Statistical and Machine Learning Downscaling Methods to Assess Changes to Rainfall Amounts and Frequency in Climate Change Context- CMIP 6

David A. Jimenez Osorio.[1], Andrea Menapace[2], Ariele Zanfei[3] Eber José de Andrade Pinto[1,4], Bruno Brentan[1]

[1]Hydraulic and Water Resources Department, Federal University of Minas Gerais, 31270-901, MG, Brazil
[2]Faculty of Science and Technologies, Free University of Bozen-Bolzano, 39100 Bolzano, Italy
[3]AIAQUA S.r.l., Via Volta 13/A, Bolzano, Italy.
[4]The Geological Survey of Brazil, 30140-002, MG, Brazil

Correspondence to: David A. Jimenez Osorio (dajimenezoo30@gmail.com)

Abstract. General Circulation Models (GCMs) simulations result on grids ranging from 50 km to 600 km, and, therefore, this coarse spatial resolution requires data processing, whereby the application of downscaling techniques has become a standard procedure. The main approaches employed are Statistical DownScaling (SDS) and Dynamic DownScaling (DDS). The former SDS consists of Linear Methods (LM), Stochastic Weather Generators, and Artificial Intelligence DownScaling techniques (IADS). Being computationally less demanding and highly portable, most studies apply LM, and IADS approaches to develop the downscaling. However, it is needed to evaluate whether these approaches allow obtaining representative, in the development of rainfall frequency analysis (RFA), in the estimative of the total precipitation (TP) and the number of rainy days (RD) both water year and multiannual level, as well as identify whether any of these approaches provide better results for the last generation of GCM's made available for CMIP 6. On this basis and considering only the models with a horizontal resolution of 100 km that participated in the SSP1-2.6 and/or SSP5-8.5 scenarios of CMIP6, the present study aim to evaluate the performance of Delta Method (DM), Quantile Mapping (QM) and Regression Trees (RT) to develop RFA, estimate the TP and RD, based on rainfall series obtained by DownScaling, respect to estimative developed with historical records. The results show that the application of DM, RT and QM does not guarantee a temporal correlation between the TP and RD estimated with DownScaling and historical series, likewise, it is observed that in the estimation of RFA, the application of RT generates better results than QM. Finally, it is evident that not applying any DownScaling technique and applying QM generates similar results.

1 Introduction

The knowledge and characterisation of maximum rainfall patterns are important because these types of events have the potential to generate flooding and destruction of urban infrastructure, leading to large economic losses, social impacts, and sometimes, loss of human life (Eekhout et al., 2018; Nashwan & Shahid, 2022; G. Wang et al., 2020). Historically, extreme rainfall has been characterised principally by the Intensity Duration and Frequency (IDF) relationships, however, in contexts

of climate change, authors such as Fadhel et al., (2017), Shahabul and Elshorbagy (2015) and Waters et al., (2003), highlight the need to develop researches that identify possible changes in these relationships under climate change contexts, since as indicated by Hassanzadeh et, al. (2014), Liu et al. (2020), Norris et al., (2020) and Tabari et al. (2021), the

35 transformation/modification of temperature and relative humidity patterns leads to the intensification of extreme weather events (Roca et al., 2019).

Most of the studies that aim to characterize and identify the impacts of climate change use the results of General Circulation Models (GCMs) simulations because these constitute the most advanced climate simulation tool available (IPCC, 2014). The GCMs have the capability to generate coherent climate estimation both physically and geographically, and like highlighted by

40 Ostad-Ali-Askari et al., (2020), they allow to examine the effect of increasing greenhouse gas emissions on climatic variables. However, due to their low spatial resolution (50-600 km), GCMs are unable to adequately reproduce the climatic variables of small areas such as basins and sub-basins (Ozbuldu & Irvem, 2021), whereby the application of downscaling techniques has become a standard procedure (Olsson et al., 2016; Worku et al., 2021).

Downscaling aims to identify the relationships between climatological variables observed at the local or regional scale and

45 those simulated by GCMs (Jimenez, 2022; Zhang & Li, 2020). Once identified and validated, those relationships are applied to GCM projections with the purpose of predicting future climatological conditions, however, it has been identified that these are not always validated, or in other cases, they are applied even presented a low performance, in this sense, it is necessary to develop studies that validate the performance of DownScaling techniques, especially in the latest generation of MCGs available in CMIP6, since the correct choice of DownSacling can lead to more reliable results.

