Revista Brasileira de Meteorologia, **v. 36, n. 3,** 339−349, 2021
DOI: http://dx.doi.org/10.1590/0102-77863630068

rbmet.org.br

Artigo

Multi-Criteria Decision Framework to Evaluate Bias Corrected Climate Change Projections in the Piracicaba River Basin

Camila Billerbeck[1], Ligia Monteiro da Silva[1], Silvana Susko Marcellini[1], Arisvaldo Méllo Junior[1]

[1]*Universidade de São Paulo, Departamento de Engenharia Hidráulica e Ambiental, Laboratório de Sistemas de Suporte a Decisões, São Paulo, SP, Brasil.*

Received: 28 May 2020 - Revised: 16 February 2021 - Accepted: 24 April 2021

Abstract

Regional climate models (RCM) are the main tools for climate change impacts assessment in hydrological studies. These models, however, often show biases when compared to historical observations. Bias Correction (BC) are useful techniques to improve climate projection outputs. This study presents a multi-criteria decision analysis (MCDA) framework to compare combinations of RCM with selected BC methods. The comparison was based on the modified Kling-Gupta efficiency (*KGE'*). The criteria evaluated the general capability of models in reproducing the observed data main statistics. Other criteria evaluated were the relevant aspects for hydrological studies, such as seasonality, dry and wet periods. We applied four BC methods in four RCM monthly rainfall outputs from 1961 to 2005 in the Piracicaba river basin. The Linear Scaling (LS) method showed higher improvements in the general performance of the models. The RCM Eta-HadGEM2-ES, corrected with Standardized Reconstruction (SdRc) method, achieved the best results when compared to the observed precipitation. The bias corrected projected monthly precipitation (2006-2098) preserved the main signal of climate change effects when compared to the original outputs regarding annual rainfall. However, SdRc produced significant decrease in monthly average rainfall, higher than 45% for July, August and September for RCP4.5 and RCP8.5 scenarios.

Keywords: Eta-RCM, bias correction, standardized reconstruction.

Análise de Decisão Multi-Critério para Avaliar Modelos de Projeção Climática com Correção de Viés na Bacia do Rio Piracicaba

Resumo

Modelos climáticos regionais (RCM) são as principais ferramentas para analisar impactos de mudanças climáticas em variáveis hidrológicas. Contudo, frequentemente apresentam vieses quando comparados a observações para o mesmo período, e técnicas de correção de viés (BC) têm sido utilizadas. Este estudo propôs uma análise de decisão multicritério (MCDA) para comparar combinações de RCM com metodologias de BC. Os critérios permitiram analisar a capacidade dos modelos em reproduzir as estatísticas gerais observadas, e obter informações relevantes para estudos hidrológicos, como aspectos relacionados à sazonalidade e de períodos secos e úmidos. A MCDA foi baseada no coeficiente modificado de Kling-Gupta (*KGE'*). Quatro métodos de BC foram aplicados em dados de precipitação mensal de quatro RCM para a bacia do rio Piracicaba. O método Linear Scaling (LS) apresentou os melhores resultados no aprimoramento da performance geral dos modelos. O método Standardized Reconstrucion (SdRc), combinado ao RCM Eta-HadGEM2-ES, obteve o melhor resultado para o período de validação (1991-2005). A análise dos indicadores da MCDA foi importante para a compreensão dos efeitos da BC. Os cenários projetados corrigidos (2006-2098) não apresentaram alterações no sinal da mudança climática do RCM com relação à precipitação total anual. No entanto, o método SdRc diminuiu em pelo menos 45% a precipitação média mensal dos meses Julho, Agosto e Setembro para os cenários RCP4.5 e RCP8.5.

Palavras-chave: Eta-RCM, correção de viés, standardized reconstruction.

Autor de correspondência: Camila Billerbeck, camila.billerbeck@usp.br.

1. Introduction

Regional climate models (RCMs) are important tools to assess climate change impacts on hydrological processes. These models are usually developed through the downscaling of global climate models (GCMs), with resolutions of approximately 100-250 km (Teutschbein and Seibert, 2010). They incorporate representations of regional climate dynamics, in resolutions that range from 5 to 50 km. Therefore, RCMs generally improve the representation of variables of interest such as precipitation and temperature at catchment scales (Argüeso *et al.*, 2013).

However, RCMs historical outputs often show different statistical properties when compared to observations for the same period. These errors, or biases, are defined as the differences in the main statistics between climate model simulations and the observed data. Deviations are caused by systematic and random RCMs' errors and by errors derived from the original GCMs such as resolution issues and atmospheric circulation and ocean dynamics (Shrestha *et al.*, 2017; Guimarães *et al.*, 2016).

Bias correction (BC) methods are used to remedy the problems with biased RCM datasets (Teutschbein and Seibert, 2012). The basic principle is that the differences between the model historical outputs and the observed data can be extrapolated for the future period. BC methods firstly identify existing biases in control or historical period and then correct the results of climate models in both control and future scenarios (Teutschbein and Seibert, 2010). Although BC has been a matter of criticism for some authors as described by Ehret *et al.* (2012) and Hempel *et al.* (2013), these techniques are widely used and recognized as helpful, since there are no other feasible alternatives for improving current models outputs (Argüeso *et al.*, 2013).

BC techniques range from simple scaling approaches to sophisticated methods related to probability mapping (Johnson and Sharma, 2012; Teutschbein and Seibert, 2012; Shrestha *et al.*, 2017). Therefore, the development of evaluation metrics is important to quantify their robustness and performance (Eum *et al.*, 2017), especially the ability to reproduce the overall statistics of observed historical data. Among the main indicators founded in research are the Kolmogorov-Smirnof nonparametric test (*KS*), the Nash-Sutcliffe efficiency (*NSE*), the ratio of the root mean square error to the standard deviation of measured data (*RSR*), among others. (Tschöke *et al.*, 2017; Akhter *et al.*, 2017).

Studies tend to assess the performance of RCMs outputs based on the mean statistics of historical data, often ignoring rainfall temporal variability. Multi-criteria decision analysis (MCDA) can capture model's ability in representing seasonal elements, as well as in reproducing dry and wet periods (Eum *et al.*, 2017).

This study presents a general MCDA framework for users that seek to evaluate and select the most suitable combination of RCM and BC method in an available inventory. We applied this framework to RCMs projected monthly rainfall and assessed their ability to replicate historical rainfall statistics related to general performance, time-series seasonality, and extreme values (such as driest and wettest months). The criteria indicators were evaluated by using the modified Kling-Gupta efficiency - *KGE'* (Kling *et al.*, 2012).

2. Material and Methods

RCMs monthly rainfall outputs were corrected based on calibrated observed data (1961-1990) by four methods: (i) Linear Scaling (LS), (ii) Standardized Reconstruction (SdRc), (iii) Empirical Quantile Mapping (EQM) and (iv) Gamma Quantile Mapping (GQM). Performance of bias corrected data from the validation period (1991-2005) was then evaluated through a multi-criteria decision analysis (MCDA) based on three criteria: general, seasonal and extreme values related performance.

