Bias correction capabilities of quantile mapping methods for rainfall and temperature variables
Maedeh Enayati, Omid Bozorg-Haddad, Javad Bazrafshan,
Somayeh Hejabi and Xuefeng Chu

ABSTRACT
This study aims to conduct a thorough investigation to compare the abilities of quantile mapping (QM) techniques as a bias correction method for the raw outputs from general circulation model (GCM)/regional climate model (RCM) combinations. The Karkheh River basin in Iran was selected as a case study, due to its diverse topographic features, to test the performances of the bias correction methods under different conditions. The outputs of two GCM/RCM combinations (ICHEC and NOAA-ESM) were acquired from the coordinated regional climate downscaling experiment (CORDEX) dataset for this study. The results indicated that the performances of the QMs varied, depending on the transformation functions, parameter sets, and topographic conditions. In some cases, the QMs’ adjustments even made the GCM/RCM combinations’ raw outputs worse. The result of this study suggested that apart from DIST, PTF:scale, and SSPLIN, the rest of the considered QM methods can provide relatively improved results for both rainfall and temperature variables. It should be noted that, according to the results obtained from the diverse topographic conditions of the sub-basins, the empirical quantiles (QUANT) and robust empirical quantiles (RQUANT) methods proved to be excellent options to correct the bias of rainfall data, while all bias correction methods, with the notable exceptions of performed PTF:scale and SSPLIN, performed relatively well for the temperature variable.

Key words | bias correction, climate change, COREDEX, quantile mapping, RCM

HIGHLIGHTS
- Quantile mapping (QM) techniques are among the most important and popular bias correction methods. This study aims to provide a comprehensive comparison to identify the potential strengths and weaknesses of these methods in coping with hydro-climatic variables.
- This study aims to shed light on the general abilities of QM methods in correcting the bias of both temperature and rainfall variables, both of which can have direct and indirect impacts on water resources. Furthermore, this study explores these capabilities in diverse topographic conditions. Coping with such topographic conditions is a known challenge for both general and regional circulation models.
- This study explores the projections of the coordinated regional climate downscaling experiment (CORDEX) dataset. The main idea behind the CORDEX project is to coordinate the results of local climatic studies. As such, the findings of this study can help expand the capabilities and applicability of this project.
The results of this study revealed that some QM methods could, in fact, worsen the accuracy of general and regional circulation models, which highlights the importance of selecting a suitable bias correction method.

According to the results obtained from the diverse topographic conditions of the sub-basins, the empirical quantiles (QUANT) and robust empirical quantiles (RQUANT) methods proved to be excellent options to correct the bias of rainfall data, while all bias correction methods, with the notable exceptions of performed PTF-scale and SSPLIN, performed relatively well for the temperature variable.

GRAPHICAL ABSTRACT

INTRODUCTION

Simulations of the present and future climate conditions often require the use of general circulation models (GCMs). Thus, employing GCMs as tools for long-term climate modeling has become common practice in most climate change studies (Zolghadr-Asli et al. 2018a). However, using these coupled atmosphere-ocean general circulation models is computationally expensive, and resultantly, it could potentially increase the computational cost and, in turn, limit the horizontal resolution of the models’ projections (Plummer et al. 2006; Romera et al. 2017). Needless to say, high-resolution predictions are of great importance for climate change impact assessment studies. To address this issue, downscaling techniques are employed by researchers to post-process and recalibrate the raw projections of GCMs. Based on the conceptual and mathematical frameworks used in these methods, there are two broad downscaling techniques: statistical downscaling and dynamic downscaling (Faiz et al. 2018). The statistical downscaling approaches attempt to link the large-scale atmospheric predictor variables to local- or regional-scale meteorological time series (Dibike & Coulibaly 2005). For more information about these approaches, readers can refer to Busuioc et al. (2001); Wilby et al. (2002); Hertig & Jacobite (2008); and Chen et al. (2012).

On the other hand, dynamic downscaling approaches use finer gridded models, technically referred to as regional climate models (RCMs), to redefine the climate change predictions at a regional scale from the coarse projection of the GCMs (e.g. Jha et al. 2004; Tadross et al. 2005; Kay et al. 2006). As RCMs operate at a relatively finer horizontal resolution, they can provide localized and high-resolution detailed hydro-climatic information that can be of great importance, especially for the regions with complex topographies. As a result, utilizing RCMs is a common practice to project hydro-climatic variables such as rainfall which is particularly sensitive to the horizontal resolution, given its strong correlation with surface topography and physiography (Samuelsson et al. 2011).

Unlike GCMs, where the ensemble projected results would be coordinated to highlight the underlying uncertainties of the outcomes (e.g. the atmospheric model intercomparison project (AMIP) or the coupled model intercomparison projects (CMIP)), RCMs projections may suffer from the lack of a vastly accepted coordinating framework.
Resultantly, most RCM-oriented studies have been isolated and tied to a specifically targeted objective, and thus a comprehensive picture of regional climate change projections based on dynamic downscaling methods is not as readily available as it should be (Giorgi et al. 2009). To overcome this significant issue, the coordinated regional climate downscaling experiment (CORDEX) has been initiated by the world climate research program (WCRP) to ensure that different available RCM projections can be coordinated based on various domains (Ghimire et al. 2018). The coordinated results have been categorized into 14 different domains, which can basically cover the entirety of the Earth. More information about these domains and the project objective is available at http://wcrp-cordex.ipsl.jussieu.fr. It should be noted that various climate change oriented studies have used and validated the CORDEX database (e.g. Jacob et al. 2014; Ayar et al. 2016; Casanueva et al. 2016; Grillakis et al. 2017; Mezghani et al. 2017).