50

Several downscaling techniques have been proposed in the literature (Xu, 1999). However, in practice, the main approaches employed are Statistical DownScaling techniques (SDS) and Dynamic DownScaling, however, into the SDS we found: Linear Methods (LM), Stochastic Weather Generators (SWG), and Downscaling based on artificial intelligence techniques (IADS).

The relationships of LM are established by linear or multiple regression methods and the predictor-predictand set can be the

55 same variable or different (e.g. both can be daily precipitation, or one can be daily precipitation and other the atmospheric pressure). On the other hand, SWG seeks to generate synthetic series that take into account the alterations produced by climate change. (Semenov & Stratonovitch, 2010; Weschenfelder et al., 2019). At the same time, DDS involves simulating GCMs predictions in high-resolution regional climate models (RCMs) to prognosticate future local climatic conditions (Adachi & Tomita, 2020). Finally, in IADS, predictor-predictand relationships are identified using artificial intelligence algorithms such

60 as Neural Networks (ANNs), Genetic Programming (GP), Vector Support Machine (VSM), among others. (Hassanzadeh et al., 2014; Niazkar et al., 2022; Sachindra, Ahmed, Rashid, et al., 2018).

Since downscaling improves the prediction accuracy of the GCMs (Ozbuldu & Irvem, 2021), several studies have proposed to evaluate the efficiency of downscaling techniques. In the case of precipitation, are highlighted the studies developed by: Hashmi et al. (2011), Li et al. (2010), Mahla et al. (2019), Sachindra et al. (2014), Sachindra et al. (2018), Sachindra et al. (2018), Salehnia et al., (2019) and Wang et al., (2016).

Salehnia et al. (2019) identified that DDS provides better results than SDS in total annual and seasonal precipitation downscaling, pointing out that SDS is computationally simpler than DDS. On the other hand, Mahla et al. (2019) indicated that downscaling of monthly precipitation based on multiple linear regressions showed good results for the study area. On the other side, Hashmi et al. (2011) identified that the PG provide better results for daily precipitation downscaling than ANNs. Finally, Sachindra et al. (2018) recommended using a Regional Vector Machine (RVM) over PG, ANNs and SVM for monthly precipitation downscaling.

Authors such as Hashmi et al. (2011), Hassanzadeh et al. (2014), e Ghasemi Tousi et al. (2021) aimed to identify the changes in IDF relationships in contexts of climate change based on GCMs in the Clutha River basin (New Zealand), in the city of Saskatoon - Canada and Tucson Arizona – USA. Hashmi et al. (2011) employed the Statistical DownScaling Model (SDSM) and the SWG LAR-WG to develop the downscaling, while Hassanzadeh et al. (2014) used the PG and Ghasemi Tousi et al. (2021) employed the CMhyd tool. In the case of Hashmi et al. (2011), models from phase 3 of the coupled model intercomparison project- CMIP3 were employed, while Hassanzadeh et al. (2014) used the third-generation coupled global climate model (CGCM3). Finally, Ghasemi Tousi et al. (2021) employed part of the CMIP6 models. In all cases was verified that the DownScaling techniques and tools employed had a good performance.

Based on the mentioned before, it is possible to perceive that there is no consensus on which downscaling technique produces better results, and as indicated by Nover et al. (2016), it is essential to evaluate several DownScaling techniques and approaches. Being computationally cheaper and simpler, most studies apply LM and artificial intelligence approaches to develop the downscaling, however, it is needed to evaluate whether approaches guarantee representative results for GCMs simulations of the CMIP6, in the development of rainfall frequency analysis (RFA), in the estimative of the total precipitation (TP) and the number of rainy days (RD) both hydrological year and multiannual level, as well as whether any of these approaches provide better results.

In the absence of studies evaluating the efficiency of DownScaling techniques in models with horizontal resolution of 100 km that participated in the simulation of scenarios SSP1-2.6 and/or SSP5-8.5 of CMIP6, the present study aims to evaluate the performance of Delta Method (DM), Quantile Mapping (QM) and Regression Trees (RT) as DownScaling techniques to: develop RFA, estimate the total precipitation and rainy days both water year (October – September) and multiannual level.

