Impacts of bias nonstationarity of climate model outputs on hydrological simulations
Yu Hui, Yuni Xu, Jie Chen, Chong-Yu Xu and Hua Chen

ABSTRACT

Bias correction methods are based on the assumption of bias stationarity of climate model outputs. However, this assumption may not be valid, because of the natural climate variability. This study investigates the impacts of bias nonstationarity of climate models simulated precipitation and temperature on hydrological climate change impact studies. The bias nonstationarity is determined as the range of difference in bias over multiple historical periods with no anthropogenic climate change for four different time windows. The role of bias nonstationarity in future climate change is assessed using the signal-to-noise ratio as a criterion. The results show that biases of climate models simulated monthly and annual precipitation and temperature vary with time, especially for short time windows. The bias nonstationarity of precipitation plays a great role in future precipitation change, while the role of temperature bias is not important. The bias nonstationarity of climate model outputs is amplified when driving a hydrological model for hydrological simulations. The increase in the length of time window can mitigate the impacts of bias nonstationarity for streamflow projections. Thus, a long time period is suggested to be used to calibrate a bias correction method for hydrological climate change impact studies to reduce the influence of natural climate variability.

Key words | bias nonstationarity, climate change signal, climate model outputs, hydrology, natural climate variability

HIGHLIGHTS

- The biases of GCM precipitation and temperature vary with time, due to natural climate variability.
- The bias nonstationarity of precipitation plays a great role in future precipitation change, while the role of temperature bias is not important.
- The bias nonstationarity of precipitation and temperature has great considerable impacts on future streamflow changes.

INTRODUCTION

The assessment of climate change impacts on the hydrological cycle has been widely investigated during recent years (Graham et al. 2007; Li et al. 2015; Chen et al. 2016a; Marhaento et al. 2017; Shen et al. 2018; Lu & Qin 2020; Ragettli et al. 2020). Global climate models (GCMs) can provide climate variables (e.g. precipitation and temperature) for the future period used as inputs of hydrological models for hydrological impact studies. However, the coarse resolution of GCMs does not meet the need of high resolution for the hydrological models (Maraun et al. 2010). In parallel,
GCMs are imperfect representations of reality with systematic biases found between climate model simulations and observations. To resolve these problems, a number of downscaling methods have been developed during the last two decades. Especially, bias correction becomes a standard procedure when using regional climate model (RCM) outputs for hydrological impact studies (van Pelt et al. 2009; Teutschbein & Seibert 2012; Olsson et al. 2015; Zhuan et al. 2019). During recent years, bias correction methods are also commonly used for GCM’s outputs. Since the resolution of climate model outputs is lower than that of observations, bias correction methods also act as downscaling methods. The usually used bias correction methods range from simple mean-based scaling to sophisticated distribution-based mapping and multivariate or/and multisite correction.

The commonly used bias correction methods are usually based on an assumption that biases of climate models outputs are stationary. In other words, these bias correction methods assume climate models outputs present the same biases in magnitude and direction between historical and future periods. However, this assumption may not be valid, as pointed out in a few recent studies (e.g. Buser et al. 2009; Ehret et al. 2012; Gutierrez et al. 2013; Maurer et al. 2013; Velázquez et al. 2015; Dixon et al. 2016). For example, Christensen et al. (2008) evaluated bias nonstationarity on simulated monthly precipitation and temperature from an ensemble of 13 RCMs over Europe. They found that biases of simulated precipitation and temperature vary as a second-order function of observations, suggesting that bias nonstationarity exists in different climatic regimes. Maraun (2012) verified the bias stationary assumption in RCMs for European seasonal mean temperature and precipitation sums using a pseudo-reality approach, which considers one climate model as pseudo observation to compare with other climate models simulations. The results showed that biases are relatively stable in general, but bias nonstationarity was identified in some regions where changes in potentially relevant physical variables are significant. This study was conducted in the climate model world, the transferability to real world needs to be further investigated. To test the bias nonstationarity in real-world climate, Chen et al. (2015) compared the biases between climate models simulations and corresponding observations over two historical periods. The results showed that biases of climate models simulated precipitation are nonstationary even for two close historical periods, while temperature biases are relatively stationary. This study attributed the bias nonstationarity of precipitation to natural climate variability (multi-decadal climate variability). A similar study was also carried out by Wang et al. (2017), who tested bias nonstationarity of precipitation in the eastern United States determined by a skill score, which compares the errors of a downscaling method over validation period with the errors of observations between calibration and validation periods. The results show that precipitation biases are nonstationary at most of the stations, especially for the annual extreme precipitation. Taking into account multi-decadal climate variability, Nahar et al. (2017) investigated the bias stationarity of six GCMs simulated monthly and seasonal precipitation and temperature over multi-decadal time periods from 1900 to 1999 in Australia. One hundred samples of the observations and GCM simulations generated using bootstrapping was used to calculate uncertainty in biases with a 95% confidence, which represents a possible range of bias if bias stationarity exists. When the actual bias goes beyond this uncertainty, the bias was assumed to be nonstationary. This study showed that biases of precipitation and temperature are not stationary for some regions in Australia (e.g. east coast of Australia), because of the natural climate variability.

For a historical period with no anthropogenic forcing, the bias nonstationarity is attributed to natural climate variability. Natural climate variability refers to variations in the mean state and other statistics of the climate, due to natural internal processes or natural external forcing. It involves a wide range of time scales, from one day to the next, as well as from one year or multi-decade to the next. For the application of a bias correction method in climate change impact studies, the decadal and multiple decadal time scales are widely and implicitly used. At the decadal or multi-decadal time scales, the natural climate variability includes some natural modes of decadal or multi-decadal climate variability, such as El Niño/Southern Oscillation (ENSO), Atlantic Multidecadal Oscillation (AMO), Pacific Decadal Oscillation (PDO), and Interdecadal Pacific Oscillation (IPO). The phase of these modes has found to be linked with changes in precipitation and temperature. For example, the PDO warm (cold) phase periods link to the...
decrease (increase) in precipitation in the majority of China (Ouyang et al. 2014). Precipitation decreases during IPO warm phase periods, while it increases during IPO cool phase periods in east Australia (Power et al. 1999). However, since the unpredictability and complexity of natural dynamical processes make it difficult for climate models to accurately capture the characteristics, GCMs from Coupled Model Intercomparison Project phase 5 (CMIP5) have been found to poorly represent the phase in natural modes of decadal or multi-decadal climate variability (Polade et al. 2013; Ruiz-Barradas et al. 2013; Bellenger et al. 2014; Fuentes-Franco et al. 2016). This poor representation for GCMs could lead to biases between simulations and observations when sampling in a finite decadal or multi-decadal time, and further lead to changes in bias among different decadal or multi-decadal periods (Nahar et al. 2017).