The study area is in the Piracicaba river basin. This basin is highly urbanized and industrialized, with an estimated population of 3.4 million people, 66.7% of them located in urban area. The basin has a historical of water conflict challenges and suffered a severe drought during the years of 2013-2015 compromising water availability (PCJ, 2020). This section describes the selected datasets, bias correction methods and the proposed evaluation framework.

2.1. Climate datasets

Observed monthly rainfall data from 1961 to 2005 were obtained from a set of 85 stations provided by the Brazilian National Water Agency (ANA). Observed data was submitted to gap-filling by using the Regional Vector Method (RVM) (Hiez, 1977). Climate change monthly rainfall from historical datasets (1961 to 2005) were obtained from the Eta RCM models provided by CPTEC/INPE. The Eta RCM used in this study are nested in four global climate models: (i) BESM (Nobre *et al.*, 2013), (ii) CanESM2 (Arora *et al.*, 2011), (iii) HadGEM2-ES (Collins *et al.*, 2011) and (iv) MIROC5 (Watanabe *et al.*, 2010). The RCM domain comprises most of Caribbean, Central and South America, with approximately 20 km of resolution in the horizontal and 38 layers in the vertical (Brazil, 2015; Chou *et al.*, 2014a, 2014b; Lyra *et al.*, 2018).

The use of Eta vertical coordinate is recommended for regions with steep orography, such as the Andes Cordillera (Chou, *et al.*, 2012; Pesquero *et al.*, 2010). The application of Eta model for climate change assessment in South America leads to satisfactory results, and monthly precipitation totals indicate that seasonal variability is reasonably reproduced (Chou *et al.*, 2005). However, some areas exhibit systematic biases and overestimates rainfall, especially near mountain regions such as southeastern Brazil (Chou *et al.*, 2012). In this paper, the Eta-RCM are designated after their original GCM source.

Figure 1 shows the Piracicaba river basin, the location of the 85 rainfall stations and the corresponding Eta RCM grid.

2.2. Bias correction methods

Four BC methods were evaluated in this study: Linear Scaling (LS), Standardized Reconstruction (SdRc), Gamma Quantile Mapping (GQM) and Empirical Quantile Mapping (EQM). LS and SdRc are statistical-based methods, while GQM and EQM are distribution-based ones. All of them were calibrated with restored observational station data for 30 years from 1961 to 1990 and evaluated for the 15-year period from 1991 to 2005.

2.2.1. Statistical-based methods

Linear Scaling (LS) is one of the simplest and most used bias correction methods. It consists of using the difference (additive) or the quotient (multiplicative) between simulated and observed data means during the calibration period to scale the model simulation (Akhter *et al.*, 2017). By definition, corrected model data presents the same monthly mean values as observed data (Teutschbein and Seibert, 2012). In this study, the multiplicative approach was applied to correct monthly rainfall at each station (Eq. (1)):

$$P'_{sim,val} = P_{sim,val} * \frac{\overline{P_{obs,cal}}}{\overline{P_{sim,cal}}} \tag{1}$$

where $P'_{sim,val}$ and $P_{sim,val}$, are the bias corrected and original RCM output precipitation for the validation period respectively; $\overline{P_{obs,cal}}$ and $\overline{P_{sim,cal}}$ are the monthly average rainfall observed and simulated values during the calibration period.

Standardized Reconstruction (SdRc) is a useful technique for application in monthly data (Acharya *et al.*, 2013; Johnson and Sharma, 2012). In SdRc, the RCM precipitation output is standardized by its monthly mean (μ) and monthly standard deviation (σ) considering the calibration period, to calculate a future rainfall anomaly $Y'_{sim,val}$ (Eq. (2)). RCM corrected data for the validation period $Z'_{sim,val}$, was obtained by Eq. (3), using the mean and standard deviation of the observed data (Akhter; *et al.*, 2017).

$$Y'_{sim,val} = \frac{Y_{sim,val} - \mu_{sim,cal}}{\sigma_{sim,cal}} \tag{2}$$

$$Z'_{sim,val} = Y'_{sim,val} * \sigma_{obs,cal} + \mu_{obs,cal} \tag{3}$$

2.2.2. Distribution-based methods

Quantile Mapping methods are statistical transformations that fit a distribution function $f(x)$ to a modeled

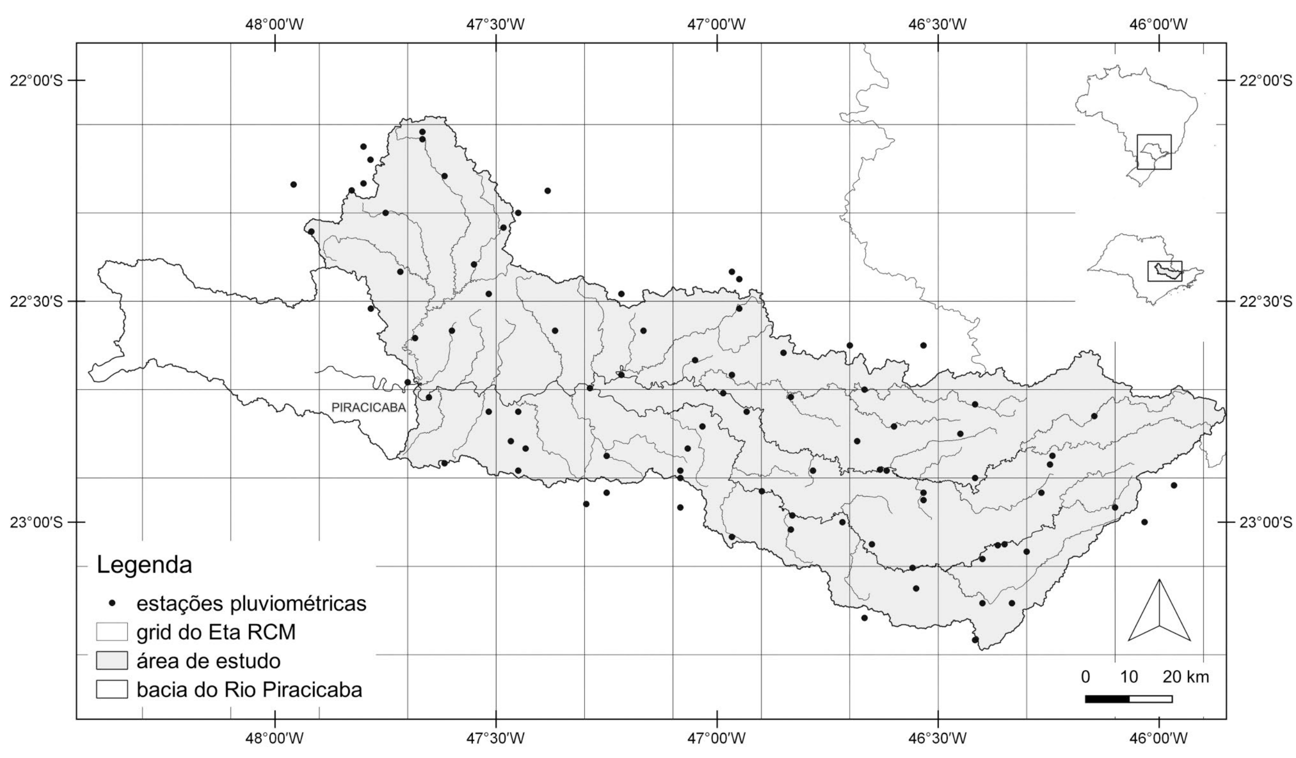

Figure 1 - Location of the study area in the Piracicaba river basin, major rivers, rainfall stations, and Eta RCM grid.

variable P_{sim} such that it equals the distribution of the observed value P_{obs} as shown in Eq. (4) (Akhter *et al.*, 2017; Gudmundsson *et al.*, 2012). If the distribution of the modeled variable P_{sim} is known, then the transformation can be obtained by Eq. (5).

$$P_{sim} = f(P_{obs}) \tag{4}$$

$$P_{obs} = F_{obs}^{-1}(F_{sim}(P_{sim})) \tag{5}$$

where F_{sim} is the Cumulative Distribution Function (CDF) of P_{sim} and F_{obs}^{-1} is the inverse CDF corresponding to P_{obs} (Gudmundsson *et al.*, 2012). Common approaches to solve Eq. (5) are the use of theoretical or empirical distributions.

