CMIP6 archive. The vertical line represents the final D_u value selected for runoff projection. Panels (c) and (d) show the dependence of RMSE of the weighted mean series on the radius of model quality D_q. Results are expressed as a fraction of the RMSE one would obtain with the MME (in the two panels D_u is set to be 0.67 times the distance between the best-performing model and observation selected in panel (b)). (c) The RMSE of the simulated area-mean summer series from the observation during 1981–2014. (d) The average RMSE of historical and future summer runoff change simulations and projections under four scenarios (SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5) relative to the baseline period (34 year moving average minus the mean state during 1981–2014) from the truth represented by an out-of-sample model (with obvious replicates removed from the ensemble). The vertical lines in panel (c) and (d) represent the final D_q value selected for runoff projection.

simple ensemble mean for the baseline period. It is obvious that taking D_q as approximately 0.55 times the distance between the best-performing model and observations could reduce the most in-sample error by about 25%–30% under comprehensive consideration (the reduction of the RMSE of the precipitation simulation can be up to about 32%). However, in-sample error is not the only factor considered. That is, even if one model in the archive can be seen as a good-performing model for the baseline period simulation, we still cannot guarantee that its projections are reliable for adaptation activities.

Thus, in order to determine whether our weighting scheme can efficiently improve the performance of the multi-model ensemble runoff projection under climate change, we conduct a 'perfect model test' following the method outlined by Sanderson *et al* (2017). It is introduced in the supplementary information in detail. In this method, a single model is selected from the archive to represent the truth and the remaining models are used to project it. For a specific range of D_q, the RMSEs of the weighted mean projection of each perfect model's projection of summer runoff changes relative to 1981–2014
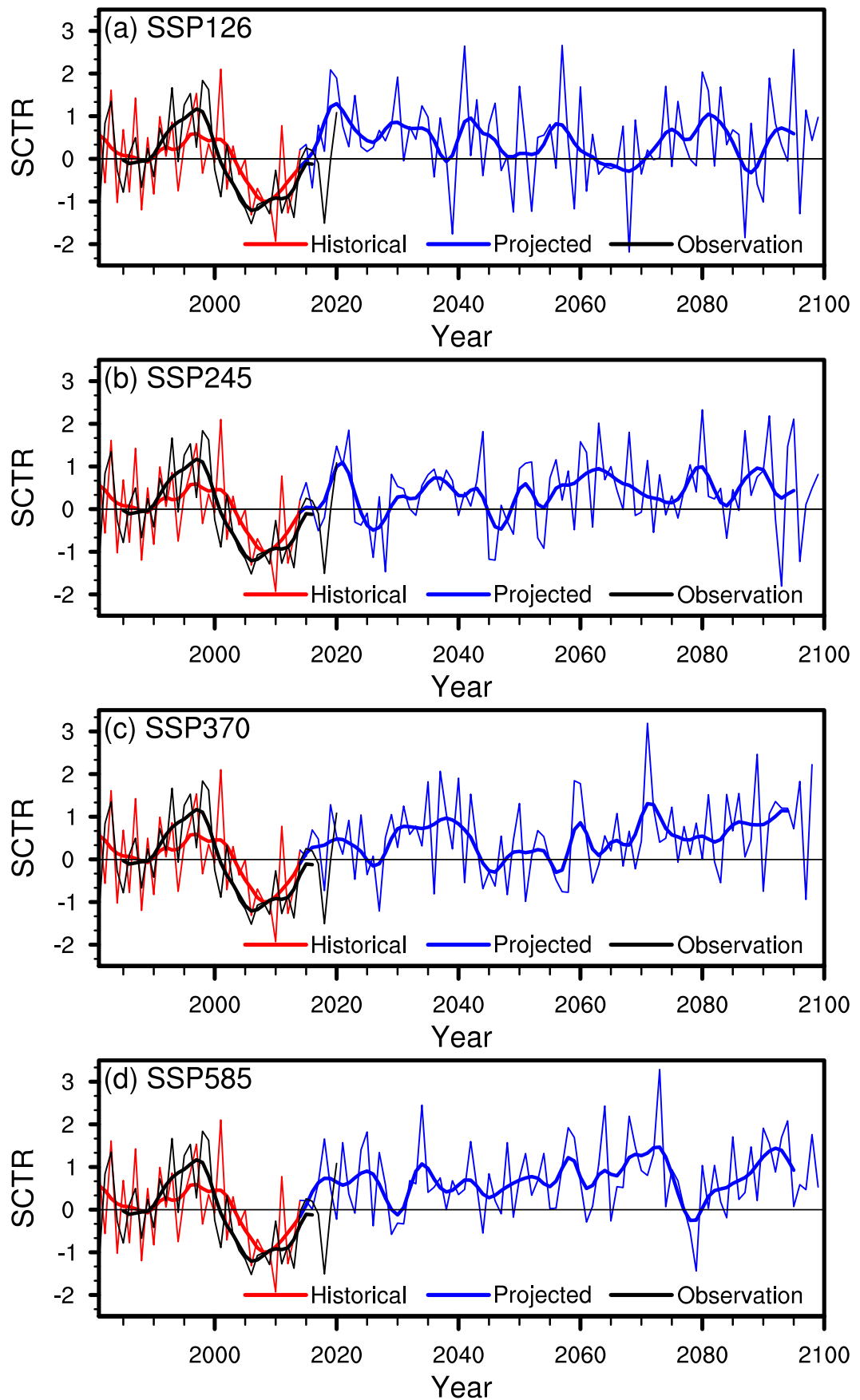
Figure 4. The observed (black line) and multi-model weighted mean projections (blue line) (standardized by the values of 1981–2010) of area-mean summer runoff in the MLYRB and its nine year low-pass filtered value (thick lines) during 2015–2100 under four scenarios (a–d: SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5). To facilitate comparison, the simulated (red line) series during the baseline period (1981–2014) are also presented.

