Received: 13 October 2018 | Revised: 26 March 2019 | Accepted: 27 March 2019 | Published on: 16 April 2019

DOI: 10.1002/joc.6086

RESEARCH ARTICLE

International Journal of Climatology RMetS

Limitations of univariate linear bias correction in yielding cross-correlation between monthly precipitation and temperature

R. Das Bhowmik[1] | **A. Sankarasubramanian**[2]

[1]Department of Civil Engineering, Indian Institute of Science, Bangalore, India

[2]Department of Civil, Construction, and Environmental Engineering, North Carolina State University, Raleigh, North Carolina

Correspondence
R. Das Bhowmik, Department of Civil Engineering, Indian Institute of Science, CV Raman Road, Bengaluru 560012, India.
Email: rajarshidb@iisc.ac.in

Funding information
National Science Foundation, Grant/Award Number: CBET- 1204368

Abstract

Statistical bias correction techniques are commonly used in climate model projections to reduce systematic biases. Among the several bias correction techniques, univariate linear bias correction (e.g., quantile mapping) is the most popular, given its simplicity. Univariate linear bias correction can accurately reproduce the observed mean of a given climate variable. However, when performed separately on multiple variables, it does not yield the observed multivariate cross-correlation structure. In the current study, we consider the intrinsic properties of two candidate univariate linear bias-correction approaches (simple linear regression and asynchronous regression) in estimating the observed cross-correlation between precipitation and temperature. Two linear regression models are applied separately on both the observed and the projected variables. The analytical solution suggests that two candidate approaches simply reproduce the cross-correlation from the general circulation models (GCMs) in the bias-corrected data set because of their linearity. Our study adopts two frameworks, based on the Fisher z-transformation and bootstrapping, to provide 95% lower and upper confidence limits (referred as the permissible bound) for the GCM cross-correlation. Beyond the permissible bound, raw/bias-corrected GCM cross-correlation significantly differs from those observed. Two frameworks are applied on three GCMs from the CMIP5 multimodel ensemble over the coterminous United States. We found that (a) the univariate linear techniques fail to reproduce the observed cross-correlation in the bias-corrected data set over 90% (30–50%) of the grid points where the multivariate skewness coefficient values are substantial (small) and statistically significant (statistically insignificant) from zero; (b) the performance of the univariate linear techniques under bootstrapping (Fisher z-transformation) remains uniform (non-uniform) across climate regions, months, and GCMs; (c) grid points, where the observed cross-correlation is statistically significant, witness a failure fraction of around 0.2 (0.8) under the Fisher z-transformation (bootstrapping). The importance of reproducing cross-correlations is also discussed along with an enquiry into the multivariate approaches that can potentially address the bias in yielding cross-correlations.

KEYWORDS

bias correction, bootstrapping, cross-correlation, Fisher z-transformation, precipitation, temperature, univariate

1 | INTRODUCTION

Bias correction techniques remove systematic biases in general circulation models (GCMs) and support climate-application studies (Hay and Clark, 2003; Hanson and Dettinger, 2005; Mejia *et al*., 2012; Singh *et al*., 2014; Seo *et al*., 2016). The most commonly employed bias correction technique is the univariate approach (Huth, 1999; Wood *et al*., 2004; Stoner *et al*., 2012), which develops an empirical relationship between the model climate variable and the observed information. Multivariate techniques, such as bivariate ranks, joint variable statistical bias correction, and asynchronous canonical correlation analysis (ACCA), have the potential to reproduce a cross-correlation structure (Zhang and Georgakakos, 2012; He *et al*., 2012; Mehrotra and Sharma, 2016; Das Bhowmik *et al*., 2017) among the observed meteorological variables in the bias-corrected data set. However, univariate bias correction remains a popular choice because of its simplicity and its ability to reproduce the observed mean and standard deviation in a bias-corrected GCM data set (Wood *et al*., 2004; Maurer and Hidalgo, 2008).

Univariate bias correction approaches vary from delta change or scaling (Mpelasoka and Chiew, 2009), multiple linear regression (Huth, 1999), quantile mapping (Wood *et al*., 2004), asynchronous regression (Stoner *et al*., 2012) to semi-parametric (Sankarasubramanian and Lall, 2003) and nonparametric approaches (Gangopadhyay *et al*., 2004). The Bureau of Reclamation (BOR), a water management agency under the U.S. Department of the Interior, used quantile mapping as the bias correction approach (bias correction and spatial disaggregation [BCSD]) to develop bias-corrected monthly precipitation (pr) and monthly average temperature (tas) data sets over the coterminous United States (CONUS) at 1/8° from Coupled Model Intercomparison Project Phase 3 (CMIP3) and Phase 5 (CMIP5) experiments (Reclamation, 2013). Linear-regression-based univariate techniques, simple linear regression and asynchronous regression, focus only on a single variable to correct the bias. Therefore, such techniques have limited ability to yield a multivariate correlation structure (He *et al*., 2012; Das Bhowmik *et al*., 2017). Although several former studies developed multivariate bias correction techniques, only a few of them estimated the intervariable dependence before and after bias correction. In recent years, two major studies have examined the cross-correlation between bias-corrected meteorological variables (Wilcke *et al*., 2013; Ivanov and Kotlarski, 2017). Both studies reported that the cross-correlation between raw climate model outputs and between bias-corrected climate model outputs have similar magnitudes. These studies considered quantile mapping as their primary approach for bias correction. Although Wilcke *et al*. (2013) considered cross-correlations between five meteorological variables, none of the two studies compared the observed and the bias-corrected/raw climate model cross-correlations. Nonetheless, to the best of our knowledge, the limitations of the univariate linear bias correction approaches have never been examined analytically.

Increasing temperature, owing to global climate change, has been documented over the last half-century in both observed records as well as GCM simulations (Plummer *et al*., 1999; Alexander *et al*., 2006; Qin *et al*., 2010; IPCC, 2013; Blunden and Arndt, 2014). This increasing temperature trend inherently changes the cross-correlation structure between precipitation and other variables (Das Bhowmik *et al*., 2017). Therefore, it is essential to correct the bias in the cross-correlation between meteorological variables. Das Bhowmik *et al*. (2017) proposed a multivariate bias correction technique, based on ACCA that reproduces the observed cross-correlation structure between monthly precipitation and monthly average temperature in bias-corrected variables. Multivariate bias correction schemes result in an improved ability to estimate the joint likelihood of precipitation and temperature. However, if the observed cross-correlation is not statistically significant, the univariate linear bias correction itself will suffice. Hence, it is essential to identify the grid points where multivariate bias correction would be of benefit in removing the model bias in cross-correlations. The grid points are identified based on the sampling distribution of the observed cross-correlation.

The primary objectives of this study are (a) to systematically show the limitations of two candidate univariate bias correction techniques, simple linear regression and asynchronous regression, in order to estimate the cross-correlation structure across multiple variables and (b) to apply two statistical frameworks that identify the grid points where the cross-correlation between raw/bias-corrected GCM variables differs significantly from that between observed variables. The limitations are derived both analytically and from the bias-corrected variables of the GCM simulations. The analytical expression developed for the bias-corrected cross-correlation is evaluated based on the CMIP5 projections of precipitation and temperature from the BCSD. We aim to develop a sampling distribution of the observed cross-correlation based on Fisher z-transformation and bootstrapping algorithms. The Fisher z-transformation allows the cross-correlation to follow a Gaussian distribution, which is later used to construct 95% upper and lower confidence limits for the GCM cross-correlation, whereas in the case of bootstrapping, we resample (with repetition) the observed precipitation and temperature to obtain a sampling distribution of the observed cross-correlation. Based on the sampling distribution, we identify the climate regions over

the CONUS where the univariate approaches fail to bias-correct the cross-correlation.

This manuscript is organized as follows: the next section presents details of the selected GCMs and the observed data set used for the study. Next, we derive the analytical expression for the downscaled cross-correlation using univariate linear techniques and then apply the Fisher z-transformation to develop the 95% upper and lower confidence limits (hereafter referred to as "permissible bound") of the observed cross-correlation. The results are presented in section 4, which is followed by a discussion and concluding remarks.

2 | DATA SOURCES AND MOTIVATION

2.1 | Data sources

We consider the observed monthly precipitation (pr) and monthly average temperature (tas) as well as the raw and bias-corrected simulations of the monthly precipitation and monthly temperature from three GCMs. Three GCM models, CNRM-CM5, IPSL-CM5-LR, and MPI-ESM-LR, from the Coupled Model Intercomparison Project Phase 5 (CMIP5) are selected. In the current analysis, we obtained the GCM and observed data for the period 1950–1999. Details of the data set are provided in Table 1. We present our analysis for nine National Climatic Data Center (NCDC) climate regions—Northwest (NW), West North Central (WNC), East North Central (ENC), Central (C), Northeast (NE), Southeast (SE), South (S), Southwest (SW), and West (W)—over the CONUS (Karl and Koss, 1984). The univariate bias-corrected monthly precipitation and temperature for the historical runs of these three GCMs' are obtained from the BOR's BCSD climate projections archive (https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/). Raw GCM outputs are also obtained from the same archive; however, raw GCM outputs can also be downloaded from CMIP5 data archive (https://esgf-node.llnl.gov/search/cmip5/). BCSD method follows three major steps to correct a model bias: (a) It re-grids the raw GCM simulations and projections to 1° spatial resolution; (b) it constructs cumulative distribution functions based on the GCM simulations and the observed data for 1950–1999 to identify the model bias; and (c) it adjusts the bias in the quantile maps of historical simulations and future projections. The CNRM has a spatial resolution of 1.4°, while the IPSL and MPI both have a spatial resolution of 1.8°. In the current study, to compare the cross-correlation values from the three GCMs, raw historical runs are further re-gridded to ~1° using bilinear interpolation. The gridded observed monthly precipitation and monthly average temperature data set over the CONUS is provided by Ed Maurer's research group (Maurer *et al*., 2007). We obtained the observed data set from BOR's climate projections archive, which provides the data set at a spatial resolution of 1° and as monthly values.

The GCM historical runs may have more than one ensemble member. We calculate the two cross-correlations between precipitation and temperature for each ensemble member: (a) raw cross-correlation—correlation between the precipitation and temperature from a raw GCM run, (b) bias-corrected cross-correlation—correlation between the bias-corrected precipitation and temperature. Additionally, we estimate the observed cross-correlation—the correlation between the observed monthly precipitation and observed monthly average temperature.

The observed cross-correlations are checked for their statistical significance using the equation $[\pm 1.96/\sqrt{(n-3)}]$ based on 95% confidence interval, where n is the number of years of observation. If a grid point exhibits a statistically significant cross-correlation between precipitation and temperature, it will henceforth be referred to as a "significant grid point."