Gamma Quantile Mapping (GQM) is a theoretical approach based on the assumption that Gamma distribution describes both observed and simulated distributions. Its application to rainfall intensities is found in several studies (Akhter *et al.*, 2017; Gudmundsson *et al.*, 2012; Piani *et al.*, 2010). Empirical Quantile Mapping (EQM) is an alternate approach based on linear interpolation. It calculates a correction factor for percentile intervals in simulated output data according to observed rainfall (Boé *et al.*, 2007).

2.3. Method evaluation

BC methods were evaluated using the corrected monthly rainfall for the validation period (1991-2005) with the corresponding observed data considering each rain gauge. The analysis was based on a multi-criteria decision (MCD) framework comprised of three criteria. These were selected to measure the ability of RCM corrected outputs in reproducing the main statistics of the observed rainfall, as well as in representing time-series and seasonality related statistics. Selected criteria indicate the reproduction of: (i) the general performance of the observed data, (ii) indicators related to time-series, i.e. annual rainfall and monthly averages and (iii) extreme monthly values, i.e. the wettest and driest months by year.

The described evaluation parameters are summarized in Table 1.

The indicators performances were evaluated using the modified Kling-Gupta efficiency (*KGE'*). *KGE* was firstly proposed by Gupta *et al.* (2009) as an improvement of the Nash-Sutcliffe efficiency (*NSE*) and later modified by Kling *et al.* (2012). The coefficient has been applied to evaluate projected rainfall in Brazil, showing satisfactory results (Baez-Villanueva *et al.*, 2018; Bozzini and Mello Junior, 2020). Eq. (6) shows *KGE'* and Eqs. (7)-(8) its parameters, respectively.

$$KGE' = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \tag{6}$$

Table 1 - Criteria and indicators of the proposed MCDA framework.

Criteria	Indicator
General	Monthly rainfall
Time-series	Annual rainfall
	Monthly average rainfall
Extremes	Wettest months rainfall series
	Driest months rainfall series

$$\beta = \frac{\mu_{sim}}{\mu_{obs}} \tag{7}$$

$$\gamma = \frac{\frac{\sigma_{sim}}{\mu_{sim}}}{\frac{\sigma_{obs}}{\mu_{obs}}} \tag{8}$$

where *KGE'* [-] is the modified *KGE*, r [-] is the correlation coefficient, β [-] is the bias ratio, γ [-] is the variability ratio, μ [mm] is the mean rainfall and σ [mm] is the standard deviation of rainfall series. *KGE'* ranges from $-\infty$ and has its optimum value in 1.

Since hydrological studies are particularly interested in evaluating temporal and variability dynamics, the combination of r β and γ in *KGE'* offers interesting diagnostic insights into the model performance. *KGE'* values corresponding to selected criteria were scored according to categories shown in Table 2 as proposed by Irving *et al.* (2018). The MCDA framework was then applied as a simple sum of the scored values, considering equal weight for the 5 indicators.

For each RCM output, five scenarios were analyzed: as simulated before bias correction (SIM), and after correction with Linear Scaling (LS), Standardized Reconstruction (SdRc), Gamma Quantile Mapping (GQM) and Empirical Quantile Mapping (EQM). The 5 *KGE'* indicators were obtained for the 85 rainfall stations (Fig. 1) and scored accordingly. Considering the minimum score is 0 and maximum is 2, the optimum performance value per scenario, per indicator, is 170 (85 times 2) and the total optimum value is 850 (170 times 5).

Next section presents the results in terms of *KGE'* obtained for the indicators, and the MCDA matrix confronting the combinations of BC techniques with each RCM output performance. The best combination was then used to generate future precipitation scenarios, considering the Representative Concentration Pathways (RCP), which reflect the pathways to reach a specific radioactive forcing by the year 2100. There are four RCP scenarios,

Table 2 - Classification of *KGE'* in categories.

Category	Condition	Score
Low	KGE' ≤ 0	0
Mid	0 < KGE' ≤ 0,4	1
High	KGE' > 0,4	2

RCP2.6, RCP4.5, RCP6 and RCP8.5, labelled after a possible range of radiative forcing values in the year 2100 (2.6, 4.5, 6, and 8.5 W/m^2, respectively). RCP4.5 and RCP8.5 scenarios are considered, respectively an intermediate and a high emission scenario in terms of radiation forcing (Moss *et al.*, 2010). Figure 2 summarizes the methodology presented in this study.

3. Results and Discussions

3.1. General performance

Four BC methods were applied to the selected RCM monthly rainfall outputs for the validation period (1991-2005). Figure 3 shows *KGE'* general performance (a) and its components r (b), β (c) and γ (d). All of them have their optimum value in 1. *KGE'* ranges from - ∞ to 1, correlation r ranges from 0 to 1, and the components β and γ range from - ∞ to ∞. Values of β and γ higher than 1 indicate that simulated outputs overestimate observed data.

Figure 3(a) shows an overall improvement of *KGE'* for all BC methods when compared to the original simulated data (SIM). LS method had the best performance in all scenarios, reaching *KGE'* values close to 0,6. SdRc showed good results with HadGEM2-ES but performed similarly to the remaining BC methods when applied to the other RCM. GQM and EQM showed improvements only when applied to BESM.

The correlation component (r) is shown in Fig. 3(b). LS also improved results when compared to other BC methods and the original simulated data (SIM). BESM presented the best correlation for the original scenario, which was only improved by LS. SdRc showed good performance when applied to HadGEM2-ES and MIROC5. For all RCM, GQM and EQM failed to improve correlation when compared to the original scenario (SIM).

Regarding the bias parameter (β), Fig. 3(c) shows that in general, HadGEM2-ES had the best performance amongst RCM outputs in the original (SIM) scenario. BC methods had similar effects in reducing the interquartile range of the monthly rainfall and in approximating the parameter to the optimum value. The variability parameter (γ), however, indicates that BC methods tended to increase dispersion of the monthly rainfall when compared to the original (SIM) outputs. Considering the optimum value of one, γ obtained with LS indicates it performed better than the other BC methods for both HadGEM2-ES and MIROC5.

