Comparison of Statistical Downscaling of Summer Daily Precipitation Through a Certain Perfect Prognosis and Bias Corrections - A Case Study Across China

Yonghe Liu
Henan Polytechnic University
Xiyue Wang
Henan Polytechnic University
Mingshi Wang
Henan Polytechnic University
Hailin Wang (whl@hpu.edu.cn)
Henan Polytechnic University

Research Article

Keywords: perfect prognosis, statistical downscaling, optimal grid box, bias corrections, generalized linear models, quantile mapping

Posted Date: December 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-880746/v1

Comparison of statistical downscaling of summer daily precipitation through a certain perfect prognosis and bias corrections - A case study across China

Yonghe Liu[1], Xiyue Wang[1], Mingshi Wang[1], Hailin Wang*[1],

1. School of Resources and Environment, Henan Polytechnic University, Jiaozuo Henan, China
*Corresponding author: whl@hpu.edu.cn

Abstract

Fewer perfect prognosis (PP) based statistical downscaling were applied to future projections produced by global circulation models (GCM), when compared with the method of model output statistics (MOS). This study is a trial to use a multiple variable based PP downscaling for summer daily precipitation at many sites in China and to compare with the MOS. For the PP method (denoted as 'OGB-PP'), predictors for each site are screened from surface-level variables in ERA-Interim reanalysis by an optimal grid-box method, then the biases in predictors are corrected and fitted to generalized linear models to downscale daily precipitation. The historical and the future simulations under the medium emission scenario (often represented as 'RCP4.5'), produced by three GCMs (CanESM2, HadGEM2-ES and GFDL-ESM2G) in the coupled model intercomparison project phase five (CMIP5) were used as the downscaling bases. The bias correction based MOS downscaling (denoted as 'BC-MOS') were used to compare with the OGB-PP. The OGB-PP generally produced the climatological mean of summer precipitation across China, based on both ERAI and CMIP5 historical simulations. The downscaled spatial patterns of long-term changes are diverse, depending on the different GCMs, different predictor-bias corrections, and the choices on selecting PP and MOS. The annual variations downscaled by OGB-PP have small differences among the choices of different predictor-bias corrections, but have large difference to that downscaled by BC-MOS. The future changes downscaled from each GCM are sensitive to the bias corrections on predictors. The overall change patterns in some OGB-PP results on future projections produced similar trends as those projected by other multiple-model downscaling in CMIP5, while the result of the BC-MOS on the same GCMs did not, implying that PP methods may be promising. OGB-PP produced more significant increasing/decreasing trends and larger spatial variability of trends than the BC-MOS methods did. The reason maybe that in OGB-PP the independent precipitation modeling mechanism and the freely selected grid-box predictors can give rise to more diverse outputs over different sites than that from BC-MOS, which can contribute additional local variability.

Key word: perfect prognosis; statistical downscaling; optimal grid box; bias corrections; generalized linear models; quantile mapping

1. Introduction

To predict future climate changes, the World Climate Research Project (WCRP) Working Group on Coupled Modelling (simulation work group) has organized the Coupled Model Intercomparison Project (CMIP) (Zhou et al., 2014; Zhou et al., 2020) to share, compare and analyze the outcomes of GCMs. So far, CMIP provides a huge amount of global climate model (GCM) simulations for the past and future climate, and has provided sufficient researches for several IPCC Assessment Reports. However, GCMs adopt coarse resolutions (about 1-2 degrees, equivalent to about 100km-200 km) for long-term climate simulation, and cannot properly describe the complex local surface conditions and some physical processes, resulting in large deviations at regional scales(Maraun et al., 2010a). Downscaling is a technique to infer the climate information on the basin / regional scale from the large-scale variables (LSVs) of coarse resolution (Wilby et al., 2002). Statistical downscaling is such a technique, which is constructed by establishing linear or nonlinear relationship between LSVs and regional-scale variables in an empirical/statistical way (Wilby et al., 2002). The corresponding regional scale variable is derived by applying the empirical relation to the output of a GCM or RCM, which is computationally economical, therefore, is more feasible than other methods (such as the dynamical downscaling) for assessing the impacts of the global climate change on local areas based on the huge-amount multi-model CMIP simulations.

Maraun et al. (2010b) classified the statistical downscaling methods into three categories according to the predictors and the model structures: perfect prognosis (PP), model output statistics (MOS) and weather generators. A PP method establishes an empirical model between an observational large-scale data and local observations, and then apply the model to the GCM simulated large-scale predictors. The large-scale observational data is usually played by the reanalysis datasets which are derived by assimilating other multi-source observational data. A MOS method is a direct empirical model between a GCM output and local observational variables. GCM long-term climate predictions are unsynchronized with historical observations, making it difficult for a MOS to use a variety of large-scale predictors as that does by a PP method. In practical applications, the most commonly used MOS is actually the bias corrections on the GCM results, such as quantile mappings (Li et al., 2010; Maraun, 2013a). For PP, the precipitation in reanalysis dataset is scarcely used in the downscaling of precipitation because the GCM simulated precipitation is unable to express large-scale circulations (Maraun et al., 2010b). However, in MOS, the GCM simulated precipitation is directly used as a large-scale predictor, therefore, MOS are equivalent to bias corrections. Due to the simplicity, bias corrections (the mostly used method is quantile mapping) have been widely used in downscaling massive multi-model climate projections under future emission scenarios (Ahmed et al., 2019; Navarro-Racines et al., 2020; Su et al., 2016; Yang et al., 2019a; Zhang et al., 2019a). In this respect, MOS is actually the mainstream method in statistical downscaling. However, Maraun (2013a) found that the quantile mapping(QM) usually produce incredible result. As pointed out by Maraun (2013b), the PP maybe more applicable, for the add value in improving the local variability.

Compared to a simple QM, PP can adopt more large-scale variable predictors, and has a more complexed model structure and is more difficult to implement (Maraun et al., 2019a). Most studies on PP models are still in the verification stage, such as Das and Akhter (2019), and the applications of PP in assessing future impacts based on multi-model GCM simulations are very limited. Except for a few traditional PP models being widely used (Al-Mukhtar and Qasim, 2019; Gebrechorkos et al., 2019; Tang et al., 2016), most new PP models are only used for downscaling of a small number of climate projections (Kaspar-Ott et al., 2019). The performance of a PP downscaling model also relies on how the model is constructed, especially for how to select predictors. The optimal grid box method is promising for screening predictors, which need no dimensionality reduction, and the final predictors have a clear physical explanation to local precipitation. Fu et al. (2018) and Zhang and Yan (2015) applied this method for downscaling monthly precipitation. For daily precipitation, this predictor-screening method has been successfully applied to a variety of large-scale variables, and has achieved a good performance in China (Liu et al., 2019a; Liu et al., 2019b, Liu et al., 2019c). Considering the predictors extracted from the reanalysis dataset may have systematic errors to that extracted from the GCM simulations, further bias corrections on the predictors are also needed.

The aim of this paper is to use a certain PP method, which is based on the optimal-grid-box predictor screening, generalized linear models and bias corrections on predictors, to downscale daily precipitation from the future projections in CMIP5, and compare the downscaled results with those downscaled by the common bias corrections (belong to the MOS category). In following paragraphs, we use the abbreviation 'OGB-PP' (here 'O' represents the optimal grid box, 'G' represents generalized linear models, 'B' represents bias corrections) to represent this PP method and the abbreviation 'BC-MOS' to represent the bias-correction methods.

Some related questions are expected to answer:1) based on the evaluation on historical samples, how the performance can be attained by OGB-PP through different options? 2) How large the influence of bias corrections on PP predictors is when compared to the influence of predictors from different GCMs? 2) for downscaling precipitation across China for the future climate projections, how future projections can be generated by OGB-PP and BC-MOS? To answer the questions, the datasets simulated by three GCMs in CMIP5 were used as the base to downscale precipitation. This study can help us understand the potential of PP models for practical downscaling on future projections.

2. Data and method

2.1 Data preparation

Dataset of daily precipitation gauged at 173 meteorological stations over China were collected from the China Meteorological Data Sharing Service System (http://data.cma.gov.cn for details) and was used to calibrate and validate the downscaling models. These related meteorological stations belong to the international climate exchange stations. Only the data observed in boreal summer (June–September) of 1979-2016 was used in this study. Very few missing values exist for the selected period, and such sporadic missing values were directly

replaced by zeros. These stations cover most of the provinces (except for Guizhou, Tibet and Taiwan) and main climatic zones in China (**Fig. 1**). Among them, seven representative sites (Beijing, Tulihe, Hanzhong, Nanjing, Yangjiang, Urumqi, Maduo) were selected to analyze with more details than other sites. These seven sites are located in the areas of North China, Northeast China, Central China, Eastern China, South China, Xinjiang (in far west China) and the Qinghai-Tibetan Plateau, respectively.

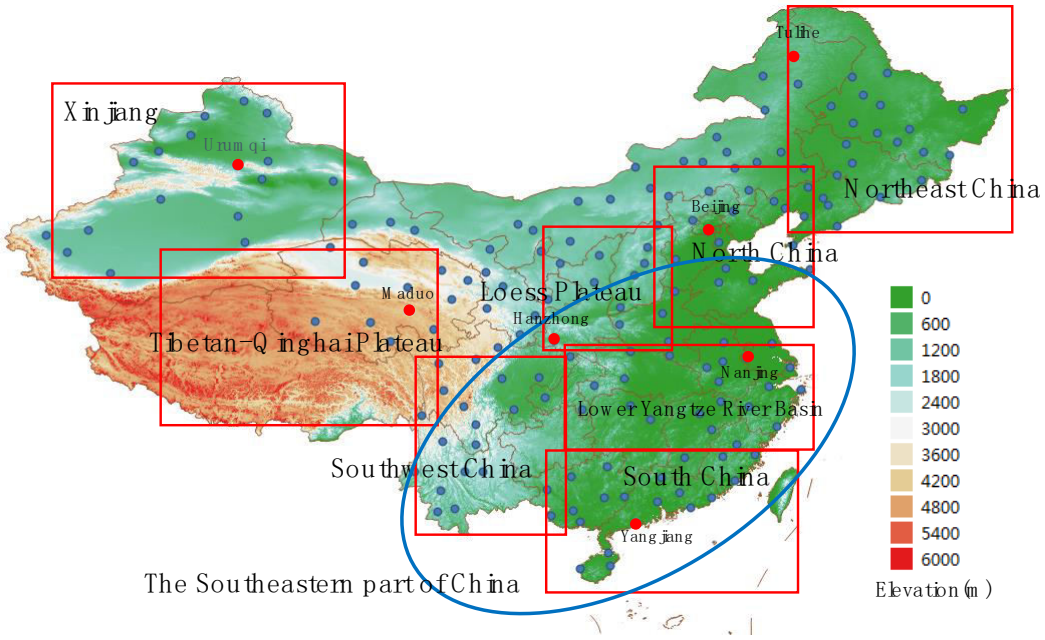


Fig.1 A sketch map of the stations/sites (in blue dots and red dots) and the referenced major zones (the different areas in the red rectangles) used in this study. The red dots represent the seven representative stations which are emphasized in this study. These major zones which are frequently mentioned in literatures are roughly divided and are in boundary-overlapped geographical areas.

The ERA-Interim Reanalysis dataset (ERAI) (Dee et al., 2011) from the European Centre for Medium-Range Weather Forecasts (ECMWF) was used to screen large-scale predictors. The daily mean variables at surface in ERAI, forecasted at 0 h UTC but for 12 h UTC in each day, covering a domain of 60°E-150°E and 10°N-65°N, were downloaded from the website of European Centre for Medium-Range Weather Forecast (ECMWF) (https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim) . Considering the future climate projections simulated by GCMs are usually in rough resolution, the original resolution (0.75 degree on average) of ERAI is unnecessary in this study. To facilitate the data processing, the ERAI dataset was degraded into a 1° × 1° resolution by the data server when downloading. Such a grid still has a finer resolution than those in GCM products. However, this resolution deference has a small influence on the model training, because the predictors used in PP downscaling models are usually the mean values of multiple neighboring grid boxes. Five surface LSVs, including 2-metre dew point temperature (T_d), mean sea level pressure (MSLP), 10-metre U wind component (U10), 10-metre V wind component (V10) and 2-metre air temperature (T2), were used to screen predictors. These LSVs are similar to that used in (Liu et al., 2019b; Liu et al., 2019b), and contain most of the circulation and moisture information which can be used in statistical downscaling, and it is generally easy to find their counterparts in

the GCM outputs. The corresponding variables in multiple pressure levels may be more predictive (Liu et al. 2019c), but the multiple-level datasets from multiple-GCM outputs are in huge sizes and inconvenient to handle.

No surface relative humidity is available in ERAI, therefore, we estimated relative humidity by transforming from 2-metre dew-point temperature (T_d) and 2-meter air temperature (T_2) (Valiantzas, 2013):

$$e^0(T) = 0.6108 \cdot \exp\left(\frac{17.27 \cdot T}{T + 273.3}\right) \quad (1)$$

$$RH = e^0(T_d)/e^0(T) \quad (2)$$

Where $e^0(T)$ is the saturation vapor pressure at the given air temperature T (℃). RH represents the surface relative humidity, which is estimated as the ratio between an actual vapor pressure and the saturation vapor pressure at a certain air temperature. Actual vapor pressure can be estimated from dew-point temperature.

Historical simulations and the future climate projections from the Phase Five of the Coupled Model Intercomparison Project (CMIP5) were used as the large-scale variables from GCM. The surface variables (precipitation, relative humidity at surface, mean sea-level air pressure, surface air temperature, eastward near-surface wind and northward near-surface wind) in CMIP5 outputs were used to extract large-scale predictors. Due to the huge data size, the data is difficult to access from the internet, only the first member of the ensemble historical simulations and future simulations under the medium emission scenario, the representative concentration pathway 4.5 (RCP4.5), from three GCMs (CanESM2, GFDL-ESM2G and HadGEM2-ES) were collected. RCP4.5 is a stabilization scenario with the total radiative forcing of 4.5 W/m2 until 2100(Moss et al., 2010; Thomson et al., 2011). According to Sperber et al. (2013), the three GCMs in CMIP5 have generally high skill scores in simulating the summer East Asia monsoon. In order to extract predictors in a convenient way for each gauge site/station, the CMIP5 datasets were also resampled into the same domain of the downloaded ERAI (1° × 1°, Domain 60°E-150°E, 10°N-66°N).