2.2 | Spatial patterns of observed cross-correlation

The motivation behind investigating the limitations of univariate linear bias correction is the significant cross-correlation between the observed precipitation and temperature for four different months over the CONUS (Figure 1). We superimpose the climate region boundaries on the CONUS map. We consider 4 months (January, April, July, and October) to exhibit the results. These 4 months represent the four seasons (summer, fall, spring, and winter) of the year over the CONUS. In January (Figure 1), a strong linear dependency exists between precipitation and temperature over the north-eastern, central, and southern parts of the United States, with

TABLE 1 Description of GCM data set

Model name	Organization	Experiment	Time length	Ensemble members
CNRM-CM5	Centre National de Recherches Météorologiques-Groupe d'étude de l'Atmosphère Météorologique) and CERFACS (Voldoire *et al*., 2013)	Historical	1950–1999	5
MRI-ESR-LR	Max-Planck-Institut Fur Meteorologie (Giorgetta *et al*., 2013)	Historical	1950–1999	1
IPSL-CM-LR	Institute Pierre' Simon Laplace (Dufresne *et al*., 2013)	Historical	1950–1999	4

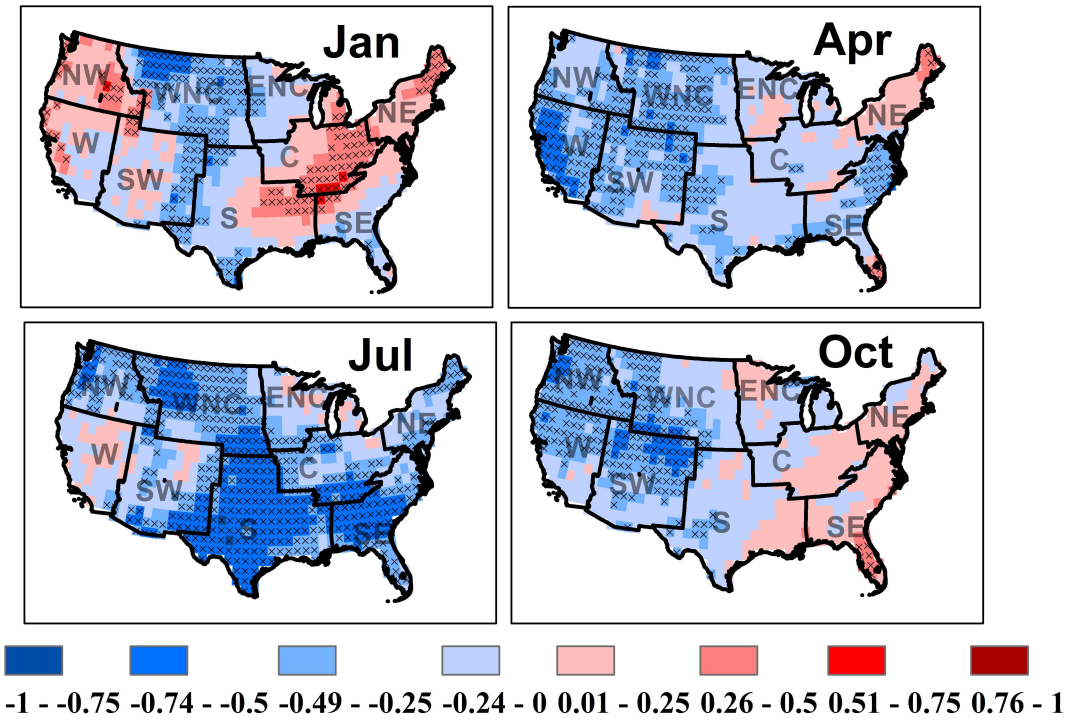

FIGURE 1 Observed cross-correlation between monthly precipitation and monthly average temperature for the period 1950–1999. Grid points with significant cross-correlation are marked. Observed cross-correlations are checked for their statistical significance using the equation [$\pm 1.96/\sqrt{(n-3)}$] based on 95% confidence interval, where n is the number of years of observation. NCDC climate region boundaries are imposed on the country map [Colour figure can be viewed at wileyonlinelibrary.com]

the cross-correlation values closer to one. West north central United States experiences statistically significant but negative cross-correlation values between −0.5 and −0.75. During April, most of the grid points across the United States experience positive cross-correlations. Western United States experiences a higher number of significant grid points during April as compared to the rest of the climate regions. During July, half of the significant grid points have negative cross-correlation values. This is expected since Trenberth and Shea (2005) reported that global P–T cross-correlation during summer is positive over the land, whereas, for the high latitudes, the cross-correlation during winter season is positive. In October, a negative dependency is observed for the grid points over western and northwestern United States. This strong interdependency between precipitation and temperature across different months and over different climate regions indicates the importance of reproducing the observed cross-correlation structure in downscaled GCM fields. For instance, if the cross-correlation of the downscaled GCM attributes is overestimated in the summer over the ENC, where a substantial fraction (around 0.7 annually) of precipitation is lost as evapotranspiration (Sanford and Selnick, 2013), it could impact the estimated evapotranspiration under potential climate change scenarios. We refer to ENC as the region experiences negative cross-correlation during the summer indicating a strong role of land-surface and atmosphere feedback, whereas for other regions with strong dependency between precipitation and temperature, the amount of precipitation lost as evapotranspiration is lower compared to the ENC. For these regions, although the cross-correlation is strong, other factors (e.g., land cover, vegetation type) are reducing the impact of P–T cross-correlation on the hydrologic process. Detailed spatial patterns of the cross-correlation could be analysed based on false discovery rate (Wilks, 2006; Ivanov *et al*., 2018a; 2018b), which explores the spatial dependence in cross-correlation. We conclude that the bias in estimating the cross-correlation is an important attribute to consider in the case of bias-corrected variables.

3 | UNIVARIATE BIAS CORRECTION: LIMITATIONS AND BOUNDS

For univariate linear bias correction, we assume two independent linear regression models that correct the bias in precipitation and temperature separately. Readers are requested to note that the current study considers simple linear and asynchronous regressions from a relatively large set of univariate approaches available. The findings of biased estimation of cross-correlation is true for only two univariate methods, linear regression and asynchronous regression (or quantile mapping), in bias-correcting GCM outputs.

This cannot be true for other methods such as resampling procedure based on univariate methods. Between the two univariate methods, asynchronous regression (also known as the quantile mapping) is extremely popular among hydroclimate community to preprocess GCM outputs for long-term hydrologic simulation (Tao *et al*., 2014; Johnson and Sharma, 2015; Reclamation, 2013; Seo *et al*., 2016). We employ two frameworks (Fisher *z*-transformation and bootstrapping) to evaluate whether the bias-corrected cross-correlation is statistically different from the observed cross-correlation (Fisher, 1925; Lall and Sharma, 1996). We combine the analytical expression from the linear regression model along with the 95% upper and lower confidence limits (permissible bounds) of the observed cross-correlation. Finally, we investigate the GCMs' ability to estimate the observed cross-correlation under univariate linear bias correction based on a combination of the analytical expression and the permissible bound.

3.1 | Limitation of univariate linear bias correction in estimating cross-correlation

Bias correction using the univariate linear method typically considers the linear regression between the observed and the GCM simulated variable over the past period. The estimated regression parameters from the past (Equations (1) and (2)) are applied to obtain the bias-corrected GCM projections for the future period. The current study implements a univariate asynchronous linear regression model at each grid point for bias correction. The original linear asynchronous regression model was proposed by Dettinger *et al*. (2004), which was later modified to an asynchronous piece-wise linear regression by Stoner *et al*. (2012). Our study uses the simplified version: asynchronous linear regression. The regression model relates the sorted (i.e., ascending/descending) observations (pr/tas) and GCMs as predictands and predictors, respectively. In general, a sorted GCM variable exhibit strong linear dependency with the sorted observed variable (see Stoner *et al*., 2012, figs 2 and 3). For example, we found that the regression slope between sorted observed precipitation and sorted GCM precipitation during January has a spatial mean of one (see Figures S1 and S2, Supporting Information). A regression model with sorted variables implies an asynchronous regression or quantile mapping for bias correction. Thus, the developed univariate models can be written as

$$\hat{P}_{BC} = \hat{a}.P_{GCM} + \hat{b}, \tag{1}$$

$$\hat{T}_{BC} = \hat{c}.T_{GCM} + \hat{d}, \tag{2}$$

where model parameters $[\hat{a}, \hat{b}]$ and $[\hat{c}, \hat{d}]$ are estimated based on the observed (P_{obs}, T_{obs}) and raw GCM projections of precipitation and temperature (P_{GCM}, T_{GCM}). $\hat{P}_{BC}$ and $\hat{T}_{BC}$ are the bias-corrected precipitation and temperature, respectively. Standard statistical property states that linear transformation does not alter the correlation between the original and the transformed variable. However, the implication of the statistical property on independent bias corrections of multiple variables and the associated cross-correlation structure has not been documented yet. We consider two independent asynchronous regressions to derive the analytical expressions for the estimated cross-correlation from the bias-corrected $\hat{P}_{BC}$ and $\hat{T}_{BC}$ (see Equations (A1)–(A6)),

$$\mathrm{corr}\left(\hat{P}_{BC}, \hat{T}_{BC}\right) = \mathrm{corr}\left(P_{GCM}, T_{GCM}\right). \tag{3}$$

The bias-corrected GCM cross-correlation will be the same as that of the raw GCM cross-correlation if regression based univariate bias correction is applied individually to multiple variables using the linear regression model structure (Equation (3)). Equation (3) confirms that the reason behind the limitation of simple linear and asynchronous regressions to yield the observed cross-correlations is the linearity of these univariate approaches, which ensures that the transfer function is linear (or close to linear) at the grid points. The relation between raw and bias-corrected cross-correlation remains unchanged as long as we apply a linear model structure for the bias correction. Thus, irrespective of an asynchronous or a regular linear regression, the findings qualify for bias correcting both climate forecasts (i.e., monthly-to-seasonal forecasts) and climate change projections.