In the estimation of RFA, TP and RD, precipitation series obtained through DownScaling were used, and the results were compared with the estimates made from historical records obtained from rain gauge stations, thus configuring an innovative approach, since most studies use the reanalysis dataset as historical data. The comparison allowed to identify if the application of these DownScaling techniques in the CMIP-6 daily precipitation simulations guarantees accurate results, respecting those

95 obtained with historical records, and to determine if there is a more appropriate technique for each case.

In the first stage, each of the downscaling techniques were applied to the simulated daily rainfall for the historical period, then, the total precipitation, maximum rainfall and the number of rainy days per hydrological year were estimated from the DownScalaing series. The frequency analysis was developed with the maximums rainfall series, the results were compared with those obtained with the historical series.

100 In order to facilitate the paper's understanding, the second section presents the study area, the data used, the Downscaling techniques considered, and the efficiency metrics used to evaluate the Downscaling techniques. The third section presents the results and discussion, and finally the fourth section presents the conclusions and final considerations.

2. Data and Methodology

2.1 Study Area and historical rainfall records

105 The study was developed in the Metropolitan Region of Belo Horizonte (RMBH), it is located between latitudes 18.0° and 20.5° south, and longitudes 43.15° and 44.75° east, in the central region of the state of Minas Gerais – Brazil.

The RMBH has 9468 km^2, is characterised by the occurrence of precipitation between the months of October and March, which can reach values higher than 300 mm/month. The RMBH monitoring network counts on more than 120 pluviometric stations, which are distributed throughout the territory (See Figure 1a).

110 The RMBH rainfall records were obtained from the hydrological information system - Hidroweb, of the Brazilian National Water Agency (https://www.snirh.gov.br/hidroweb/serieshistoricas). Once the rainfall information was downloaded, its consistency was verified by building double mass curves with total precipitation by hydrological year, selecting the rainfall stations that presented more than 20 years of consistent records with missing data lower than 10%, emphasising that in no case were the missing records filled in, because this could increase the uncertainties of the results.

115 Double mass curves were processed to perform consistency analysis on the collected data. Stations with distances less than 44 km and a correlation equal to or greater than 0.7 from each reference station were selected to perform this calculation. The analysis allowed identifying that only 32 presented consistent series. Thus, the study was developed from the rainfall information of the 32 stations presented in Figure 1b.