2.2 Statistical downscaling Method

The statistical downscaling techniques based on OGB-PP and BC-MOS were performed for each single site/station independently. The flow charts for the two downscaling methods are displayed in **Fig.2**. For OGB-PP, the process involves several procedures: 1) establish statistical models by relating ERAI datasets to the observed precipitation for each site, by screening predictors and constructing generalized linear models (GLM); 2) extract predictor values in the CMIP5 simulations, including both historical simulations and future projections under the RCP4.5 scenario; 3) correct the bias in the extracted CMIP5 predictors relative to the predictor values in ERAI with different methods, and feed the PP statistical models with the bias-corrected CMIP5 predictors to get downscaled precipitation. For BC-MOS, only one step is needed: correcting the bias in the CMIP5-GCM simulated precipitation relative to the observed site precipitation.

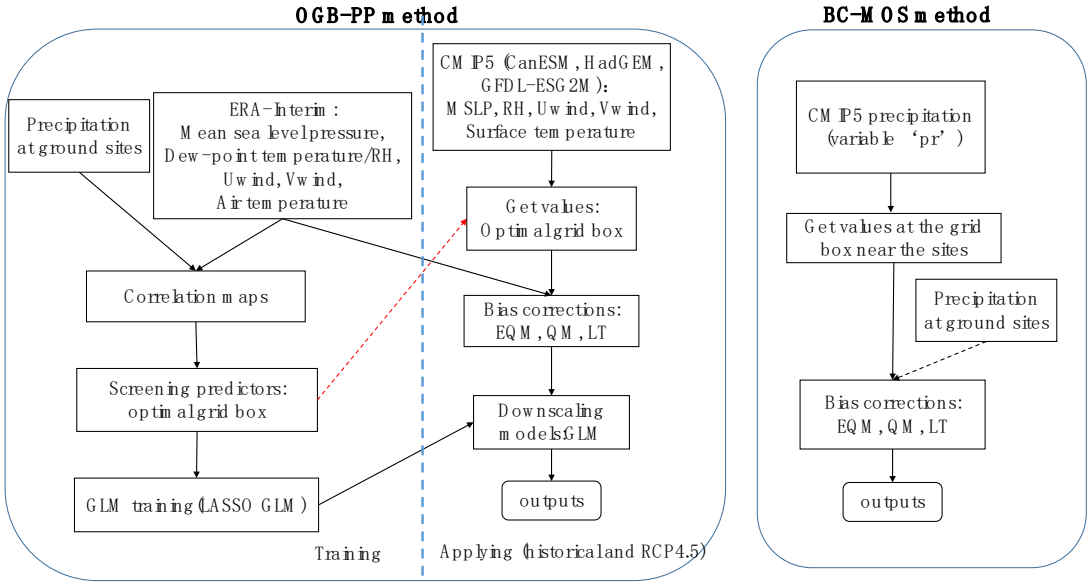


Fig.2 Streamlines for the OGB-PP method and the BC-MOS method

2.2.1 Screening predictors

Screening predictors comprise two steps: selecting the appropriate large-scale variables and finding the optimal values across different locations (grid boxes). Although many large-scale variables can be used as candidate predictors, we do not plan to screen predictors through a lot of large-scale variables, since the sizes of daily GCM products are usually too huge to obtain. In this study, only five surface variables were considered: relative humidity (RH), latitudinal direction wind (V), longitudinal direction wind (U), air temperature (T) and MSLP, since such surface variables are more portable than those in multiple pressure levels. These variables are generally similar to that used in (Liu et al., 2019c) and can represent both atmosphere circulation and humidity. The same method for automatic searching optimal grid boxes for each site/station as that in (Liu et al., 2019b; Liu et al., 2019c) was used in this study. Here, the optimal grid boxes are the boxes with the highest correlations (both negative and positive correlations) to the observed precipitation. To avoid the skewed precipitation distribution which is not suitable to use Pearson's correlations (CC), the precipitation series were logarithmic transformed after adding a small value (0.25 was used here) (Liu et al., 2019a).

Someone may concern that this automatic algorithm may find some fake/erroneous grid boxes to be used. Actually, this is not true, since the high-correlation grid boxes always has a center which is surrounded by multiple sub-high-correlation grid boxes, and there is no abrupt single high-correlation grid box alone. The high-correction areas in the grid are always concentrated at one positive-correlation center and one negative-correlation center. Meanwhile, each high-correlation center must have CCs exceeding 0.2 to the observed precipitation, which can easily pass the significant test at the 0.01 level. Through such an algorithm, the most correlated grid box values were directly used as predictors, and the rest of the grid boxes with less correlations were discarded. Therefore, with this method, it is unnecessarily to use the widely used method of principal components analysis (PCA) to reduce dimensions Comparatively, PCA is mathematically 'complex' method, which makes the final extracted

predictors/principal components uneasy to interpret. For large-scale variable RH, only the positive correlation center in the domain was used as a predictor location, while for other LSVs (P, U, V and T), both the positive correlation and the negative correlation centers were used. With each center, the mean value at a 3×3 window around the selected grid box was used as the final predictor. The window was determined by our inspection on the high-correlation area of many sites. The LSV values and the correlation coefficients change smoothly across the grid boxes in high-correlation areas, and the modeled result is generally not sensitive to the window size for selection.

2.2.2 Generalized linear models

GLMs are widely used in PP for downscaling daily precipitation (San-Martin et al., 2017). In this study, two schemes of downscaling were considered. The first scheme is a single-model scheme, which uses a single model to simulate both the precipitation occurrence and the precipitation amount at wet days, as used in (Liu et al., 2019a; Liu et al., 2019c). Such a model is a common GLM under the assumption of Tweedie distribution of precipitation. When a modeled precipitation value exceeds the threshold of the corresponding site, the value is regarded as a valid precipitation amount value, otherwise, the day is regarded as dry day. Here the threshold for the GLM output of each site is determined as follows: feed the GLM with all the historical samples and sort the GLM outputs (precipitation amount) ascendingly, then use the *N*-th (*N* is the number of dry days in historical observation) value as the threshold. In this way, the smallest *N* values simulated by GLM are regarded as dry-day values. Such GLMs are fitted by the historical data of all days whenever the day is wet or not, which can get a better skill than the GLM fitted by only wet-day samples which are usually used in the traditional two-model scheme, according to our experience. The direct output of such a single model is deterministic and is more convenient for further evaluation.

The second scheme is a traditional two-model scheme which consists of a logistic model and a common GLM, for precipitation occurrence and precipitation amount respectively. The outputs of these two models need to be combined to get stochastic simulations (San-Martin et al., 2017). This two-model scheme in this study is used to compare the results with the single-model scheme, and the latter is the main scheme we used to downscale precipitation from the CMIP5 simulations. Each stochastic value was obtained from a two-step process: first, the wet or dry status for this day is obtained by generating a uniform-distribution random number: if this number lies in 0-γ (γ is the probability of precipitation occurrence), the status of this day is regarded as a wet day, otherwise, a dry day. Second, for a wet day, another random number of Tweedie distribution is generated and act as the precipitation amount of that day. For the single-model scheme, the deterministic output is directly used as daily precipitation, but the values smaller than the thresholds are replaced by zeros.

The models' parameters were estimated through the GLM based on the least absolute shrinkage and selection operator (LASSO) (Hammami et al., 2012). Here, the aim of using LASSO is not to simply reduce the number of parameters, but to eliminate the overfitting problem which is caused by the potential collinearity in predictors. During the traditional parameter-estimation process of a GLM or a linear regression, one parameter (usually a factor

coefficient) of a predictor may get a contrary sign (positive value or negative value) to the sign of the correlation coefficient between this predictor and the predictand. Such contrary-sign phenomena is a reflection of overfitting due to the collinearity in predictors and the limited samples. The LASSO algorithm relies on a parameter λ which controls the number of parameters/coefficients to be replaced by zeros. The parameters estimated with overfitting problem are more likely to be set to zeros by the LASSO algorithm. In this study, the LASSO-based GLMs were trained in two steps. In this first step, the GLMs were trained by setting a small λ value (for example, 0.1). After this step, the parameter (the factor which multiplies with the predictor) of each predictor should be checked manually for whether the sign of this parameter is contrary to the sign of correlation coefficient between the predictor and the observed precipitation. This checking should be done for each site. If true, this parameter is wrongly estimated due to overfitting, then the GLM parameters for this station should be estimated in a further step by setting a larger λ value (for example, 0.2) to set more parameters to zeros. If the wrongly estimated parameters still exist, the LASSO-GLM training should be repeated by setting another larger λ value than before.

2.2.3 Predictor-bias correction in OGB-PP

In OGB-PP, due to the systematic difference of predictors between the ERAI and those extracted from CMIP5 outputs, the latter (the values extracted from the OGB) need bias correction relative to the former before applying a PP model. In this study, three methods of bias corrections were used. The first one is equidistance quantile-matching method (EQM) which was described in Li et al. (2010) and Su et al. (2016):

$$X_{c1} = X_{RCP45} + F_{ERA}^{-1}[F_{RCP45}(X_{RCP45})] - F_{HIS}^{-1}[F_{RCP45}(X_{RCP45})] \quad (3)$$

where x_{c1} is the bias corrected series by the EQM, x_{RCP45} is values simulated by the CMIP5 models and is to be bias corrected. $F_{RCP45}(\cdot)$ is the cumulative distribution function of the values simulated for the RCP4.5 scenario. $F_{ERA}^{-1}(\cdot)$, $F_{RCP45}^{-1}(\cdot)$ are the inverse function of the cumulative distribution function of the values in ERAI reanalysis and in RCP4.5 simulation, respectively.

The second one is the standard quantile mapping (QM)(Yang et al., 2018).

$$X_{c2} = F_{ERAI}^{-1}[F_{HIS}(X_{RCP45})] \quad (4)$$

The third one is a simple linear transformation (LT) based on mean and standard deviation.

$$X_{c2} = \sigma_{ERAI} \frac{X_{RCP45} - m_{HIS}}{\sigma_{HIS}} + m_{ERAI} \quad (5)$$

Where m_{HIS} and σ_{HIS} are the mean value and the standardized deviation of the CMIP5 historical simulation, while m_{ERAI} and σ_{ERAI} are the mean and standardized deviation of the values in ERAI reanalysis. LT is a kind of change factor methods (Maraun et al., 2017).

Therefore, four choices of the predictors in CMIP5 are used in the PP downscaling: the

original predictors without bias correction, the predictors corrected by EQM, the predictors corrected by QM and the predictors corrected by LT. The bias corrections were performed directly on the extracted optimal grid-box predictor values, against the corresponding daily series (1979-2005) in ERAI, for each station individually.

2.2.4 Bias corrections acting as the BC-MOS

Direct bias corrections on the GCM simulated daily precipitation act as the MOS. Such bias corrections were performed on the daily precipitation series extracted for the corresponding site locations from the GCM simulated precipitation in CMIP5. The bias corrections were conducted against the observed precipitation during 1961-2005, considering that during this period, both the CMIP5 historical simulations and the observed precipitation are available. The same EQM and QM as that used for the above predictor-bias corrections in OGB-PP were used here. LT was not used in the BC-MOS, because we think LT does not influence the long-term relative changes in the original GCM precipitation, and can also produce the completely same pattern of climatological mean as that in observation.

2.2.5 Abbreviations for different combinations of options

According to the above descriptions, there are three GCMs (CanESM2, GFDL-ESM2G and HadGEM2-ES), two classes of models (OGB-PP and BC-MOS) and four bias-corrections (NC(No correction), EQM, QM and LT) used in this study, therefore, multiple combinations of the downscaled outputs can be obtained. For simplicity, we use 'CanESM', 'GFDL' and 'HadGEM' to represent the three GCMs, respectively. We represent the different model outputs in such a similar way: a GFDL-ESM2G based PP downscaling with the EQM bias correction is denoted as GFDL-PP/EQM, while the direct linear transformation on large-scale precipitation of GFDL-ESM2G is denoted as GFDL-MOS/LT.

2.2.6 Validation of OGB-PP

In this study, the detailed performance on cross-validation, such as the credibility of the finally trained models and the climate related stationarity, are not discussed, since similar analysis were presented in previous literatures (Liu et al., 2019b; Liu et al., 2019c) which demonstrate that such problems need not be concerned any more. The validations for similar methods were done in Liu et al. (2019c). The only difference between the method in this study to that in (Liu et al.(2019c) is that only the surface-level LSVs were used here, while the latter used the same LSVs but in multiple pressure levels. In this study, the CC between the downscaled output (in logarithmic transformation) and the observed precipitation (logarithmic transformed after adding a small value of 0.25) was used as a main validation metric. Here, the CC is calculated based on the logarithmic transformed values other than on the original precipitation values, just to reduce the effect caused by the strong skewness in the original values. Adding the small value of 0.25 to the original precipitation values is to avoid logarithmic transformation on zeros.

A simple cross validation was used to train and validate the models: the summer samples of 38 years (a total of 4636 days) were concatenated into a large sample set and then were divided into four subsets of the same size (each subset has 1159 daily samples); the model training was repeated four times by changing the subsets: three subsets were used for training and the remaining set was used for validation (calculating CC).

The performance of OGB-PP driven by ERAI predictors was evaluated against the precipitation observation during 1979-2016. The performances obtained by some related options of different schemes were also compared: the option of using surface-level predictors versus the option of using multiple-level predictors, the option of using single-model scheme (deterministic) versus the option of using two-model scheme (stochastic). The stochastic outputs cannot be evaluated on the daily scale, therefore, the means of ensemble simulations were used to compare on the annual scale. In order to assess the simulated statistical distribution of daily precipitation, quantile-quantile plots were used to compare with the observed precipitation, on a daily scale. Meanwhile, linear trends during 1979-2005 for all sites were used to compare the simulated with the observed.

2.3 Evaluation and comparison of the downscaled results in CMIP5

The correlation between the results of OGB-PP and that of the original large-scale GCM simulated precipitation were also analyzed.

In order to compare the change of precipitation at different stations, the analysis was done on a 40-year long-term basis: 1960-1999 in historical simulation and 2060-2099 in RCP4.5 simulations. The 40-year period is long enough to eliminate most interannual variations. Four indices, the relative changes (R_{q95}, unit: percent) of daily precipitation at the 0.95 quantile (q95), the relative change (R_{tsp}, unit: percent) of the annual total summer precipitation (tsp), the change (C_{dds}, unit: day) of duration of the longest dry spells (dds) and the relative change (R_{nrd}, unit: day) of the number of annual summer rainy days (nrd), were summarized from the downscaled time series. Here, the dds were calculated by averaging the four longest durations of dry spells during the 40-year period. They were calculated respectively by

$$R_{q95}=(q95_{rcp} - q95_{his})*100/\ q95_{his}$$

$$R_{tsp}=(tsp_{rcp} - tsp_{his})*100/\ tsp_{his}$$

$$C_{dds}=dds_{rcp} - dds_{his}$$

$$R_{nrd}=(nrd_{rcp} - nrd_{his.})*100/nrd_{his}$$

where the subscription ‘rcp’ and ‘his’ represents RCP4.5 and historical simulations, respectively.