In climate change impact studies, natural climate variability used to be estimated based on multi-decadal climate model simulations in the absence of anthropogenic-induced climate change (Hulme et al. 1999; Arnell 2003; Chen et al. 2016b). In these methods, the differences among several separated multi-decadal periods obtained from a long time series of unforced simulations were supposed to represent the multi-decadal natural climate variability. Similarly, to evaluate bias nonstationarity of climate model outputs, the estimate of bias nonstationarity is depicted as the range of the biases among multi-decadal historical periods, while the timescale of each period may have an effect on the results of estimation. The different temporal scales of sampling can lead to changes in the magnitude and direction of bias, due to different phases between GCM outputs and observations. Therefore, the impacts of different temporal scales of natural climate variability on bias nonstationarity of climate model outputs need to be investigated, especially for their propagation in hydrology.

Even though few studies (Maraun 2012; Chen et al. 2015; Velázquez et al. 2015; Nahar et al. 2017) have investigated the bias nonstationarity of climate model outputs, the role of bias nonstationarity in future climate change was not investigated, especially in hydrological impact studies. Accordingly, the objectives of this study are to investigate: (1) the bias nonstationarity of climate models outputs in the context of natural climate variability, (2) the role of bias nonstationarity in future climate change, and (3) the propagation of bias nonstationarity in hydrological impact studies. The first objective is achieved by separating a long historical climate series (100 years) to decadal and multi-decadal time periods to calculate the difference in bias among different periods. Precipitation and temperature, which are primary variables for hydrological simulations, are investigated. The second objective is achieved by comparing the range of biases among multiple periods to climate change signals between future and historical periods. The last objective is achieved by analyzing streamflow time series simulated by a hydrological model using the above two approaches.

The rest of the paper is organized as follows: the following section introduces the study area and hydro-climate data followed by the methodology. The main results are then presented, and the discussion and conclusions are presented in the final section.

**STUDY AREA AND DATA**

**Study area**

The case study was conducted at the Hanjiang River basin (Figure 1), located in south-central China, which is the largest tributary of the Yangtze River basin. The Hanjiang River flows through Shaanxi and Hubei Provinces with a drainage area of 159,000 km² and a length of 1,567 km. The watershed above the Danjiangkou Reservoir, with a sub-basin area of 95,200 km² and a length of 918 km, was used in this study. The Danjiangkou reservoir is the water source of the Middle Route of the South-to-North Water Transfer Project in China. The watershed has a subtropical monsoon climate. The mean annual precipitation is about 840 mm, of which 70–80% of the total amount falls in the wet season from May to September. The average maximum and minimum temperatures are 24–29 and 0–3 °C, respectively. The daily mean discharge of the Hanjiang River is approximately 1,150 m³/s.

**Data**

This study used both observed and GCM-simulated precipitation and temperature for the Hanjiang watershed. The
inflow to the Danjiangkou reservoir calculated using a water mass balance method was also used for hydrological model calibration and validation. The streamflow time series covering the 1961–2000 period was provided by the Bureau of Hydrology of the Changjiang Water Resources Commission.

The observed precipitation and temperature were obtained from the gridded Climatic Research Unit (CRU) Time-series (TS) data, produced by CRU at the University of East Anglia, UK. This dataset provides monthly gridded data with a high resolution of 0.5° × 0.5°. The latest version CRU TS version 4.01 (CRU et al. 2017) covering the period 1901–2000 was used in this study.

The GCM-simulated daily precipitation and maximum and minimum temperatures were extracted from the database of CMIP5. In order to capture the uncertainty related to climate models, 17 GCMs were used in this study (Table 1). The time period covers from 1901 to 2100. All GCMs’ simulations during the 1901–2005 period were driven by historical climate forcing (including natural solar, volcanic variations and anthropogenic radiative forcing), while those during the 2006–2100 period were generated under the representative concentrations pathway (RCP) 4.5. RCP4.5 corresponds to the medium anthropogenic forcing for the 21st century. The other greenhouse gas emission scenarios (RCP8.5) are also presented and discussed in the Discussion section. Since this study only used monthly and annual data, the daily precipitation and temperature were summed (for precipitation) or averaged (for temperature) to monthly and annual values. In order to run a lumped hydrological model, all gridded precipitation and temperature within and surrounding the Hanjiang watershed were averaged to a single time series using the Thiessen polygon method for both observations and GCM simulations.

**METHODOLOGY**

A climate time series for the past–present period may contain both natural climate variability and anthropogenic climate change signal. In order to investigate the impacts of natural climate variability on bias correcting climate model outputs, the climate change signal needs to be first removed. A simple linear detrending method of Zhuan et al. (2018) was used to remove the anthropogenic climate change signal for observed and GCM-simulated monthly precipitation and temperature for the 1901–2000 period. This detrending
method first detects the trend and breakpoint using the Mann–Kendall test (Mann 1945; Kendall 1975). If a trend exists, it is removed for the period after the breakpoint using a linear method. To preserve the seasonality, the detrending method was applied for each month, respectively.