While both GCMs and RCMs can project climatic variables, RCMs alone can directly simulate hydrologic components such as surface runoff. However, since such simulations are not usually in synchronization with the observed streamflow, it has been strongly suggested to avoid using such hydrologic projections in a direct manner (Teutschbein & Seibert 2012). In the meantime, utilizing the modified RCM projections of climatic variables (e.g. rainfall and temperature), especially those that have been coordinated under the CORDEX program, can allow hydrologic simulations to be carried out separately. However, the systematic modeling errors caused by imperfect conceptualization, discretization, and spatial averaging within grids may result in significant biases in the climatic projections of the RCMs (Teutschbein & Seibert 2012). Thus, to use the RCMs’ predictions for hydro-climatic impact studies, an additional post-processing procedure is often needed. Various bias correction techniques, such as statistical downscaling, histogram equalizing, rank matching, and quantile mapping (QM), have been proposed. Most of these methods use cumulative distribution functions (CDFs) of the observed and simulated climatic variables to form a bias correction function (Piani et al. 2010a).

Previous studies have highlighted the role of using bias correction techniques to modify the RCMs’ projections. Dosio & Paruolo (2011), for instance, demonstrated a successful implementation of a bias correction technique to enhance the RCMs’ rainfall and temperature projections. Their results also showed the ability of this technique to improve the statistics of the simulated results that depended strongly on the temporal sequence of the original fields, such as the number of consecutive dry days and the total amount of rainfall following heavy rainfall events, as well as improve the probability distribution functions (PDF) of the simulated rainfall and temperature variables. In another case, Teutschbein & Seibert (2012) compared the performance of some bias correction techniques including linear scaling, local intensity scaling, power transformation, variance scaling, distribution transfer, and the delta-change approach to adjust the RCMs’ rainfall and temperature projections, indicating that while all of these methods would, more or less, improve the RCMs’ predictions, some performed better. Rajczak et al. (2016) illustrated the presence of a systematic bias in the transition probability and determined the lengths for raw RCM outputs, in which the bias was immensely mitigated by using the QM methods as bias correction techniques.

As demonstrated, numerous potential bias correction techniques have been proposed, which perform differently under different conditions. This notion highlights the importance of investigating the potential roles of various bias correction techniques, and in turn, selecting suitable ones to post-process the RCMs’ projections. Several studies have already attempted to address the raised issue. Themessl et al. (2011), for example, compared the performance of several bias correction methods and concluded that the QM methods were potentially the best option available to adjust the raw RCMs’ outputs. Lafon et al. (2013) found that the QMs’ corrected results were highly accurate. It should be noted that some studies even attempted to use the QM methods to correct the bias associated with the CORDEX dataset (e.g. Štěpánek et al. 2016).

While it seems that the QM techniques, in general, can be considered as potentially valid techniques to correct the bias, it should be noted that the QM is referred to a family of methods. It is worth investigating which of the QMs can be the most suitable one for a given domain. Thus, it is crucial to conduct a thorough comparison over the effectiveness of the QM methods in the bias correction procedure, especially for the climatic variables (e.g. rainfall and temperature) that play the most crucial role in water
resources impact assessment. Consequently, this study aims to cover a comprehensive set of potentially essential QM methods and provide an insight into their roles in the bias correction procedure.

Identifying suitable bias correction techniques for the CORDEX database for different domains is an important topic (Jacob et al. 2014). The main objective of the CORDEX project is to coordinate the RCM-oriented climate change studies at regional scales. Therefore, investigating the role of the QM methods in bias correction of the gathered dataset can be the next milestone along the project development line. This could be crucially important for Iran, a country that has already faced a series of mild to severe water crises and, accordingly, the situation may even get worse due to the climate change phenomena (Zolghadr-Asli et al. 2018a). Therefore, the Karkheh River basin in Iran has been selected as a representative and strategically important case study. Furthermore, the Karkheh River basin is known for its diverse range of topographic features (e.g. both flat plains and mountains), which proved to be challenging characteristics for both GCMs and RCMs. Thus, this basin is ideal to test whether the bias correction techniques are able to cope with the challenges for such a topographically diverse landscape. To the best of our knowledge, the CORDEX dataset has never been implemented in this region, and thus, the results can also verify the validity of this dataset. According to the CORDEX, Iran is located in the domain of Middle East and North Africa (MENA) and also the domain of South Asia, but this study concentrates on the MENA domain only. The results obtained from two GCM/RCM combinations (ICHEC (Irish center for high-end computing) and NOAA-ESM (national oceanic and atmospheric administration Earth system models)) with a spatial resolution of 0.22° and a monthly temporal resolution are used. It should be noted that for the MENA region, the CORDEX datasets only provide two GCM/RCM combinations, both of which are employed in this study.