3 Result

3.1 OGB-PP models

3.1.1 Optimal grid boxes

Similar patterns were obtained for the correlation maps (not displayed here) with different LSVs of ERA-Interim reanalysis as that obtained by Liu et al. (2019a) and Liu et al. (2019c). The best correlations obtained from multiple LSVs for each station are shown in **Fig.3**. For each station, the CCs shown in **Fig.3** is the highest negative or positive correlations obtained among all the grid-box time series of the LSVs. The relative humidity gets CCs ranging from 0.15-0.3 in most stations, and the CCs larger than 0.3 only appeared at some stations in the west part of Inner Mongolia, Qinghai-Tibetan Plateau and the north of Xinjiang.

For MSLP, the highest positive correlations are at the stations in the area of south Gansu, Sichuan, north Yunnan and east Qinghai. This large area lies along the east outer side of the Qinghai-Tibetan plateau. These positive correlations represent the remote correlation of precipitation to the west pacific Subtropical High (WPSH) (Liu et al., 2019c). When the WPSH moves westward to the south-eastern China, there is a high probability that the east outer side of the Qinghai-Tibetan plateau is covered by the rain belt. Apart from this area, the stations in the rest of China has very small positive correlations to MSLP. For the best negative correlations to MSLP, the highest correlations (ranging from -0.3 to -0.4) lie in Northeast China and Qinghai, and the lower values are in Central China, especially in Sichuan and South Shaanxi.

The distributions of correlations obtained from the rest of the LSVs will not be described here. Overall, each LSV can provide large correlations to some stations. However, comparatively, the correlations obtained from each single LSV is small (mostly is smaller than 0.4), therefore combining multiple-LSV based predictors into one downscaling model is necessary to improve the modeling skills.

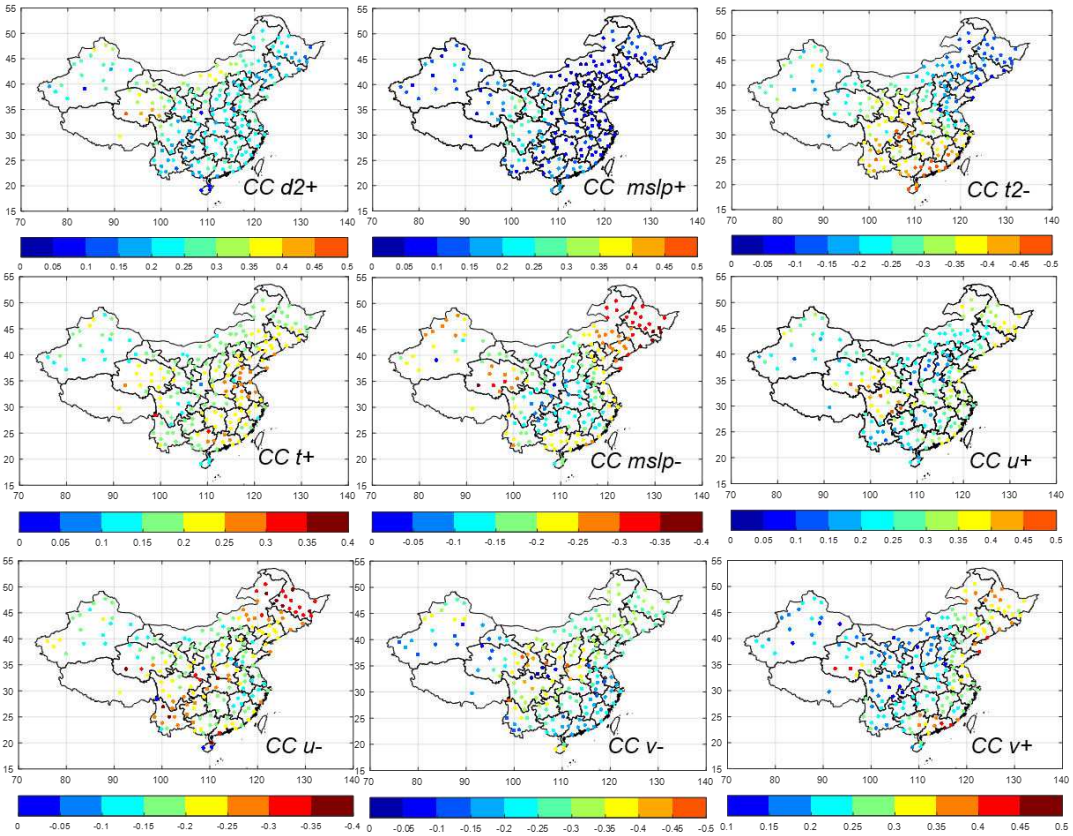


Fig.3 Optimal correlation coefficients (CC) at different sites, between observed daily precipitation (logarithmic transformed, 1979-2016) and the large-scale predictors in ERAI. Each colored dot represents a site, and the positions of the best grid boxes of the large-scales variables are not displayed here. Both positive correlations and negative correlations were used. Here, the subscriptions d2, mslp, t2, u, and v denote dew point pressure at 2 meters above surface, mean sea level pressure, air temperature at 2 meters above surface, surface wind velocity at U direction and surface wind velocity at V direction. The symbols '+' and '-' represent positive correlation and negative correlation, respectively. Most of the CCs displayed in the maps pass the significance tests at the 0.05 level (the CC threshold for this test is close to CC=0.06 for 4636 samples), and only the positive correlations at two sites and the negative correlation at one site to mean sea-level pressure do not pass the significance tests of 0.05 level.

3.1.2 GLM based PP models

The predictor series were extracted from the optimal grid boxes. The two temperature based predictors including the positive center and the negative center, but during the model training, the coefficients (model parameters) of these two temperature based predictors have different values, which implies that the combination of both predictors may give rise to an artificial trend, considering that air temperature are rising in future projections. Therefore, instead of using both the air temperature values at the two centers as independent predictors, the values of the two centers were combined into one predictor by using the difference between the two. Such a treatment can cancel out the increasing trend of the rising global air temperature. The common GLM used for all the sites is as follows:

$$Log(Pr+0.25)=x_0+x_1*P_s+x_2*R_h+x_3*U_{Max}+x_4*U_{min}+x_5*V_{max}+x_6*V_{min}+x_7*(T_{max}-T_{min}) \quad (6)$$

Where Pr is the logarithmic transformation of precipitation amount (mm). Adding a small value 0.25 is to avoid logarithmic transformation on zeros. P_s is mean sea level pressure (hPa); R_h is relative humidity; U_{max} and U_{min} are the values at the maximum (postive) correlation coefficient and the minimum (negative) correlation coefficients from U-directional wind (m/s), V_{max} and V_{min} are the values from V-directional wind (m/s), having similar meanings as that of U_{max} and U_{min}; T_{max} and T_{min} are the values at the maximum (positive) correlation center and the minimum (negative) correlation center of 2-meter surface air temperature (K). In following paragraphs, we use ΔT to represent the difference T_{max}-T_{min}. x_1, x_2, …x_7 are the coefficients/parameters to be estimated, and the estimated values for some representative stations are shown in **Table 1**.

Table 1. Values of the coefficients (parameters) of predictors at the seven representative sites, estimated by the LASSO-GLM training. Zeros indicate the parameters were removed by the LASSO algorithm to avoid overfitting problem.

	x_1 (msl)	x_2(rh)	x_3(u_{max})	x_4(u_{min})	x_5(v_{max})	x_6(v_{min})	x_7(T_{max}-T_{min})
Beijing	-0.00097	0.912	0.051	-0.073	0.021	-0.124	0.152
Hanzhong	-0.00029	0.926	0.087	-0.147	0.149	-0.340	0.118
Ulumq	0	5.107	0.050	0	0.049	-0.128	0.088
Yangjiang	-0.00088	1.659	0	-0.274	0.051	0	0.479
Tulihe	-0.00056	1.625	0	-0.083	0.137	-0.064	0.069
Maduo	-0.00077	2.214	0.019	-0.093	0	-0.046	0.096
Nanjing	-0.00066	0	0.268	-0.122	0.099	0.055	0.297

By combing multiple predictors, the GLM can produce output with larger correlations to the observed precipitation than that obtained from any single predictor. The variation of CCs (calculated from the validation samples of the cross validation) with different sample sets is generally very small, for each site. Therefore, finally we trained the models with all the 38-year samples to get a set of parameters, these parameters were used to downscale the precipitation in CMIP5. During this cross validation, the CCs (only those obtained from the samples for validation) are larger than 0.55 for most sites (**Fig.4** (a)). The correlations obtained from these surface LSVs are generally smaller than that obtained from the LSVs in multiple–pressure levels (**Fig. 4**(b)) as used in (Liu et al., 2019b).

The CCs obtained by most sites in Northwest China are comparatively low. The reason is that the rainy-day samples are very scarce in the dry areas, where the precipitation is usually sporadic and difficult to model. Meanwhile, the GLMs usually underestimated the annual total precipitation, which is very significant for the sites in Northwest China (in supplementary information: **Fig.S1, S3**). Such underestimation/biases is caused by the training process which uses both the dry-day samples and rainy-day samples, herein, the values in rainy days are underestimated while the values in dry-days are simulated as small precipitation values. By this reason, the less rainy days, the more significant underestimation has. According to our experience, the model training based on both rainy-day samples and dry-day samples, used in this study, obtains better CCs than does the traditional GLM training which uses only wet-day samples for precipitation amount. That is, training the GLMs with all the samples instead of

using only wet-day samples has the advantage of obtaining better correlation coefficients, but also bring about a disadvantage of underestimating the wet-day precipitation amounts.

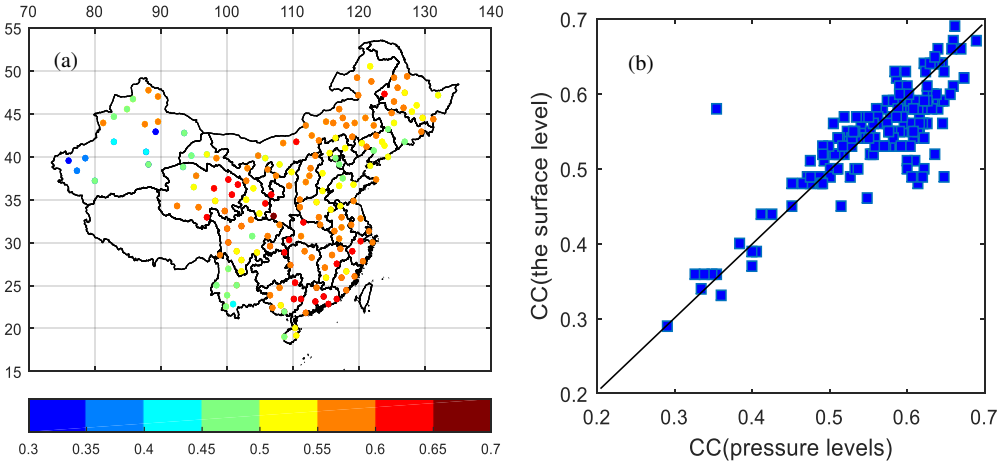


Fig.4 Pearson's correlation coefficients (calculated based on the daily logarithmic transformed series) obtained from the GLM based simulations for all the sites in China. (a) the distribution of CCs (based on the subsets of samples for validation, averaged for the four sample sets of cross validation) for the sites across China; (b) the CCs obtained by the models using surface-level LSVs versus the CCs obtained by the models using multiple-level LSVs (Liu et al., 2019b).

3.2 Downscaling on historical simulations in CMIP5

3.2.1 Spatial pattern of the climatology during the cross validation

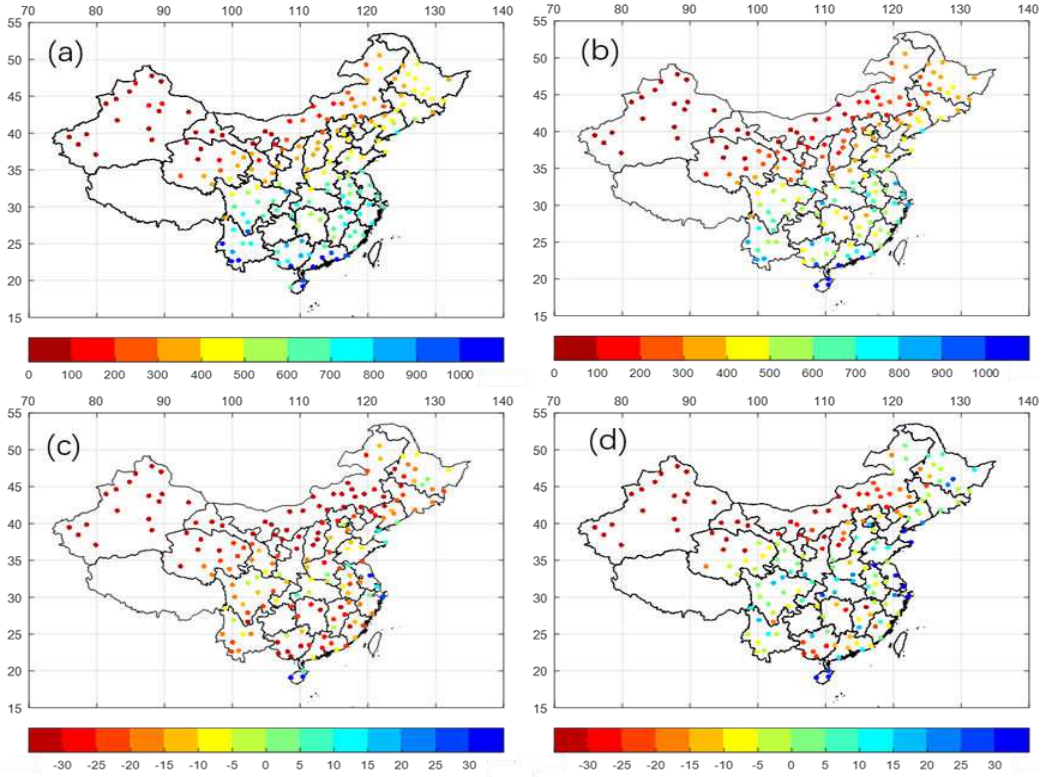


Fig.5 Climatological mean summer precipitation (mm) during 1961-2016: (a) observed annual totals (mm), (b) downscaled annual totals (mm) by the OGB-PP based on ERAI; (c) relative bias (%); (d) relative bias (%) after multiplying the downscaled values by an inflation factor of 1.2. The maps for climatological number of summer rainy days are not shown here, because the same number of simulated rain days can be obtained as the observed for each site.