### Calculation of difference in bias

After detrending, the variation of observed and GCM-simulated precipitation and temperature within the 1901–2000 period is considered to be only attributed to the natural climate variability. Since natural climate variability is inherently complex and manifests itself over various temporal and spatial scales, this study only investigates its impacts on bias correcting climate model outputs over decadal and multi-decadal temporal and watershed spatial scales. In order to investigate the impacts of natural climate variability at the temporal scale, the 100-year precipitation and temperature time series are divided into several independent decadal and multi-decadal periods. The time window consists of 10, 20, 33 and 50 years for the 100-year period. For example, when the time window is 20 years, the 100-year period is divided into five 20-year periods (1901–1920, 1921–1940, 1941–1960, 1961–1980 and 1981–2000). The first period of each time window (e.g. 1901–1910 for 10-year window) is used as the baseline period.

For each decadal or multi-decadal period, the biases of GCM-simulated monthly and annual precipitation ($BP$) and temperature ($BT$) relative to observations are calculated using Equations (1) and (2), respectively:

$$BP_i = \frac{(P_{\text{mod},i} - P_{\text{obs},i})}{P_{\text{obs},i}}$$

$$BT_i = T_{\text{mod},i} - T_{\text{obs},i}$$

### Table 1: Basic information of the used 17 CMIP5 models

| ID | Model name       | Modeling center | Institution                                                                 | Resolution (Longitude × Latitude) |
|----|------------------|-----------------|------------------------------------------------------------------------------|-----------------------------------|
| 1  | ACCESS1.0        | CSIRO-BOM       | Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia | $1.875 \times 1.25^\circ$          |
| 2  | ACCESS1.3        |                 |                                                                              | $1.875 \times 1.25^\circ$          |
| 3  | BCC-CSM1.1(m)    | BCC             | Beijing Climate Center, China Meteorological Administration                  | $1.125 \times 1.25^\circ$          |
| 4  | CNRM-CM5         | CNRM-CERFACS    | Centre National de Recherches Méteorologiques/Centre Européen de Recherche et Formation Avancée en Calcul Scientifique | $1.4 \times 1.4^\circ$             |
| 5  | CSIRO-Mk3.6.0    | CSIRO-QCCCE     | Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence | $1.9 \times 1.9^\circ$             |
| 6  | CanESM2          | CCCMA           | Canadian Centre for Climate Modelling and Analysis                           | $2.8 \times 2.8^\circ$             |
| 7  | GFDL-CM3         | NOAA GFDL       | NOAA Geophysical Fluid Dynamics Laboratory                                   | $2.5 \times 2.0^\circ$             |
| 8  | GFDL-ESM2G       |                 |                                                                              | $2.5 \times 2.0^\circ$             |
| 9  | GFDL-ESM2M       |                 |                                                                              | $2.5 \times 2.0^\circ$             |
| 10 | INM-CM4          | INM             | Institute for Numerical Mathematics                                          | $2.0 \times 1.5^\circ$             |
| 11 | IPSL-CM5A-LR     | IPSL            | L’Institut Pierre-Simon Laplace                                              | $3.75 \times 1.9^\circ$             |
| 12 | IPSL-CM5A-MR     |                 |                                                                              | $2.5 \times 1.25^\circ$             |
| 13 | IPSL-CM5B-LR     |                 |                                                                              | $3.75 \times 1.9^\circ$             |
| 14 | MIROC-ESM-CHEM   | MIROC           | Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies | $2.8 \times 2.8^\circ$             |
| 15 | MIROC-ESM        |                 |                                                                              | $2.8 \times 2.8^\circ$             |
| 16 | MIROC5           | MIROC           | Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology | $1.4 \times 1.4^\circ$             |
| 17 | MRI-CGCM3        | MRI             | Meteorological Research Institute                                            | $1.1 \times 1.1^\circ$             |
where $i$ indicates the number of decadal or multi-decadal historical periods, $P_{\text{mod}}$ and $P_{\text{obs}}$ indicate the mean values of simulated and observed precipitation over the period, and $T_{\text{mod}}$ and $T_{\text{obs}}$ indicate the mean values of simulated and observed temperature.

The differences in bias of precipitation ($DBP$) and temperature ($DBT$) between the baseline period and all other historical periods are then calculated for each GCM. The use of the difference in bias is beneficial to compare with the results from different GCMs. The differences in bias represent the possible changes in GCM bias due to natural climate variability at decadal and multi-decadal scales. The range of difference in biases is used to judge the bias stationarity of climate model outputs. The impacts of natural climate variability on bias nonstationarity is investigated by using multiple time windows. The main steps are illustrated in Figure 2 for a 20-year time window.

**Calculation of climate change signals relative to differences in bias**

To identify the role of bias nonstationarity in future climate change, the signal-to-noise ratio (SNR) is used as a criterion to quantify the extent of bias nonstationarity relative to future climate change. The SNR is defined as the ratio of climate change signal (CCS) of future periods and the range of bias ($RB$) among multiple historical periods.

$RB$ is calculated based on the differences in bias among multiple periods using Equation (3):

$$RB = \begin{cases} \max (|DB_{\max}|, |DB_{\min}|) & \text{if} \ DB_{\max} \cdot DB_{\min} > 0 \\ |DB_{\max} - DB_{\min}| & \text{if} \ DB_{\max} \cdot DB_{\min} \leq 0 \end{cases}$$

where $DB_{\text{max}}$ and $DB_{\text{min}}$ indicate the maximum and minimum values of differences in bias between historical periods and the baseline period, respectively. The outliers were removed when calculating the $RB$ to ensure the reasonability of estimation. The outliers are detected if they are larger than $Q_{25} + 1.5 \times (Q_{75} - Q_{25})$ or smaller than $Q_{25} - 1.5 \times (Q_{75} - Q_{25})$, where $Q_{25}$ and $Q_{75}$ indicate the 25th and 75th percentiles, respectively. It is worth noting that when the time window length is 50 years, the $RB$ is just equal to the absolute value of difference in bias between the 1951–2000 period and the 1901–1950 baseline period.