3.2. Time-series and extremes performance

The analysis of time series and extremes related indicators based on *KGE'* presented other insights on BC performances, which may be useful depending on the application of RCM output data. Figure 4(a) and (b), show the annual rainfall and the monthly average rainfall indicators (time-series analysis), while Fig. 4(c) and (d) show the wettest and driest months indicators (extremes analysis).

Regarding the annual rainfall, results indicate that the original simulated (SIM) and their respective corrected outputs failed to reproduce the observed data. In terms of monthly average values, BC methods significantly improved the model outputs. All BC methods showed similar improvements for BESM and CanESM2, while LS and SdRc performed better for HadGEM2-ES and MIROC5, which can be explained by the fact that both are based on monthly statistics. Although BC was applied on a monthly basis, this result indicates that the chosen methods were more capable of representing an overall monthly average than preserving the original seasonality in terms of total annual rainfall.

In terms of extremes indicators, the application of BC methods resulted in small improvements of *KGE'* compared to the SIM data for both dry and wet months, although the results for the wettest months were, overall, higher. Despite the high dispersion in all BC methods, BESM and HadGem2-ES had better *KGE'* results for the wettest months. Considering the driest months, both uncorrected (SIM) and corrected data showed poor results and, in general, EQM showed better performance for all RCM. These results indicate that BC methods applied on a

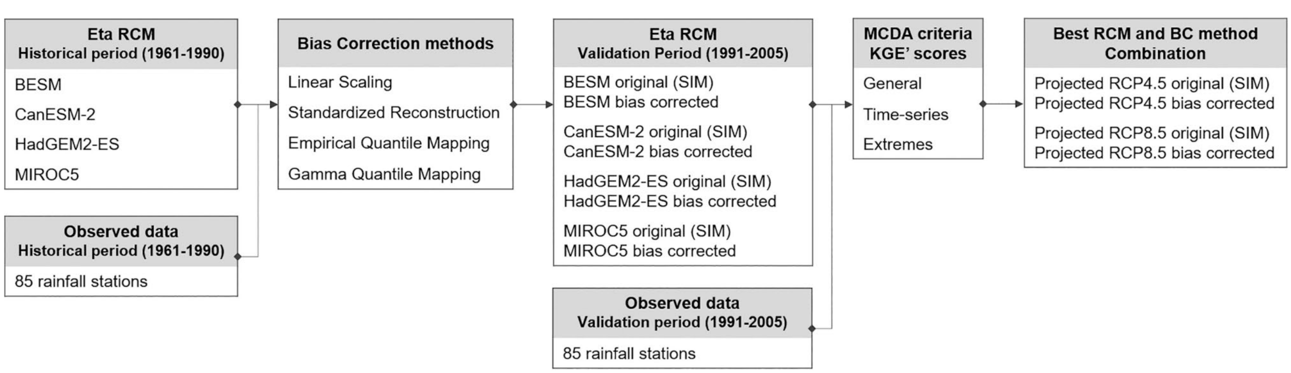

Figure 2 - Overview of proposed methodology.

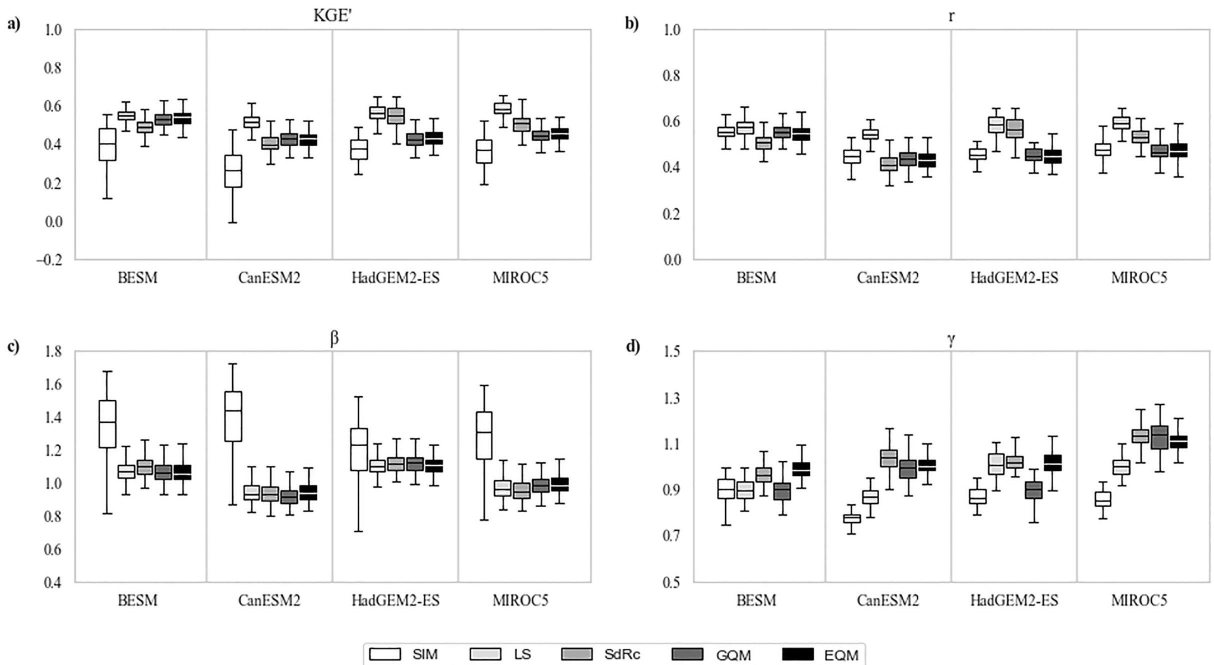

Figure 3 - *KGE'* and its components for the validation period (1991-2005) for monthly rainfall analysis: a) general performance, b) correlation (r), c) bias ratio (β) and d) variability ratio (γ). Legend presents original simulated data (SIM) and the BC methods used: Linear Scaling (LS), Standardized Reconstruction (SdRc), Gamma Quantile Mapping (GQM) and Empirical Quantile Mapping (EQM).