We verify the analytical expression (Equation (3)) after comparing the raw GCM cross-correlation obtained from the CNRM-CM5 model with the bias-corrected cross-correlation for the same GCM obtained from the BOR's BCSD database (Figure 2). Bias-corrected GCM outputs and raw GCM variables of 815 grid points over the CONUS are gathered for 4 months (January, April, July, and October). We examine the raw and the bias-corrected GCM cross-correlation individually for each of the 4 months. The results show that the bias-corrected GCM cross-correlation is almost equal to the raw GCM cross-correlation, thereby not reducing the bias in cross-correlation in the latter. The average absolute differences from the identity line (1:1 line), to yield the raw cross-correlation by the bias-corrected BCSD products from the BOR, are 0.03, 0.03, 0.05, and 0.04 for January, April, July, and October, respectively. Deviations of the bias-corrected cross-correlations from the identity line are due to the nonlinear transfer functions between the raw and bias-corrected GCM cross-correlations. Figure 2 confirms that the univariate linear bias-corrected variables simply reproduce the raw GCM cross-correlations without correcting for the observed cross-correlation. This limitation is critical since

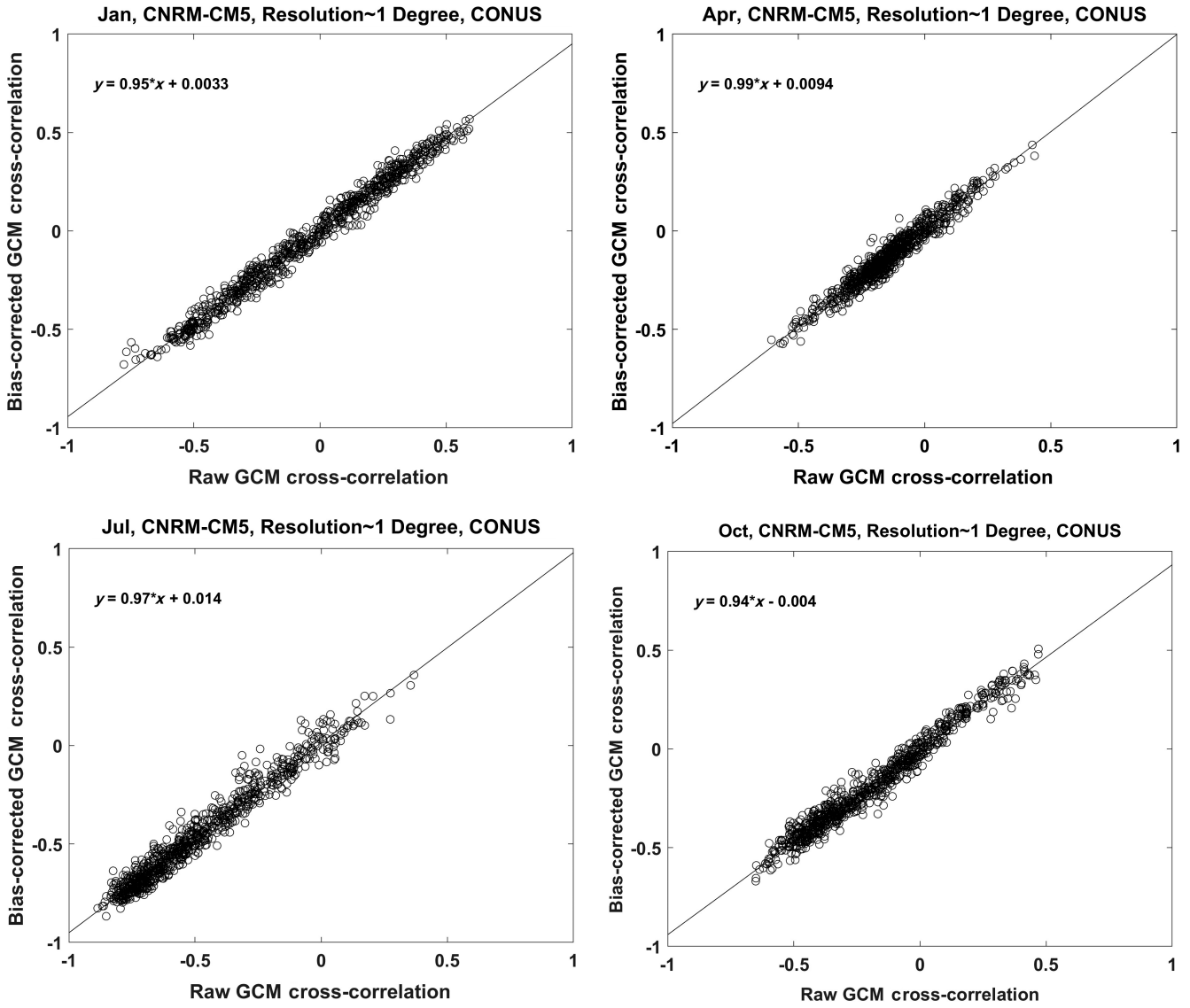

FIGURE 2 Comparison between raw GCM cross-correlation and bias-corrected GCM cross-correlation. Each dot represents a grid point on the CONUS, *X* value indicates the raw GCM cross-correlation at that grid point, and the *Y* value indicates the bias-corrected GCM cross-correlation at that grid point. Bias-corrected and raw CNRM-CM5 values for 1950–1999, over the CONUS, are collected from the BOR's achieve (BCSD climate projections). A single ensemble member from the CNRM-CM5 historical run is used for demonstration. The linear fit equation is provided in the inset

the univariate linear bias-corrected cross-correlation will not be the same as the observed cross-correlation on the significant grid points (Figure 1). Accordingly, assessing whether the raw cross-correlation is statistically different from the observed cross-correlation will provide us an insight into when and where the univariate procedure is likely to produce biased cross-correlations.

3.2 | Sampling distribution of cross-correlation based on Fisher *z*-transformation

The Fisher *z*-transformation transforms the correlation distribution into a distribution that is closer to a Gaussian distribution. In the current study, the Fisher *z*-transformation is considered to estimate the sampling distribution of the correlation. R. A. Fisher proposed the implementation of the *z*-transformation of sample correlation (Fisher, 1925; 1950) when the sampling distribution of a correlation is skewed and the population correlation is not equal to zero. The Fisher *z*-transformation is commonly used to test the null hypothesis that two independent correlations are significantly different from one another (Devineni and Sankarasubramanian, 2010a; 2010b). We incorporate the Fisher *z*-transformation to evaluate whether the estimated cross-correlation is significantly different from the observed cross-correlation.

The current study assumes that the bias-corrected/raw (ρ^{BC} or ρ^{GCM}) and observed cross-correlations (ρ^{obs}) follow a bivariate normal distribution and applies the Fisher z-transformation [Z_{obs} and Z_{GCM}, based on $Z = 0.5 \ln(\frac{1+\rho}{1-\rho})$] on them. This method forces the cross-correlations to follow a normal distribution with standard deviation equal to $1/\sqrt{(n-3)}$, where n is the number of data points used to calculate the cross-correlation. The test statistic, $\hat{Z}$ ($\hat{Z} = \frac{Z_{obs} - Z_{GCM}}{\sigma_{Z_{obs} - Z_{GCM}}}$), to retain or reject the null hypothesis (which states that two independent correlations are significantly different from one another) follows a standard normal distribution. Our interest lies in determining whether the two cross-correlations can be considered the same for the chosen confidence interval. We calculate both $\rho_{high}^{BC/GCM}$ and $\rho_{low}^{BC/GCM}$ (see Appendix B), which are the permissible values of the bias-corrected/raw cross-correlation, given ρ^{obs}, n and the chosen confidence interval (Equations (4a) and (4b)),

$$\rho_{low}^{BC/GCM} = \left[\frac{(C-1)+\rho^{obs}(C+1)}{(C+1)+\rho^{obs}(C-1)}\right], \quad (4a)$$

$$\rho_{high}^{BC/GCM} = \left[\frac{1-C+\rho^{obs}(1+C)}{1+C+\rho^{obs}(1-C)}\right], \quad (4b)$$

where $C = 1.96 \times \frac{2\sqrt{2}}{\sqrt{(n-3)}}$ (for 95% CI).

3.3 | Sampling distribution of cross-correlation based on bootstrapping

The current study also undertakes a bootstrapping algorithm to obtain the sampling distribution of cross-correlation for the grid points where pr and tas do not follow a bivariate normal distribution. The Fisher z-transformation assumes that the sample data follows a bivariate normal distribution. However, meteorological variables may not follow a bivariate Gaussian distribution and may experience a high magnitude of skewness. When tested, we found that the multivariate skewness coefficient value (Mardia, 1970) for a grid point on CONUS can be either substantial and statistically significant from zero or small but statistically insignificant from zero (results are not included in the current study). Therefore, we design a bootstrapping algorithm to construct 95% upper and lower confidence limits (permissible bounds) of cross-correlations for the grid points that experience a statistically significant from zero and substantial multivariate skewness coefficient value. Note that a multivariate skewness coefficient value higher than 1.13 (approximately) for a 50-year long pr and tas data set is generally considered to be statistically significant from zero. Therefore, we consider a multivariate skewness coefficient of 1.13 as a threshold for the current data set. If a grid point experiences a multivariate skewness lesser than this threshold, we apply the Fisher z-transformation to construct the sampling distribution of cross-correlation; otherwise, the bootstrapping algorithm is followed to obtain the confidence limits $-\left[\rho_{high}^{BC/GCM}, \rho_{low}^{BC/GCM}\right]$. We follow a nonparametric resampling with replacement to synthetically generate a time series for the pr and tas data set (Lall and Sharma, 1996; Hamill, 1999; Devineni and Sankarasubramanian, 2010b; Libera and Sankarasubramanian, 2018). The equally likely samples of pr and tas are simultaneously obtained from the observed data set. For each synthetic series, the cross-correlation between pr and tas is estimated to form the sampling distribution of cross-correlation. The null hypothesis assumes that the observed cross-correlation and the GCM cross-correlation are from the same population. Details of the bootstrapping sampling distribution are provided in Appendix C.

Following the two approaches, Fisher z and bootstrapping, we demonstrate the shape of the confidence limits with respect to the observed cross-correlation magnitudes. Figure 3a (Fisher z) and Figure 3b (bootstrapping) indicate the approaches that the current study follows to count the grid points that have a raw GCM cross-correlation beyond the confidence limits. Equations (4a) and (4b) are used to estimate the permissible bounds of the raw cross-correlation by varying observed cross-correlations between −1 to 1 at a 95% confidence interval (Figure 3a) for $n = 50$. On the other hand, bootstrapping is applied to construct confidence limits based on the observed cross-correlation values for July. Therefore, the range of observed cross-correlation (the X-axis) in Figure 3b is restricted to the actual observed cross-correlation values for July. Figure 3a,b also shows the raw GCM cross-correlations for July from the CNRM-CM5 historical runs. For the current demonstration, we apply a fourth order polynomial on the bootstrapping confidence limits to obtain a smoother curve. For the Fisher z-transformation, the observed cross-correlation values closer to zero have a broader bandwidth of permissible bound (Figure 3a) as compared to the observed cross-correlation values closer to −1 or 1. The bound becomes narrower as the length of the observation record increases (figure not shown). The raw cross-correlation from a 30-year simulation has a higher sampling variability as compared to the cross-correlation estimated using the 50-year data set. In bootstrapping, the 95% upper and lower confidence limits do not follow a regular shape as compared to the confidence limits of the Fisher z. However, the shape should become finite as the number of synthetic simulations increase. Nevertheless, Figure 3b shows that the confidence bounds become narrower as the absolute values of the observed cross-correlation increase. Additionally, we infer that the width of the confidence limits