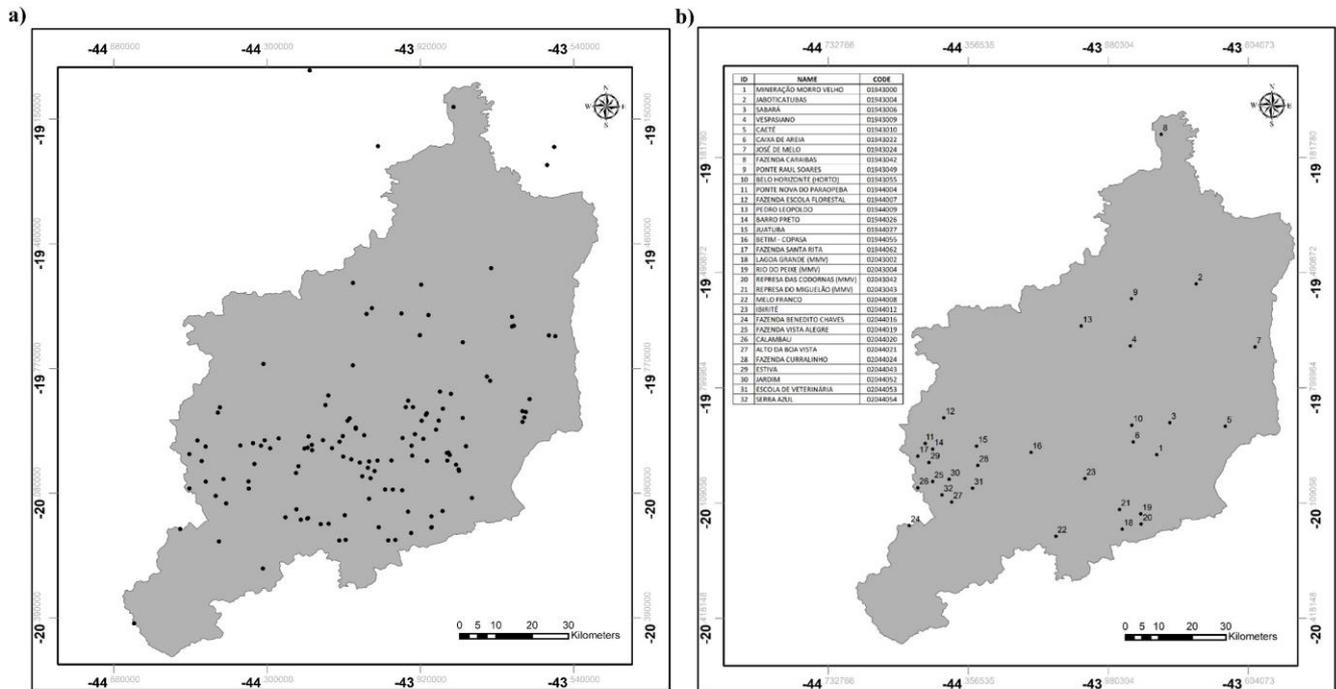



Figure 1 Pluviometric stations of RMBH.
a) Monitoring Network of pluviometry stations. B) Pluviometric stations used in the present study.

2.2 Simulation of rainfall conditions

The GCMs simulations provide the daily historical precipitation downloaded by ref (https://esgf-node,llnl,gov/search/cmip6/) for the emission scenarios SSP1-2.6 and SSP5-8.5. Only the GCMs with 100 km horizontal resolution, which present simulations of precipitation at the daily scale, both in the past (1850-2014) and in the future (2015-2100), were considered, it should be noted that only the simulations of the historical period were used, however, it was sought to ensure that the models used have future projections, so that the results of this study could serve as a basis for future research. On the other hand it is important to note that the historical period of CMIP6 was defined until 2014, and that from that year forward, atmospheric conditions are disturbed according to greenhouse gas concentration projected in each emission scenario.

SSP5-8.5 scenario predicts that the economic and social development of humankind until the end of the 21st century will be governed by: i) high exploitation of resources, ii) intensive use of fossil fuels, iii) high global energy demand. All these factors lead to high greenhouse gas concentrations, leading to a radiative forcing of 8.5 W m^{-2} by the end of the 21st century (Riahi et al., 2016). On the other hand, SSP1-2.6 scenario considers that: i) the world is turning towards sustainability, ii) there is a commitment by nations to reduce social inequalities, iii) consumption is oriented towards low material growth and low resource

and energy consumption. All these factors were combined with a radiative forcing of 2.6 W m^{-2} (Riahi et al., 2016). The simulations contemplated are presented in table 1.

Table 1 Overview of the CMIP6 GCM ensemble used in this study (r –realisation or ensemble member; i –initialisation method; p–physics; f –forcing).