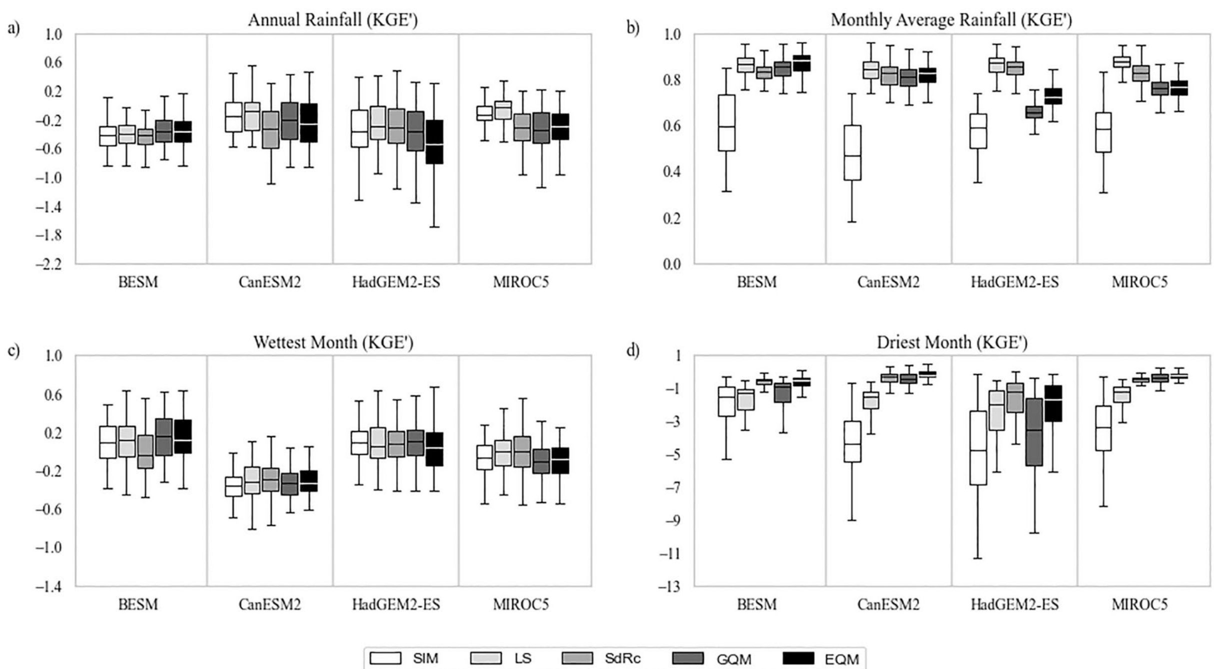

Figure 4 - *KGE'* for the validation period (1991-2005) for Time-series related criteria: a) Annual rainfall and b) Monthly average rainfall; and extremes: c) Wettest months, and d) Driest months. Legend presents original simulated data (SIM) and the BC methods used: Linear Scaling (LS), Standardized Reconstruction (SdRc), Gamma Quantile Mapping (GQM) and Empirical Quantile Mapping (EQM).

monthly scale are better at representing wet months over dry months, as shown by Shresta *et al.* (2017). The poor performance regarding the extremes indicator analysis for all BC applications might be explained by the uncorrected RCM biased outputs. As pointed out by Pastén-Zapata (2020), these may be unable to represent extremes.

3.3. *KGE'* scored classification

Figure 5 presents the categorical distribution of the 85 rainfall stations as presented in Table 2, regarding the obtained *KGE'* scores for each indicator. Results showed that rainfall gauge stations located in the same area respond differently when same BC method is applied, as found by Tschöke *et al.* (2017). In terms of general performance, all BC methods improved *KGE'* classification in comparison to the uncorrected (SIM) data. LS method achieved the greatest number of stations classified as high for all RCM outputs. SdRc improved the results significantly for BESM, HAdGEM2-ES and MIROC5. GQM and EQM presented better results when applied to BESM model, which has the majority of stations classified as high for the original (SIM) data.

The percentage of stations achieving high *KGE'* scores for the other four indicators showed that all BC methods were better at representing average monthly values, compared to preserving annual rainfall seasonality. It is also notable that BC methods had poor improvement in representing dry months statistics, compared to the wet months.

Table 3 summarizes the MCDA results, obtained by simple sum of *KGE'* scores classified according to Table 2. The optimum score is 850, which represents the *KGE'* classified as 2 (high) for all 85 rainfall stations, considering the five indicators. Low uncorrected data (SIM) scores indicate that original RCMs are a source of uncertainty and biases, as previous studies point out (Teutschbein and Seibert, 2012, Guimarães *et al.*, 2016; Shrestha *et al.*,

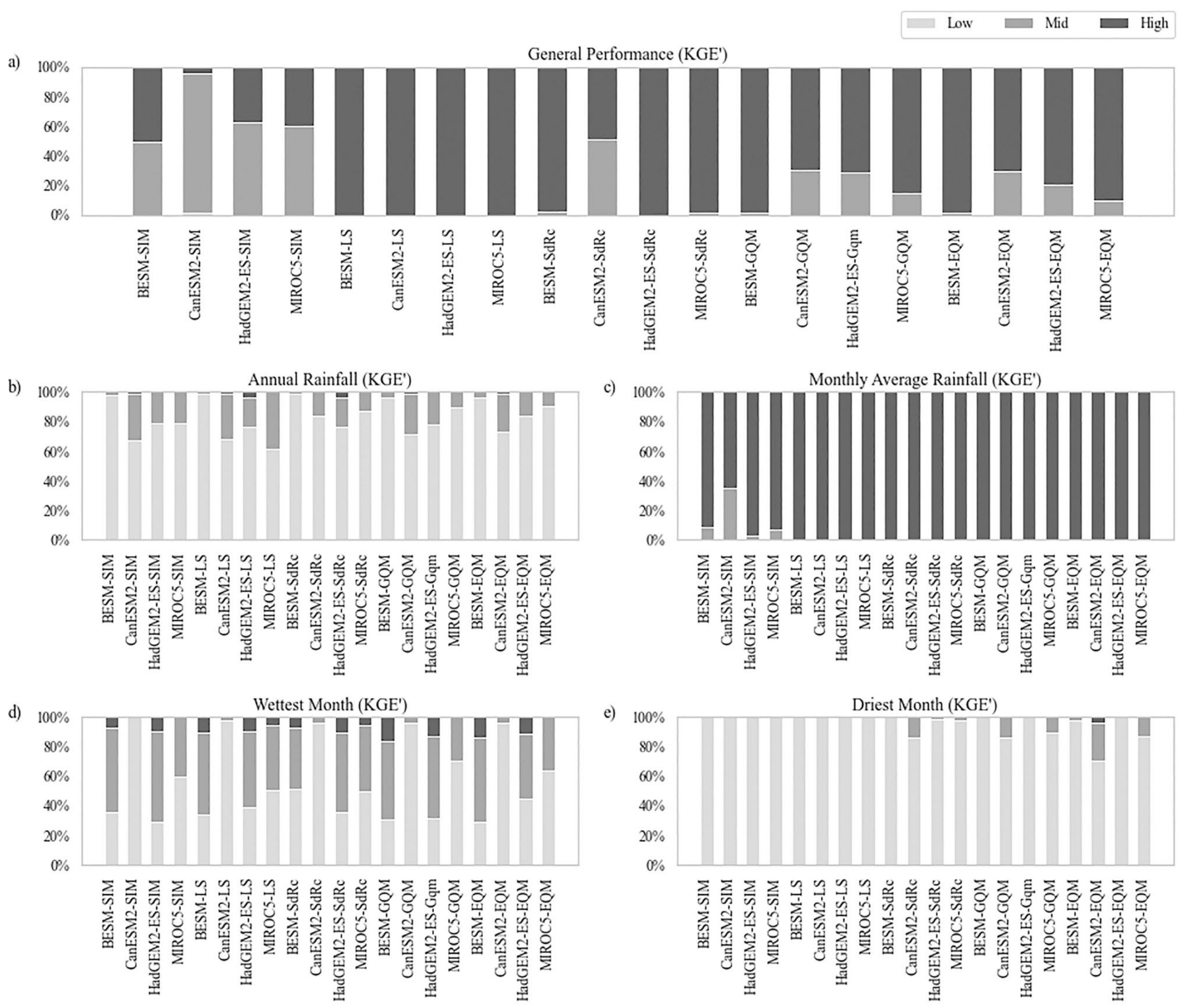

Figure 5 - Percentage distribution of the 85 gauge stations classified with *KGE'* scores. Horizontal axis present, for each RCM, the original data (SIM) and the BC methods: Linear Scaling (LS), Standardized Reconstruction (SdRc), Gamma Quantile Mapping (GQM) and Empirical Quantile Mapping (EQM).