**METHODOLOGY**

As aforementioned, the main objective of this study is to explore the potential capabilities of different QM methods under various conditions. The flowchart of this study is illustrated in Figure 1. The following section provides the theoretical background behind the tested QM methods, as well as the statistical performance measurement criteria that are used in this study.

![Flowchart of this study.](http://iwaponline.com/jwcc/article-pdf/12/2/401/865958/jwc0120401.pdf)
Quantile mapping

In general, the QM methods implement statistical transformations for post-processing of climate modeling outputs. The statistical transformations involve transforming the distribution functions of the modeled variables into the observed ones using a mathematical function, which can be mathematically expressed as (Piani et al. 2010b):

\[ x^o = f(x^m) \]  

(1)

in which \( x^o \) = observed variable; \( x^m \) = modeled variable; and \( f() \) = transformation function.

Given that the QM methods use the quantile-quantile relation to converge the simulated variables’ distribution function to the observed one, one should note that with the CDFs of both observed and simulated variables’ time series, their quantile relation can also be determined, as shown below (Ringard et al. 2017):

\[ x^o = F_o^{-1}[F_m(x^m)] \]  

(2)

where \( F_m(x^m) \) = CDF of \( x^m \); and \( F_o^{-1}[] \) = inverse form of the CDF of \( x^o \), which is technically referred to as the quantile function.

As stated earlier, various frameworks have been proposed to form the transformation function. Hence, different QM methods, as detailed below, are used in this study for bias corrections.

Parametric transformation functions (PTF)

This QM method is featured with a parameter-oriented framework, which employs parametric transformation functions to form the quantile-quantile relation. A series of predefined transformation functions are proposed for this purpose, and they can be mathematically expressed as (Maraun et al. 2010):

\[ P_o = bP_m \]  

(3)

\[ P_o = a + bP_m \]  

(4)

\[ P_o = (a + bP_m) \left(1 - e^{-(\frac{P_m}{c})}\right) \]  

(5)

\[ P_o = bP_m^c \]  

(6)

\[ P_o = b(P_m - y)^c \]  

(7)

in which \( P_o \) and \( P_m \) = probability of the observed and modeled variables, respectively; \( a, b, c, y, \) and \( r \) = method-related parameters that are subject to calibration. Equations (3)–(8) are different variations of the PTF method or parametric transformation functions. They are PTF:scale (Equation (3)), PTF:linear (Equation (4)), PTF:expasymp (Equation (5)), PTF:power (Equation (6)), PTF:power. \( \times 0 \) (Equation (7)), and PTF:expasymp. \( \times 0 \) (Equation (8)). For more information on the theoretical background of these transformation functions and the process of determining the related parameters, the readers can refer to Maraun et al. (2010) and Piani et al. (2010b).

Distribution derived transformations (DIST)

While in general the CDFs of the observed and simulated variables can be fitted by either empirical or theoretical distributions (Ines & Hansen 2006), according to the DIST method, a suitable theoretical CDF can be used to derive the transfer function (Equation (2)). It should be noted that this particular QM method can only be used to mitigate the bias of rainfall variable (Li et al. 2010; Teutschbein & Seibert 2012). In the DIST method, the theoretical CDF \( [F_m()] \) is formed by using a mixture of the Bernoulli and the Gamma distributions, where the Bernoulli distribution is used to compute the occurrence probability of the rainfall variable, while the Gamma distribution is used to model the rainfall intensities (Ines & Hansen 2006).

Empirical quantiles (QUANT)

Unlike PTF, QUANT is a non-parametric QM method using a non-parametric transformation function. QUANT estimates values of the empirical CDFs of observed and modeled time series for regularly spaced quantiles. Accordingly, QUANT uses interpolations to adjust a datum with unavailable quantile values (Osuch et al. 2017).

Robust empirical quantiles (RQUANT)

RQUANT is another non-parametric QM method, which uses the local linear least squares regression to estimate the values of the quantile-quantile relation of the observed
and modeled time series for regularly spaced quantiles. Similarly, the unavailable quantile values are estimated by using interpolation of the fitted values (Villani et al. 2015).

**Smoothing splines (SSPLIN)**

Like QUANT and RQUANT, SSPLIN is another non-parametric QM method. In this method, a smoothing spline is fitted to the quantile-quantile plot of the observed and modeled time series (Kouhestani et al. 2016).

The QM family of methods are considered to be the most important, popular, and promising bias correction methods (Villani et al. 2015; Kouhestani et al. 2016). Some of these parametric methods (e.g. PTF:scale, PTF:linear, PTF:expasymp, PTF:power, PTF:power. × 0, and PTF:expasymp. × 0) depend on the predetermined functions that need to be parameter-tuned for each specific case. Furthermore, as shown by their equations, such methods are regression-based, indicating that inherently a specific theoretical distribution is assumed for the data in these methods. In non-parametric methods (e.g. QUANT, RQUANT, and SPLIN), however, the empirical distributions of the data are used, and thus there is no initial assumption of specific theoretical distributions about the data. In addition, unlike their parametric alternatives, they are not, particularly, bound by any predetermined functions, which makes such methods more flexible. However, few attempts have been made to compare a variety of these methods under various topographic conditions for different hydroclimatic variables such as temperature and rainfall that inherently have dissimilar behaviors and characteristics.