Meanwhile, the CCS of each future period is calculated relative to the historical baseline period. The 100-year period from 2001 to 2100 is also divided into several decadal or multi-decadal periods as the same time window as the historical period. By doing this, the SNR can be calculated for each time scale. The climate change signal for precipitation (CCSP) and temperature (CCST) was calculated using Equations (4) and (5):

$$CCSP_j = \left( \frac{P_{\text{mod,}j} - P_{\text{mod,b}}}{P_{\text{mod,b}}} \right)$$

$$CCST_j = \left( \frac{T_{\text{mod,}j} - T_{\text{mod,b}}}{T_{\text{mod,b}}} \right)$$

where $j$ indicates the number of decadal or multi-decadal periods from 2001 to 2100, and subscript $b$ indicates the baseline period (e.g. 1901–1910, 1901–1920, 1901–1933 or 1901–1950). These precipitation and temperature time series without detrending were used to retain climate change tendency.

Then, SNR is calculated as the absolute ratio of CCS and $RB$ for both monthly and annual precipitation and temperature. If the CCS is relatively smaller than the $RB$, e.g. the SNR being smaller than 1, the $RB$ is considered to have large impacts on future climate changes and their hydrological impacts and vice versa.

**Hydrological modeling**

The hydrological simulations were carried out using a lumped hydrological model named as the Two-Parameter Monthly Water Balance Model (Xiong & Guo 1999). It is a simple and monthly lumped rainfall-runoff model with only two physical parameters. The first parameter is $\alpha$,
which represents a coefficient to take account of the effect of the time scale change. The second parameter is SC, which represents the field capacity of a watershed. The monthly runoff \((Q)\) was simulated using Equations (6) and (7):

\[
E_t = c \times EP_t \times \tanh (P_t/EP_t) \quad (6)
\]

\[
Q_t = (S_{t-1} + P_t - E_t) \times \tanh [(S_{t-1} + P_t - E_t)/SC] \quad (7)
\]

where \(t\) indicates the number of months, \(E\) and \(EP\) indicate the monthly actual and potential evapotranspiration, respectively, \(P\) indicates the monthly areal precipitation, and \(S\) indicates the water content in the soil. The monthly potential evapotranspiration was calculated by the Thornthwaite method (Thornthwaite 1948; Xu & Singh 2001), using the monthly mean temperature. This model has been tested and proved to be highly efficient in several watersheds with the monsoon-rainfall-dominated climate in the south of China (Xiong & Guo 1999; Guo et al. 2002; Chen et al. 2007; Bai et al. 2015).

The observed monthly streamflow was used to model calibration (1961–1980) and validation (1981–2000). Model calibration was done automatically using shuffled complex evolution-University of Arizona (SCE-UA) algorithm (Duan et al. 1992). The optimal combination of parameters was chosen on the basis of the Nash–Sutcliffe efficiency (NSE) coefficient (Nash & Sutcliffe 1970). The values of NSE are 0.79 for both calibration and validation periods for the Hanjiang watershed, indicating a good performance.

Using the calibrated hydrological model, the streamflow time series are generated for each time window using the detrended observed and GCM-simulated precipitation and temperature for the historical period. Meanwhile, the streamflow time series are also generated for each time window of the future periods. The propagation of precipitation and temperature bias nonstationarity in hydrology was investigated using the same methods as presented above under ‘Calculation of difference in bias’ and ‘Calculation of climate change signals relative to differences in bias’.

**RESULTS**

**Bias nonstationarity of precipitation and temperature**

Figure 3 presents the difference in bias of mean annual precipitation between each historical period and the baseline period for four different time window lengths: 10, 20, 33 and 50 years. For each time window length, the dash lines represent differences in bias of all historical periods relative to the baseline period for all 17 GCMs, and the thick line is their median value. The x axis represents the midpoints of the time window. The results show that the difference in bias of mean annual precipitation varies with time over historical periods, indicating that the bias of precipitation is not stationary. The difference in bias reflects the impacts of
natural climate variability on bias nonstationarity, because the trend of precipitation change was removed. The results in Figure 3 show that the natural climate variability has great impacts on bias nonstationarity for mean annual precipitation, especially for the short time window. However, this problem becomes less important as the time window gets longer, as indicated by the fact that the difference in bias decreases with the extension of the time window. For example, in terms of the median value over all GCMs, the difference in bias changes between −11.5 and 3.8% for the 10-year window, between 0.5 and 5.5% for the 20-year window, between −2.7 and −0.4% for the 33-year window, and it is −4.4% for the 50-year window. This indicates that the decadal variability has more impacts on bias nonstationarity than multi-decadal variability. Not only the median of difference in bias but also the uncertainty (the difference between maximum and minimum values over the whole period) related to GCMs decreases with the longer time window. The uncertainty of difference in bias is 56.4% for the 10-year window, 27% for the 20-year window, 17.5% for the 33-year window, and 15.8% for the 50-year window. Therefore, using a long reference period to filter out low frequency modes of variability may lessen the impacts of natural climate variability on bias correction methods. However, if just looking at a single climate model, the difference in bias can reach 10%, even for the 50-year time window.

Figure 4 presents the difference in bias of mean annual temperature between each historical period and the baseline period for four different time windows. The results show that the differences in temperature bias vary with time. In particular, the magnitude and uncertainty of differences in bias are larger for short than long time windows. For example, the median value of difference in temperature bias changes between −0.7 and −0.05 °C for the 10-year window, between −0.5 and −0.05 °C for the 20-year window, between −0.28 and −0.06 °C for the 33-year window, and it is −0.15 °C for the 50-year window. The negative values of the difference in bias indicate the bias over the first baseline period is larger than that over the following several historical periods. The large uncertainty related to climate models is observed, especially for the short time window. Since the temperature time series were detrended before calculating the difference in bias, the variation resulted from natural climate variability. Even though the bias varies with time due to natural climate variability, its impacts on future climate change need to be investigated by comparing future climate change signals.