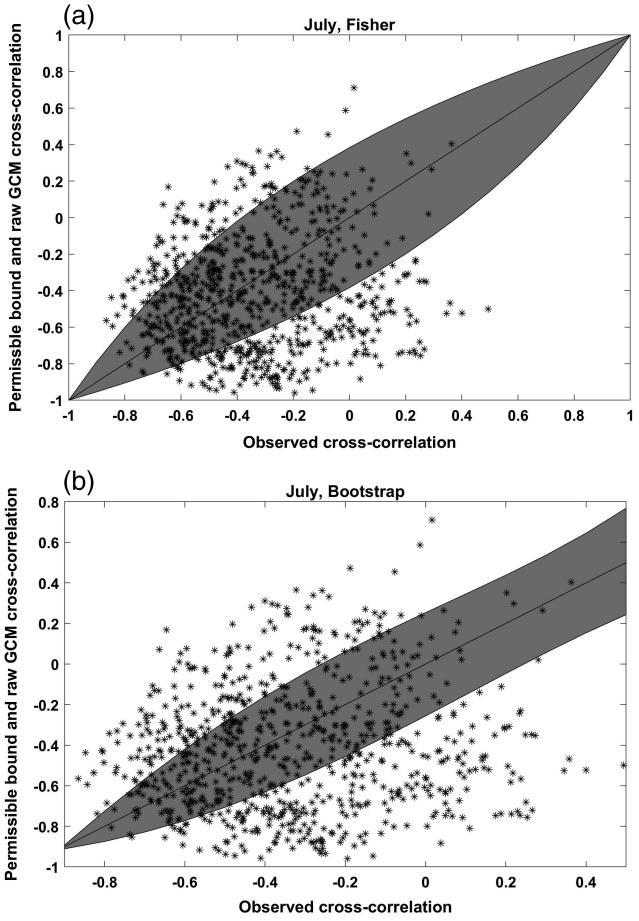

FIGURE 3 Raw GCM cross-correlations from CNRM-CM5 historical runs (July) plotted based on the observed cross-correlation and its 95% upper and lower confidence limits (permissible bounds). Permissible bounds are estimated using Fisher *z*-transformation (a) and bootstrapping (b). Each dot represents a grid point on the CONUS, *X* value indicates the observed cross-correlation at that grid point, and the *Y* value indicates the raw GCM cross-correlation at that grid point. The study used Fisher *z*-transformation (bootstrapping) to create the permissible bound for a grid point whose multivariate skewness coefficient is lesser (higher) than the threshold value. A multivariate skewness coefficient value greater than the threshold value is statistically significant at zero, and vice versa. The permissible bound from bootstrapping is constructed by fitting a fourth degree polynomial

is narrower in the case of bootstrapping than the Fisher *z*. In each plot (Figure 3a,b), each dot represents a raw GCM cross-correlation value for a particular grid point. Thereafter, we count the number of grid points that are outside the 95% upper and lower confidence limits.

Finally, we assess the statistical difference between the two cross-correlations (observed and raw) by calculating the monthly failure fraction. Failure fraction denotes the fraction of grid points over the CONUS or a climate region that witness raw cross-correlation values beyond the permissible bound (Equation (5)).

$$\text{Failure fraction } f^m = \left[\frac{\sum_{i=1}^{k}\sum_{j=1}^{J} I^{i,m,j}}{k \times L}\right], \qquad (5)$$

where $(m = 1,\ldots, 12)$, $i(i = 1,\ldots, k)$, and $j(j = 1,\ldots, L)$ denote month, grid point, and ensemble member respectively, with $I^{i,m,j} = 0$ if either $\rho^{i,m}_{\text{low}} > \rho^{i,m,j}_{\text{GCM}}$ or $\rho^{i,m,j}_{\text{GCM}} > \rho^{i,m}_{\text{high}}$ which is otherwise one.

4 | RESULTS

For sections 4.1 and 4.2, we consider only the significant grid points. Non-significant grid points exhibit a smaller magnitude of cross-correlation values as compared to significant grid points, resulting in a broader permissible bound (from Fisher *z*-transformation). Therefore, the results related to non-significant grid points are excluded from sections 4.1 and 4.2. Furthermore, if the multivariate skewness coefficient for a grid point is higher (lower) than its threshold value, bootstrapping (Fisher *z*-transformation) is applied to obtain the 95% upper and lower confidence limits. From Equation (3) it is clear that the bias-corrected GCM cross-correlation (using univariate linear bias correction) is same as the raw GCM cross-correlation. Therefore, following sections 4.1–4.3 consider raw GCM cross-correlation for the analysis.

4.1 | Comparison across raw GCM outputs in estimating the observed cross-correlation

The failure fraction values from Fisher *z* and bootstrapping for three GCMs (CNRM-CM5, IPSL, and MPI) over a period of 12 months over the CONUS are calculated and their results are shown in Figure 4. The secondary axis presents the number of grid points considered during the calculation. Under the Fisher *z*-transformation, the failure fraction values from CNRM (around 0.3) and MPI (around 0.5) historical runs are higher during April than the other months. The IPSL exhibits the highest magnitude of failure fraction (approximately 0.4) during December. During July, when half of the total grid points are statistically significant, the univariate linear bias correction could potentially yield cross-correlations that differ significantly from the observed ones in up to 40% of the grid points. However, the failure fraction values related to bootstrapping exhibit that almost all grid points (failure fraction values around 0.9) fail to reproduce the observed cross-correlation, irrespective of the GCM or the month of the year. Overall, all the three climate models exhibit a more or less similar performance to yield the observed cross-correlation. However, the IPSL exhibits slightly lower failure fraction values under the Fisher

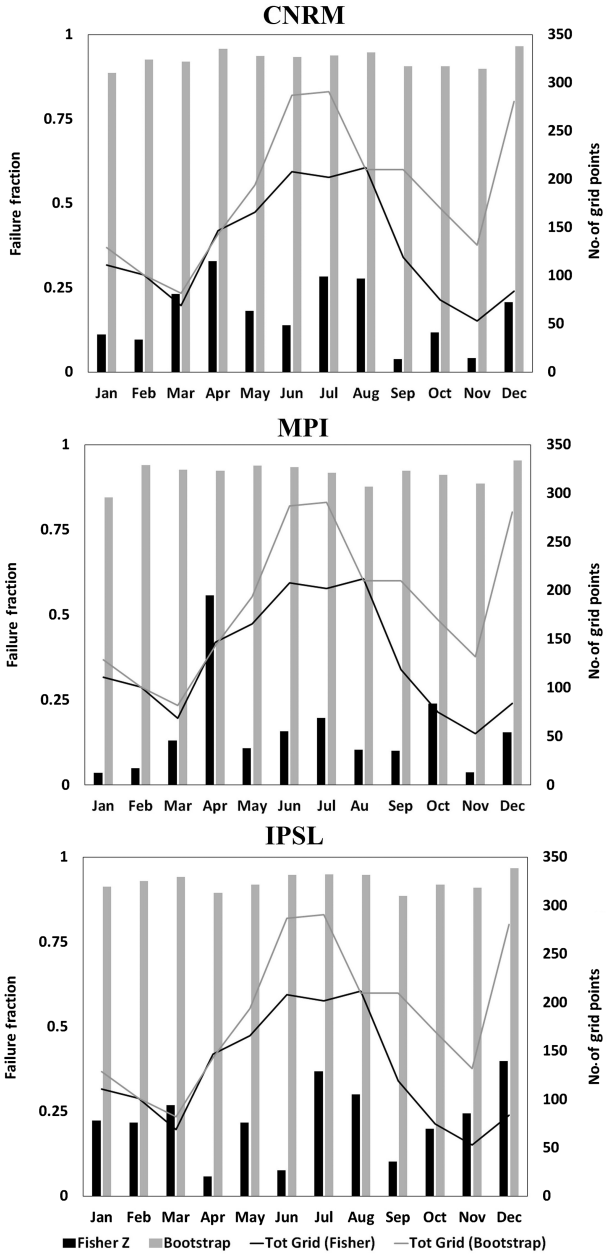

FIGURE 4 Failure fraction for three GCMs' (CNRM-CM5, MPI, and IPSL) historical runs, calculated over the CONUS. Failure fraction is calculated using Fisher *z*-transformation and bootstrapping. Fisher *z*-transformation (bootstrapping) is considered for statistically significant grid points that have multivariate skewness coefficient values lesser (greater) than the threshold. Raw GCM cross-correlation (which is same as the bias-corrected GCM cross-correlation under the univariate linear bias corrections) is estimated for the current analysis

z-transformation as compared to the values exhibited by the other two models. Summer months witness lower values of failure fraction under the Fisher *z*-transformation, which signifies that GCMs efficiently estimate the negative cross-correlation between precipitation and temperature. All three models imprecisely estimate the observed cross-correlation over the CONUS, emphasizing that the application of simple linear regression or asynchronous regression could result in biased estimations of the observed cross-correlation. For further analysis in sections 4.2 and 4.3, we consider only one GCM: the CNRM-CM5. Since our current analysis shows that the performances of the three GCMs exhibit similar patterns across months, conclusions from the following subsections can be extended to the other GCMs as well.

4.2 | Performance of CNRM-CM5 in reproducing cross-correlations across climate regions

The CNRM-CM5's performance is evaluated by estimating the failure fraction. Table 2a (Fisher *z*-transformation) and Table 2b (bootstrapping) show the fraction of grid points over the NCDC climate regions whose cross-correlation estimates fall outside the permissible bound for the period 1950–1999. The number of grid points considered for the analysis is presented in parenthesis.

Under the Fisher *z*-transformation, the WNC and S have a higher number of grid points as compared to the other regions (Table 2a). However, the CNRM-CM5's performances on WNC and S do not remain uniform across months. For example, it exhibits a failure fraction of 0.14 over the WNC during January, efficiently estimating the positive observed cross-correlation (Figure 1). However, during July, the CNRM-CM5 performs poorly over the WNC, with a failure fraction value of 0.58. For a particular month, the CNRM-CM5 exhibits varying failure fraction values across climate regions depending on the GCM's ability to capture the linear dependency between precipitation and temperature. For example, during April, the CNRM-CM5 historical runs have failure fraction values of 0.38 and 0.48 over the NW and SW, respectively. Results from Table 2a confirm the CNRM-CM5's varying ability under the Fisher *z*-transformation to capture the interdependency of precipitation and temperature over different months and across the nine climate regions. However, the failure fraction values related to bootstrapping remain uniform across months and across climate regions (Table 2b). Although the grid points are not common between Fisher *z* and bootstrapping, the latter rejects more grid points in comparison to the Fisher *z*-transformation, with failure fraction values of around or more than 0.9. Since the sampling variability of the null distribution from bootstrapping is smaller than the Fisher *z*-transformation, the raw GCM cross-correlation typically rests outside the 95% upper and lower confidence limits. Overall, we conclude that the failure fraction varies across months and

TABLE 2 Failure fraction from (a) Fisher *z*-transformation and (b) bootstrapping for raw GCM cross-correlation across NCDC climate regions (CNRM, historical simulations, 1950–1999)