**Performance evaluation criteria**

Like any other modeling process, the QMs outputs also need to be validated by using the observed data. The Taylor diagram enables one to evaluate various results by using three different performance evaluation criteria, including linear correlation coefficient (R) (Equation (9)), normalized standard deviation (σ) (Equation (10)), and centralized root mean squared deviation (CRMSD) (Equation (11)) (Taylor 2001; Jolliff et al. 2009; Liang et al. 2018):

\[
R = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{x_i^o - \bar{x}}{\text{σ}_o \text{σ}_m} \right)
\]

\[
\sigma = \frac{\sigma_m}{\text{σ}_o}
\]

\[
\text{CRMSD} = \left( \frac{1}{N} \sum_{i=1}^{N} \left( \frac{x_i^o - \bar{x}}{\text{σ}_o \text{σ}_m} \right)^2 \right)^{0.5}
\]

in which \(σ_o\) and \(σ_m\) = standard deviation of the observed and modeled datasets, respectively; and \(N\) = number of the data. \(R\) is a dimensionless criterion that ranges from \(-1\) to \(+1\). As the correlation between \(x_o\) and \(x_m\) gets stronger, the computed \(R\) gets closer to its boundary values of \(-1\) or \(+1\) (Taylor 1990). \(σ\) is also dimensionless, and in an ideal case it is equal to \(+1\), which indicates that the standard deviation values of the observed and simulated datasets are identical. The closer to \(+1\) the computed \(R\) is, the better the performance of the model. Lastly, as the performance of the model improves, the \(\text{CRMSD}\) value would decrease (Jolliff et al. 2009). The lower the \(\text{CRMSD}\), the better the performance of the model; and in an ideal case it is equal to \(0\). It should be noted that the reference field point, which consists of the statistics generated from the observed dataset itself, is corresponding to the polar coordinates \((1.0, 0.0)\). Overall, the main idea behind the Taylor diagrams is to provide a visual premise that facilitates the comparison of three quantitative measures (i.e. \(\text{CRMSD}\), \(R\), and \(σ\)).

In addition to the Taylor diagram, in this study the mean bias error (MBE) is also used as a performance evaluation criterion to quantify the magnitude and direction of the bias. While a zero value signifies the absence of bias in the generated results, positive and negative MBE values respectively indicate under- and over-estimation. MBE can be mathematically expressed as (Willmott & Matsuura 2009):

\[
\text{MBE} = \frac{1}{n} \sum_{i=1}^{n} (x_i^o - x_i^m)
\]

**CASE STUDY**

The Karkheh River basin, with an area of 51,000 km², is located in the south-western region of Iran (latitude: 30°–35° N; longitude: 46°–49° E). Based on the topographic characteristics, the Karkheh River basin can be divided
distinctly into mountain and plain sections. Two-thirds of the basin has elevations higher than 3,500 meters above sea level (m.a.s.l.), which is mostly in the north and north-east portion of the basin. The remaining one-third of the basin in the southern part has elevations lower than 10 m.a.s.l. The Karkheh River is one of the three largest rivers in Iran. As shown in Figure 2, the basin has five major sub-basins: Gamasiab Karkheh-e-Jonobi, Kashkan, Qarasou, and Saymareh.

The Karkheh River basin is mainly characterized by arid to semi-arid climate. According to the records, the basin also experiences a high level of spatiotemporal variations in rainfall and air temperature. The southern part of the basin receives an average annual rainfall of approximately 150 mm, while the average annual rainfall for the northern and northeastern parts is as high as 1,000 mm (Zolghadr-Asli et al. 2018b). It should also be noted that more than 50% of the rainfall occurs during winters (December, January, and February), while it hardly rains during summers (June, July, and August). The temperature shows similar spatial variations. The temperature in the southern part of the basin is usually higher than that in the northern parts. The average maximum temperature during the summer, for instance, has been reported to be 45 °C in the southern part of the basin, while in the northern part of the basin it is only 35 °C (Muthuwatta et al. 2010). Thus, the northern section of the Karkheh River basin has a semi-arid climate, while the southern part of the basin can be categorized as an arid region. Figures 3 and 4 respectively illustrate the variations in the average monthly rainfall and temperature in the Karkheh River basin in the baseline period (1975–2005).

As shown in Table 1, the elevations of the sub-basins vary drastically. The northern parts of the basin are mostly dense-forest mountains, while the southern parts are relatively flat plains. For such conditions, RCMS can be potentially valuable tools for climate change projections. Accordingly, the hilly topography is often an obstacle for GCMs with coarsely gridded networks. In contrast, RCMS, which are represented in a finer gridded network, can be a better representation of the highly variable surface topography (Zolghadr-Asli et al. 2018a). In this study, the CORDEX datasets (MENA domain) were used to predict rainfall and temperature with a monthly time step. All the raw data were downloaded from the ESGF Datanode (https://esgf-dn1.nsc.liu.se/search/cordex/).

The CORDEX provides two different sets of resolution for the MENA domain, MNA22i (a finer resolution with a 0.22° gridded network) and MNA44i (a coarser resolution with a 0.44° gridded network). Since previous studies (e.g. Prein et al. 2016) have already demonstrated that a finer resolution could immensely mitigate the RCM bias, the high-resolution MNA22i dataset was selected for this study.