**Impacts of bias nonstationarity on future climate changes**

In order to investigate the role of bias nonstationarity in future climate change, the $SNR$ of mean annual precipitation and temperature is calculated for all four time windows. Figure 5 presents the $SNR$ of mean annual precipitation and temperature over multiple time periods from 2001 to 2100 for 10-, 20-, 33- and 50-year windows. Each boxplot depicts the distribution of $SNR$ from an ensemble of 17 GCMs. The larger $SNR$ values suggest that the bias range is relatively less important compared with the climate

![Figure 4](add)
The unity of SNR means that the range of bias is equal to the climate change signal. In this case, the calculated climate change signal may be just the difference in bias between two periods. Figure 5(a) shows that SNR values of precipitation are consistently less than one for the 10-year window of all future periods. In this case, it is hard to judge the existence of climate change signal. The calculated climate change signal may be just a part of the range of bias caused by the natural climate variability, especially taking into account the fact that the SNR values are mostly constant over time. With the increase in the length of time window, the bias nonstationarity of mean annual precipitation becomes less important, as indicated by some GCMs presenting SNR values being greater than one. In particular, the uncertainty related to GCMs becomes larger with the increase in the length of time window. However, the median values of SNR are still less than one, even for the 33- and 50-year time window. The above results indicate that great attention should be paid to natural climate variability when investigating the change of precipitation, in the face of the increase in the greenhouse gas emission scenario.

However, different results are observed for annual mean temperature. Even though the difference in bias varies with time, the anthropogenic climate change is significantly greater than the bias nonstationarity for annual mean temperature, especially for the future periods (Figure 5(b)). In other words, even though the bias nonstationarity can affect the detection of the real anthropogenic climate change of annual mean temperature, the influence is limited. Comparing to future climate change signal, the temperature bias can be considered as stationary. However, for the first
and second decades, the natural climate variability still plays an important role in anthropogenic climate change or the climate change signal may not exist, as indicated by the relatively small SNR. The SNR values increase with time, which is caused by the increase in the anthropogenic climate change, as the range of bias is supposed to be constant. Meanwhile, the dispersion of boxplot is gradually amplified with time, resulting from the increase in the GCM uncertainty in the far future. Additionally, the magnitude and dispersion of the SNR become larger with the extension of time window. This implies that the impacts of bias nonstationarity on future climate changes are important for the long time window.

The SNR of precipitation and temperature is also calculated at monthly scale. Figure 6 shows the portrait diagram of the median value of the SNR for mean monthly precipitation and temperature over multiple time periods from 2001 to 2100 with 10-, 20-, 33- and 50-year windows, respectively. The y-axis represents the 12 months from January to December. The median value of the SNR was calculated over 17 GCMs for each period. Generally, the bias of monthly precipitation is nonstationary, while the temperature is stationary compared to future climate change. In other words, the impacts of bias nonstationarity on future climate changes are large for monthly precipitation, but it is limited for monthly temperature. Similar to annual precipitation and temperature, the impacts of bias nonstationarity on future climate changes lessen for the longer time window, as indicated by the increase in the SNR. However, the monthly variability is observed for these impacts. In terms of the monthly precipitation (Figure 6(a)), the SNR in the wet season (May–September) is larger than in the dry season (October–April), especially for the mid and far future, while the SNR is still smaller than one for 10-, 20- and 33-year windows. When the time window increases to 50 years, the SNR in the wet season becomes larger than 1. This indicates that the impacts of bias nonstationarity on future climate changes

![Figure 6](http://iwaponline.com/hr/article-pdf/doi/10.2166/nh.2020.254/769146/nh2020254.pdf)
are less important in the wet season than in the dry season. Therefore, the impacts of bias nonstationarity in the dry season need to be paid more attention than in the wet season and annual scale. In terms of the mean monthly temperature (Figure 6(b)), the SNR from May to November is larger than other months for the mid and far future, and the maximum values of the SNR present between May and August over each decadal or multi-decadal period. In addition, the comparison of the SNR between annual and monthly temperature shows that the SNR at the monthly scale is smaller than that at the annual scale, especially for short time windows, while for the 50-year window, the SNR between June and October is comparable to or even larger than that at the annual scale. In the near future (2001–2030), the SNR in each month is comparable to or smaller than one for the short time window, because of the weak climate change signal.

**Propagation of bias nonstationarity in hydrology**

The propagation of bias nonstationarity of climate models simulated precipitation and temperature in hydrology is investigated by running a hydrological model over the Hanjiang watershed. Figure 7 presents the difference in bias of mean annual streamflow between each historical period and the baseline period for 10-, 20-, 33- and 50-year time windows. The difference in bias between GCM-driven and observation-driven streamflow also varies with time, which reflects the impacts of natural climate variability on hydrological simulations. The results show that the impacts of natural climate variability are amplified from climate world to hydrological world. For example, the difference in bias for all GCMs varies between −30 and 20% for mean annual precipitation, while that varies between −60 and 60% for mean annual streamflow for the 10-year time window. This is because of the non-linear process from climate to hydrology. Similar to precipitation and temperature, the extent of the impacts of natural climate variability gradually mitigates as the time window gets longer. For example, the median value of difference in streamflow bias changes between −12.3 and 25.6% for the 10-year window, between 8.5 and 20.2% for the 20-year window, between −0.9 and 5.9% for the 33-year window, and it is −6.2% for the 50-year window. The uncertainty of difference in streamflow bias related to GCMs also reduces as the extension of the time window. The uncertainty is 121.6% for the 10-year window, 58.2% for the 20-year window, 35% for the 33-year window and 27.1% for the 50-year window.