are compared with the MME. Results are shown in figure 3(d). Out-of-sample errors in runoff change projections can be reduced most when taking D_q as approximately 65% of the distance between the best-performing model and observation. Considering model skills during historical, baseline and future periods, we finally determine D_q to be 0.7 times the smallest distance between models and observation, and then, each model's skill weight and overall weight can be estimated (supplementary figure 6).

It is obvious that our weighting scheme can effectively constrain the multi-model biases and really improve the performance of the multi-model ensemble runoff simulation over the MLYRB during the baseline period (figure 2 and supplementary figures 3(b) and (e)). This is mainly reflected in the reduction of the RMSE compared with the MME. Although we take the MLYRB region as a whole and just consider its area-mean runoff variations in the weighting scheme, results show significant improvement in the weighted mean simulation of the spatial distribution of two key runoff features: the climatology and trend over the MLYRB (supplementary figures 3(b) and (e)). The RMSEs of the two features have been reduced by 76.8% and 88.6%, respectively, compared with the MME. Figure 2 shows the improvement in the simulation of both interannual and decadal variations. The weighted mean has an ACC of 0.34 (passes the 95% significance level) (0.88) and an RMSE of 1.44 (1.12) for the interannual (decadal) time scale. It has obviously better skills than almost all individual models and the MME. Considering the improvement and good performance (especially in simulating runoff's decadal variations), the set of overall model weights produced in this section are finally used to investigate the characteristics of interdecadal variation of summer runoff in the MLYRB during the future period (2015–2100) under climate change.