(a) Month	NW	WNC	ENC	C	NE	SE	S	SW	W
Jan	0.05 (2)	0.14 (43)	0 (4)	0.29 (34)	0 (4)	0.2 (1)	0.17 (24)	0.05 (15)	0 (1)
Apr	0.38 (3)	0.26 (22)	0 (0)	0 (2)	0.47 (1)	0.22 (4)	0 (17)	0.48 (47)	0.7 (47)
Jul	0.1 (44)	0.58 (84)	0.25 (6)	0.1 (29)	0.28 (1)	0.13 (14)	0.22 (89)	0.46 (22)	0 (1)
Oct	0.02 (43)	0.02 (36)	0 (0)	0 (0)	0 (0)	0.2 (3)	0.1 (8)	0.38 (52)	0 (27)
(b) Month	**NW**	**WNC**	**ENC**	**C**	**NE**	**SE**	**S**	**SW**	**W**
Jan	0.8 (2)	0.88 (43)	0.95 (4)	0.91 (34)	0.75 (4)	0.6 (1)	0.96 (24)	0.87 (15)	1 (1)
Apr	1 (3)	0.98 (22)	0 (0)	1 (2)	1 (1)	0.95 (4)	0.95 (17)	0.95 (47)	0.94 (47)
Jul	0.89 (44)	0.94 (84)	1 (6)	0.96 (29)	1 (1)	0.96 (14)	0.95 (89)	0.99 (22)	1 (1)
Oct	0.88 (43)	0.92 (36)	0 (0)	0 (0)	0 (0)	0.93 (3)	0.93 (8)	0.94 (52)	0.9 (27)

Note: The number of grid points considered for the analysis is shown within brackets.

across climate regions for grid points that have multivariate skewness coefficient values lesser than the threshold. However, for grid points that have multivariate skewness coefficient values greater than the threshold, the CNRM-CM5's performance is poor and the performance remains uniform across months as well as climate regions.

4.3 | Performance across different cross-correlation ranges

We expect that univariate bias correction based on regression techniques' inability to yield observed cross-correlations to change with a change in the magnitude of the observed cross-correlation values. The shape of a permissible bound,

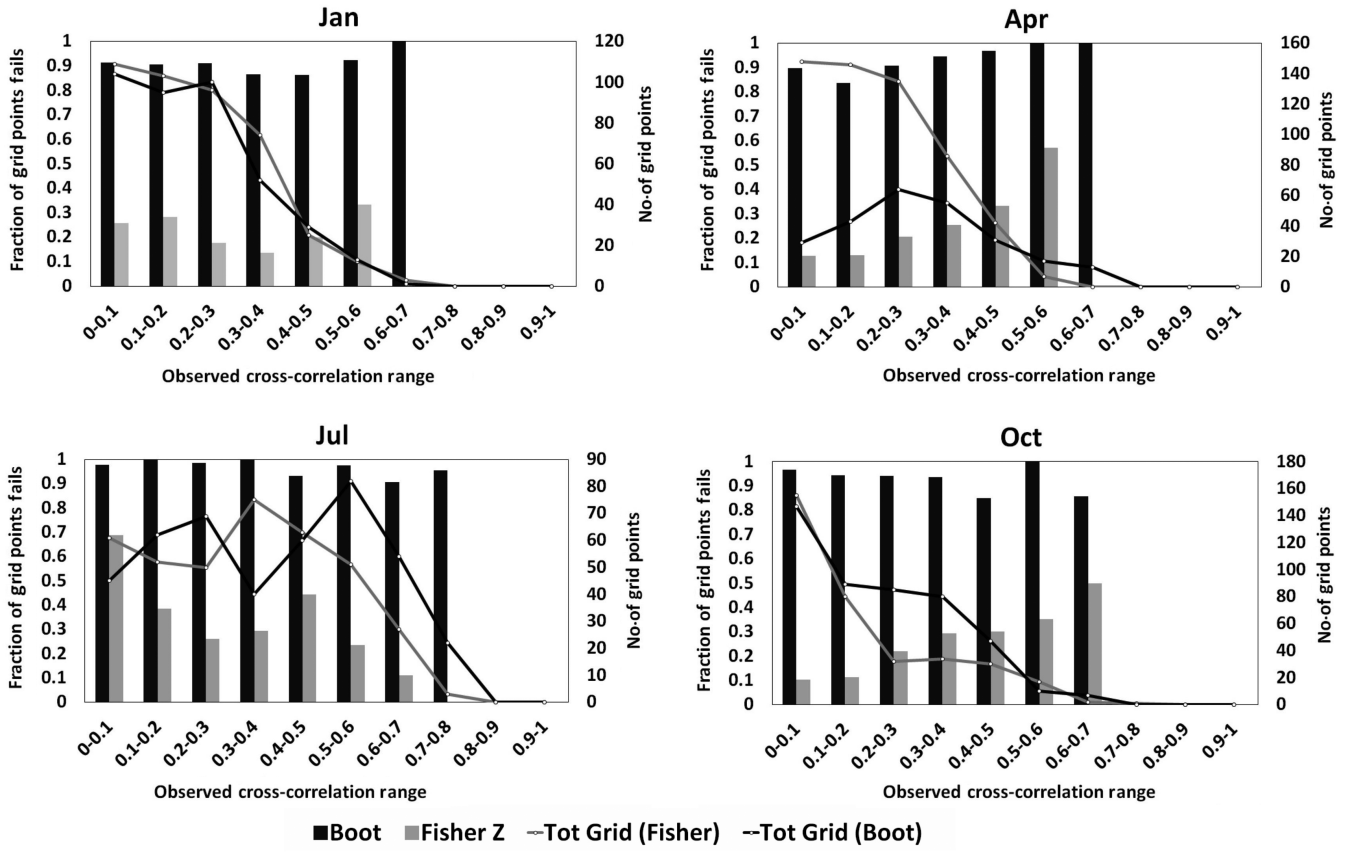

FIGURE 5 Fraction of grid points yield raw GCM cross-correlations that differ significantly from observations for different intervals of absolute observed cross-correlation. Number of grid point within an interval plotted on the secondary axis. Fisher *z*-transformation (bootstrapping) is considered grid points that have multivariate skewness coefficient values lesser (greater) than the threshold. Single ensemble member from CNRM-CM5 is considered for the study. Vertical lines indicate that observed cross-correlation values to the right of the line are statistically significant

especially for the Fisher z-transformation, suggests that the failure is expected to be higher when the observed cross-correlation values are closer to ± 1. Note that for the current analysis, we used raw GCM outputs without any bias correction. In Figure 5, we show the CNRM-CM5's performance in estimating the absolute values of the observed cross-correlation across different intervals. The total number of grid points within a range is shown on the secondary axis. The vertical lines at the observed cross-correlation range of 0.3–0.4 indicate that the observed cross-correlation values towards their right are statistically significant. Similar to sections 4.1 and 4.2, for a particular grid point, 95% confidence limits are constructed by applying the Fisher z-transformation (bootstrapping) when the multivariate skewness coefficient at the grid point is lower (higher) than the threshold value.

Under the Fisher z-transformation, the CNRM-CM5 fails to yield the observed cross-correlation on 25% of the grid points (approx.) if the absolute observed cross-correlations are within the 0.3–0.4 range. During July, the CNRM-CM5's performance under the Fisher z-transformation improves with the increase in the absolute values of the observed cross-correlation. However, during April and October, the CNRM-CM5's performance under the Fisher z-transformation decreases with the increase in the absolute values of observed cross-correlation, following our hypothesis. On the other hand, the confidence limits from the bootstrapping do not change substantially with the change in the absolute value of the observed cross-correlation. Therefore, the CNRM-CM's performance under the bootstrapping remains almost constant across different ranges of observed cross-correlations. The performance under bootstrapping algorithm exhibits similar trends as the performance under Fisher z-transformation, but almost 90% of the grid points fail to yield the observed cross-correlation under bootstrapping. These findings indicate that if statistically significant cross-correlation (e.g., negative correlation between precipitation and temperature over the Midwest) exhibits over a region, then it would be prudent to consider a multivariate bias correction scheme or a complex univariate bias correction scheme to preserve the cross-correlation.

5 | DISCUSSION AND CONCLUDING REMARKS

This study demonstrated the limitations of univariate linear bias correction (in particular, simple linear regression and asynchronous regression) in reproducing observed cross-correlations between precipitation and temperature in a bias-corrected data set. We found that two univariate linear approaches simply reproduce the GCM-estimated cross-correlation (Raw cross-correlation) in the bias-corrected data set, since the transfer function between the raw and the bias-corrected GCM data sets is linear across the grid points. We validated this finding with a popular univariate bias correction product from the BOR archive, which followed quantile mapping to correct the model bias. We calculated 95% lower and upper confidence limits (permissible bounds) in the observed cross-correlation based on the Fisher z-transformation and bootstrapping. Permissible bounds investigate whether the GCM-estimated (raw or bias-corrected) cross-correlation is statistically different from the observed cross-correlation. Since the permissible bounds are independent of GCMs and bias correction procedures, the procedure to calculate the bounds can be extended to evaluate any bias-corrected GCM cross-correlation structure. Failure of a grid point is largely dependent on GCMs; therefore, permissible bounds help to compare between multiple GCMs. It is expected since the performance of bias correction procedure depends on the raw GCM's values. During the analysis, we considered only those grid points where the observed cross-correlations are statistically significant. The significant grid points were further divided into two groups (a) significant grid points whose multivariate skewness coefficient values are lesser than its threshold value, and (b) the remaining significant grid points where the multivariate skewness coefficient are substantial, hence higher than the threshold value. A Fisher z-transformation assumes that dependent variables follow a bivariate Gaussian distribution. Therefore, a Fisher z-transformation is applied to the first group of grid points while bootstrapping is considered for the second group of grid points to construct the permissible bounds. Based on the permissible bounds, we identified the grid points and climate regions whose raw GCM cross-correlations are significantly different from the observed ones. The current study considered only the raw GCM cross-correlation since it remains unaltered by univariate linear bias correction approaches.

Results show that a high amount of bias exists between the observed and raw GCM cross-correlations from 20th century simulations. GCM historical runs are unable to reproduce the observed cross-correlation on 30–50% (over 90%) of the first (second) group of grid points over the CONUS, indicating that the univariate linear bias correction is not suitable for these grid points. Although Fisher z and bootstrapping were not applied for a common set of grid points, we infer that the framework related to the Fisher z-transformation is a conservative one as compared to the bootstrapping framework. Since all three GCMs exhibit similar failure fraction values, we decided to consider only one representative GCM (CNRM-CM5) for further analysis. Under the Fisher z-transformation, the performance of the CNRM-CM5 in yielding the observed cross-correlation varies across climate regions and over months. For example, the WNC

exhibits a failure fraction value of 0.14 during January, whereas almost half of the grid points over the WNC fail to yield the observed cross-correlation during July. However, under bootstrapping, the performance of the CNRM-CM5 in reproducing the observed cross-correlation remains almost same across climate regions and over different months, with a failure fraction value higher than 0.8. Following this, we examined the CNRM-CM5's ability to reproduce the observed cross-correlation against varying magnitudes of absolute observed cross-correlations. Typically, under the Fisher *z*-transformation, the number of grid points fails to reproduce the observed cross-correlation, which increases with an increase in the absolute magnitude of the observed cross-correlation. We found that up to 25% of the grid points, where the absolute observed cross-correlation values are around 0.3–0.4, fail to reproduce the observed cross-correlation. Since the sampling variability of cross-correlation under bootstrapping is narrower as compared to the Fisher *z*-transformation, the framework related to the latter provides a conservative evaluation of univariate bias correction.