ID	Model	Ensamble	SSP1-2.6 future	SSP5-8.5 future
1	CESM2	r11i1f1p1	X	✓
2	CESM2	r4i1f1p1	✓	X
3	CESM2-WACCM	r1i1f1p1	X	✓
4	CESM2-WACCM	r2i1f1p1	X	X
5	CESM2-WACCM	r3i1f1p1	X	✓
6	CMCC-CM2-SR5	r1i1f1p1	✓	✓
7	CMCC-ESM2	r1i1f1p1	✓	✓
8	EC-Earth3-CC	r1i1f1p1	X	✓
9	EC-Earth3	r101i1p1f1	✓	✓
10	EC-Earth3	r102i1p1f1	✓	✓
11	EC-Earth3	r103i1p1f1	✓	✓
12	EC-Earth3	r104i1p1f1	✓	✓
13	EC-Earth3	r105i1p1f1	✓	✓
14	EC-Earth3	r106i1p1f1	✓	✓
15	EC-Earth3	r107i1p1f1	✓	✓
16	EC-Earth3	r108i1p1f1	✓	✓
17	EC-Earth3	r109i1p1f1	✓	✓
18	EC-Earth3	r110i1p1f1	✓	✓
19	EC-Earth3	r111i1p1f1	✓	✓
20	EC-Earth3	r112i1p1f1	✓	✓
21	EC-Earth3	r113i1p1f1	✓	✓
22	EC-Earth3	r114i1p1f1	✓	✓
23	EC-Earth3	r115i1p1f1	✓	✓
24	EC-Earth3	r116i1p1f1	✓	✓
25	EC-Earth3	r117i1p1f1	✓	✓
26	EC-Earth3	r118i1p1f1	✓	✓
27	EC-Earth3	r119i1p1f1	✓	✓
28	EC-Earth3	r11i1f1p1	✓	✓
29	EC-Earth3	r121i1p1f1	✓	✓
30	EC-Earth3	r122i1p1f1	✓	✓
31	EC-Earth3	r123i1p1f1	✓	✓
32	EC-Earth3	r124i1p1f1	✓	✓
33	EC-Earth3	r125i1p1f1	✓	✓
34	EC-Earth3	r126i1p1f1	✓	✓
35	EC-Earth3	r127i1p1f1	✓	✓
36	EC-Earth3	r128i1p1f1	✓	✓
37	EC-Earth3	r129i1p1f1	✓	✓
38	EC-Earth3	r130i1p1f1	✓	✓
39	EC-Earth3	r131i1p1f1	✓	✓
40	EC-Earth3	r132i1p1f1	✓	✓
41	EC-Earth3	r133i1p1f1	✓	✓
42	EC-Earth3	r134i1p1f1	✓	✓
43	EC-Earth3	r135i1p1f1	✓	✓
44	EC-Earth3	r136i1p1f1	✓	✓

ID	Model	Ensamble	SSP1-2.6 future	SSP5-8.5 future
45	EC-Earth3	r137i1p1f1	✓	✓
46	EC-Earth3	r138i1p1f1	✓	✓
47	EC-Earth3	r139i1p1f1	✓	✓
48	EC-Earth3	r13i1p1f1	✓	✓
49	EC-Earth3	r140i1p1f1	✓	✓
50	EC-Earth3	r141i1p1f1	✓	✓
51	EC-Earth3	r142i1p1f1	✓	✓
52	EC-Earth3	r143i1p1f1	✓	✓
53	EC-Earth3	r144i1p1f1	✓	✓
54	EC-Earth3	r145i1p1f1	✓	✓
55	EC-Earth3	r146i1p1f1	✓	✓
56	EC-Earth3	r147i1p1f1	✓	✓
57	EC-Earth3	r148i1p1f1	✓	✓
58	EC-Earth3	r149i1p1f1	✓	✓
59	EC-Earth3	r150i1p1f1	✓	✓
60	EC-Earth3	r15i1p1f1	✓	✓
61	EC-Earth3	r1i1f1p1	✓	✓
62	EC-Earth3	r3i1f1p1	X	✓
63	EC-Earth3	r4i1f1p1	✓	✓
64	EC-Earth3	r6i1f1p1	✓	✓
65	EC-Earth3-Veg	r1i1f1p1	✓	✓
66	EC-Earth3-Veg	r2i1f1p1	X	✓
67	EC-Earth3-Veg	r3i1f1p1	✓	✓
68	EC-Earth3-Veg	r4i1f1p1	✓	✓
69	EC-Earth3-Veg	r6i1f1p1	✓	✓
70	GFDL-CM4	r1i1f1p1	X	✓
71	GFDL-ESM4	r1i1f1p1	✓	✓
72	INM-CM4-8	r1i1f1p1	✓	✓
73	INM-CM5-0	r1i1f1p1	✓	✓
74	MPI-ESM1-2-HR	r1i1f1p1	✓	✓
75	MPI-ESM1-2-HR	r2i1f1p1	✓	✓
76	MRI-ESM2-0	r1i1f1p1	✓	✓
77	NorESM2-MM	r1i1f1p1	✓	✓
78	TaiESM1-R1	r1i1f1p1	✓	✓

140

2.3 Downscaling

A pixel-station extraction data was made, thus, for each station, it was located and identified the pixel in which it was located. Thereafter, it was extracted from this the daily precipitation simulated, in all cases the temporal uniformity between daily precipitation observed and simulated was guaranteed, for example, if there was a historical record for 01/01/2000, the simulated daily precipitation for that day was extracted. Once it was obtained the simulated series, the Downscaling techniques were applied for each rain gauge.