3.4. Future projection with constraints of model biases

In this section, after the multi-model biases during the future period are constrained using the model weighting scheme, we aim to provide a relatively objective and reasonable runoff projection in the MLYRB. In this study, we focus on interdecadal variation characteristics of mean summer runoff in the MLYRB under the SSP1-2.6, SSP2–4.5, SSP3–7.0 and SSP5–8.5 scenarios. Figure 4 shows interannual and decadal series (blue lines) of runoff projected by multi-model weighted mean during 2015–2100. For the convenience of comparison, the simulated series (red lines) during 1981–2014 are also presented. Following Yao *et al* (2020), all series are standardized by the values of 1981–2010, which represents the World Meteorological Organization climatological standard normal (Kappelle *et al* 2021). It can be seen that the weighted mean can well simulate

the observed interdecadal variation during 1981–2020, which demonstrates that our model weighting scheme is applicable and reliable. It can also be reflected by evaluating the skill of the weighted and unweighted mean simulation of climatology, trend and variation of area-average summertime runoff over the MLYRB (supplementary table 2). After the model biases are constrained with the weighting scheme, great improvement can be found in not only past simulations, but also in the future projection during 2015–2020 under different climate scenarios. This is due to the consideration of the model's future performance by conducting the 'perfect model test' in the scheme. It shows a certain degree of credibility of the runoff projection presented in this paper. Under the climate change scenario, positive standardized values indicate that summer runoff in the MLYRB tends to be larger in the future than the mean state during the present day. It is worth noting that the decadal series of runoff under the four SSPs all indicate that the peak value occurs in 2020, which to some extent can explain the rainstorm and flood disasters that occurred in the MLYRB in that year. According to the changes in runoff and precipitation during 2021–2040 relative to the mean state during 1981–2010 (supplementary figure 7), we projected that the MLYRB and its surrounding regions are going to experience a new pluvial period in the coming 20 years. Under the direct impacts of the precipitation, summertime runoff in the middle and south of the MLYRB tends to increase significantly under all economy pathways. The high consistency of the spatial patterns of runoff and precipitation changes demonstrates that precipitation is still the major climate variable influencing summertime runoff in the south of China under future climate change.

4. Conclusion and discussion

This study reveals that the summer runoff in the MLYRB experienced an obvious interdecadal oscillation during 1981–2020 with significantly high runoff during 1991–1999 and low runoff during 2003–2013. According to the results of the composite and correlation analysis, we found that the appearance of the western North Pacific anomalous anticyclone and the strengthening of zonal westerly in the upper troposphere are main climate causes through their impacts on precipitation, which is confirmed to be the direct factor leading to decadal changes of summer runoff in the MLYRB. Because of its close linkage with climate change, long-term decadal changes in summer runoff over the MLYRB were then evaluated with 28 GCMs participating in CMIP6.

Due to large multi-model biases and wide model spread, the MME poorly simulates the climatology, trend, interannual and decadal variations of summer runoff in the MLYRB during the baseline period. In this situation, no matter whether an individual model

or the MME is selected for runoff projection, the result will have considerable biases and cannot provide a valuable reference. Therefore, in this paper, a model weighting scheme is conducted with the simulations of 28 GCMs from the CMIP6 archive. Following the framework outlined by Sanderson *et al* (2017), each model is weighted on the basis of its skill in simulating the historical observations as well as its independence from the other models. The computational process determines that this scheme can constrain the multi-model biases during both the historical and future period under different climate change scenarios. It is confirmed that the weighted mean has much better performance in simulating the historical observed climatology, trend, interannual and especially decadal variations of summer runoff in the MLYRB compared with the MME or individual models in the CMIP6 archive. In that case, the multi-model weighted mean runoff simulation and projection were finally used to detect the decadal variations of summer runoff in the MLYRB during a longer term under climate change.

Compared with the mean state during 1981–2010, summer runoff in the MLYRB is projected to be higher in the future. It is worth noting that the peaks of summer runoff in the MLYRB under different scenarios all occur in 2020, which is likely to be followed by increased summer runoff until 2040. Specifically, the weighted projections indicate that in the following 20 years, the MLYRB and its surrounding regions are projected to experience a new pluvial period. And in the future extreme precipitation is still the major climate hazard impacting south of China. The weighting scheme in this paper can be reliable for runoff projections in other regions at multiple time scales.