Since our study reported that univariate linear bias correction has a limited role to play in reproducing observed cross-correlations, the focus is now shifted towards

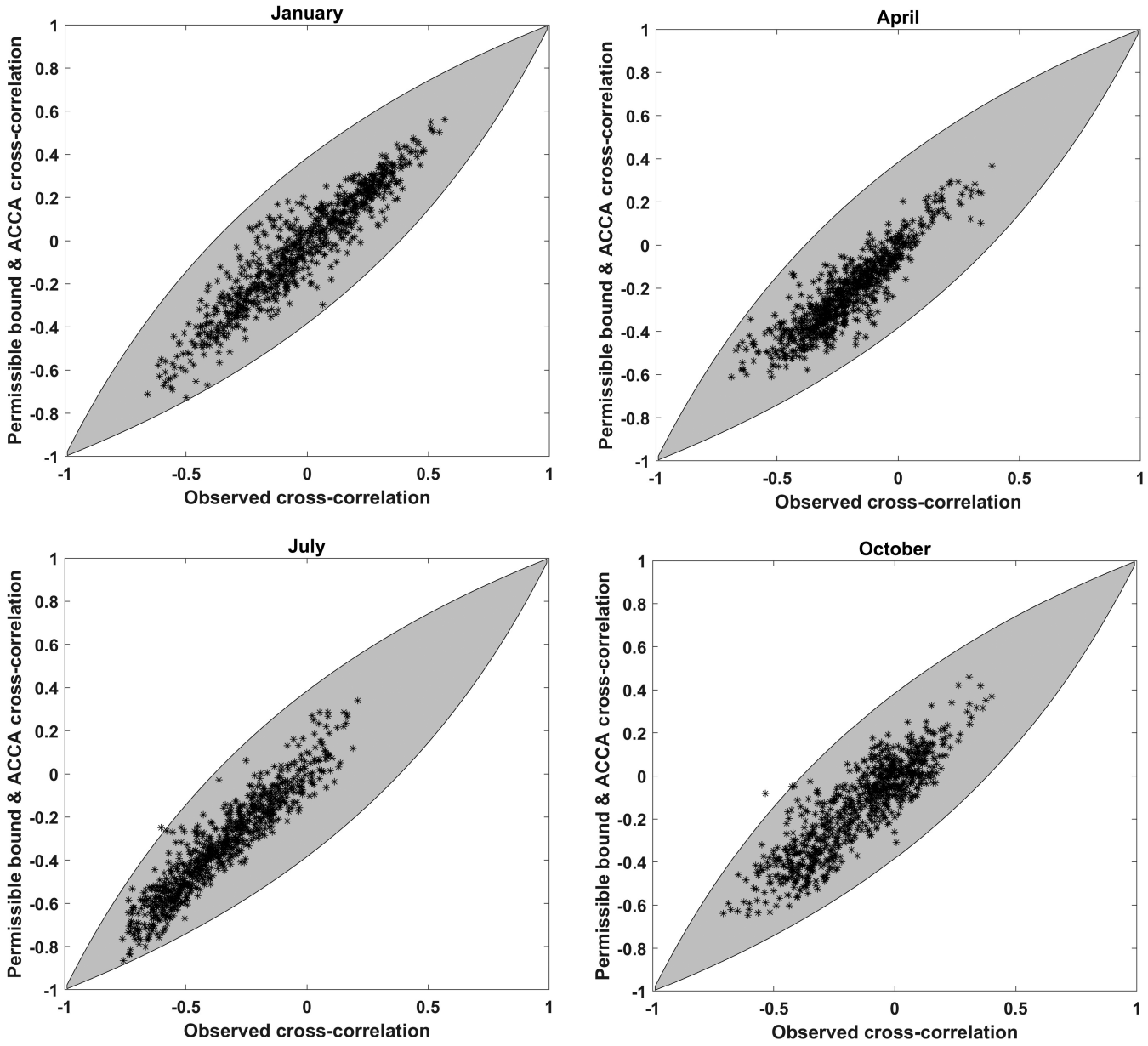

FIGURE 6 Bias-corrected GCM cross-correlations, using a multivariate bias-correcting technique (ACCA), from CNRM-CM5 historical runs overlaid on permissible bounds. Each dot represents a grid point on the CONUS, *X* value indicates the observed cross-correlation at that grid point, and the *Y* value indicates the bias-corrected GCM cross-correlation at that grid point. Permissible bounds are calculated using Fisher *z*-transformation. Only one ensemble member is considered for plotting

multivariate and nonlinear univarite bias correction techniques. We briefly investigate the potential of a multivariate bias correction algorithm (Das Bhowmik *et al*., 2017), an ACCA, to reproduce the observed cross-correlation. We apply the permissible bounds from the Fisher *z*-transformation on the CNRM-CM5 cross-correlations estimated by the ACCA for the selected 4 months (Figure 6). The multivariate bias-corrected data set is obtained by first applying a bivariate sorting, followed by developing a canonical correlation regression. Multivariate bias-corrected GCM cross-correlations are within the sampling variability of the observed cross-correlations. Multivariate bias corrections have a better ability to yield observed cross-correlations than the univariate linear approach. Hence, our analysis emphasize on applying multivariate bias correction techniques, which can also be trusted to estimate the joint likelihood of the relevant variables better than the univariate bias correction approach. Besides ACCA, there are also other multivariate bias correction approaches that can reduce bias in multiple variables. A recently developed technique, MACA (Hidalgo *et al*., 2008; Abatzoglou and Brown, 2012), corrects the bias in GCM simulations by identifying the patterns between the GCM and observed fields using constructed analogs. He *et al*. (2012) proposed a bivariate approach by extending the concept of quantile mapping to bivariate asynchronous measurements. Bias correction and statistical downscaling procedures are modified to yield a joint dependence among multiple variables (Zhang and Georgakakos, 2012; Mehrotra and Sharma, 2016). In recent years, efforts were made to develop multivariate bias correction techniques focusing particularly on the intervariable dependence between climate variables (He *et al*., 2012; Piani and Haerter, 2012; Li *et al*., 2014; Cannon, 2016; Das Bhowmik *et al*., 2017). Li *et al*. (2014) showed that their proposed joint bias correction (JBC) methodology corrects the bias in the cross-correlation along with reducing the biases in mean and in the variance. Piani and Haerter (2012) proposed a 2D bias correction technique that improves precipitation and temperature copula during validation. A recent study (Guo *et al*., 2019) proposed two stage quantile mapping that is capable of reproducing intervariable rank correlations. Nevertheless, these studies did not explore the reason behind univariate linear approach's limited performance to yield the observed cross-correlation. Traditionally, multivariate approaches were considered because either they exhibit an improved performance metric as compared to univariate approaches or they ease the computation when multiple variables are to be downloaded. The permissible bounds we suggested could be potentially applied to any multivariate methods and evaluated for preserving observed cross-correlations.

The main novelty of the current study is that it states that the inability of univariate linear bias correction approach to yield the observed correlation is because of its linear model structure. In addition, the study provides a spatiotemporal investigation of raw GCM outputs' performance to yield the observed cross-correlation, which can be extended further for hydroclimate modelling. However, criticism of the current study might arise primarily from three perspectives: (a) assumptions related to the Fisher *z*-transformation, (b) temporal resolution of the data, and (c) the linear dependency measure. First, for hypothesis testing, the Fisher *z*-transformation assumes that the two variables used to estimate the correlation follow a bivariate normal distribution. However, the precipitation and temperature values on a grid point may not follow a bivariate normal distribution. Hence, a prior examination of multivariate normality, such as Cox-Small test, should improve the precision in estimating the bias-corrected cross-correlation. Apart from the normality assumption, 95% upper and lower confidence limits from the Fisher *z*-transformation are not constant across different climate model runs, since the limits are the function of the respective sample lengths. For instance, the hindcast runs (Taylor *et al*., 2012) are typically 10-/30-year-long runs; therefore, the confidence interval of a hindcast is wider than that of a historical run. Thus, the GCMs' performance in reproducing observed cross-correlations should be carefully examined depending on the sample length of the data.

Second, the current study did not evaluate the impact of univariate bias correction on a daily data set. Daily data tends to be affected by lag-correlation, for which additional consideration is required during bias correction. For example: variations of univariate linear bias correction techniques need to be adopted to account zero precipitation values. The derivation of the bias in a cross-correlation (Equation (3)) is useful even for daily data while the framework (Equations (4a) and (4b)) is significant even for the application of daily data, if it exhibits a temporal independence. Therefore, the limitations of univariate linear bias correction procedures, simple linear regression, and asynchronous regression, discussed in this study is invariant of the timescale of the GCM data.

Finally, our study assumes the Pearson correlation coefficient to calculate the dependence between two variables, which measures the linear dependency. Nonparametric dependency measures such as Spearman or Kendall's Tau could have been considered to capture the monotonic dependence between precipitation and temperature. For instance, we could have applied Spearman's rank correlation for the resampled precipitation and temperature from the bootstrapping algorithm. Since we find that GCM estimated Pearson correlation and GCM estimated Spearman's rank correlations exhibits similar values (see Figure S3), we did

not consider Spearman's rank correlation. Furthermore, Fisher z-transformation may not be applicable on Spearman's rank correlation since underlying assumptions related to Fisher z-transformation is based on linear dependency (Fisher, 1950). Hence, we expressed the dependency between precipitation and temperature using Pearson's correlation. An extension of the current research should include the nonlinear dependency between precipitation and temperature to assess the performance of GCMs and bias correction techniques.

The current study has a major implication on hydroclimate studies, in projecting future changes in hydrologic regimes. A linear dependency between precipitation and temperature has a direct influence on the estimation of evapotranspiration. Earlier studies have reported that around 60% of the precipitation could be lost in the form of evapotranspiration (Oki and Kanae, 2006; Sanford and Selnick, 2013). We infer that an inaccurate estimation of evapotranspiration by empirical approaches, resulting from the bias in GCM estimated cross-correlations, could influence the water-balance model at a watershed scale. Additionally, simulations/projections of various hydrologic fluxes (e.g., soil moisture and overland flow) rely on cross-correlations among climate variables for preserving cross-correlations across land-surface fluxes. Hydrologic models forced with climate variables that have a bias in their cross-correlation will result in a biased simulation of land-surface attributes (Seo *et al*., 2019; Chen *et al*., 2018). Thus, the emphasis laid here on bias-correcting GCM outputs could also potentially apply to bias-correcting land-surface models and semi-distributed watershed model outputs (Libera and Sankarasubramanian, 2018). A recent study (Seo *et al*., 2019) considered two sets of monthly climate forcing to run long-term simulation of hydrologic fluxes. One set of monthly forcing preserve the observed cross-correlation, while the other set ignores it. They found that hydrologic simulation of subsurface variable is improved by preserving the linear dependence. These findings confirm a substantial influence of linear dependence between climate variables on long-term hydrologic simulations. Raw cross-correlations with low-bias have also improved the joint likelihood of relevant attributes (Das Bhowmik *et al*., 2017). Thus, it is important to consider multivariate bias correction procedures for the dissemination of bias-corrected products.