RESULTS AND DISCUSSION

As stated earlier, the bias correction techniques can play a crucial role in adjusting the RCMs’ results. This study, in particular, aims to shed light on the potential of the QM methods to mitigate the embedded bias of the CORDEX dataset for rainfall and temperature through a topographically diverse case study in the Karkheh River basin, Iran. The related modeling and computations were conducted by using R programming language. As for the CORDEX dataset, the results of two GCM/RCM combinations (i.e. ICHEC and NOAA-ESM, hereafter simply referred to as GCM/RCM combination I and N, respectively) were obtained with a resolution of 0.22° (MNA22i).

Rainfall

Figure 5 shows the performances of the QM methods to adjust the rainfall bias. The Taylor diagrams in Figure 5 revealed that the performances of the QM methods in all sub-basins and for both GCM/RCM combinations followed a similar pattern. DIST was the worst rainfall bias correction method, which made the results even worse compared to the non-corrected data. While the rest of the tested results were relatively more satisfactory, QUANT and RQUANT seemed to be the best options to correct the bias of the rainfall variable, for they not only improved the \( R \) and \( CRMSD \) criteria, but generated a value of \( \sigma \) close to 1 (desirable value). Further analysis of the evaluation criteria also indicated that while PTF:scale and PTF:linear performed significantly better than DIST, the rest of the methods outperformed them.
Figure 2 | Karkheh River basin and its five major sub-basins.
In addition to $R$, $\sigma_x$, and CRMSD, which were covered by the Taylor diagram, the abilities of the models to adjust the average of monthly rainfall time series were evaluated. As shown in Figure 6, while in the non-corrected condition, GCM/RCM combination I performed better than GCM/RCM combination N, both GCM/RCM combinations underestimated the rainfall in all sub-basins. For the Qarasou sub-basin, for instance, the average rainfall values estimated by GCM/RCM combinations I and N were 20.1 and 18.0 mm, respectively, both of which were much lower than the observed average rainfall (35.9 mm) in the time frame, indicating significant underestimations by these non-corrected GCM/RCM combinations.

Figure 6, however, also indicates that all QM methods were able to, more or less, modify the performances of these GCM/RCM combinations in the rainfall simulation.
In the Qarasou sub-basin, for example, the simulated rainfall by GCM/RCM combination N and using SSPLIN as the bias correction method was 36.1 mm. This value is significantly closer to the observed value compared to the non-corrected condition, which signifies the role of the QM methods in adjusting the raw data of GCM/RCM.
combinations. According to the results for the Qarasou sub-basin, both QUANT and RQUANT methods were able to outperform others as they accurately estimated the average rainfall variable in this sub-basin. The results for other sub-basins followed a similar pattern. Overall, the results indicated that QUANT and RQUANT methods stood out for rainfall bias correction. On the other hand, PTF:scale and DIST were the least capable methods to cope with the rainfall bias. For instance, according to the results for the Qarasou sub-basin, the rainfall values obtained from GCM/RCM combination N was 29.4 mm (PTF:scale) and 38.6 mm (DIST), which had the greatest differences from the observed value (35.9 mm) compared to the results from other bias correction methods.

Figure 7 shows the performances of these bias correction methods using the MBE criterion. According to the results, both GCM/RCM combinations suffered from under-estimation of the rainfall variable. For the Kashkan sub-basin, for instance, the MBE values of GCM/RCM combinations I and N in the non-corrected form were 28.9 and 28.5 mm, respectively. Meanwhile, all methods were able to, more or less, modify the raw results. For the Kashkan sub-basin, for instance, RQUANT and QUANT \( (MBE = 0.0 \text{ for both}) \) stood out as the best options to cope with the rainfall bias. On the other hand, while both DIST and PTF:scale were able to improve the MBE criterion relatively, they were the worst test options for this task. Furthermore, one should note that PTF:scale managed to change the direction of the bias. Overall, the results for the rest of the sub-basins followed a similar pattern. Thus, it can be concluded that QUANT and RQUANT are the best QM methods to correct the bias of the rainfall variable, while
not only may PTF:scale be the worst available option for this purpose, but it seems to change the direction of the bias in all tested cases. DIST was also one of the methods that did not perform satisfactorily.