It should be noted that there are several limitations of the current research. First, only the linkage between the interdecadal variation of summer runoff in the MLYRB and atmospheric circulation anomalies during 1981–2020 is analyzed. It is necessary to use available long-term observations or long-term reconstruction data to obtain more samples and further investigate whether the relationship could last for a longer time during the historical period. Second, the problem of trade-off between model skill and model independence outlined by Sanderson *et al* (2017) still exists. That is, in the CMIP6 archive, those models with higher model skill weights also tend to have more duplicates and thus have lower model independence weights. Due to the unavoidable trade-off problem of the CMIP6 archive, in this paper the model skill and independence are likely to have compensating effects in the weighting scheme, which leads to there still being much room to further reduce the model biases and uncertainties. Researchers should use more scientific ensemble methods to deal with the trade-off problem in future runoff projections. Moreover, the weighting scheme just considers local runoff and its direct meteorological impact factor (i.e. precipitation). In fact, it has been pointed out by many previous studies that summer runoff in the MLYRB is significantly influenced by individual remote circulation patterns through their impacts on the East Asia summer monsoon (e.g. El Niño/Southern Oscillation, Pacific Decadal Oscillation) (Xiao *et al* 2015, Su *et al* 2017). We propose that in future studies, more optimal and physically relevant variables as well as the climate patterns should be selected and their local and nonlocal effects should be considered to acquire better runoff projections. Finally, human activities can have a great impact on river hydrology in many catchments. With the increasing population, dam capacity and gross domestic product, together with the warming trend induced by human activities, runoff may be influenced significantly. However, in the weighting scheme we did not consider any impact of human factors. To some satisfaction, runoff projected by this scheme tends to increase under global warming and to be larger under stronger warming forcing except for SSP1-2.6 (figure not shown). Additionally, Chen *et al* (2014) indicated that the annual runoff in the YRB has shown little response to human activities since the late 1970s. This implies that the characteristics of runoff variability at annual and seasonal time scale are distinct, which should be paid more attention in future studies.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

Wang H J and He S P designed the research. He S P and Zhao J Z performed research. Zhao J Z prepared the manuscript with contributions from all co-authors.

ORCID iD

Shengping He ⓘ https://orcid.org/0000-0003-4245-357X

References

Caldwell P, Bretherton C, Zelinka M, Klein S, Santer B and Sanderson B 2014 Statistical significance of climate sensitivity predictors obtained by data mining *Geophys. Res. Lett.* **41** 1803–8

Chen J, Wu X, Finlayson B, Webber M, Wei T, Li M and Chen Z
2014 Variability and trend in the hydrology of the Yangtze
River, China: annual precipitation and runoff *J. Hydrol.*
513 403–12

Chen Z, Zhang S-R, Coster A and Fang G 2015 EOF analysis and
modeling of GPS TEC climatology over North America *J.
Geophys. Res. Space. Phys.* **120** 3118–29

Chen Z, Zhou T, Zhang L, Chen X, Zhang W and Jiang J 2020
Global land monsoon precipitation changes in CMIP6
projections *Geophys. Res. Lett.* **47** e2019GL086902

Duchon C 1979 Lanczos filtering in one and two dimensions *J.
Appl. Meteorol.* **18** 1016–22

Eyring V, Bony S, Meehl G, Senior C, Stevens B, Stouffer R and
Taylor K 2016 Overview of the Coupled Model
Intercomparison Project Phase 6 (CMIP6) experimental
design and organization *Geosci. Model Dev.* **9** 1937–58

Gao H *et al* 2020 Assessing glacier retreat and its impact on water
resources in a headwater of Yangtze River based on CMIP6
projections *Sci. Total Environ.* **765** 142774

Hagemann S *et al* 2013 Climate change impact on available water
resources obtained using multiple global climate and
hydrology models *Earth Syst. Dyn.* **4** 129–44