ACKNOWLEDGEMENTS

This research was supported in part by the National Science Foundation grant (CBET-1204368). Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

We acknowledge the World Climate Research Programme's Working Group for their Coupled Modelling, which is responsible for the CMIP, and we thank the climate modelling groups (listed in Table 1) for producing and making their model output available. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provided their coordinating support and led in the development of the software infrastructure in partnership with the Global Organization for Earth System Science Portals. "Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections" is archived at http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/.

We also thank the two anonymous reviewers and the Editor, Dr. Radan Huth, for their comments and detailed reviews, which led to substantial improvements in the quality of the manuscript.

ORCID

R. Das Bhowmik https://orcid.org/0000-0002-2849-3180

REFERENCES

Abatzoglou, J.T. and Brown, T.J. (2012) A comparison of statistical downscaling methods suited for wildfire applications. *International Journal of Climatology*, 32(5), 772–780.

Alexander, L.V., Zhang, X., Peterson, T.C., Caesar, J., Gleason, B., Klein Tank, A.M., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F. and Tagipour, A. (2006) Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research: Atmospheres*, 111(D5), D05109. https://doi.org/10.1029/2005JD006290

Blunden, J. and Arndt, D.S. (Eds.). (2014) State of the climate in 2013. *Bulletin of the American Meteorological Society*, 95(7), S1–S257.

Cannon, A.J. (2016) Multivariate bias correction of climate model output: matching marginal distributions and intervariable dependence structure. *Journal of Climate*, 29(19), 7045–7064.

Chen, J., Li, C., Brissette, F.P., Chen, H., Wang, W. and Essou, G.R.C. (2018) Impacts of correcting the inter-variable correlation of climate model outputs on hydrological modeling. *Journal of Hydrology*, 560, 326–341. https://doi.org/10.1016/J.JHYDROL.2018.03.040.

Das Bhowmik, R., Sankarasubramanian, A., Sinha, T., Mahinthakumar, G. and Kunkel, K.E. (2017) Multivariate downscaling approach preserving cross-correlations across climate variable for projecting hydrologic fluxes. *Journal of Hydrometeorology*, 18, 2187–2205. https://doi.org/10.1175/JHM-D-16-0160.1.

Dettinger, M.D., Cayan, D.R., Meyer, M.K. and Jeton, A.E. (2004) Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American River basins, Sierra Nevada, California, 1900–2099. *Climatic Change*, 62(1–3), 283–317.

Devineni, N. and Sankarasubramanian, A. (2010a) Improved categorical winter precipitation forecasts through multimodel combinations of coupled GCM. *Geophysical Research Letters*, 37, L24704. https://doi.org/10.1029/2010GL044989.

Devineni, N. and Sankarasubramanian, A. (2010b) Improving the prediction of winter precipitation and temperature over the continental United States: role of the ENSO state in developing multimodel combinations. *Monthly Weather Review*, 138(6), 2447–2468.

Dufresne, J.L., Foujols, M.A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R. and Bony, S. (2013) Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Climate Dynamics*, 40(9–10), 2123–2165.

Fisher, R.A. (1925) *Statistical Methods for Research Workers*. Edinburgh: Oliver & Boyd. Available at: http://psychclassics.yorku.ca/Fisher/Methods/ [Accessed 1st January 2018].

Fisher, R.A. (1950) *Statistical Methods for Research Workers*, 11th edition. London: Oliver & Boyd.

Gangopadhyay, S., Clark, M., Werner, K., Brandon, D. and Rajagopalan, B. (2004) Effects of spatial and temporal aggregation on the accuracy of statistically downscaled precipitation estimates in the Upper Colorado River basin. *Journal of Hydrometeorology*, 5(6), 1192–1206.

Giorgetta, M.A., Jungclaus, J., Reick, C.H., Legutke, S., Bader, J., Böttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K. and Glushak, K. (2013) Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *Journal of Advances in Modeling Earth Systems*, 5(3), 572–597.

Guo, Q., Chen, J., Zhang, X., Shen, M., Chen, H. and Guo, S. (2019) A new two-stage multivariate quantile mapping method for bias correcting climate model outputs. *Climate Dynamics*, 1–21.

Hamill, T.M. (1999) Hypothesis tests for evaluating numerical precipitation forecasts. *Weather and Forecasting*, 14(2), 155–167.

Hanson, R.T. and Dettinger, M.D. (2005) Ground water/surface water responses to global climate simulations, Santa Clara-Calleguas basin, Ventura, California. *JAWRA Journal of the American Water Resources Association*, 41(3), 517–536.

Hay, L.E. and Clark, M.P. (2003) Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western United States. *Journal of Hydrology*, 282(1–4), 56–75.

He, X., Yang, Y. and Zhang, J. (2012) Bivariate downscaling with asynchronous measurements. *Journal of Agricultural, Biological, and Environmental Statistics*, 17(3), 476–489.

Hidalgo, H.G., Dettinger, M.D. and Cayan, D.R. (2008) *Downscaling with constructed analogues: daily precipitation and temperature fields over the United States*. California Energy Commission PIER Final Project Report CEC-500-2007-123.

Huth, R. (1999) Statistical downscaling in central Europe: evaluation of methods and potential predictors. *Climate Research*, 13(2), 91–101.

IPCC. (2013) *Climate Change 2013: The Physicalscience Basis. Working Group I Contribution to the IPCC Fifth Assessment Report*. Cambridge and New York, NW: Cambridge University Press. Available at: www.ipcc.ch/report/ar5/wg1 [Accessed 1st January 2018].

Ivanov, M.A. and Kotlarski, S. (2017) Assessing distribution-based climate model bias correction methods over an alpine domain: added value and limitations. *International Journal of Climatology*, 37(5), 2633–2653. https://doi.org/10.1002/joc.4870.

Ivanov, M.A., Warrach-Sagi, K. and Wulfmeyer, V. (2018a) Field significance of performance measures in the context of regional climate model evaluation. Part 1: temperature. *Theoretical and Applied Climatology*, 132(1), 219–237. https://doi.org/10.1007/s00704-017-2100-2.

Ivanov, M.A., Warrach-Sagi, K. and Wulfmeyer, V. (2018b) Field significance of performance measures in the context of regional climate model evaluation. Part 2: precipitation. *Theoretical and Applied Climatology*, 132(1), 239–261. https://doi.org/10.1007/s00704-017-2077-x.

Johnson, F. and Sharma, A. (2015) What are the impacts of bias correction on future drought projections? *Journal of Hydrology*, 525, 472–485.

Karl, T.R. and Koss, W.J. (1984) *Regional and National Monthly, Seasonal and Annual Temperature Weighted by Area, 1895–1983*. *Historical Climatology Series 4-3*. Asheville, NC: National Climatic Data Center, 38 pp.

Lall, U. and Sharma, A. (1996) A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resources Research*, 32 (3), 679–693.

Li, C., Sinha, E., Horton, D.E., Diffenbaugh, N.S. and Michalak, A.M. (2014) Joint bias correction of temperature and precipitation in climate model simulations. *Journal of Geophysical Research: Atmospheres*, 119(23), 13–153.

Libera, D. and Sankarasubramanian, A. (2018) Multivariate bias correction of streamflow and loading predictions from mechanistic models. *Journal of Hydrology*, 564, 529–541.

Mardia, K.V. (1970) Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3), 519–530. https://doi.org/10.1093/biomet/57.3.519.

Maurer, E.P., Brekke, L., Pruitt, T. and Duffy, P.B. (2007) Fine-resolution climate projections enhance regional climate change impact studies. *Eos, Transactions American Geographical Union*, 88(47), 504.

Maurer, E.P. and Hidalgo, H.G. (2008) Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrology and Earth System Sciences*, 12, 551–563.

Mehrotra, R. and Sharma, A. (2016) A multivariate quantile-matching bias correction approach with auto-and cross-dependence across multiple time scales: implications for downscaling. *Journal of Climate*, 29(10), 3519–3539.

Mejia, J.F., Huntington, J., Hatchett, B., Koracin, D. and Niswonger, R.G. (2012) Linking global climate models to an integrated hydrologic model: using an individual station downscaling approach. *Journal of Contemporary Water Research & Education*, 147(1), 17–27.

Mpelasoka, F.S. and Chiew, F.H. (2009) Influence of rainfall scenario construction methods on runoff projections. *Journal of Hydrometeorology*, 10(5), 1168–1183.

Oki, T. and Kanae, S. (2006) Global hydrological cycles and world water resources. *Science*, 313(5790), 1068–1072.

Piani, C. and Haerter, J.O. (2012) Two dimensional bias correction of temperature and precipitation copulas in climate models. *Geophysical Research Letters*, 39(20), L20401.

Plummer, N., Salinger, M., James, N.N., Ramasamy, S., Hennessy, K. J., Leighton, R.M., Blair, T., Page, C.M. and Lough, J.M. (1999) Changes in climate extremes over the Australian region and New Zealand during the twentieth century. *Climate Change*, 42(1), 183–202.

Qin, N., Chen, X., Fu, G., Zhai, J. and Xue, X. (2010) Precipitation and temperature trends for the Southwest China: 1960–2007.

Geoenvironmental Disaster, 24, 3733–3744. https://doi.org/10.1002/hyp.7792.

Reclamation. (2013) *Downscaled CMIP3 and CMIP5 Climate Projections: Release of Downscaled CMIP5 Climate Projections, Comparison with Preceding Information, and Summary of User Needs*. Denver, CO: U.S. Department of the Interior, Bureau of Reclamation, Technical Service Center, 116 p. Available at: http://gdodcp.ucllnl.org/downscaled_cmip_projections/techmemo/downscaled_climate.pdf [Accessed 1st January 2018].

Sanford, W.E. and Selnick, D.L. (2013) Estimation of evapotranspiration across the conterminous United States using a regression with climate and land-cover data 1. *Journal of the American Water Resources Association*, 49(1), 217–230.

Sankarasubramanian, A. and Lall, U. (2003) Flood quantiles and changing climate: seasonal forecasts and causal relations. *Water Resources Research*, 39(5), 1134.