Similar spatial patterns were obtained between the OGB-PP results and the observed ones across different sites in China (**Fig.5**), according to the validation on OGB-PP downscaling based on the large-scale predictors of ERAI. In order to visually compare with the observations, we multiply the simulated values by a factor of 1.2 which is an approximated average factor in all the 173 sites across China. After such a rough correction, biases still exist in different sites and ranges from -40% to 30%. For the northern and northwest areas of China, underestimations exist, and the overestimation exists in some sites of the eastern part of China.

3.2.2 Climatological mean for the CMIP5 historical simulations

The daily precipitation downscaled by OGB-PP are correlated to the results by BC-MOS (see supplementary information: **Text.S2, Fig.S4**). In following sections, we mainly focus on analysis of the climatological mean.

The spatial patterns in the observed precipitation across China can be roughly captured by the PP downscaled results from the GCM historical simulations. Here, only the relative biases of summer total precipitation and the absolute biases of the number of rainy days are presented (**Fig.6**). The changes of both mean summer total precipitation and the mean number of summer rainy days share more inner similarities among the different GCMs, while the simulated spatial patterns of climatological mean have more differences to the observed patterns(**Fig.7(d)**). All the downscaling results based on the three GCMs produced a wetter Southwest China, a drier South China and Central China. In Northwestern and Northern China, more wet days are in the simulations, but the relative bias are generally different: the most overestimation is produced by the CanESM-PP/EQM, while the least overestimation is produced by the HadGEM-PP/EQM; the areas with underestimation of wet days in the three PP/EQM simulations are very different. Similar patterns were generally obtained by the PP downscaling with predictors from QM and LT bias corrections.

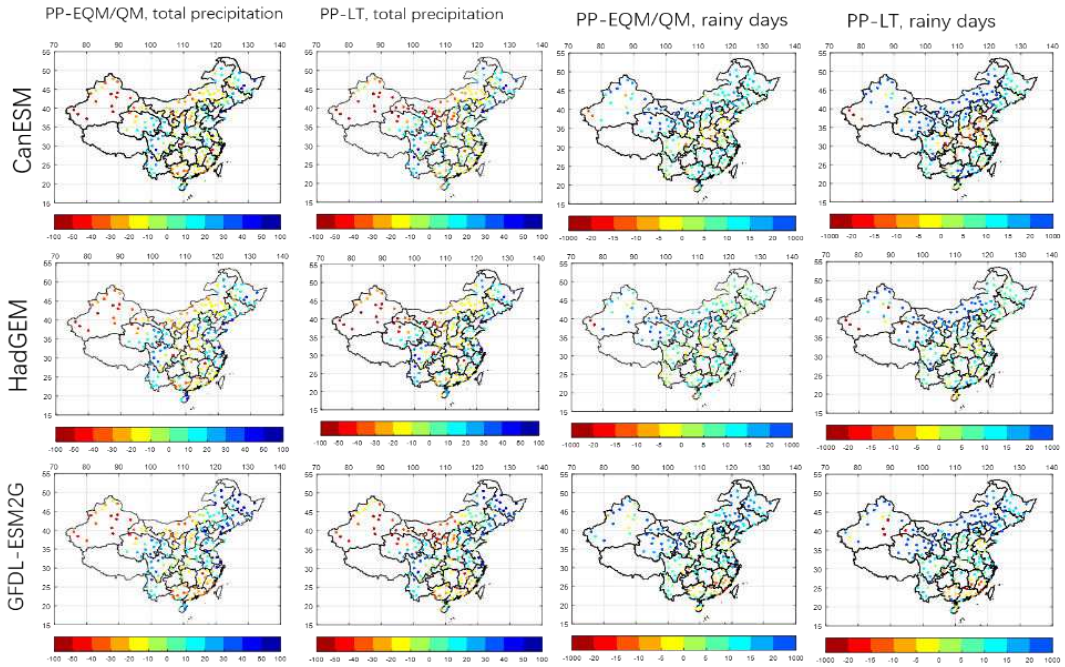


Fig.6 Relative biases (%) of annual summer total precipitation (in the first two columns) and the change of rainy-day number downscaled by OGB-PP driven by CMIP5 Historical simulation (1962-2005). The downscaled results from the predictor-bias correction by QM have small differences to that displayed here for EQM.

The climatological means simulated by BC-MOS are inherited from the GCM simulated precipitation. All the three produced a wetter Sichuan province and a drier area in Eastern China: Shandong, Jiangsu and Anhui. Among the three GCMs, HadGEM2-ES with bias corrections produce the closest pattern to the observed. Comparatively, GCMs produced small underestimation for heavy daily precipitation but also produced too many slight-precipitation days (**Fig.7**(e)). The bias corrections can improve the statistical distribution of daily precipitation amount for each site (**Fig.7** (f)).

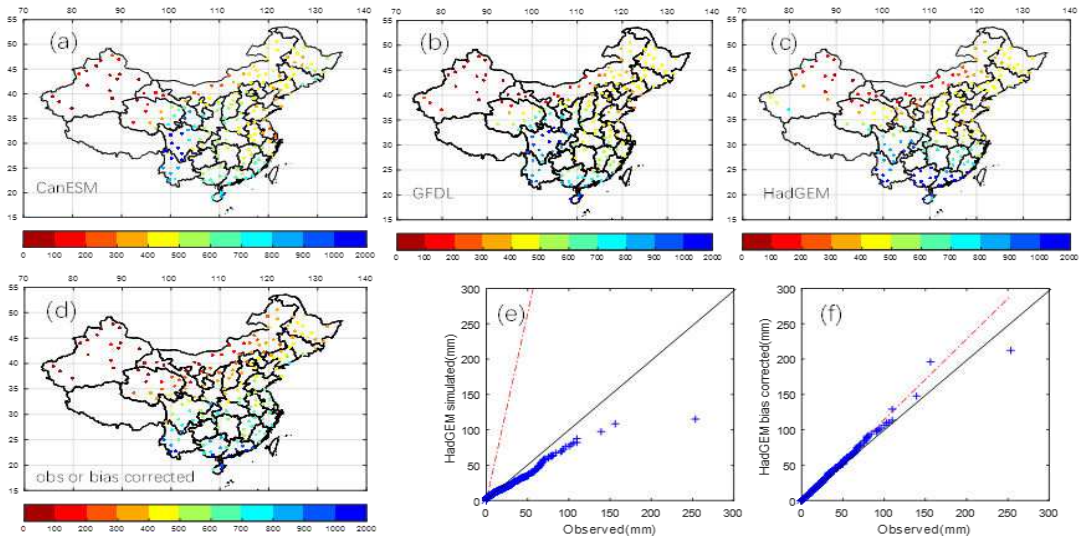


Fig.7 Comparison between the results of BC-MOS and the observation. (a-c) Climatological mean of summer total precipitation downscaled by BC-MOS based on CMIP5 Historical simulations (1962-2005) of three GCMs; (d) The observed climatological mean of observed summer total precipitation (the same pattern must be obtained by the different bias corrections on the GCM simulated precipitation, theoretically). (e) Quantile-quantile (QQ) plot on daily precipitation (Beijing), the HadGEM2-ES simulated large-scale precipitation versus the observed; (f) QQ plot for the bias corrected HadGEM2-ES precipitation (Beijing) versus the observed. Similar effects were obtained for the other two GCMs (CanESM2, GFDL-ESM2G) and for other sites.

3.2.3 Interannual variations for the CMIP5 historical simulations

For most of the sites, the different predictor-bias corrections have small influence on the precipitation downscaled by OGB-PP, therefore, here only the OGB-PP downscaling results with the predictor-bias correction of QM are presented.

It is reasonable that the simulated variations of annual total precipitation at each site cannot be directly compared with the observations, since the GCM simulations are not synchronized with the observations. However, the comparison of the smoothed interannual variations between the simulated and the observation may be feasible (**Fig.8**). For Beijing and Tulihe, the downscaled result from CanESM-PP/QM got the best trend. The lower tails of annual precipitation for Tulihe and Beijing after 1998 are only reproduced by CanESM-PP/QM. GFDL-PP/QM performed poorly at the long-term variations at Beijing. For the site Urumqi, CanESM-MOS/QM performs well. For Yangjiang, GFDL-MOS/QM performs the best. In some cases, the results from HadGEM/PP also perform well. Nevertheless, whether these interannual variations are able to actually reflect the driving emission forces or are just the internal variations is difficult to justify here.

For the sites with the Pearson's CC (based on the smoothed ones) larger than 0.4 across China, the numbers of these sites are: 45 (CanESM/PP), 37(HadGEM/PP), 55(GFDL/PP), 45(CanESM/QM), 30(HadGEM/QM), 45(GFDL/QM). For CanESM, the same numbers are

obtained by the PP and the QM. But for the other two GCMs, the numbers obtained by the PP is larger than that from the QM. It is difficult to conclude that the PP is better than the QM.

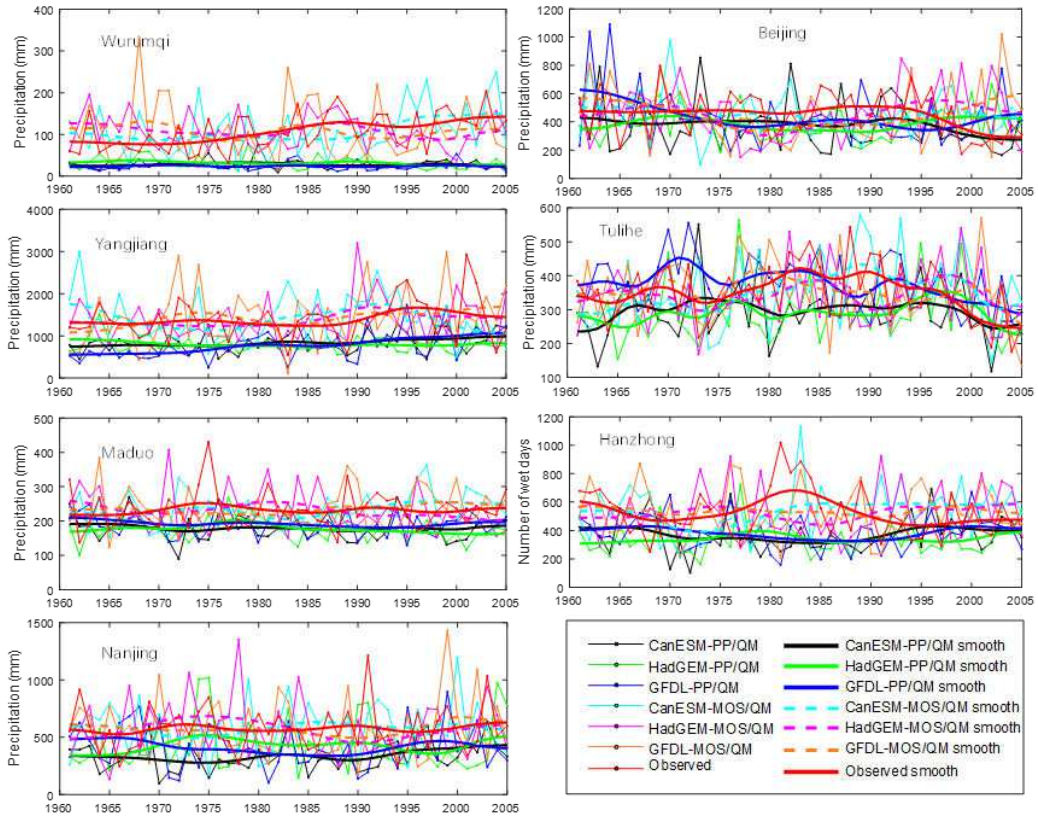


Fig.8 Annual total summer precipitation downscaled by different technical options for the seven representative stations

3.2.5 Linear trends for CMIP5 historical simulations

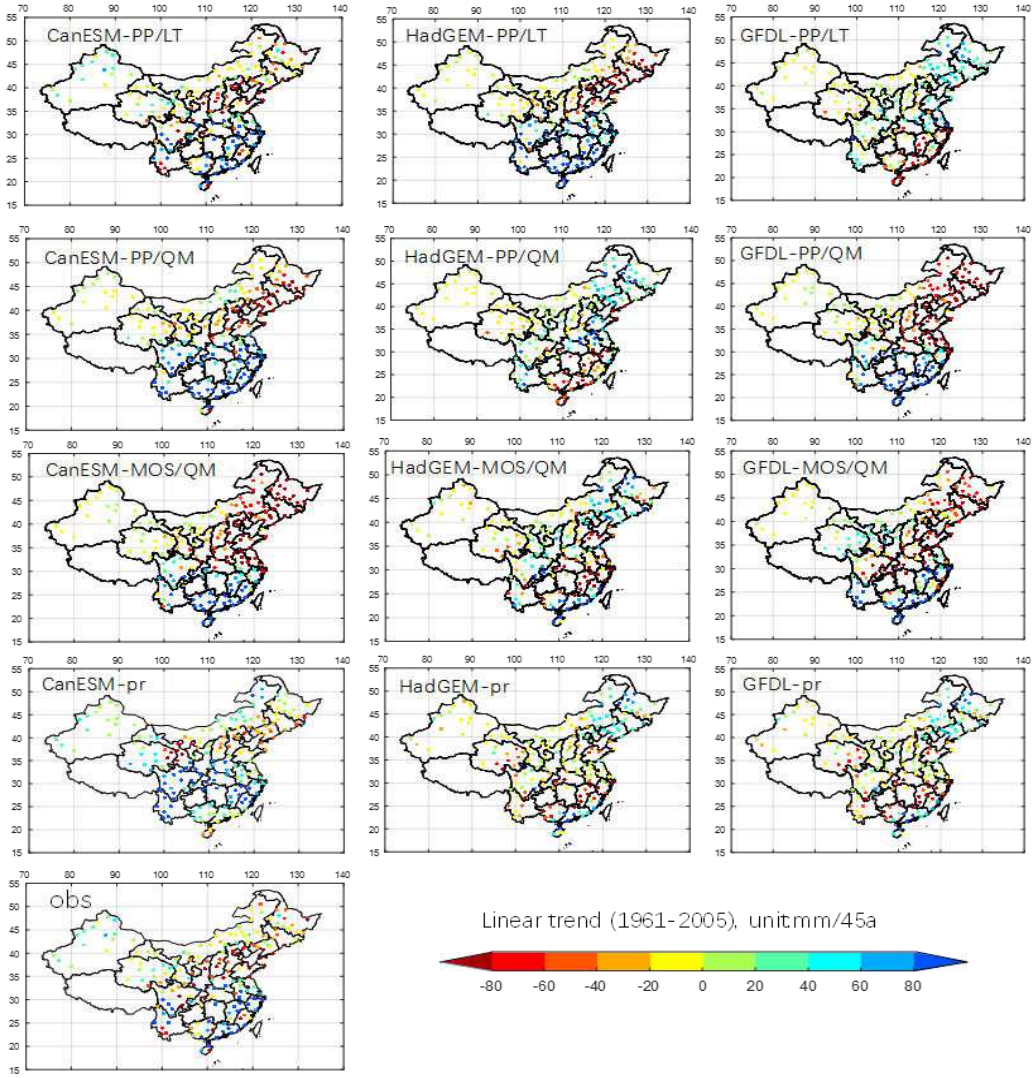
During 1961-2005, the observed summer precipitation has different trends across China: decrease in the south-eastern part of China at most sites, but increasing trends also exist at some sites (mostly around the Yangtze River basin) which are mixed with other sites of decreasing trends. In Xinjiang and most of the sites in Northwestern China show light increasing trends.