**Temperature**

The abilities of QMs to adjust the raw temperature projections of the CORDEX were also evaluated. Figure 8 shows the simulated temperature values in both non-corrected and corrected forms which were compared against the observed dataset using $R$, $\sigma$, and CRMSD. According to Figure 8, by comparing the Taylor diagrams on the left column (GCM/RCM combination I) and those on the right column (GCM/RCM combination N), it can be concluded that in the non-corrected form these GCM/RCM combinations did not outperform each other in a statistically meaningful manner. Furthermore, while some studies suggested that the QM methods could unquestionably improve the RCMs’ raw outputs, as indicated by the diagrams in Figure 8, this may not be the case for this study. Evidently, the performances of the QM methods to modify the temperature variable depended on various conditions including, but not limited to, the transformation function and the topographic condition of the region. Figure 8(a), for instance, indicated that implementing the PTF:scale method could, in fact, worsen the accuracy of temperature projections, even compared to the non-corrected condition. The rest of the QM methods, however, performed relatively better in correcting the bias of the temperature variable. As shown in Figure 8(e) and 8(f), a similar pattern was observed for the rest of the sub-basins, with the notable exception of the Kashkan sub-basin. Unlike the results for the rest of the sub-basins, analyzing the performance criteria revealed that the SSPLIN method did not perform as well as others.
Figure 9 demonstrates the QMs’ capability in modifying the average temperature in each sub-basin. First, it can be concluded that both GCM/RCM combinations I and N were equally accurate in simulating the average temperature for all sub-basins. For the Karkheh-e-Jonobi, for instance, both GCM/RCM combinations simulated the average temperature to be approximately 20.7 °C, which was close to the observed value of 23.0 °C. Chiefly, this is due to the particular nature of temperature and the better ability of the RCMs to simulate this climatic variable in comparison to rainfall (Alexander et al. 2006). As for the performance of QM methods, the results obtained for the Gamasiab sub-basin, for example, showed that except for the PTF:scale, all other tested methods were able to eliminate the bias. That is, the adjusted average temperature for this basin was equal to the observed one (11.3 °C). PTF:scale, however, was able to improve this estimation by transforming from 7.3 °C (non-corrected average temperature) to 9.9 °C. Similar results were obtained for other sub-basins, with the notable exception of SSPLIN for the Kashkan sub-basin. The average temperature for the Kashak sub-basin in the baseline period was 14.7 °C, while the raw outputs from GCM/RCM combinations I and N were 12.5 and 12.4 °C, respectively. According to the results of the SSPLIN method for the Kashkan sub-basin, however, the adjusted average temperatures were 40.3 °C (GCM/RCM combination I) and 39.7 °C (GCM/RCM combination N). It should be noted, however, that according to Figure 9, the performance of the SSPLIN method was satisfactory for the remaining sub-basins.

Figure 10 displays the performances of different methods to correct the temperature bias using the MBE criterion. As shown in Figure 10, both GCM/RCM combinations I and N suffered from underestimation. In the Gamasiab sub-basin, for instance, the computed MBE values for GCM/RCM combinations I and N were 5.6 and 5.2 °C, respectively. Evidently, the QM methods were able to overcome this issue.
With two notable exceptions, all other QM methods were able to reduce MBE to zero, implying that they completely remove the temperature bias. The exceptions were the results obtained from the PTF:scale method for all sub-basins and the SSPLIN results for the Kashkan sub-basin. For the Gama-siab sub-basin, for example, not only did PTF:scale fail to

Figure 8 | Comparisons of the performances of QM methods to adjust the temperature bias using $R$, $\sigma$, and CRMSD criteria. (Continued.)
control or reduce the bias, but it changed the direction of the bias, resulting in an MBE of \(-1.5^\circ C\). The SSPLIN methods drastically worsened MBE. The computed MBE values for GCM/RCM combinations I and N in the non-corrected condition were 3.2 and 3.0 °C, respectively, while the SSPLIN-adjusted values were 25.6 °C (GCM/RCM combination I) and 24.9 °C (GCM/RCM combination N).

**SUMMARY AND CONCLUSIONS**

Although nowadays advanced regional climate models (RCMs) are more than capable of characterizing climatic conditions on a regional scale, they may still suffer from systematic errors. Bias correction procedures such as the quantile mapping (QM) methods can help overcome this critical problem. However, one should bear in mind that there are a variety of QM methods that may not necessarily have a similar capability in the correction of the bias of the RCMs’ outputs. Addressing this issue is crucially important, especially for the cases with distinct topographic features, since it has been an already challenging task for RCMs to simulate the climatic variables under such conditions. This study aimed to conduct a thorough investigation to compare the performances of ten different QM methods and quantify their capabilities in modifying the bias of rainfall and temperature variables. The results from two general circulation model (GCM)/RCM combinations were obtained from the coordinated regional climate downscaling experiment (CORDEX) dataset for the Karkheh River basin in Iran.
The results indicated that the performances of these bias correction methods could vary significantly from one another. For the rainfall variable, for instance, distribution derived transformations (DIST) had the worst performance among all the tested methods. For the temperature variable, however, parametric transformation functions (PTF:scale) led to the worst results. It should be noted that in both cases the PTF:scale method even worsened the outputs compared to the non-corrected form. Additionally, PTF:scale changed the direction of the rainfall and temperature biases, while other PTF methods performed as satisfactorily as their non-parametric counterparts. This can be attributed to the simplistic nature of the parametric transformation function of this method. For the DIST method, the overly simplified mathematical transformation function disables it to sufficiently emulate the observed cumulative distribution functions (CDFs) of climatic variables.

On the other hand, the overall analysis of the obtained results demonstrated that in the correction of the bias of the rainfall variable, the empirical quantiles (QUANT) and robust empirical quantiles (RQUANT) were the best available options. The outstanding performances of these non-parametric methods (i.e. QUANT and RQUANT) can be attributed to their flexible natures and the fact that unlike their parametric alternatives, they are not, particularly, bound by any predetermined functions. As for the temperature variable, with the notable exception of PTF:scale, the performances of all other methods were sufficiently acceptable. Chiefly, this is due to the particular nature of temperature and the better ability of the GCM/RCM combinations to simulate this climatic variable in comparison to rainfall.