Seo, S.B., Das Bhowmik, R., Sankarasubramanian, A., Mahinthakumar, G. and Kumar, M. (2019) The role of cross-correlation between precipitation and temperature in basin-scale simulations of hydrologic variables. *Journal of Hydrology*, 570, 304–314. https://doi.org/10.1016/j.jhydrol.2018.12.076.

Seo, S.B., Sinha, T., Mahinthakumar, G., Sankarasubramanian, A. and Kumar, M. (2016) Identification of dominant source of errors in developing streamflow and groundwater projections under near-term climate change. *Journal of Geophysical Research: Atmospheres*, 121, 7652–7672. https://doi.org/10.1002/2016JD025138.

Singh, H., Sinha, T. and Sankarasubramanian, A. (2014) Impacts of near-term climate change and population growth on within-year reservoir systems. *Journal of Water Resources Planning and Management*, 141(6), 04014078.

Stoner, A.M.K., Hayhoe, K., Yang, X. and Wuebbles, D.J. (2012) An asynchronous regional regression model for statistical downscaling of daily climate variables. *International Journal of Climatology*, 33, 2473–2494.

Tao, B., Tian, H., Ren, W., Yang, J., Yang, Q., He, R., Cai, W. and Lohrenz, S. (2014) Increasing Mississippi river discharge throughout the 21st century influenced by changes in climate, land use, and atmospheric CO_2. *Geophysical Research Letters*, 41(14), 4978–4986.

Taylor, K.E., Stouffer, R.J. and Meehl, G.A. (2012) An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93, 485–498.

Trenberth, K.E. and Shea, D.J. (2005) Relationships between precipitation and surface temperature. *Geophysical Research Letters*, 32 (14), L14703.

Voldoire, A., Sanchez-Gomez, E., y Mélia, D. S., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A. and Chevallier, M. Déqué, M. (2013). The CNRM-CM5.1 global climate model: description and basic evaluation. *Climate Dynamics*, 40(9–10), 2091-2121.

Wilcke, R.A.I., Mendlik, T. and Gobiet, A. (2013) Multi-variable error correction of regional climate models. *Climate Change*, 120(4), 871–887.

Wilks, D.S. (2006) On "field significance" and the false discovery rate. *Journal of Applied Meteorology*, 45(9), 1181–1189. https://doi.org/10.1175/JAM2404.1.

Wood, A.W., Leung, L.R., Sridhar, V. and Lettenmaier, D.P. (2004) Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, 62, 189–216.

Zhang, F. and Georgakakos, A.P. (2012) Joint variable spatial downscaling. *Climatic Change*, 111(3), 945–972.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Bhowmik RD, Sankarasubramanian A. Limitations of univariate linear bias correction in yielding cross-correlation between monthly precipitation and temperature. *Int J Climatol*. 2019;39:4479–4496. https://doi.org/10.1002/joc.6086

APPENDIX A: ANALYTICAL ANALYSIS OF UNIVARIATE LINEAR APPROACH

Linear model is fitted between [P^{GCM}, T^{GCM}] and [P^{obs}, T^{obs}]. Model parameters are a, b, c, and d. ξ^p and ξ^t are the model errors of precipitation and temperature, respectively,

$$P^{obs} = aP^{GCM} + b + \xi^p, \tag{A1}$$

$$T^{obs} = cT^{GCM} + d + \xi^t. \tag{A2}$$

Regression slopes can be expressed as using ordinary least squares

$$\hat{a} = \frac{\mathrm{cov}\left(P^{GCM}, P^{obs}\right)}{\mathrm{var}\left(P^{GCM}\right)}, \hat{c} = \frac{\mathrm{cov}\left(T^{GCM}, T^{obs}\right)}{\mathrm{var}\left(T^{GCM}\right)},$$

where cov(.,.) denotes the temporal covariance between the specified variables and var(.) denotes the temporal variance of the variable within the parenthesis. Bias-corrected precipitation ($\hat{P}^{BC}$) and temperature ($\hat{T}^{BC}$) variables are estimated using model parameters,

$$\hat{P}^{BC} = \hat{a}P^{GCM} + \hat{b}, \tag{A3}$$

$$\hat{T}^{BC} = \hat{c}T^{GCM} + \hat{d}. \tag{A4}$$

Hence, the covariance between the bias-corrected variables can be written as

$$\begin{aligned}\mathrm{cov}\left(\hat{P}^{BC}, \hat{T}^{BC}\right) &= \mathrm{cov}\left(\hat{a}P^{GCM} + \hat{b}, \hat{c}T^{GCM} + \hat{d}\right) \\ &= \hat{a}.\hat{c}.\mathrm{cov}\left(P^{GCM}, T^{GCM}\right).\end{aligned}$$

Writing the cross-correlation between bias-corrected variables as

$$\mathrm{corr}\left(\hat{P}^{BC}, \hat{T}^{BC}\right) = \frac{\mathrm{cov}\left(\hat{P}^{BC}, \hat{T}^{BC}\right)}{\sigma_{\hat{P}}^{BC}\sigma_{\hat{T}}^{BC}} = \frac{\hat{a}.\hat{c}.\mathrm{cov}\left(P^{GCM}, T^{GCM}\right)}{\sigma_{\hat{P}}^{BC}\sigma_{\hat{T}}^{BC}}. \tag{A5}$$

Since, $\mathrm{var}(\hat{P}^{BC}) = \hat{a}^2\mathrm{var}\left(P^{GCM}\right)$ and $\mathrm{var}(\hat{T}^{BC}) = \hat{c}^2\mathrm{var}\left(T^{GCM}\right)$ from the model formulation, we substitute them by taking the square roots in the denominator for the standard deviations of the bias-corrected variables,

$$\mathrm{corr}\left(\hat{P}^{BC}, \hat{T}^{BC}\right) = \frac{\hat{a}.\hat{c}.\mathrm{cov}\left(P^{GCM}, T^{GCM}\right)}{\hat{a}.\hat{c}.\sigma_P^{GCM}\sigma_T^{GCM}} = \mathrm{corr}\left(P^{GCM}, T^{GCM}\right). \tag{A6}$$

This results in Equation (A6) show analytically that the bias-corrected cross-correlations will always be equal to the GCM estimated cross-correlation at a grid point.

APPENDIX B: FISHER z-TRANSFORMATION

ρ^{obs} and ρ^{BC} are the observed and bias-corrected cross-correlations.

Null hypothesis: $\rho^{obs} = \rho^{BC}$.

Z-transformation:

$$Z^{obs} = 0.5.\ln\left[\frac{1+\rho^{obs}}{1-\rho^{obs}}\right], \tag{B1}$$

$$Z^{BC} = 0.5.\ln\left[\frac{1+\rho^{BC}}{1-\rho^{BC}}\right]. \tag{B2}$$

$$\text{Test statistic}: \tilde{z} = \frac{Z^{obs} - Z^{BC}}{\sigma_{Z^{obs}-Z^{BC}}}. \tag{B3}$$

Test statistics $\tilde{z}$ follows standard normal distribution, $\tilde{z} \sim N(0,1)$ and $\sigma_{Z^{obs}-Z^{BC}} = \sqrt{\frac{2}{n-3}}$. "$n$" is the number of observations. Bounds for test statistic $\tilde{z}$ are [−1.96, 1.96] and [−2.58, 2.58] to retain the null hypothesis with 95 or 99% CI, respectively. Let us assume, the allowable range for bias-corrected cross-correlation is $[\rho_{high}^{BC}, \rho_{low}^{BC}]$. Considering the lower confidence interval for $\tilde{z}$, we back-calculated the highest value of allowable bias-corrected cross-correlation,

$$\tilde{z}_{low} = -\mathrm{CI\,or}, \frac{1}{2.\sigma_{Z^{obs}-Z^{BC}}}\ln\left[\frac{\left(1+\rho^{obs}\right)\left(1-\rho_{high}^{BC}\right)}{\left(1+\rho_{hign}^{BC}\right)\left(1-\rho^{obs}\right)}\right]$$

$$= -\mathrm{CI\,or}, \frac{\left(1+\rho^{obs}\right)\left(1-\rho_{high}^{BC}\right)}{\left(1+\rho_{high}^{BC}\right)\left(1-\rho^{obs}\right)} = \exp\left(-\mathrm{CI}\times 2\times\sigma_{Z^{obs}-Z^{BC}}\right)$$

$$= C\left[\text{where } C = \exp\left(-\mathrm{CI}\times\frac{2\times\sqrt{2}}{\sqrt{n-3}}\right)\right]\mathrm{or}, \frac{\left(1-\rho_{high}^{BC}\right)}{\left(1+\rho_{high}^{BC}\right)}$$

$$= C\times\frac{\left(1-\rho^{obs}\right)}{\left(1+\rho^{obs}\right)}\,\mathrm{or}, \frac{1}{2} + \frac{\rho_{high}^{BC}}{2} = \frac{1+\rho^{obs}}{(1+C)+\rho^{obs}(1-C)}\,\mathrm{or},$$

$$\rho_{high}^{BC} = \frac{1-C+\rho^{obs}(1+C)}{1+C+\rho^{obs}(1-C)}.$$

Similarly, back-calculating for positive bound of CI gives us ρ_{low}^{BC}, which is

$$\rho_{\text{low}}^{\text{BC}} = \frac{C-1+\rho^{\text{obs}}(C+1)}{1+C+\rho^{\text{obs}}(C-1)}.$$

APPENDIX C: BOOTSTRAPPING ALGORITHM

1. Resample monthly observed raw precipitation and temperature of 50 years to generate synthetic time series of precipitation and temperature. Generate a random number from a uniform distribution between 1 and 50; whatever is the number that we obtain, select X and Y from observed precipitation and temperature data set for that year for the given grid point "i" at time step "t."
2. Repeat step 1 for 50 times (corresponds to 50 years of data) to obtain $[\boldsymbol{X^i},\boldsymbol{Y^i}]$. Calculate the cross-correlation ρ_1^{XY}.
3. Repeat steps 1 and 2 for 1,000 times to obtain the sampling distribution of $\boldsymbol{\rho}$,

$$\boldsymbol{\rho} = \left[\rho_1^{XY}, \rho_2^{XY} \ldots \rho_{1,000}^{XY}\right].$$

Sampling distribution of $\boldsymbol{\rho}$ must be centred on ρ_{obs}^i.

4. Assume the null hypothesis,

$$\text{Null hypothesis}: \rho_{\text{obs}}^i = \rho_{\text{GCM}}^i,$$

$$\text{Alternate hypothesis}: \rho_{\text{obs}}^i \neq \rho_{\text{GCM}}^i.$$

Consider the sampling distribution of ρ to construct 95% confidence limits of ρ_{GCM}^i which will retain the null hypothesis,

$$\rho_{\text{BC/GCM}}^{\text{high/low}} = \bar{\boldsymbol{\rho}} \pm 1.96 \times \sigma_{\boldsymbol{\rho}},$$

where $\bar{\boldsymbol{\rho}}$ is the mean of the sampling distribution and $\sigma_{\boldsymbol{\rho}}$ is the sampling variability.