The downscaled patterns by different methods and different GCMs are different (**Fig.9**). The trends in the downscaled precipitation based on different technical options are listed in **Table 2**. Generally, the results from CanESM2 are more consistent to the trend of the observed than those from other GCMs. For HadGEM2-ES, the OGB-PP/LT has consistent pattern to the observed precipitation, but the OGB-PP/QM has many contrary trends and the similar contrary trends were also produced in the outputs of BC-MOS/QM downscaling. For GFDL-ESM2G, the GFDL-PP/LT and the original large-scale precipitation produced contrary pattern to the observed, but GFDL-PP/QM and GFDL-MOS/QM produced general consistent trends. Above results demonstrates that the trends are sensitive to the different technical options (bias correction methods, in both OGB-PP or BC-MOS), and it is difficult to tell which option is

559 better than other options. In summary, most of the contrary trends were produced by
560 HadGEM-PP/QM, HadGEM-MOS/QM, GFDL-PP/LT and GFDL-PR. The output of every
561 GCM has one or two precipitation downscaling results (with different technical options) which
562 can reproduce the observed pattern.
563
564
565 **Table 2** Long-term trends of the downscaled precipitation from CMIP5 and the observed
566 precipitation during 1961-2005. The outputs with significant contrary trend to the observed are
567 marked by asterisks (*).

	Northeast China	North China	Xinjiang	Loess Plateau	Lower Yangtze River Basin	South China	Southwest China
CanESM-PP/LT	Lightly drier	Drier	Slightly wetter	drier	Highly wetter	wetter	Drier and wetter, mixed
CanESM-PP/QM	Highly drier	Drier	No change	drier	Highly wetter	Highly wetter	Highly wetter
CanESM-MOS/QM	Highly drier	Highly drier	No change	drier	wetter	Highly wetter	wetter
CanESM-PR	Wetter north, drier south	Drier	Slightly wetter	Drier and wetter, mixed	wetter	Slightly wetter	wetter
HadGEM-PP/LT	Highly drier	Drier	No change	drier	Wetter	Highly wetter	Wetter
HadGEM-PP/QM	Wetter*	Wetter*	No change	Drier and wetter, mixed	Drier*	Highly drier*	wetter
HadGEM-MOS/QM	Wetter*	Mixed	No change	Wetter*	Drier*	wetter	Mixed
HadGEM-PR	Wetter*	Drier	Mixed	Slightly drier	Drier*	Wetter	Drier*
GFDL-PP/LT	Wetter*	Wetter*	Slightly drier*	Slightly wetter*	Drier*	Highly wetter	Wetter
GFDL-PP/QM	Highly drier	Highly drier	Wetter east, drier west	Drier	Wetter	Highly wetter	wetter
GFDL-MOS/QM	Highly drier	Drier	Wetter east, drier west	Mixed	Mixed, drier west, wetter east	Wetter	Wetter
GFDL-PR	Wetter*	Mixed, mainly drier	mixed	mixed	Drier*	wetter	Drier*
Observed	Drier	Highly drier	wetter	Highly drier	Wetter	Wetter	Mixed, wetter

568
569



570

571 **Fig.9** Downscaled Linear trend (mm/45a) of annual summer precipitation driven by CMIP5
572 historical simulation (1961-2005). The marks 'PP' and 'MOS' indicating OGB-PP and BC-MOS,
573 respectively. The results obtained by EQM and QM are identical for the downscaling of
574 historical simulation.

575

576 **3.3 Future projections**

577 **3.3.1 Sensitivity to different bias corrections**

578 **a. OGB-PP downscaling**

579 For the OGB-PP downscaling on historical simulations, there is almost no significant
580 differences caused by the different predictor-bias corrections, but for the downscaling on
581 RCP4.5, some differences appeared in the results (**Fig.10**). For the results of some stations
582 (Beijing, Hanzhong and Yangjiang) based on CanESM and GFDL, the EQM correction
583 produces an apparent downward shift from the historical simulations to the RCP4.5 simulations

584 in 2006 from which the simulation of future projections started. For the Maduo station, the shift
585 of EQM is upward. This shift is apparent throughout the period from 2006 to the end of the
586 century. One may think that the differences also can be presumed as contrary shifts caused by
587 QM and LT. However, theoretically, the LT correction is impossible to produce any downward
588 or upward shift between the historical and RCP4.5 simulations. From **Fig.10**, QM produces no
589 difference to that of LT in the several years after 2006, but EQM produced such apparent
590 difference. This demonstrate that the EQM, instead of QM and LT, can produce this shift.
591 For the Yangjiang station and the PP downscaling from CanESM, the difference of EQM to
592 QM is large in the early years in RCP4.5 simulations, but become close by the end of the
593 century. EQM produces an increasing trend, and QM gives no trend. The output of LT produces
594 a strong increasing trend in the RCP4.5 simulations, which becomes more different to that of
595 EQM and QM. The output from GFDL also has the similar different trends as that existed from
596 CanESM, for the three bias corrections, while the output from HadGEM has no large differences
597 caused by the three bias corrections. Similar comparison effect was also obtained for the site
598 Guangzhou (not displayed), which is very close to the station Yangjiang. Overall, for most of
599 the stations and the PP downscaled outputs driven by the three CMIP5 simulations, the bias
600 correction of EQM and QM on predictors produce very similar trend and variations of annual
601 total summer precipitation.
602
603 **b. BC-MOS downscaling**
604
605 The BC-MOS downscaling by the EQM and QM are identical in this study, which have
606 different variations to the LT based correction (**Fig.10**). The CanESM based downscaling for
607 the site Urumqi is an extreme case: a different trend is produced by the QM/EQM correction.
608 Such different trends may imply that bias corrections are not always credible (Maraun, 2013b;
609 Maraun et al., 2017). For Beijing, almost all the downscaled results show increasing trend,
610 although the differences are also large. For other sites, the trends simulated are different, which
611 means there is a large uncertainty in the projections.
612

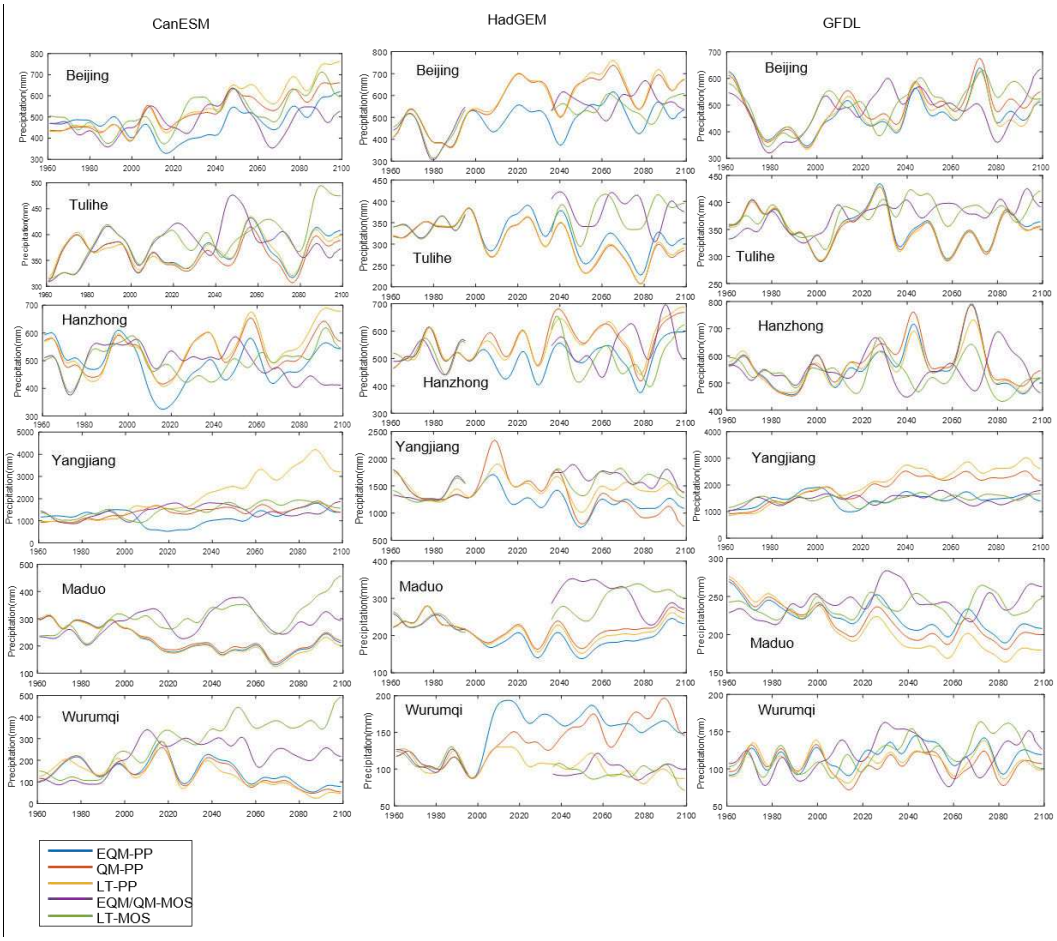


Fig.10 Interannual variations of total precipitation amounts (smoothed version) simulated by OGB-PP and BC-MOS, from historical simulation to the future projection of the RCP4.5 scenario, based on three CMIP5 simulations. In BC-MOS, EQM and QM has almost the same result.

3.3.2 Future Changes

The OGB-PP downscaling without any corrections on the predictors can give rise to very large annual summer precipitation, which is 4-5 folds of the observed. This indicates the biases between the GCM simulations and the ERAI need to be corrected before downscaling, when the absolute values of precipitation are our main concern. Nevertheless, such a downscaling result may also be valuable for analyzing future relative changes. After all, any bias correction can also give rise to further uncertainties. The relative changes produced by this no-correction scheme reflect large spatial differences in different areas of China: some areas become 40% wetter, but some places become 40% drier (The spatial patterns are not presented in this paper). The changes in this no-correction scheme is very different from that in OGB-PP based on bias-corrected predictors, for some areas. After the bias correction, the drier projection in some sites can be changed to wetter projection, and vice versa. For example, the sites in Southeastern China in CanESM based OGB-PP downscaling, and the same area in GFDL based OGB-PP downscaling (**Fig.11-12**).

For the five stations of Beijing, Tulihe, Hanzhong, Yangjiang and Maduo, the annual summer total precipitation downscaled based on different GCM simulations has varied interannual trends (**Fig.10**). At Beijing station, the result based on CanESM has a stable increasing trend in the 21 century, but for HadGEM and GFDL, the trends are not significant. At Tulihe station, the result based on HadGEM and GFDL show decreasing trend, but that of CanESM has an insignificant trend.

a. Patterns of the four indices

The downscaled patterns are generally similar among the four indices (see **Fig.11-12** and **Fig.S5-S6** in supplementary information). For example, a site having larger increase of 'q95' generally get larger increase of 'tsp', shorter 'dds' and less 'nrd'. Almost equivalent patterns are obtained by BC-MOS methods with QM and EQM bias corrections, which are different to the patterns in the original GCM simulated large-scale precipitation. The OGB-PP results based on different bias corrections all obtained different change patterns, but also generally share more common characteristics when compared to the patterns obtained by BC-MOS. The changes obtained by OGB-PP are more drastic (usually >30%) than the changes obtained by the BC-MOS and that in the original GCM precipitation. Among the different GCMs, the changes from GFDL/PP are comparatively mild (usually <20%) when compared to the other two GCMs, and the most drastic changes are produced by HadGEM/PP.

Here, based on the case of the 'q95'(**Fig.S5** in supplementary information), the change patterns are described as follows:

CanESM/MOS: The two bias corrections have changed the trend. In the original CanESM simulation, most of the stations show increasing trend. Especially in western China and northeastern China, the relative change of total annual summer precipitation is larger than 20%. After the QM bias correction, the increasing trends become smaller than the original at many stations in northern China, and the trends in southeastern China become decreasing, although the decreasing is generally slight.

CanESM/PP: Unlike the BC-MOS downscaling based results which have generally increasing trends, the results of OGB-PP downscaling has both large increasing and decreasing in different sites. The downscaled results based on different bias-corrected predictors are different. The LT corrected predictors produce very significant increases in the north part of eastern China (mainly in Shandong and Jiangsu) and southern China (Guangdong and Guangxi), and the increasing is mostly larger than 30%. Southwest China has an increase larger than 10%. Very significant decrease (>20%) is produced in Hunan. In the results based on EQM and QM, Eastern China and Southeastern China, the west part of Inner Mongolia get an increase change (>10%) and the areas in Central China, Northwest China, and Northeastern China have a slight decreasing or no significant change (within 10%).

HadGEM/MOS: The original increasing is smaller than the two bias-corrected results, and the latter two are generally similar to each other. The relative change (increasing) along a long belt from South China to Northwest China is enhanced by the bias corrections.

HadGEM/PP: The main increasing happens in the areas of Central China north to the Yangtze River, Gansu Corridor and the northeastern part of Qinghai. This increase in the results

based on QM is larger than 30%. The results based on the three bias corrections are very different, and the QM based one got the largest increase. The EQM get a significant decreasing area in Eastern China in the west to Shanghai. The LT get a large decreasing (change larger than 10%) area in south to the Yangtze River.

GFDL/MOS: The increasing is slightly reduced for most stations in China, except for the stations in Xinjiang and the northern part of Southwest China, where the significant increasing is lowered by the two bias corrections.

GFDL/PP: All the three bias corrections on predictors produced an increasing trend (20%-30%) in the large areas of the Yangtze River basin and the Huai River basin. A slight increasing trend (>10%) is produced in Northern Xinjiang. Meanwhile, a slight drier trend (<10%) is produced in most areas of the northern part of China, including Southwest Xinjiang, Qinghai, Inner Mongolia, Northeastern China. Large difference is produced in South China: slightly drier trend (<10%), highly wetter trend (>40%) and slightly wetter trend (20%-30%) are produced by the predictor corrections of EQM, QM and LT, respectively.