Table 3 - *KGE'* scored sum. Columns present, for each RCM, the original data (SIM) and the BC methods: Linear Scaling (LS), Standardized Reconstruction (SdRc), Gamma Quantile Mapping (GQM) and Empirical Quantile Mapping (EQM). Values in bold highlight the best BC for each Eta model. Minimum value is 0 and optimum value is 850.

Model	SIM	LS	SdRc	GQM	EQM
BESM	354	406	386	415	416
CanESM2	257	370	326	354	370
HadGEM2-ES	371	423	428	404	394
MIROC5	335	420	400	371	382

2017). We found that, in general, all BC methods were able to improve the climate projection model outputs to some extent, although there are clear differences in their performance on each criterion (Teutschbein and Seibert, 2012).

The best BC method for each RCM is shown in bold. For BESM, EQM showed the best results, while for CanESM2 outputs, both LS and EQM presented the highest scores. For HadGEM2-ES outputs, the best BC method was SdRc, and for MIROC5 was LS. HadGEM2-ES model, corrected with SdRc method, gived the best results for the chosen indicators. Regarding the BC methods' complexity, we found that methods with relatively simpler application such as LS and SdRc performed well when compared to the statistical EQM and GQM methods. This reinforces the findings of Shresta *et al.* (2017), who points that not necessarily the most complex BC methods produces the best results in a monthly timestep analysis. CanESM2 model, either corrected or not, was the least representative among the observed data for the study area.

We point that this study considered indicators with equal weights, although the individual indicators assessment discussed before showed that BC methods perform better regarding wet months and monthly average rainfall. Therefore, in applications where accurate representations of total annual precipitation of the behavior of dry months are required, other BC methods should be analyzed, or weight distribution in the MCDA should be considered differently.

At large-scale areas, the most suitable RCM model varies according to the region, as concluded by Almagro *et al.* (2020). Thus, we highlight that the best combination of BC and RCM may vary depending on the chosen location and scale. However, the presented MCDA framework and criteria can be used in other spatial and temporal scales.

3.4. Climate projection

To evaluate the effects of BC when applied to RCM simulated outputs, we used SdRc to correct future HadGEM2-ES rainfall scenarios, which was considered the best performance according to the MCDA framework. Correction was applied under the RCP4.5 and RCP8.5 scenarios, from 2006 to 2098. The bias corrected results were compared to the original simulated (SIM) and evaluated in terms of the monthly average rainfall and the ensemble annual rainfall regarding the 85 rainfall stations.

Figure 6 shows the observed precipitation monthly averages from 1961 to 2005, the original simulated precipitation data (SIM) from 2006 to 2098 monthly averages, and the effects of SdRc. For both RCP scenarios, the simulated average precipitation is lower than the observed average from November to May, and higher from June to September. The average rainfall of bias corrected model projection in July, August and September were significantly lower, with a mean reduction of 56% (from

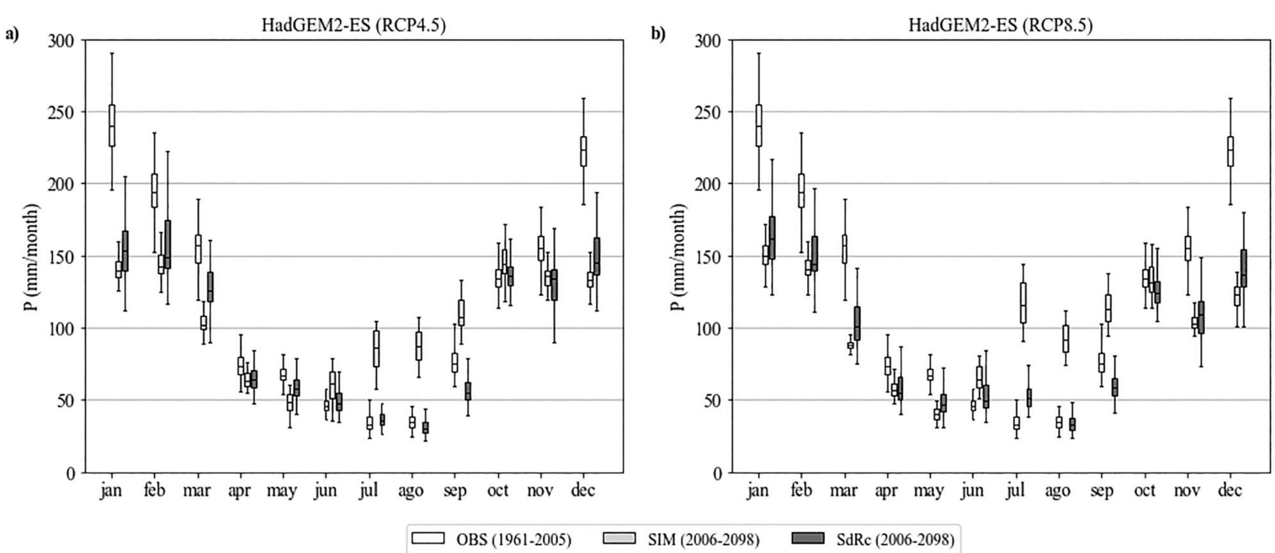

Figure 6 - Monthly average rainfall observed from 1961 to 2005 and projected by HadGEM2-ES from 2006 to 2098 under RCP4.5 (a) and RCP8.5 (b) scenarios. Legend presents observed data (OBS), original simulated data (SIM) and the Standardized Reconstruction (SdRc) method.

81,5 mm to 35,9 mm), 61% (from 85,1 mm to 33,0 mm) and 47% (from 108,1 mm to 57,6 mm) for RCP4.5, respectively. For RCP8.5, the bias corrected model projections showed an average decrease of 54% (from 113,8 mm to 52,7 mm), 60% (from 90,4 mm to 35,8 mm) and 45% (from 111,2 mm to 61,0 mm) for the same months.

Figure 7 shows, for the 85 rainfall stations, the total average annual precipitation results for simulated (SIM) and corrected HadGEM2-ES outputs with SdRc. Stations' minimum and maximum corrected total annual outputs compose the lower and upper boundaries indicated by the shaded area. Results indicated a slight decrease of bias corrected outputs compared with the corresponding original results for both RCP. Overall, SdRc reduced the total average annual precipitation by approximately 6% (from 1246,1 mm to 1169,4 mm) under RCP4.5 and by 8% (from 1217,1 mm to 1119,1 mm) under RCP8.5 scenario.