Overall, while non-parametric methods such as smoothing splines (SSPLIN) can be potentially good options, as evidenced by the performance to cope with the biases of rainfall and temperature, implementation of these methods, especially in the cases of diverse topographic features, should be done with great caution. For the Kashkan sub-basin with significant topographic variations (mixed high mountains and flat plains),
for instance, SSPLIN could not efficiently adjust the temperature variable, while this method provided good results if only one topographic feature was relatively dominant.

**ACKNOWLEDGEMENT**

The authors thank Iran’s National Science Foundation (INSF) for its financial support for this research.

**DATA AVAILABILITY STATEMENT**

The data that support the findings of this study are available from the corresponding author upon request.

**REFERENCES**

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F. & Tagipour, A. 2006 Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res. Atmos.* 111 (D5), D05109.

Ayar, P. V., Vrac, M., Bastin, S., Carreau, J., Déqué, M. & Gallardo, C. 2016 Intercomparison of statistical and dynamical downscaling models under the EURO-and MED-CORDEX initiative framework: present climate evaluations. *Clim. Dyn.* 46 (3–4), 1301–1329.

Busuioc, A., Chen, D. & Hellström, C. 2001 Performance of statistical downscaling models in GCM validation and regional climate change estimates: application for Swedish precipitation. *Int. J. Climatol.* 21 (5), 557–578.

Casanueva, A., Kottlarski, S., Herrera, S., Fernández, J., Gutiérrez, J. M., Boberg, F., Colette, A., Christensen, O. B., Goergen, K., Jacob, D. & Keuler, K. 2016 Daily precipitation statistics in a EURO-CORDEX RCM ensemble: added value of raw and bias-corrected high-resolution simulations. *Clim. Dyn.* 47 (3–4), 719–737.
Chen, H., Xu, C. Y. & Guo, S. 2012 Comparison and evaluation of multiple GCMs, statistical downscaling and hydrological models in the study of climate change impacts on runoff. *J. Hydrol.* 434, 36–45.

Dibike, Y. B. & Coulibaly, P. 2005 Hydrologic impact of climate change in the Saguenay watershed: comparison of downscaling methods and hydrologic models. *J. Hydrol.* 307 (1–4), 145–163.

Dosio, A. & Paruolo, P. 2011 Bias correction of the ENSEMBLES high-resolution climate change projections for use by impact models: evaluation on the present climate. *J. Geophys. Res. Atmos.* 116 (D16), D16106.

Faiz, M. A., Liu, D., Fu, Q., Li, M., Baig, F., Tahir, A. A., Khan, M. I., Li, T. & Cui, S. 2018 Performance evaluation of hydrological models using ensemble of General Circulation Models in the northeastern China. *J. Hydrol.* 565, 599–613.

Ghimire, S., Choudhary, A. & Dimri, A. P. 2018 Assessment of the performance of CORDEX-South Asia experiments for monsoonal precipitation over the Himalayan region during present climate: part I. *Clim. Dyn.* 50 (7–8), 2311–2334.

Giorgi, F., Jones, C. & Asrar, G. R. 2009 Addressing climate information needs at the regional level: the CORDEX framework. *World Meteorol. Org. (WMO) Bull.* 58 (3), 175.

Grillakis, M. G., Koutourous, A. G., Daliakopoulos, I. N. & Tsanis, I. K. 2017 A method to preserve trends in quantile mapping bias correction of climate modeled temperature. *Earth Syst. Dyn.* 8 (3), 889–900.

Hertig, E. & Jacob, J. 2008 Downscaling future climate change: temperature scenarios for the Mediterranean area. *Glob. Planet. Change* 63 (2–3), 127–131.

Ines, A. V. & Hansen, J. W. 2006 Bias correction of daily GCM rainfall for crop simulation studies. *Agric. For. Meteorol.* 138 (1–4), 44–53.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M. & Georgoulis, E. 2014 EURO-CORDEX: New high-resolution climate change projections for European impact research. *Reg. Environ. Change* 14 (2), 563–578.

Jha, M., Pan, Z., Takle, E. S. & Gu, R. 2004 Impacts of climate change on streamflow in the Upper Mississippi River Basin: a regional climate model perspective. *J. Geophys. Res. Atmos.* 109 (D9), D09105.

Jolliff, J. K., Kindle, J. C., Shulman, I., Penta, B., Friedrichs, M. A., Helber, R. & Aronne, R. A. 2009 Summary diagrams for coupled hydrodynamic-ecosystem model skill assessment. *J. Mar. Syst.* 76 (1–2), 64–82.

Kay, A. L., Jones, R. G. & Reynard, N. S. 2006 RCM rainfall for UK flood frequency estimation. II. Climate change results. *J. Hydrol.* 318 (1–4), 163–172.

Kouhestani, S., Esfahanian, S. S., Abedi-Koupai, J. & Besalatpour, A. A. 2016 Projection of climate change impacts on precipitation using soft-computing techniques: a case study in Zayandeh-rud Basin, Iran. *Glob. Planet. Change* 144, 158–170.

Lafon, T., Dadson, S., Buys, G. & Prudhomme, C. 2013 Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. *Int. J. Climatol.* 33 (6), 1367–1381.