Overall, for the simulated precipitation by the three GCMs, the two bias-corrected results are generally similar although they have different changes at some specific stations. For most of the stations, the direct CMIP5 precipitation show wetter trend areas and generally no significant drier trend areas.

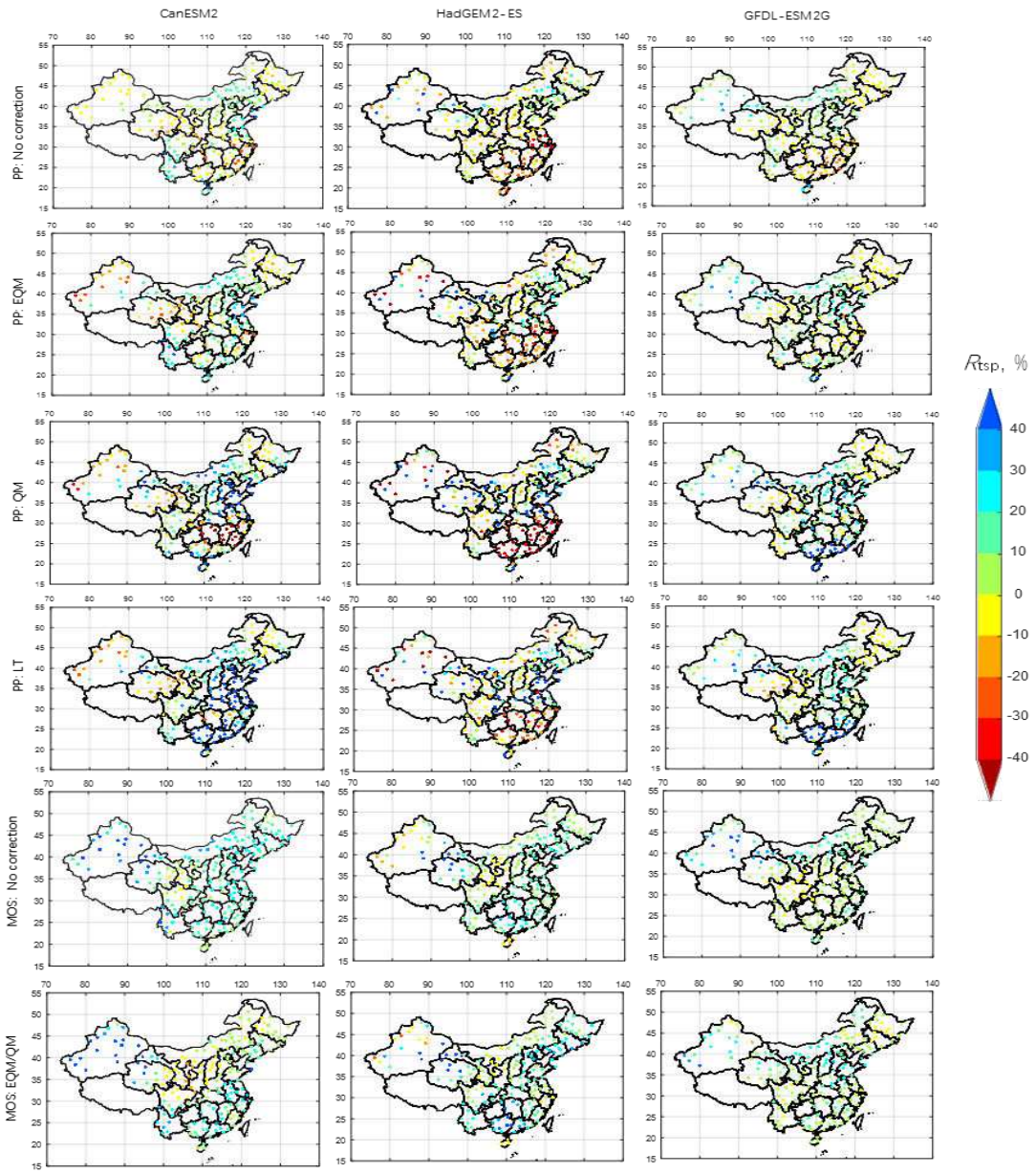


Fig.11 Relative change of annual summer total precipitation (R_{tsp}, unit: %) of 2060-2099 relative to 1960-1999, downscaled from RCP4.5 simulations. The MOS/LT is not listed here considering that the linear correction will produce the same percentage change as the original precipitation.

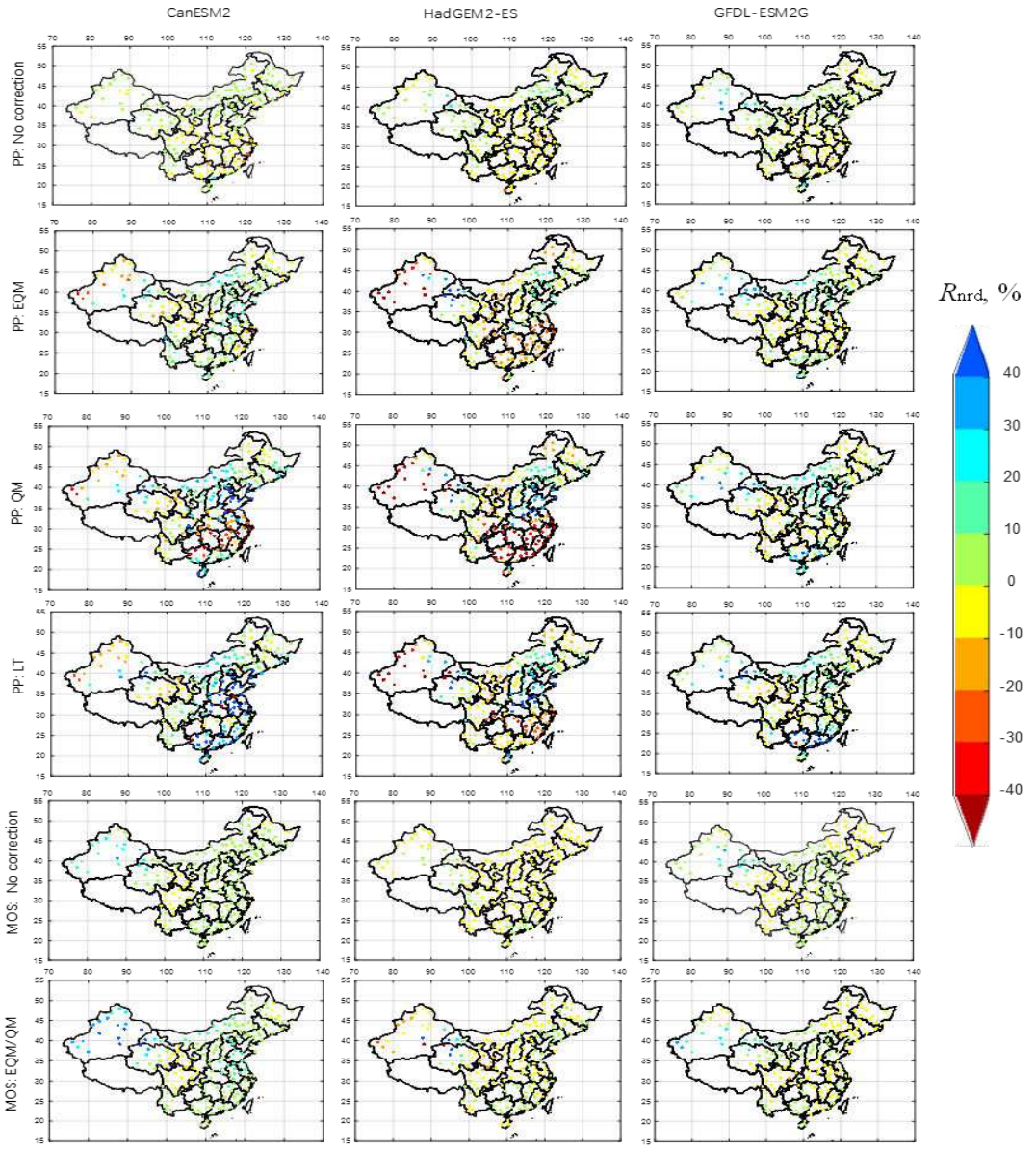


Fig. 12 Similar to Fig.11, but for the change of the numbers of summer rainy days (R_{nrd}, unit: %).

b. Comparison to previous studies

To the end of 21st century, the BC-MOS projected wetter changes for the annual total summer precipitation across all the sites in China, although the changes are different in different areas. OGB-PP produce wetter changes at a large number of sites, but also produce drier changes in some areas which are different among different GCM simulations. However, some similarities are also shared by different downscaled results. For example, all the downscaling results show the slightly drier trend on the west part of Loess Plateau, and no wetter changes are produced for the Sichuan Basin. According to most simulations in this study, west Xinjiang has a high probability of drier change but the large area of Eastern

Xinjiang, west Gansu Corridor and Northwest Qinghai has a contrary trend; North China has a high probability of wetter change; Inner Mongolia and Northeastern China has a high probability of no significant change. There is also a high probability that in South China, an area to the south of Yangtze River, will become drier, which is simulated by OGB-PP, while BC-MOS produce slightly wetter changes.

The above changes in different areas is similar to that obtained by previous studies: the downscaling using bias corrections on multiple-model CMIP5 ensembles (**Table 3**). Such ensemble MOS downscaling results can be regarded as a reference to test the performance of new PP downscaling models, since there is no observation of future climate. Particularly, the decreasing precipitation in south to Yangtze River is supported by the synthesis of multiple CMIP5 models (Zhou et al., 2017). Nevertheless, these previous studies simulated the decreasing trend in the entire area of South China including Guangdong and Guangxi provinces. In this study, the overall PP downscaling produces the increasing precipitation in Guangdong and Guangxi. Overall, the OGB-PP downscaled result in this study can be regarded as general well, if we consider the previous ensemble results are credible.

Table 3 Long-term trends of downscaled precipitation in different areas of China in previous studies

PP-This study	QM-this study	(Yang et al., 2018)	(Yang et al., 2019b)	(Zhang et al., 2019b)	(Hui et al., 2018).	(Bao et al., 2015)
Downscaling method		Bias correction	Bias correction	Bias correction	Dynamical downscaling	Dynamical downscaling
Drying Central South China	Slightly wetter or drier	drier	drier	drier	drier	drier
Drying Loess Plateau	Slightly drier or wetter	drier	Partially drier	Slightly wetter	Partially drier	wetter
Drying Sichuan Basin	Slightly wetter	drier	drier	drier	drier	drier
Small change in Inner Mongolia and Northeastern China	Slightly wetter	Slightly wetter	Slightly wetter	Slightly wetter	Small changes	Slightly wetter
Drying west Xinjiang	2/3 support drier	Slightly wetter or drier	Slightly wetter	Slightly wetter	Drying areas embedded in wetter areas	wetter
Wetter Eastern Xinjiang, west Gansu Corridor and Northwest Qinghai	Wetter	Wetter	Wetter	Wetter	wetter	wetter
Wetter North China	wetter	wetter	Slight wetter	wetter	wetter	wetter

The GFDL/PP with no bias-corrections on predictors projected a drying change in Central-South China, but the GFDL/PP using three bias-corrections all projects a slight

wetter or drier change of this area. Based on the GFDL-ESM2G simulation in CMIP5, the dynamical downscaling by Bao et al. (2015) has also projected the decreasing of summer precipitation of this area in the near future (2031-2050).

c. Trends in predictors of OGB-PP

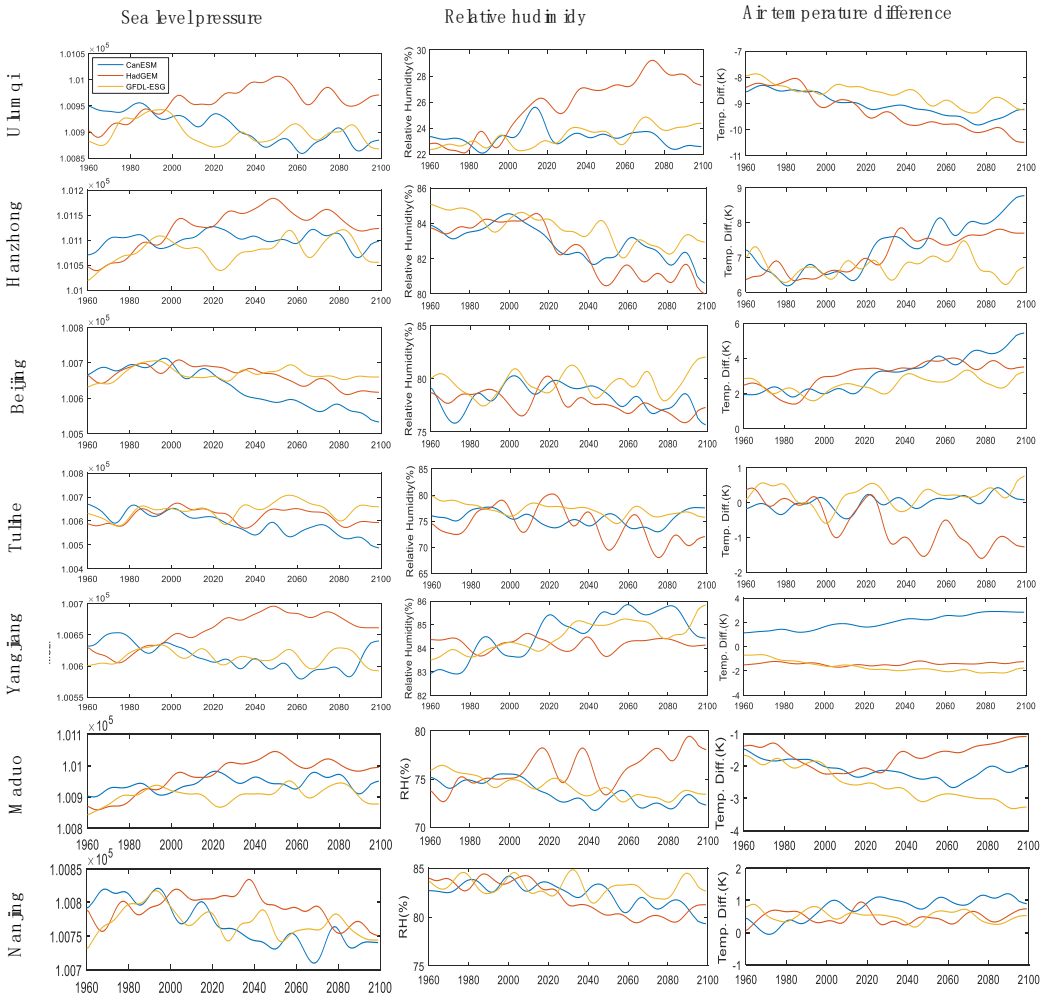


Fig. 13 Five-year smoothing variations in predictors (mean sea level pressure (P_s), relative humidity (Rh) and the difference of air temperature ($\Delta T = T_{max}\text{-}T_{min}$)) of the OGB-PP models

Compared to the CanESM and the GFDL-ESM2G, for most of the cases, the HadGEM presents more eccentric trends of predictors (**see Fig.13-14**). For a few cases, the HadGEM also generally produced similar trends as that produced based on the other two GCMs.