Hawkins E and Sutton R 2011 The potential to narrow
uncertainty in projections of regional precipitation change
Clim. Dyn. **37** 407–18

Hersbach H *et al* 2020 The ERA5 global reanalysis *Quart. J. R.
Meteorol. Soc.* **146** 1999–2049

Huang D, Dai A, Yang B, Yan P, Zhu J and Zhang Y 2019
Contributions of different combinations of the IPO and
AMO to recent changes in Winter East Asian Jets *J. Clim.*
32 1607–26

Jiang J, Zhou T, Chen X and Zhang L 2020 Future changes in
precipitation over Central Asia based on CMIP6 projections
Environ. Res. Lett. **15** 054009

Kappelle M, Kennedy J, Baddour O and Wang Y 2021 *The State of
the Global Climate 2020*

Kim H-J, Wang B and Ding Q 2008 The global monsoon
variability simulated by CMIP3 coupled climate models *J.
Clim.* **21** 5271–94

Kitoh A, Ose T, Kurihara K, Kusunoki S, Sugi M and KAKUSHIN
Team-3 Modeling Group 2009 Projection of changes in
future weather extremes using super-high-resolution global
and regional atmospheric models in the KAKUSHIN
Program: results of preliminary experiments *Hydrol. Res.
Lett.* **3** 49–53

Knutti R, Masson D and Gettelman A 2013 Climate model
genealogy: generation CMIP5 and how we got there
Geophys. Res. Lett. **40** 1194–9

Knutti R, Sedláček J, Sanderson B, Lorenz R, Fischer E and
Eyring V 2017 A climate model projection weighting scheme
accounting for performance and interdependence: model
projection weighting scheme *Geophys. Res. Lett.*
44 1909–18

Koirala S, Hirabayashi Y, Mahendran R and Kanae S 2014 Global
assessment of agreement among streamflow projections
using CMIP5 model outputs *Environ. Res. Lett.*
9 064017

Lee J-Y and Wang B 2014 Future change of global monsoon in the
CMIP5 *Clim. Dyn.* **42** 101–19

Li L, Geng S, Wu C, Song K, Sun F, Visvanathan C, Xie F and
Wang Q 2019 Microplastics contamination in different
trophic state lakes along the middle and lower reaches of
Yangtze River Basin *Environ. Pollut.* **254** 112951

Li X, Zhang K, Gu P, Feng H, Yin Y, Chen W and Cheng B 2021
Changes in precipitation extremes in the Yangtze River Basin
during 1960–2019 and the association with global warming,
ENSO, and local effects *Sci. Total Environ.* **760** 144244

Mizuta R *et al* 2012 Climate simulations using MRI-AGCM3.2
with 20-km grid *J. Meteorol. Soc. Japan* **90A** 233–58

Merrifield A, Brunner L, Lorenz R and Knutti R 2020 A weighting
scheme to incorporate large ensembles in multi-model
ensemble projections *EGU General Assembly 2020 4–8 May
2020* EGU2020–4524

Nam D, Duong P, Thuan D, Mai D and Dung N 2018 Assessment
of near-term runoff response at a River Basin scale in central
Vietnam using direct CMIP5 high-resolution model outputs
Water **10** 477

Nam D, Hoa T, Duong P, Thuan D and Mai D 2019 Assessment of
flood extremes using downscaled CMIP5 high-resolution
ensemble projections of near-term climate for the upper
Thu Bon Catchment in Vietnam *Water* **11** 634

NOAA National Centers for Environmental Information 2021
State of the Climate: Global Climate Report for Annual 2020
NOAA

O'Neill B *et al* 2016 The scenario model intercomparison project
(ScenarioMIP) for CMIP6 *Geosci. Model Dev.*
9 3461–82

Sanderson B, Knutti R and Caldwell P 2015 A representative
democracy to reduce interdependency in a multimodel
ensemble *J. Clim.* **28** 5171–94