4. Conclusions

This study presented a multi-criteria decision analysis (MCDA) to evaluate the performance of four bias correction (BC) methods, Linear Scaling (LS), Standardized Reconstruction (SdRc), Gamma Quantile Mapping (GQM) and Empirical Quantile Mapping (EQM), to correct monthly precipitation outputs from climate change projections in the Piracicaba river basin area. RCM data were generated by CPTEC/INPE and derived from Eta RCM nested in four GCM (i) BESM, (ii) CanESM2, (iii) HadGEM2-ES and (iv) MIROC5. The MCDA analysis comprised three criteria: (i) general performance, (ii) seasonal aspects related (represented by the annual rainfall and monthly average rainfall) and (iii) extremes rainfall related (evaluated by the wettest and driest months). Indicators were measured in terms of the modified Kling-Gupta efficiency (*KGE'*) for the validation period (1991-2005).

Regarding the general performance, BC methods improved all RCM outputs, although individual *KGE'* components (σ, β and γ) were affected differently. Linear Scaling (LS) provided a better correlation (r) and variability (γ) fit in comparison to other BC methods. Annual rainfall results, for both corrected and uncorrected outputs, were unable to reproduce adequately the observed data. The monthly average results, however, showed that all BC methods improved the RCM outputs, with LS providing the overall best performance. The wettest and driest months indicators showed small improvements of indicators, although the wettest months resulted in higher *KGE'* than the driest months. The extreme indicators performance can be explained by poor similarity between uncorrected RCM (SIM) outputs with the observed rainfall.

The individual indicators assessment showed that BC methods impacted differently amongst the series statistics. It also indicated that BC methods alone can't guarantee improvements of RCM performance, when the uncorrected data presents poor *KGE'* indicators. Thus, if the RCM application requires a better fit of indicators such as annual rainfall or extremes months analysis, other models must be considered individually. Combining HadGEM2-ES with Standardized Reconstruction (SdRc) BC method presented the best outcome for the study area.

The effects of SdRc on climate change projected scenarios (2006-2098) were significant when evaluated in terms of monthly average rainfall under RCP4.5 and RCP8.5 scenarios. In months of July, August and September, the average precipitation was at least 45% lower than the original RCM results. Regarding annual rainfall, the

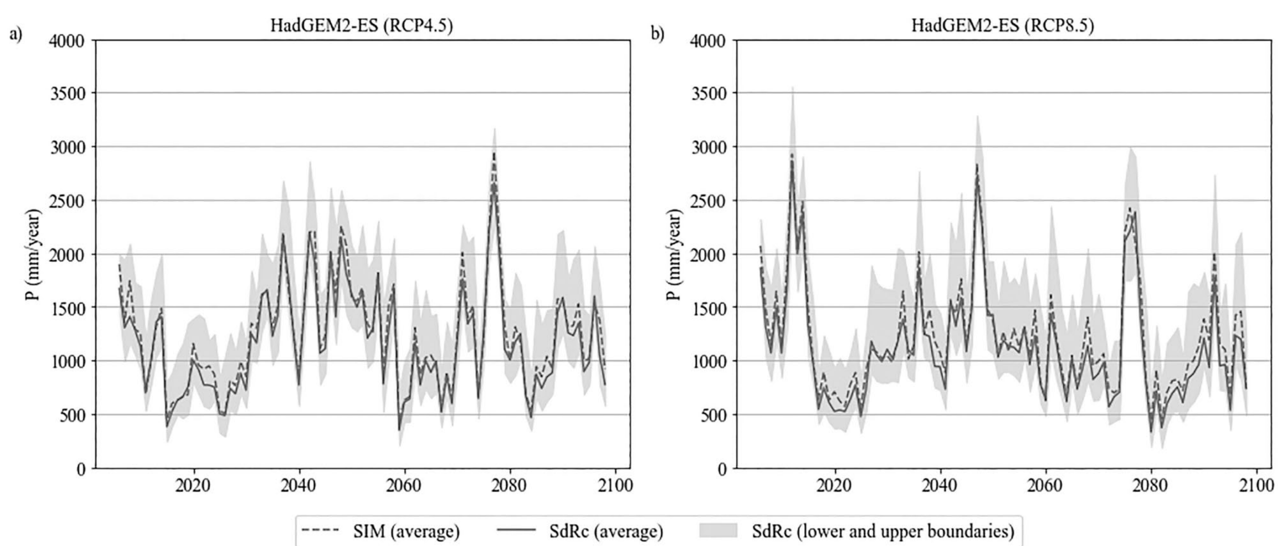

Figure 7 - Annual average rainfall projected by HadGEM2-ES from 2006 to 2098 under RCP4.5 (a) and RCP8.5 (b) scenarios and the corrected average, lower and upper boundaries. Legend presents original simulated data (SIM) and the *Standardized Reconstruction* (SdRc) method.

bias corrected scenarios did not indicate changes in the signal of climate change when compared to the original outputs, but showed an average decrease of 6% for RCP4.5 and by 8% for RCP8.5.

The MCDA proposed in this study provides a useful framework for performance evaluation of both RCM precipitation outputs and BC methods. *KGE'* parameter analysis together with its components provides a broader understanding of the RCM data in relation with observed data. Although this study application is based on a local catchment region and monthly data timestep, the proposed MCDA methodology framework and criteria is suitable for application at other spatial and temporal scales.

References

ACHARYA, N.; CHATTOPADHYAY, S.; MOHANTY, U.C.; DASH, S.K.; SAHOO, L.N. On the bias correction of general circulation model output for Indian summer monsoon. **Meteorological Applications**, v. 20, n. 3, p. 349-356, 2013.

AKHTER, J.; DAS, L.; DEB, A. CMIP5 ensemble-based spatial rainfall projection over homogeneous zones of India. **Climate Dynamics**, v. 49, n. 5-6, p. 1885-1916, 2017.

ALMAGRO, A.; OLIVEIRA, P.T.S.; ROSOLEM, R.; HAGEMANN, S.; NOBRE, C.A. Performance evaluation of Eta/HadGEM2-ES and Eta/MIROC5 precipitaion simulations over Brasil. **Atmospheric Research**, v. 244, p. 105053, 2020.

ARGÜESO, D.; EVANS, J.P.; FITA, L. Precipitation bias correction of very high-resolution regional climate models. **Hydrology and Earth System Sciences**, v. 17, n. 11, p. 4379-4388, 2013.

ARORA, V.K.; SCINOCCA, J.C.; BOER, G.J.; CHRISTIAN, J.R.; DENMAN, K.L. *et al*. Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. **Geophysical Research Letters**, v. 38, n. 5, 2011.

BOÉ, J.; TERRAY, L.; HABETS, F.; MARTIN, E. Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. **International Journal of Climatology**, v. 27, n. 12, p. 1643-1655, 2007.

BAEZ-VILLANUEVA, O.M.; Zambrano-Bigiarini, M.; RIBBE, L.; NAUDITT, A.; GIRALDO-OSORIO, J.D.; THINH, N.X. Temporal and spatial evaluation of satellite rainfall estimates over different regions in Latin-America. **Atmospheric Research**, v. 213, p. 34-50, 2018.

BOZZINI, P.L.; MELLO JUNIOR, A.V. Previsões de Precipitação de Modelos Atmosféricos como Subsídio à Operação de Sistemas de Reservatórios. **Revista Brasileira de Meteorologia**, v. 35, n. 1, p. 99-109, 2020.