For the site of Beijing, almost all the PP downscaling variants based on all the three GCMs produced significant increasing trend in annual precipitation (**Fig.10**). The trends of predictors from different GCMs are also generally similar: decreasing MSLP, no change /slight change of RH, and increasing ΔT (**Fig.13**). The predictors for the site Hanzhong is somehow similar to that

for Beijing, except that the CanESM gave a contrary trend of MSLP: increasing. Tulihe also got similar downscaling trends (the only exception is the contrary trends of ΔT), but the results are more complex than that for Beijing. These three sites have similar trends of predictors may be attributed to that they generally share the same summer monsoonal system: the rain belt in northern China. Beijing and Tulihe has decreasing MSLP predictors and increasing ΔT predictors, which means the low-pressure cyclone and the convergence of warm air mass happens more frequently in the rest of this century. Nevertheless, most of the GCMs provided the decreasing RH at the 3 sites, which may be due to that the rising air temperature reduces the relative humidity in most part of the northern China. However, the interannual fluctuations of predictors and the downscaled precipitation in these sites becomes more complex in the rest of the century, implying that more extreme precipitation and drought may happen due to the climate change.

For the site Yangjiang in South China, the CanESM and the GFDL produced downscaling of increasing trends of summer precipitation, while the HadGEM produced a contrary trend (**Fig.10**). CanESM and GFDL provides the predictors of decreasing MSLP predictor, which may be related to more frequent Typhoon or monsoonal precipitation in South China, but the HadGEM provides a contrary trend of MSLP (**Fig.13**). All the 3 GCMs provides the increasing RH at Yangjiang. The ΔT predictors for this site has slight changes: slight increasing (CanESM) and slight decreasing (HadGEM and GFDL).

In Xinjiang, very different precipitation trends were produced by the GCM basis for the site Urumqi: decreasing (CanESM), increasing(HadGEM), no change or slightly increasing (GFDL) (**Fig.10**). Such projections also can be reflected by the complex trends of predictors. The predictors from the HadGEM are in more exceptional trends than others: increasing MSLP, increasing RH, while in the other two is decreasing MSLP, small changes of RH (**Fig.13**). This contradiction of trends presents an uncertainty. All the GCMs produced decreasing ΔT, which may imply that cold waves from the north pole become less.

For the site Maduo, at the northern part of Qinghai-Tibet plateau, decreasing precipitation were produced by the PP models based on all the three GCMs (**Fig.10**). For CanESM and GFDL, the decreasing is mainly attributed to the increasing MSLP predictors, decreasing RH and decreasing ΔT (**Fig.13**). For HadGEM, the increasing MSLP, increasing RH and the increasing ΔT (**Fig.13**) makes the precipitation trend very unclear and complex, and the extreme precipitation and drought are more likely to happen.

4 Discussions

4.1 Why the downscaling variants produce diverse trends?

The different downscaling variants in this study produce diverse trends of precipitation for future projections (**Fig.11-12**). Here, we divide the difference of trends into two groups: the difference between BC-MOS and OGB-PP and the difference caused by bias corrections on predictors. Comparatively, the former difference is more significant than the latter. Here, we assume the reason for the former one.

BC-MOS produced mild precipitation increasing at all the sites in China, just as that directly produced by the original GCMs, while OGB-PP produced comparatively large increase/decrease trends. Meanwhile, the trends from different versions (with different bias corrections on predictors) of OGB-PP downscaling are also largely different in some areas of China.

The reason of such difference may be that the multiple parameters in OGB-PP can utilizing more information than BC-MOS can. GCMs always produce an area mean of precipitation and can only reflect a low spatial variability due to limited spatial resolution, therefore, precipitation produced by GCMs are always in mild trends. For BC-MOS, such mild long-term trends are inherited from the GCM outputs, since the predictor and predictand are the same variable. By adjusting the distributions of data, bias correction is unlikely to change the trend to a large extent. As Maraun et el. (2019a,b) pointed out, bias corrections also cannot produce additional local variability and MOS methods rely highly on the GCM modeled precipitation.

For a PP, it is another case: it uses multiple variables from a GCM and modeling precipitation by its independent mechanism. More importantly, the OGB-PP in this study freely utilizes the predictor values at the grid boxes which usually lie in the locations inconsistence to the site. This mechanism for modeling precipitation in OGB-PP models can be regarded as an analog of that in GCM, since the predictors used by both of them are in the same resolution. In another words, the PP models has a different and independent precipitation modeling mechanism from GCM, which can enrich the diversity of outputs. At the locations of some neighboring sites, although the large-scale variables at these grid boxes have relatively close values, the models trained for the sites generally have varying parameter values and freely use multiple predictors from different grid boxes other than the grid box close to the site. For example, the model of one site maybe mainly sensitive to the predictors from sea level pressure, but the model of a neighboring site maybe relatively sensitive to the predictors from air temperature. This characteristic give rise to more diverse outputs over different sites than that from the MOS, thus can contribute additional local variability.

Of course, the multiple PP predictors should also give rise to more uncertainty than the MOS. For OGB-PP, not only the model parameters should produce uncertainty, but the preliminary predictor-bias corrections also can contribute uncertainty. Therefore, taking the ensemble of multiple outputs of PP downscaling variates should be more meaningful than taking one downscaling result.

4.2 Whether the OGB-PP downscaling is sensitive to different GCM bases?

Different trends of future precipitation were produced by OGB-PP downscaling based on the three GCMs (see Fig.13-17). Here we only discuss the predictors MSLP, relative humidity and air temperature difference (**Fig.13**), since the U-wind and V-wind are also correlated with these three variables. For example, the projections for Beijing based on all the GCMs produced significant increasing trends, however, OGB-PP driven by the three GCMs presents different predictor trends for Beijing: decreasing MSLP and increasing air temperature difference are given in CanESM and HadGEM, while the trends in GFDL based predictors are not so significant. For the site Yangjiang, the projections based on CanESM and GFDL have increasing trends, which can be mainly attributed to the increasing relative humidity and air temperature

difference. The downscaling based on HadGEM for Yangjiang has a decreasing trend, which can be attributed to the increasing MSLP and a slight decreasing of relative humidity. Here, it need to be pointed out that the similarity produced by CanESM and GFDL for Beijing and Yangjiang are not always true for other sites (not listed). According to the above facts, the PP downscaled precipitation trends in this study are sensitive to the predictors from GCMs, which require that the large-scale projections be reasonably produced by the GCMs.

4.3 Which downscaling options are more promising?

The options for downscaling are multiple: the choice of GCM in CMIP5, the choice of using PP downscaling or MOS downscaling, the choice of predictor bias corrections for PP downscaling and the choice of bias corrections for MOS downscaling. Actually, there is no ideally perfect combination of different options, because uncertainty exists everywhere.

For GCMs which provide LSVs for downscaling, we prefer multiple GCM simulations other than one simulation of a certain GCM, because no GCM is superior to all other GCMs according to many previous studies.

As for choosing which bias correction method for the predictors, we prefer the LT first in both PP downscaling and MOS downscaling, since LT only modify the mean and standard deviation, while quantile mapping modify the statistical distribution in a more inexplicable way. We prefer the no-correction method in PP downscaling to be the second better one, if only the relative changes are our main concern. However, it does not imply that the QM are completely unusable. In most cases, there is some biases or incorrect statistical distribution in the original predictors, the QM based bias corrections can give a better distribution to reduce the overall biases.

Is the OGB-PP better than the MOS (bias corrections)? We guess the answer is yes, based on several reasons. Generally, the BC-MOS do not change the trend in the coarse resolution GCMs. Therefore, the downscaled relative changes are mostly inherited from the driving GCM simulations(Maraun et al., 2019b), which can explain why the MOS downscaling results in this study show increasing precipitation at all the sites across China. Although the trend at a few specific sites can be changed from wetter to drier or vice versa by the bias correction, this transformation is based on a pure statistical process other than on a physical basis. This problem of quantile mapping was criticized by Maraun et al. (2017). On the contrary, the PP downscaling uses physically based multiple predictors which can represent the main large-scale circulations (Maraun et al., 2019b), and theoretically, more signals can be utilized than the MOS.

Of course, PP is also very limited in producing additional local variability, since all the predictors used are derived from the driving GCM(Maraun et al., 2019b). Nevertheless, the OGB-PP models are independent to each other among different sites, so each model has a unique set of parameters which are locally estimated. These parameters themselves carry some local variability to both daily precipitation and long-term precipitation trend. Note that the BC-MOS methods are also specific to each site, but theoretically carry smaller local variability than the PP, because the latter utilize more predictors.

Additionally, some of the PP downscaled results in this study can produce the decreasing precipitation in the areas south to Yangtze River (from Central China to South China) as manifested in many previous studies (as listed in Table 1), while the MOS not.

4.4 Physical meanings reflected by the PP-model predictors

One advantage of PP downscaling over MOS is the use of multiple physically based predictors, which is helpful to understand the physical mechanism of the future changes. Nevertheless, the predictors for downscaling daily precipitation is the localized signals in LSVs and cannot directly reflect the physical mechanisms in the global scales simulated by the GCMs, although these local signals are also affected by the mechanisms in the global scales. In this study, it is unsuitable to tangle the influence of changes in global scales, due to the many sites and multiple regions in this study. Here we only present some discussions on the changes of such localized predictors.

The three base CMIP5/GCMs simulations provided different downscaling predictors on which it is difficult to get a good conclusion, since the number of samples is too small. Generally, we can get several characteristics for the trends of these predictors:

Firstly, the different GCMs generally provided similar predictor trends for some sites. For example, Nanjing, Beijing, Tulihe and Hanzhong. Notice that the sites in the North China (Hanzhong, Beijing) and Northeast China (Tulihe) are in the same northern part of the summer monsoonal area in China, while Nanjing is in the southern part of the monsoonal area. For Beijing and Tulihe, the downscaled increasing trends of precipitation at are mainly related to the decreasing MSLP and increasing ΔT.

Secondly, for some cases, the contribution of the trends in different predictors to the trends of downscaled precipitation is contradictory. For Beijing, Tulihe, the RH predictors are in decreasing trends, which is contrary to the trends of MSLP and ΔT.

Thirdly, different trends were provided by different GCMs. for example, at Yangjiang and Ulumqi, the HadGEM has different trends in MSLP and RH when compared to the CanESM and GFDL. Similar exceptional effects also exist for the RH at Maduo, ΔT at Tulihe and Maduo.

All the above trends in predictors for all the sites are actually the reflection of the change patterns of MSLP, RH and air temperature over the large domain around China. The patterns provided by the three base GCMs have large differences (**Fig.14**). Generally, the MSLP patterns from the three GCMs show increasing changes over South China and a small area in Southwest China (corresponding to the source region of the Yangtze River), decreasing changes over Central Asia, but the changes in North China have some differences between the GCMs. The increasing MSLP in the southern areas of East Asia may be related to the northward movement of subtropical high, the weakened South Asia Monsoon, and the weakened Mongol-Siberia High. The RH patterns from the three GCMs are also different, but all of them show increasing RH over the northern part of North China (corresponding to an area in Inner Mongolian) and an area across the Eastern Xinjiang and Qinghai-Tibetan plateau, and decreasing trends are shown in the areas around the Yangtze River basin. These RH change patterns also may be related to the northward movement of the subtropical high. For air temperature, both the HadHEM-E2 and the CanESM2 provided large increasing change (larger than 4.0 degree) over the north part of Asia and small increasing change (smaller than 2 degree) over the south part of Asia. The GFDL-ESM2G only provided a mild increasing in Central Asia, Xinjiang and Mongolia, and small increasing in the north part of Asia, which is largely different to the mainstream opinion

that large increasing will happen in the areas around the north pole (IPCC, 2021). This implies that the GFDL-ESM2G simulation may be more incredible than the other two GCM simulations.

Overall, the downscaling of daily precipitation can utilize the local signals around the sites in the GCM outputs, but the reliability of such signals is vital to the downscaling results. This requires that the GCM outputs which are used as the base for downscaling should be carefully evaluated and ensembles of multiple GCM outputs should be used to assess the uncertainty.

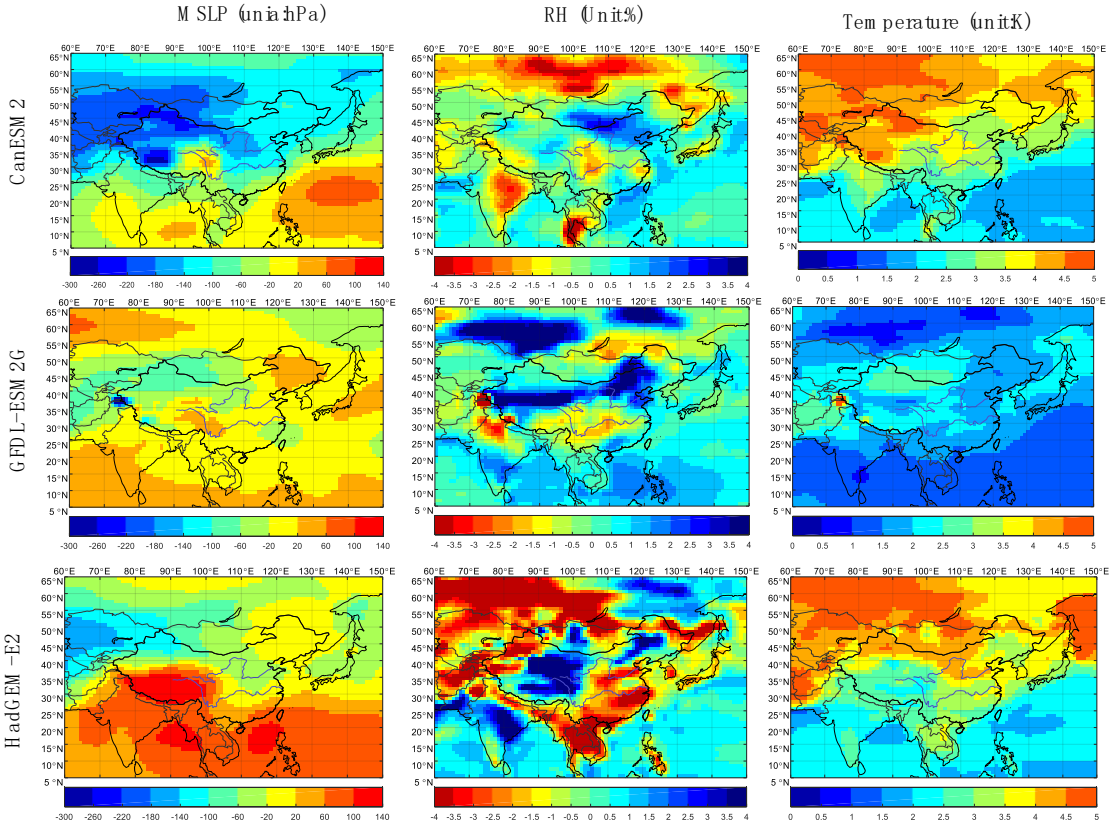


Fig.14 Changes of mean sea level pressure, surface relative humidity and surface air temperature in summer, projected by the RCP45 scenario for the period 2060-2099, by the three CMIP5 models, relative to the period 1960-1999. All the changes are calculated by subtracting the mean values of 1960-1999 from the mean values of 2060-2099.