Sanderson B, Wehner M and Knutti R 2017 Skill and
independence weighting for multi-model assessments
Geosci. Model Dev. **10** 2379–95

Schewe J *et al* 2014 Multimodel assessment of water scarcity
under climate change *Proc. Natl Acad. Sci. USA*
111 3245–50

Sillmann J, Kharin V, Zwiers F, Zhang X and Bronaugh D 2013
Climate extremes indices in the CMIP5 multimodel
ensemble: part 2. Future climate projections *J. Geophys. Res.
Atmos.* **118** 2473–93

Sohoulande C, Djebou S and Singh V 2016 Impact of climate
change on the hydrologic cycle and implications for society
Environ. Soc. Psychol. **1** 36–49

Su Z, Hao Z, Yuan F, Chen X and Cao Q 2017 Spatiotemporal
variability of extreme summer precipitation over the Yangtze
River Basin and the associations with climate patterns *Water*
9 873

Tang J, Li Q and Chen J 2018 Summertime runoff variations and
their connections with Asian summer monsoons in the
Yangtze River Basin *J. Water Clim. Change* **9** 89–100

Tao L, He X and Qin J 2021 Multiscale teleconnection analysis of
monthly total and extreme precipitations in the Yangtze
River Basin using ensemble empirical mode decomposition
Int. J. Climatol. **41** 348–73

Wang B, Yim S-Y, Lee J-Y, Liu J and Ha K-J 2014 Future change of
Asian-Australian monsoon under RCP 4.5 anthropogenic
warming scenario *Clim. Dyn.* **42** 83–100

Wang C, Yao Y, Wang H, Sun X and Zheng J 2021 The 2020
summer floods and 2020/21 winter extreme cold surges in
China and the 2020 typhoon season in the Western North
Pacific *Adv. Atmos. Sci.* **38** 896–904

Wilby R, Beven K and Reynard N 2008 Climate change and fluvial
flood risk in the UK: more of the same? *Hydrol. Process.*
22 2511–23

Xiao D and Li J 2007 Spatial and temporal characteristics of the
decadal abrupt changes of global atmosphere-ocean system
in the 1970s *J. Geophys. Res. Atmos.* **112** D24S22

Xiao Z, Shi P, Jiang P, Hu J, Qu S, Chen X, Chen Y, Dai Y and
Wang J 2018 The spatiotemporal variations of runoff in the
Yangtze River Basin under climate change *Adv. Meteorol.*
2018 5903451

Xing W, Wang W, Zou S and Deng C 2018 Projection of future
runoff change using climate elasticity method derived from
Budyko framework in major basins across China *Glob.
Planet. Change* **162** 120–35

Xiao M, Zhang Q and Singh V 2015 Influences of ENSO, NAO,
IOD and PDO on seasonal precipitation regimes in the
Yangtze River Basin, China *Int. J. Climatol.* **35** 3556–67

Yao N, Li L, Feng P, Feng H, Liu D-L, Liu Y, Jiang K, Hu X and Li Y
2020 Projections of drought characteristics in China based
on a standardized precipitation and evapotranspiration
index and multiple GCMs *Sci. Total Environ.* **704** 135245

Zhang G, Zeng G, Li C and Yang X 2020 Impact of PDO and
AMO on interdecadal variability in extreme high
temperatures in North China over the most recent 40-year
period *Clim. Dyn.* **54** 3003–20

Zhou T, Lu J, Zhang W and Chen Z 2020 The sources
of uncertainty in the projection of global land
monsoon precipitation *Geophys. Res. Lett.*
47 e2020GL088415

Zhu J, Huang D-Q, Zhang Y-C, Huang A-N, Kuang X-Y and
Huang Y 2013 Decadal changes of Meiyu rainfall around
1991 and its relationship with two types of ENSO *J. Geophys.
Res. Atmos.* **118** 9766–77