The impacts of bias nonstationarity of climate models outputs on future streamflow changes are also evaluated using the SNR as a metric. Figure 8 shows the SNR of mean annual streamflow over multiple periods from 2001 to 2100 for 10-, 20-, 33- and 50-year windows. In the future, the magnitude of the streamflow change between the future and baseline period is much smaller than the range of bias, as indicated by the small SNR values, especially for the 10- and 20-year time windows. This
indicates that the natural climate variability plays an important role in future streamflow changes, especially for the short time window. In other words, previous studies that attribute the streamflow change to anthropogenic climate change may not be correct when taking into account the impacts of climate change. Similar to precipitation and temperature, the importance of natural variability reduces with the extension of the time window, as indicated by the increase in the SNR. For example, the median values of the SNR change between 0.07 and 0.21 for the 10-year window, between 0.09 and 0.3 for the 20-year window, between 0.38 and 0.85 for the 33-year window, and between 0.61 and 1.01 for the 50-year window. The uncertainty related to GCMs becomes greater for the longer time window. With the exception of the second period of the 50-year time window, the median values of SNR are smaller than one, indicating that the natural climate variability is still more important than the anthropogenic climate change for future streamflow simulations. These results of streamflow are similar to those of mean annual precipitation, but the SNR of the former is smaller than the latter, due to the larger range of bias and lower streamflow change.

Figure 9 presents the median of SNR of mean monthly streamflow across 17 GCMs for 10-, 20-, 33- and 50-year windows. The results show that the bias nonstationarity has large impacts on the change of future monthly streamflow, even though the impacts mitigate with the extension of the time window. For example, the values of SNR are consistently smaller than one for 12 months and all future periods in 10- and 20-year windows. This is also
the case for most months and future periods in the 33-year window. For the 50-year window, the values of SNR increase to larger than 1 in the wet season, indicating that the anthropogenic climate change may be more important than the bias nonstationarity resulting from natural climate variability for the wet season. In addition, the impact of bias nonstationarity on future climate change is amplified in hydrological simulations, as indicated by the fact that the SNR of streamflow is smaller than that of precipitation and temperature.

DISCUSSIONS AND CONCLUSION

Results summary

This study first investigated the bias nonstationarity of GCM-simulated precipitation and temperature in the context of natural climate variability for multiple time windows. The role of bias nonstationarity in future climate changes was estimated by comparing the range of bias in the historical periods to climate change signal in the future periods. Finally, the propagation of bias nonstationarity in hydrological simulations is investigated by running a hydrological model. The main findings are as follows: (1) The biases of GCM precipitation and temperature vary with time, due to natural climate variability, especially for the short time windows; (2) Precipitation bias nonstationarity plays an important role in future precipitation changes at annual and monthly scales, while the importance of temperature bias nonstationarity in future temperature change is not significant; (3) The bias nonstationarity of climate model outputs is amplified when driving a hydrological model for hydrological simulations. In other words, the bias nonstationarity of precipitation and temperature has considerable impacts on future streamflow changes for this specific watershed. This implies that the usually predicted hydrological change in the future may be just the result of bias nonstationarity due to natural climate variability. In addition, the increase in the length of time window mitigates the impacts of bias nonstationarity on streamflow projections. Thus, it suggests using a long period (e.g. 50 years) for calibrating a bias correction method in hydrological climate change impact studies. The climatology of 20 or 30 years may not be enough for calibrating a bias correction method for impact studies.

Uncertainty of baseline observations

In this study, the CRU dataset was used as the baseline to calculate the bias of GCM-simulated precipitation and temperature. The CRU data is gridded data obtained by interpolating gauged precipitation and temperature to regular grids. Due to limited gauges used for interpolation, the CRU dataset may also be biased. In order to investigate the uncertainty related to the baseline dataset, another dataset produced by the University of Delaware (version 4.01, available at www.esrl.noaa.gov/psd/data/gridded/data.UDel_AirT_Precip.html), with a spatial resolution of 0.5° from 1901 to 2000, the same as the CRU dataset, is also used to estimate the difference in bias between each historical period and the baseline period. Figure 10 shows the difference in bias for mean annual precipitation and temperature obtained from CRU and the University of Delaware (USA) dataset for 10-, 20-, 33- and 50-year windows. Similarly, the difference in bias varies with time for all time windows. Meanwhile, the range of difference in bias decreases with the extension of the time window. Generally, the range of difference in bias is similar for two datasets, even though the magnitude is not exactly the same.

Uncertainty of greenhouse gas emission scenarios

The role of bias nonstationarity on future climate changes may rely on greenhouse gas emission scenarios, as different scenarios predict different future climate change signal. All the above results are based on the scenario of RCP4.5. In order to investigate the uncertainty related to the greenhouse gas emission scenario, the climate change signal of mean annual precipitation and temperature predicted by RCP4.5 and RCP8.5 are compared (Figure 11). The results showed that climate change signals predicted by RCP4.5 and RCP8.5 are comparable for precipitation, while the former is much smaller than the latter for temperature in terms of both median value and uncertainty related to GCMs for all time windows. Thus, the use of different greenhouse gas emission scenarios would not change the conclusion that the role of precipitation bias nonstationarity is important in future precipitation change. However, the impacts of temperature bias nonstationarity
may depend on emission scenarios. For a high emission scenario (e.g. RCP8.5), the bias nonstationarity will become even less important, while for a low emission scenario (e.g. RCP4.5), it may become important for near future periods.

**Methods to estimate natural climate variability**

For the historical period, the difference in bias between two periods is caused by natural climate variability, because the anthropogenic climate change is pre-removed. Thus, to investigate the role of bias nonstationarity in future climate change, the range of bias is defined as ‘noise’, i.e. the difference between maximum and minimum values of biases across multiple periods with outliers deleted. An alternative approach using the standard deviation of bias over multiple periods can also be used to estimate natural climate variability. Table 2 presents a comparison of the median of range of bias and standard deviation of bias over historical periods. The median value is obtained from all 17 GCMs. The results show that the range of bias is much larger than the standard deviation of bias for both precipitation and temperature, especially over the short time window. When using the standard deviation of bias instead of the range of bias, the lower values of ‘noise’ can lead to larger SNR, indicating the impacts of bias nonstationarity would be less important. However, the impacts of precipitation bias are still great for most GCMs in the near and mid future, as indicated by the SNR being smaller than 1.