Li, H., Sheffield, J. & Wood, E. F. 2010 Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. *J. Geophys. Res. Atmos.* 115 (D10), D10101.

Liang, Z., Fang, T., Li, B., Liu, T., Wang, J. & Hu, Y. 2018 Long-term streamflow forecasting using SWAT through the integration of the random forests precipitation generator: case study of Danjiangkou Reservoir. *Hydrol. Res.* 49 (5), 1513–1527.

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. 2010 Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.* 48 (3), RG3003.

Mezghani, A., Dobler, A., Haugen, J. E., Benestad, R. E., Parding, K. M., Piniewski, M., Kardel, I. & Kundzewicz, Z. W. 2017 CHASE-PL Climate Projection dataset over Poland–bias adjustment of EURO-CORDEX simulations. *Earth Syst. Sci. Data* 9 (2), 905–925.

Muthuwapatta, L. P., Bos, M. G. & Rientjes, T. H. M. 2010 Assessment of water availability and consumption in the Karkheh River Basin, Iran using remote sensing and geostatistics. *Water Resour. Manage.* 24 (3), 459–484.

Osuch, M., Lawrence, D., Meresa, H. K., Napiorkowski, J. J. & Romanowicz, R. J. 2017 Projected changes in flood indices in selected catchments in Poland in the 21st century. *Stoch. Environ. Res. Risk Assess.* 31 (9), 2435–2457.

Piani, C., Haerter, J. O. & Coppola, E. 2010 Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor. Appl. Climatol.* 99 (1–2), 187–192.

Piani, C., Weedon, G. P., Best, M., Gomes, S. M., Viterbo, P., Hagemann, S. & Haerter, J. O. 2010 Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *J. Hydrol.* 395 (3–4), 199–215.

Plummer, D. A., Caya, D., Frigon, A., Côté, H., Giguère, M., Paquin, D., Biner, S., Harvey, R. & De Elia, R. 2006 Climate and climate change over North America as simulated by the Canadian RCM. *J. Clim.* 19 (13), 3112–3132.

Prein, A. F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., Maule, C. F., Van Meijgaard, E., Déqué, M., Nikulin, G. & Vautard, R. 2016 Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: high resolution, high benefits? *Clim. Dyn.* 46 (1–2), 383–412.

Rajczak, J., Kotlarski, S. & Schär, C. 2016 Does quantile mapping of simulated precipitation correct for biases in transition probabilities and spell lengths? *J. Clim.* 29 (5), 1605–1615.

Ringard, J., Seyler, F. & Lingnet, L. 2017 A quantile mapping bias correction method based on hydroclimatic classification of the Guiana Shield. *Sensors* 17 (6), 1413.
Romera, R., Gaertner, M. Á., Sánchez, E., Domínguez, M., González-Alemán, J. J. & Miglietta, M. M. 2017 Climate change projections of medicanes with a large multi-model ensemble of regional climate models. *Glob. Planet. Change* 151, 134–143.

Samuelsson, P., Jones, C. G., Willén, U., Ullerstig, A., Gollvik, S., Hansson, U. L. F., Jansson, E., Kjellström, C., Nikulin, G. & Wyser, K. 2011 The rossby centre regional climate model RCA3: model description and performance. *Tellus A Dyn. Meteorol. Oceanogr.* 63 (1), 4–23.

Štepánek, P., Zahradníček, P., Farda, A., Skalák, P., Trnka, M., Meitner, J. & Rajdl, K. 2016 Projection of drought-inducing climate conditions in the Czech Republic according to Euro-CORDEX models. *Clim. Res.* 70 (2–3), 179–193.

Tadross, M., Jack, C. & Hewitson, B. 2005 On RCM-based projections of change in southern African summer climate. *Geophys. Res. Lett.* 32 (23), L23713.

Taylor, R. 1990 Interpretation of the correlation coefficient: a basic review. *J. Diagn. Med. Sonogr.* 6 (1), 35–39.

Taylor, K. E. 2001 Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* 106 (D7), 7183–7192.

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, 12–29.

Themeßl, M. J., Gobiet, A. & Leuprecht, A. 2011 Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *Int. J. Climatol.* 31 (10), 1530–1544.

Villani, V., Rianna, G., Mercogliano, P. & Zollo, A. L. 2015 Statistical approaches versus weather generator to downscale RCM outputs to slope scale for stability assessment: a comparison of performances. *Electron. J. Geotech. Eng.* 20 (4), 1495–1515.

Wilby, R. L., Dawson, C. W. & Barrow, E. M. 2002 SDSM – a decision support tool for the assessment of regional climate change impacts. *Environ. Model. Softw.* 17 (2), 145–157.

Willmott, C. J. & Matsuura, K. 2005 Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* 30 (1), 79–82.

Zolghadr-Asli, B., Bozorg-Haddad, O., Sarzaeim, P. & Chu, X. 2018a Investigating the variability of GCMs’ simulations using time series analysis. *J. Water Clim. Change* 10 (3), 449–463.

Zolghadr-Asli, B., Bozorg-Haddad, O. & Chu, X. 2018b Effects of the uncertainties of climate change on the performance of hydropower systems. *J. Water Clim. Change* 10 (3), 591–609.