2.3.1 Delta Method

The Delta method employs the relationships between observed and simulated local climate variables. The method has been employed in various research due to its simplicity and easy implementation (e.g. Golkar Hamzee et al., (2019); Salehnia et al., (2019), Salehnia et al., (2020); Ullah et al., (2018)). The mathematical equation employed by the Delta method is presented below:

$$P_{SD}^{Delta} = P_{Mod,daily} \left(\frac{\bar{P}_{obs}}{\bar{P}_{Mod}} \right)_{Monthly} \tag{2}$$

Where: $\boldsymbol{P_{SD}^{Delta}}$ Precipitation with Downscaling, $P_{Mod,daily}$ precipitation simulated by the GCM, $\bar{P}_{obs}$ average monthly precipitation of the station, $\bar{P}_{Mod}$ average monthly precipitation simulated by GCM.

2.3.2 Quantile Mapping

Quantile Mapping employs the empirical probability distributions of the observed and simulated series. Its use is a little more complex than the DM, but in general terms it is easy to implement. Due to its low complexity, it has been employed in several studies, among which we highlight those developed by Cannon *et al.* (2015), Enayati *et al.* (2021) and Heo *et al.* (2019). The following is a mathematical description of the method:

$$P_{SD}^{QQ} = F_o^{-1}[F_M(P_M)] \tag{1}$$

where P_{SD}^{QQ} is the precipitation with *downscaling,* F_o^{-1} is the inverse empirical probability function of daily precipitation for the historic period, F_M is the empirical probability function of simulated precipitation, and P_M is the simulated precipitation by MCG.

2.3.3 Regression Trees

165 Regression trees is a machine learning technique employed in the generation of predictive models. These are obtained from the recursive partitioning of the sample space and the adjustment of predictive models to each subdivision (Loh, 2011). The development of downscaling processes from regression trees has been used in several studies, such as those developed by Khalid e Sitanggang (2022), Hutengs e Vohland (2016), Im *et al.* (2016) e Pouteau *et al.* (2011).

Thus, the general goal of the technique is to divide the sample space into k units and fit a predictive model to each subspace
170 in such a way that the prediction of the variable of interest, Y, can be performed using a piecewise function of the type:

$$Y = \begin{cases} f_{E_0}(x), & x \in E_0 \\ f_{E_1}(x), & x \in E_1 \\ \quad \dots \\ f_{E_k}(x), & x \in E_k \end{cases} \tag{3}$$

Where: Y is the predicted variable, $f_{E_i}(x)$ is the predictive model of the sample subspace E_i, and x is the predictor variable.

Downscaling from RT requires the use of observed and simulated data, such as records of temperature, atmospheric pressure and precipitation, among others. However, considering that the uncertainties of downscaling increase with the number of
175 predictors, we chose to use only simulated daily precipitation as the predictor variable.

It was decided to train and validate the model based on the observed and simulated precipitation quantiles since it was not evident temporal correlation between the magnitudes of rainfall events; that is, the observed and simulated precipitation heights did not coincide, nor were they close in most cases, thus there were days when the GCMs simulated rainfall, while the historical records showed dry days.

180 In the training stage, 85% of the records were used, while in the validation stage, 15% were used. The training and validation of the models were performed in the Matlab 2020[b] software using the default settings offered by the software.

2.4 Frequency Analysis

Once the downscaling was performed, we proceeded to develop the frequency analysis based on the historical and downscaling series. First, the maximum rainfall per water year was extracted from the historical records and Downscaling series. Then, the
185 stationarity and homogeneity of the maximum series were verified by applying the statistical tests of Spearman (NERC, 1975) and Mann-Whitney (1947). The tests were applied with a significance level of 5%, as presented by Naghettini and Pinto (2007).