**Future work**

This study only investigated the bias nonstationarity of precipitation and temperature for the historical period. Thus, the bias nonstationarity only resulted from natural climate variability. However, for a future period, bias nonstationarity resulted...
from both natural climate variability and climate model sensitivity (Chen et al. 2015; Velázquez et al. 2015; Hui et al. 2019). The climate model sensitivity is more significant than natural climate variability, especially for precipitation bias nonstationarity (Hui et al. 2019). Climate models were proposed using different structures and parameterization schemes to represent the atmospheric physics, due to the limitation of inadequate knowledge of climate system, such as cloud and convective precipitation mechanisms, surface albedo feedback and land-atmosphere interactions. Thus different models simulate somewhat different responses to the same external forcing. This may lead to different biases between different climate models and observations. As the attribution of both natural climate variability and climate model sensitivity, the performance of a bias correction method may be more deteriorated in the future climate change than historical period. This will transfer to hydrological simulations, implying that hydrological simulations forced by bias-corrected precipitation and temperature may perform even worse when taking into account the impacts of climate model sensitivity. The impacts of climate model sensitivity on bias nonstationarity can be investigated in future studies.
Due to the limited length of observed data, 100 years data were divided into 10, five, three and two non-overlapping periods respectively corresponding to four different window lengths (10, 20, 33 and 50 years). For each time window, the biases over these finite periods were used to represent the range of biases (RB) caused by natural climate variability. However, small sample size, e.g. two non-overlapping periods for the 50-year window, may lead to large uncertainty of the estimated RB. To investigate the impacts of sample size on the estimation of bias variability, the use of multiple member GCMs may be a solution (Chen & Brissette 2019). In addition, this study only used the Two-Parameter Monthly Water Balance Model to simulate monthly runoff. The hydrological model uncertainty was not considered. All these can be avenues for future studies.

ACKNOWLEDGEMENTS

This work was partially supported by the National Natural Science Foundation of China (Grant No. 51779176 and 51539009) and the National Key R&D Program of China (2019YFC0408903). The authors would like to acknowledge the contribution of the World Climate Research Program Working Group on Coupled Modelling, which is responsible for CMIP. We wish to thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

REFERENCES

Arnell, N. W. 2005 Relative effects of multi-decadal climatic variability and changes in the mean and variability of climate due to global warming: future streamflows in Britain. J. Hydrol. 270 (3–4), 195–213.

Bai, P., Liu, X. M., Liang, K. & Liu, C. M. 2015 Comparison of performance of twelve monthly water balance models in different climatic catchments of China. J. Hydrol. 529, 1030–1040.

Bellenger, H., Guiyvardi, E., Leloup, J., Lengaigne, M. & Vialard, J. 2014 ENSO representation in climate models: from CMIP3 to CMIP5. Clim. Dyn. 42 (7–8), 1999–2018.

Buser, C. M., Kunsch, H. R., Luthi, D., Wild, M. & Schar, C. 2009 Bayesian multi-model projection of climate: bias assumptions and interannual variability. Clim. Dyn. 33, 849–868.

Chen, J. & Brissette, F. P. 2009 Reliability of climate model multi-member ensembles in estimating internal precipitation and temperature variability at the multi-decadal scale. Int. J. Climatol. 39, 843–856.

Chen, H., Guo, S. L., Xu, C.-Y. & Singh, V. P. 2007 Historical temporal trends of hydro-climatic variables and runoff response to climate variability and their relevance in water resource management in the Hanjiang basin. J. Hydrol. 344 (3–4), 171–184.

Chen, J., Brissette, F. P. & Lucas-Picher, P. 2015 Assessing the limits of bias-correcting climate model outputs for climate change impact studies. J. Geophys. Res. Atmos. 120 (3), 1123–1136.

Chen, J., Brissette, F. P. & Lucas-Picher, P. 2016a Transferability of optimally-selected climate models in the quantification of climate change impacts on hydrology. Clim. Dyn. 47, 3359–3372.

Chen, J., St-Denis, B. G., Brissette, F. P. & Lucas-Picher, P. 2016b Using natural variability as a baseline to evaluate the performance of bias correction methods in hydrological climate change impact studies. J. Hydrometeorol. 17, 2155–2174.

Christensen, J. H., Boberg, F., Christensen, O. B. & Lucas-Picher, P. 2008 On the need for bias correction of global climate projections of temperature and precipitation. Geophys. Res. Lett. 35, L20709.

Dixon, K. W., Lanzante, J. R., Nath, M. J., Hayhoe, K., Stoner, A., Radhakrishnan, A., Balaji, V. & Gaitan, C. F. 2016 Evaluating the stationarity assumption in statistically downscaled climate projections: is past performance an indicator of future results? Clim. Change 135, 395–408.

Duan, Q., Sorooshian, S. & Gupta, V. K. 1992 Effective and efficient global optimization for conceptual rain-runoff models. Water Resour. Res. 28 (4), 1015–1031.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K. & Liebert, J. 2012 HESS opinions ‘Should we apply bias correction to global and regional climate model data?’. Hydrol. Earth Syst. Sci. 16, 3391–3404.

Fuentes-Franco, R., Giorgi, F., Coppola, E. & Kucharski, F. 2016 The role of ENSO and PDO in variability of winter precipitation over North America from twenty first century CMIP5 projections. Clim. Dyn. 46 (9–10), 3259–3277.

Graham, L. P., Andreassen, J. & Carlsson, B. 2007 Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods – a case study on the Lule River basin. Clim. Change 81, 293–307.

Guo, S. L., Wang, J. X., Xiong, L. H., Ying, A. W. & Li, D. F. 2002 A macro-scale and semi-distributed monthly water balance model to predict climate change impacts in China. J. Hydrol. 268 (1–4), 1–15.

Gutierrez, J. M., San-Martin, D., Brands, S., Manzanas, R. & Herrera, S. 2015 Reassessing statistical downscaling
techniques for their robust application under climate change conditions. J. Clim. 26 (1), 171–188.