CHOU, S.C.; BUSTAMANTE, J.F.; GOMES, J.L. Evaluation of Eta Model seasonal precipitation forecasts over South America. **Nonlinear Processes in Geophysics**, v. 12, n. 4, p. 537-555, 2005.

CHOU, S.C.; LYRA, A.; MOURÃO, C.; DERECZYNSKI, C.P.; PILOTTO, I. *et al.* Assessment of climate change over South America under RCP 4.5 and 8.5 downscaling scenarios. **American Journal of Climate Change**, v. 3, n. 5, p. 512-535, 2014.

CHOU, S.C.; LYRA, A.; MOURÃO, C.; DERECZYNSKI, C.P.; PILOTTO, I. *et al.* Evaluation of the Eta simulations nested in three global climate models. **American Journal of Climate Change**, v. 3, n. 5, p. 438-454, 2014.

CHOU, S.C.; MARENGO, J.A.; LYRA, A.A.; SUEIRO, G.; PESQUERO, J.S. *et al.* Downscaling of South America present climate driven by 4-member HadCM3 runs. **Climate Dynamics**, v. 38, n. 3-4, p. 635-653, 2012.

COLLINS, W.J. *et al.* Development and evaluation of an Earth-System model-HadGEM2. **Geoscientific Model Development Discussions**, v. 4, n. 2, p. 997-1062, 2011.

COMITÊS PCJ, AGÊNCIA DAS BACIAS PCJ. **Relatório Final: Plano de Recursos Hídricos das Bacias Hidrográficas dos Rios Piracicaba, Capivari e Jundiaí, 2020 a 2035**. PCJ, 2020.

EHRET, U.; ZEHE, E.; WULFMEYER, V.; WARRACH-SAGI, K.; LIEBERT, J. Should we apply bias correction to global and regional climate model data? **Hydrology and Earth System Sciences**, v. 16, n. 9, p. 3391-3404, Sept. 2012.

EUM, H.I.; CANNON, A.J.; MURDOCK, T.Q. Intercomparison of multiple statistical downscaling methods: multi-criteria model selection for South Korea. **Stochastic Environmental Research and Risk Assessment**, v. 31, n. 3, p. 683-703, 2017.

GUDMUNDSSON, L.; BREMNES, J.B.; HAUGEN, J.E.; ENGEN-SKAUGEN, T. Downscaling RCM precipitation to the station scale using statistical transformations: a comparison of methods. **Hydrology and Earth System Sciences**, v. 16, n. 9, p. 3383-3390, 2012.

GUIMARÃES, S.O.; COSTA, A.A.; VASCONCELOS JÚNIOR, F.C.; SILVA, E.M.; SALES, D.C. *et al.* Projeções de mudanças climáticas sobre o nordeste Brasileiro dos modelos do CMIP5 e do CORDEX. **Revista Brasileira de Meteorologia**, v. 31, n. 3, p. 337-365, 2016.

GUPTA, H.V.; KLING, H.; YILMAZ, K.K.; MARTINEZ, G.F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. **Journal of Hydrology**, v. 377, n. 1-2, p. 80-91, 2009.

HEMPEL, S.; FRIELER, K.; WARSZAWSKI, L.; SCHEWE, J.; PIONTEK, F. A trend-preserving bias correction: the ISI-MIP approach. **Earth System Dynamics**, v. 4, n. 2, p. 219-236, 2013.

HIEZ, G. L'homogénéité des données pluviométriques. **Cahiers ORSTOM, série Hydrologie**, v. 14, n. 2, p. 129-173, 1977.

IRVING, K.; KUEMMERLEN, M.; KIESEL, J.; KAKOUEI, K.; DOMISCH, S.; JÄHNIG, S.C. A high-resolution streamflow and hydrological metrics dataset for ecological modeling using a regression model. **Scientific data**, v. 5, n. 1, p. 1-14, 2018.

JOHNSON, F.; SHARMA, A. A nesting model for bias correction of variability at multiple time scales in general circulation model precipitation simulations. **Water Resources Research**, v. 48, n. 1, p.1504, 2012.

KLING, H.; FUCHS, M.; PAULIN, M. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. **Journal of Hydrology**, v. 424, p. 264-277, 2012.

LYRA, A.; TAVRES, P.; CHOUS, S.C.; SUEIRO, G.; DEREC-ZYNSKI, C.P. *et al.* Climate change projections over three metropolitan regions in Southeast Brazil using the non-hydrostatic Eta regional climate model at 5-km resolution. **Theoretical and applied climatology**, v. 132, n. 1-2, p. 663-682, 2018.

MOSS, R.H.; EDMONDS, J.A.; HIBBARD, K.A.; MANNING, M.R.; ROSE, S.K. *et al.* The next generation of scenarios for climate change research and assessment. **Nature**, v. 463, n. 7282, p. 747-756, 2010.

NOBRE, P.; SIQUEIRA, L.S.P.; ALMEIDA, R.A.F.; MALA-GUTTI, M.; GIAROLLA, E. *et al.* Climate simulation and change in the Brazilian climate model. **Journal of Climate**, v. 26, n. 17, p. 6716-6732, 2013.

PESQUERO, J.F.; Chou, S.C.; NOBRE, C.F.; MARENGO, J.A. Climate downscaling over South America for 1961-1970 using the Eta Model. **Theoretical and Applied Climatology**, v. 99, n. 1-2, p. 75-93, 2010.

PIANI, C.; HAERTER, J.O.; COPPOLA, E. Statistical bias correction for daily precipitation in regional climate models over Europe. **Theoretical and Applied Climatology**, v. 99, n. 1-2, p. 187-192, 2010.

SHRESTHA, M.; ACHARYA, S.C.; SHRESTHA, P.K. Bias correction of climate models for hydrological modelling - are simple methods still useful? **Meteorological Applications**, v. 24, n. 3, p. 531-539, 2017.

TEUTSCHBEIN, C.; SEIBERT, J. Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. **Journal of Hydrology**, v. 456, p. 12-29, 2012.

TEUTSCHBEIN, C.; SEIBERT, J. Regional climate models for hydrological impact studies at the catchment scale: A review of recent modeling strategies**. Geography Compass**, v. 4, n. 7, p. 834-860, 2010.

TSCHÖKE, G.V.; KRUK, N.S.; DE QUEIROZ, P.I.B.; CHOU, S.C.; DE SOUSA JUNIOR, W.C. Comparison of two bias correction methods for precipitation simulated with a regional climate model. **Theoretical and Applied Climatology**, v. 127, n. 3-4, p. 841-852, 2017.

PASTÉN-ZAPATA, E.; JONES, J.M.; MOGGRIDGE, H.; WID-MANN, M. Evaluation of the performance of Euro-COR-DEX Regional Climate Models for assessing hydrological climate change impacts in Great Britain: A comparison of different spatial resolutions and quantile mapping bias correction methods. **Journal of Hydrology**, v. 584, p. 124653, 2020.

WATANABE, M.; SUZUKI, T.; OISHI, R.; KOMURO, Y.; WATANABE, S. *et al.* Improved climate simulation by MIROC5: mean states, variability, and climate sensitivity. **Journal of Climate**, v. 23, n. 23, p. 6312-6335, 2010.