2 Conclusion

For downscaling precipitation from multiple-model simulated future projections across China, few PP downscaling has been applied, when compared to the MOS. This study is a trial to use a multiple variable based, perfect prognostic downscaling for daily precipitation in China and compare the result with the MOS (bias correction). The predictors for the PP downscaling were screened from the large-scale variables by an optimal grid-box method: finding the most representative grid box area which has the most correlations to the sites'

precipitation, then combining multiple predictors from different variables (air temperature, geopotential height, etc.) into one GLM.

Daily large-scale variables in CMIP are usually packaged in huge-sized datasets, which is difficult to collect through the internet. Thus, in this study, the surface-level variables other than the multiple-level variables were used to extract predictors. The model fitting process of this study indicates that predictors from the surface variables get slightly inferior modeling performances to that from the multiple-level variables at some sites.

The PP downscaling can generally produce the climatological mean of summer precipitation across China, both for the simulation based on ERAI and CMIP5 historical simulations. Based on the historical simulation (1961-2005), the downscaled spatial patterns of linear trends are different, between the three GCM simulations, as well as between the PP and the QM. The downscaled results based on CanESM2 get the most similar patterns to the observed as regard to all the sites. Nevertheless, each downscaling can also generally reproduce the trends in most areas of China except for some specific regions. For HadGEM2-ES and GFDL-ESM2G, at least one downscaled pattern can produce the observed spatial pattern, though most of the downscaled results not. CanESM-PP/QM has reproduced the lower tail of annual variation at the sites in Northern China (Beijing is a one case), but failed reproducing those variations/trends in other areas. Whether this is a true modeling performance is difficult to certify.

The annual variations downscaled by OGB-PP have comparatively small differences among the three bias corrections on predictors, but have large difference to that downscaled by the BC-MOS, according to the representative sites (Beijing, Tulihe, Hanzhong, etc.). The trends downscaled from the three GCMs are also diverse, which can be attributed to the different trends in the predictors provided by the GCMs. The quantile mapping based bias correction can also change the trend of precipitation at a few sites.

The change patterns of the OGB-PP downscaled precipitation over China are different, depending on different GCM bases and different bias corrections for predictors, and are also different to the results of the BC-MOS. Therefore, we cannot conclude which bias correction on predictors or which GCM is better than the others, and also cannot tell whether the OGB-PP is more credible than the BC-MOS. Considering the diverse results downscaled based on each GCM with different bias corrections on predictors, one may think that PP is useless. However, the diverse future projections produced by the OGB-PP shows larger spatial variability and more significant increasing/decreasing trends in different areas, than that produced by the BC-MOS or the original GCM simulated precipitation. This maybe an advantage of PP due to that PP can freely utilize multiple predictors with flexibly selected grid boxes.

Considering the multiple downscaling results as a whole, the OGB-PP downscaling can produce the trends similar to those projected by many multiple-model results in CMIP5, while the bias corrections (BC-MOS) on precipitation of the same CMIP5 models not. It is worth noting that these previous studies were mostly based on the synthesis on original multiple CMIP5 simulations or the MOS downscaling based on multiple GCMs. This fact implies that the PP downscaling in this study may be promising. This precipitation downscaling is made on a daily basis and also involves a complex data processing, but is economical when compared to any dynamical downscaling models.

In this study, due to the huge-sized datasets of daily multiple variables in CMIP5, only those from three GCMs under the RCP4.5 emission scenario were used and the projection under other emission scenarios (for example, RCP8.5) are unavailable for us, therefore, the result cannot give a very convincing conclusion. Further studies are needed for downscaling precipitation from more GCM simulations in CMIP5 or CMIP6.

Acknowledgments. This work is mainly funded by the key scientific research projects of universities in Henan (Grant 19A170007).

Declarations

Funding

Funded by the key scientific research projects of universities in Henan (Grant 19A170007).

Conflicts of interest/Competing interests

No conflicts of interests exist.

Availability of data and material

The data and the program code for data processing support the results of our study and comply with the related field standards.

Code availability

All code is written by the first author (Yonghe Liu), and run under Python environment.

References

Ahmed K, Shahid S, Nawaz N, Khan N (2019) Modeling climate change impacts on precipitation in arid regions of Pakistan: a non-local model output statistics downscaling approach. THEORETICAL AND APPLIED CLIMATOLOGY 137:1347-1364

Al-Mukhtar M, Qasim M (2019) Future predictions of precipitation and temperature in Iraq using the statistical downscaling model. ARABIAN JOURNAL OF GEOSCIENCES 12

Bao J, Feng J, Wang Y (2015) Dynamical downscaling simulation and future projection of precipitation over China. JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES 120:8227-8243

Das L, Akhter J (2019) How well are the downscaled CMIP5 models able to reproduce the monsoon

precipitation over seven homogeneous zones of India? INTERNATIONAL JOURNAL OF CLIMATOLOGY 39:3323-3333

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Holm EV, Isaksen L, Kallberg P, Koehler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thepaut JN, Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY 137:553-597

Fan L, Yan Z, Chen D, Fu C (2015) Comparison between two statistical downscaling methods for summer daily rainfall in Chongqing, China. INTERNATIONAL JOURNAL OF CLIMATOLOGY 35:3781-3797

Fu G, Charles SP, Chiew FHS, Ekstrom M, Potter NJ (2018) Uncertainties of statistical downscaling from predictor selection: Equifinality and transferability. ATMOSPHERIC RESEARCH 203:130-140

Gebrechorkos SH, Huelsmann S, Bernhofer C (2019) Statistically downscaled climate dataset for East Africa. SCIENTIFIC DATA 6

Hammami D, Lee TS, Ouarda TBMJ, Lee J (2012) Predictor selection for downscaling GCM data with LASSO. JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES 117

Hertig E, Jacobeit J (2013) A novel approach to statistical downscaling considering nonstationarities: application to daily precipitation in the Mediterranean area. JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES 118:520-533

Hui P, Tang J, Wang S, Niu X, Zong P, Dong X (2018) Climate change projections over China using regional climate models forced by two CMIP5 global models. Part II: projections of future climate. INTERNATIONAL JOURNAL OF CLIMATOLOGY 381:E78-E94

IPCC (2021) Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press.

Kaspar-Ott I, Hertig E, Kaspar S, Pollinger F, Ring C, Paeth H, Jacobeit J (2019) Weights for general circulation models from CMIP3/CMIP5 in a statistical downscaling framework and the impact on future Mediterranean precipitation. INTERNATIONAL JOURNAL OF CLIMATOLOGY 39:3639-3654

Li H, Sheffield J, Wood EF (2010) Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES 115

Liu Y, Feng J, Liu X, Zhao Y (2019a) A method for deterministic statistical downscaling of daily precipitation at a monsoonal site in Eastern China. THEORETICAL AND APPLIED CLIMATOLOGY 135:85-100

Liu Y, Feng J, Yang Z, Hu Y, Li J (2019b) Gridded Statistical Downscaling Based on Interpolation of Parameters and Predictor Locations for Summer Daily Precipitation in North China. JOURNAL OF APPLIED METEOROLOGY AND CLIMATOLOGY 58:2295-2311

Liu Y, Feng J, Shao Y, Li J (2019c) Identify optimal predictors of statistical downscaling of summer daily precipitation in China from three-dimensional large-scale variables. ATMOSPHERIC RESEARCH 224:99-113

Liu Y, Feng J, Yang Z, Hu Y, Li J (2019b) Gridded Statistical Downscaling Based on Interpolation of Parameters and Predictor Locations for Summer Daily Precipitation in North China. JOURNAL OF APPLIED METEOROLOGY AND CLIMATOLOGY 58:2295-2311

Maraun D (2013a) Bias Correction, Quantile Mapping, and Downscaling: Revisiting the Inflation Issue. Journal of Climate 26:2137-2143

Maraun D, Shepherd TG, Widmann M, Zappa G, Walton D, Gutierrez JM, Hagemann S, Richter I, Soares PMM, Hall A, Mearns LO (2017) Towards process-informed bias correction of climate change simulations. NATURE CLIMATE CHANGE 7:764-773

Maraun D, Wetterhall F, Ireson AM, Chandler RE, Kendon EJ, Widmann M, Brienen S, Rust HW, Sauter T, Themessl M, Venema VKC, Chun KP, Goodess CM, Jones RG, Onof C, Vrac M, Thiele-Eich I (2010a) PRECIPITATION DOWNSCALING UNDER CLIMATE CHANGE: RECENT DEVELOPMENTS TO BRIDGE THE GAP BETWEEN DYNAMICAL MODELS AND THE END USER. REVIEWS OF GEOPHYSICS 48

Maraun D, Huth R, Gutierrez JM, San Martin D, Dubrovsky M, Fischer A, Hertig E, Soares PMM, Bartholy J, Pongracz R, Widmann M, Casado MJ, Ramos P, Bedia J (2019a) The VALUE perfect predictor experiment: Evaluation of temporal variability. INTERNATIONAL JOURNAL OF CLIMATOLOGY 39:3786-3818

Maraun D, Widmann M, Gutierrez JM (2019b) Statistical downscaling skill under present climate conditions: A synthesis of the VALUE perfect predictor experiment. INTERNATIONAL JOURNAL OF CLIMATOLOGY 39:3692-3703

Moss RH, Edmonds JA, Hibbard KA, Manning MR, Rose SK, van Vuuren DP, Carter TR, Emori S, Kainuma M, Kram T, Meehl GA, Mitchell JFB, Nakicenovic N, Riahi K, Smith SJ, Stouffer RJ, Thomson AM, Weyant JP, Wilbanks TJ (2010) The next generation of scenarios for climate change research and assessment. NATURE 463:747-756

Navarro-Racines C, Tarapues J, Thornton P, Jarvis A, Ramirez-Villegas J (2020) High-resolution and bias-corrected CMIP5 projections for climate change impact assessments. Scientific data 7:7

Qian C, Zhou W, Fong SK, Leong KC (15) Two Approaches for Statistical Prediction of Non-Gaussian Climate Extremes: A Case Study of Macao Hot Extremes during 1912 – 2012. Journal of Climate 28:623-636

Salvi K, Ghosh S, Ganguly AR (2016) Credibility of statistical downscaling under nonstationary climate. CLIMATE DYNAMICS 46:1991-2023

San-Martin D, Manzanas R, Brands S, Herrera S, Gutierrez JM (2017) Reassessing Model Uncertainty for Regional Projections of Precipitation with an Ensemble of Statistical Downscaling Methods. JOURNAL OF CLIMATE 30:203-223

Sperber KR, Annamalai H, Kang IS, Kitoh A, Moise A, Turner A, Wang B, Zhou T (2013) The Asian summer monsoon: an intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century. CLIMATE DYNAMICS 41:2711-2744

Su B, Huang J, Gemmer M, Jian D, Tao H, Jiang T, Zhao C (2016) Statistical downscaling of CMIP5 multi-model ensemble for projected changes of climate in the Indus River Basin. ATMOSPHERIC RESEARCH 178:138-149

Tang J, Niu X, Wang S, Gao H, Wang X, Wu J (2016) Statistical downscaling and dynamical downscaling of regional climate in China: Present climate evaluations and future climate projections. JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES 121:2110-2129

Thomson AM, Calvin KV, Smith SJ, Kyle GP, Volke A, Patel P, Delgado-Arias S, Bond-Lamberty B, Wise MA, Clarke LE, Edmonds JA (2011) RCP4.5: a pathway for stabilization of radiative forcing by 2100. CLIMATIC CHANGE 109:77-94

Valiantzas JD (2013) Simplified forms for the standardized FAO-56 Penman-Monteith reference evapotranspiration using limited weather data. JOURNAL OF HYDROLOGY 505:13-23

Wilby RL, Dawson CW, Barrow EM (2002) SDSM - a decision support tool for the assessment of regional climate change impacts. ENVIRONMENTAL MODELLING & SOFTWARE 17:147-159

Yang X, Wood EF, Sheffield J, Ren L, Zhang M, Wang Y (2018) Bias Correction of Historical and Future Simulations of Precipitation and Temperature for China from CMIP5 Models. Journal of Hydrometeorology 19:609-623

Yang Y, Tang J, Xiong Z, Wang S, Yuan J (2019a) An intercomparison of multiple statistical downscaling methods for daily precipitation and temperature over China: future climate projections. CLIMATE DYNAMICS 52:6749-6771

Zhang Q, Shen Z, Xu C, Sun P, Hu P, He C (2019a) A new statistical downscaling approach for global evaluation of the CMIP5 precipitation outputs: Model development and application. SCIENCE OF THE TOTAL ENVIRONMENT 690:1048-1067

Zhang X, Yan X (2015) A new statistical precipitation downscaling method with Bayesian model averaging: a case study in China. CLIMATE DYNAMICS 45:2541-2555

Zhou T, Chen X, Wu B, Guo Z, Sun Y, Zou L, Man W, Zhang L, He C (2017) A Robustness Analysis of CMIP5 Models over the East Asia-Western North Pacific Domain. ENGINEERING 3:773-778

Zhou T, Chen Z, Zou L, Chen X, Yu Y, Wang B, Bao Q, Bao Y, Cao J, He B, Hu S, Li L, Li J, Lin Y, Ma L, Qiao F, Rong X, Song Z, Tang Y, Wu B, Wu T, Xin X, Zhang H, Zhang M (2020) Development of Climate and Earth System Models in China: Past Achievements and New CMIP6 Results. JOURNAL OF METEOROLOGICAL RESEARCH 34:1-19

Zhou T, Zou L, Wu B, Jin C, Song F, Chen X, Zhang L (2014) Development of earth/climate system models in China: A review from the Coupled Model Intercomparison Project perspective. JOURNAL OF METEOROLOGICAL RESEARCH 28:762-779

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterials.pdf