Harris, C. R. U., & Jones, I. C. & D. P. 2017 Climatic Research Unit (CRU) Time Series (TS) Version 4.1 of High-Resolution Gridded Data of Monthly-Monthly Variation in Climate (Jan. 1901–Dec. 2016). Centre for Environmental Data Analysis. Oxon, UK.

Hui, Y., Chen, J., Xu, C.-Y., Xiong, L. H. & Chen, H. 2019 Bias nonstationarity of global climate model outputs: the role of internal climate variability and climate model sensitivity. Int. J. Climatol. 39 (4), 2278–2294.

Hulme, M., Barrow, E. M., Arnell, N. W., Harrison, P. A., Johns, T. C. & Downing, T. E. 1999 Relative impacts of human-induced climate change and natural climate variability. Nature 397 (6721), 688–691.

Kendall, M. G. 1975 Rank Correlation Methods. Griffin, London, UK.

Li, L., Diallo, I., Xu, C.-Y. & Stordal, F. 2015 Hydrological projections under climate change in the near future by RegCM4 in Southern Africa using a large-scale hydrological model. J. Hydrol. 528, 1–16.

Lu, W. & Qin, X. S. 2020 Integrated framework for assessing climate change impact on extreme rainfall and the urban drainage system. Hydrol. Res. 51, 77–89.

Mann, H. B. 1945 Nonparametric tests against trend. Econometrika 13, 245–259.

Maraun, D. 2012 Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums. Geophys. Res. Lett. 39, L06706.

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brien, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M. & Thiele-Eich, I. 2010 Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. Rev. Geophys. 48 (3), RG3003.

Marhaento, H., Buij, M. J. & Hoekstra, A. Y. 2017 Attribution of changes in stream flow to land use change and climate change in a mesoscale tropical catchment in Java, Indonesia. Hydrol. Res. 48 (3–4), 1143–1155.

Maurer, E. P., Das, T. & Cayan, D. R. 2013 Errors in climate model daily precipitation and temperature output: time invariance and implications for bias correction. Hydrol. Earth Syst. Sci. 17 (6), 2147–2159.

Nahar, J., Johnson, F. & Sharma, A. 2017 Assessing the extent of non-stationary biases in GCMs. J. Hydrol. 549, 148–162.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models. J. Hydrol. 10, 282–290.

Olsson, T., Jakkila, J., Veijalainen, N., Backman, L., Kaurola, J. & Velvilaivanen, B. 2015 Impacts of climate change on temperature, precipitation and hydrology in Finland-studies using bias corrected Regional Climate Model data. Hydrol. Earth Syst. Sci. 19 (7), 3217–3238.

Ouyang, R., Liu, W., Fu, G., Liu, C., Hu, L. & Wang, H. 2014 Linkages between ENSO/PDO signals and precipitation, streamflow in China during the last 100 years. Hydrol. Earth Syst. Sci. 18 (9), 3561–3566.

Polade, S. D., Gershunov, A., Cayan, D. R., Dettinger, M. D. & Pierce, D. W. 2013 Natural climate variability and teleconnections to precipitation over the Pacific-North American region in CMIP5 and CMIP5 models. Geophys. Res. Lett. 40 (10), 2296–2301.

Power, S., Casey, T., Folland, C., Colman, A. & Mehta, V. 1999 Inter-decadal modulation of the impact of ENSO on Australia. Clim. Dyn. 15, 319–324.

Ragettli, S., Tong, X., Zhang, G., Wang, H., Zhang, P. & Stähli, M. 2020 Climate change impacts on summer flood frequencies in two mountainous catchments in China and Switzerland. Hydrol. Res. https://doi.org/10.2166/nh.2019.118.

Ruiz-Barradas, A., Nigam, S. & Kavvada, A. 2013 The Atlantic Multidecadal Oscillation in twentieth century climate simulations: uneven progress from CMIP3 to CMIP5. Clim. Dyn. 41, 3301–3315.

Shen, M., Chen, J., Zhuang, M., Chen, H., Xu, C.-Y. & Xiong, L. 2018 Estimating uncertainty and its temporal variation related to global climate models in quantifying climate change impacts on hydrology. J. Hydrol. 556, 10–24.

Teutschbein, C. & Seibert, J. 2012 Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. J. Hydrol. 456–457, 12–29.

Thornthwaite, C. W. 1948 An approach toward a rational classification of climate. Geograph. Rev. 38 (1), 55–94.

Van Pelt, S. C., Kabat, P., Ter Maat, H. W., Van den Hurk, B. J. J. M. & Weerts, A. H. 2009 Discharge simulations performed with a hydrological model using bias corrected regional climate model input. Hydrol. Earth Syst. Sci. 13 (12), 2387–2397.

Velázquez, J. A., Troin, M., Caya, D. & Brissette, F. 2015 Evaluating the time-invariance hypothesis of climate model bias correction: implications for hydrological impact studies. J. Hydrometeorol. 16 (5), 2013–2026.

Wang, Y., Sivandran, G. & Bielicki, J. M. 2017 The stationarity of two statistical downscaling methods for precipitation under different choices of cross-validation periods. Int. J. Climatol. 38 (S1), e330–e348.

Xiong, L. H. & Guo, S. L. 1999 A two-parameter monthly water balance model and its application. J. Hydrol. 216 (1–2), 111–123.

Xu, C.-Y. & Singh, V. P. 2001 Evaluation and generalization of temperature-based methods for calculating evaporation. Hydrol. Process. 15 (2), 305–319.

Zhuang, M. J., Chen, J., Shen, M. X., Xu, C.-Y., Chen, H. & Xiong, L. H. 2018 Timing of human-induced climate change emergence from internal climate variability for hydrological impact studies. Hydrol. Res. 49 (2), 421–437.

Zhuang, M. J., Chen, J., Xu, C.-Y., Zhao, C., Xiong, L. H. & Liu, P. 2019 A method for investigating the relative importance of three components in overall uncertainty of climate projections. Int. J. Climatol. 39, 1853–1871.

First received 2 May 2020; accepted in revised form 25 August 2020. Available online 7 October 2020.