Once the homogeneity, randomness and homogeneity of the series had been verified, the frequency analysis was performed, in which the following probability distributions were considered: Exponential, Gamma, Gumbel, GEV, Log-Normal, Pearson III and Log-Pearson III, their parameters being estimated by means of the L-moments (Hosking, 1997). The adherence of the series to the probability distributions was performed by applying the nonparametric Kolmogorov-Smirnov test, with a significance level of 5%. For each station the quantiles of precipitation associated to return times of 2, 5, 10, 15, 15, 30, 35, 45, 50, 60, 70, 80, 90 and 100 years were estimated with the distribution that presents the best fit.

2.5 Comparison between estimates made with historical series and downscaling

Using the metrics of Nash-Sutcliffe (NSE), Kling-Gupta (KGE), root-mean-square error (RMSE) and correlation coefficient of Pearson (R), the results of the number of rainy days and total precipitation per hydrological year, estimated with the DownScaling series, were compared respect to the historical series.

Nash-Sutcliffe (1979) and Gupta et al. (2009) point out that values of NSE and KGE equal to 1 indicate a perfect fit between the observed and simulated values, while in the case of RMSE, it is represented with a value of 0, on the other hand, values of R between 0 and 1 represent a positive correlation, values between -1 and 0 indicate negative correlation, finally, values near of 0 suggests the non-correlation. The equations employed in estimating NSE, KGE, RMSE and R are presented below.

$$NSE = 1 - \frac{\sum_{i=1}^{n}(X_i - X_i')^2}{\sum_{i=1}^{n}(X_i - \bar{X}_i)^2} \qquad (4)$$

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_i'}{\sigma_i} - 1\right)^2 + \left(\frac{\bar{X}_i'}{\bar{X}_i} - 1\right)^2} \qquad (5)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_i - X_i')^2}{n}} \qquad (6)$$

$$R = \frac{n(\sum X_i X_i') - (\sum X_i * \sum X_i')}{\sqrt{[n(\sum X_i^2) - (\sum X_i)^2] * [n(\sum X_i'^2) - (\sum X_i')^2]}} \qquad (7)$$

With X_i and X_i' being respectively the observed and simulated values, $\bar{X}_i$ and $\bar{X}_i'$ are the mean of the observed and simulated values, respectively, n is the number of simulated data, σ_i' being the standard deviation of the simulated values, σ_i is the standard deviation of the observed records, and R is the correlation coefficient between the observed and simulated records.

In addition to the evaluation per water year, the total rainfall and the number of rainy days were compared at the multiannual level by means of the percentage error; the equation used to estimate the percentage error is described below.

$$\% \, Error = \frac{|X_i' - X_i|}{X_i} * 100 \tag{4}$$

where X_i and X_i' are the observed and simulated values, respectively. A percentage error near 0 indicates a perfect match between the observed and simulated values.

3. Results and discussions

3.1 Total precipitation and number of rainy days per hydrological year

210 Considering that each station had 156 analyses, 78 associated to the total precipitation per hydrological year and 78 to the number of rainy days, the mean values of NSE, KGE, RMSE and R were estimated to facilitate the analysis and interpretation of the results presented in Figure 2 and Figure 3.

For total precipitation and the number of rainy days per hydrologic year, the high RMSE, low NSE and KGE, and R less than 0.6 and greater than -0.6 show that there is no temporal correlation between total precipitation and the number of rainy days

215 per hydrologic year, estimated from the precipitation simulated by the MCGs with respect to that estimated from the historical records, even if downscaling with the DM, QM and RT are applied like downscaling techniques (See Figure 2 and Figure 3).

On the other hand, it is observed that the efficiency metrics in the cases in which no downscaling technique was applied are similar to those obtained when the DM is applied, showing that the application of this technique may offer results that are not very representative for the study region, when seeking to estimating the number of rainy days by hydrologic year from the

220 daily precipitation simulated by the GCMs participating of SSP1-2.6 and SSP5-8.5 scenarios.

Compared with DM, the application of the QM and RT allows obtaining a better performance of the efficiency metrics. However, it was shown that the application of these techniques either guarantee the temporal correlation between the series estimated from the daily precipitation simulated by the GCMs respect to those estimated with historical records, which makes it evident that, to estimating the number of rainy days and total precipitation per hydrological year, it will be necessary to

225 evaluate DownScaling approaches different from those presented in this article.

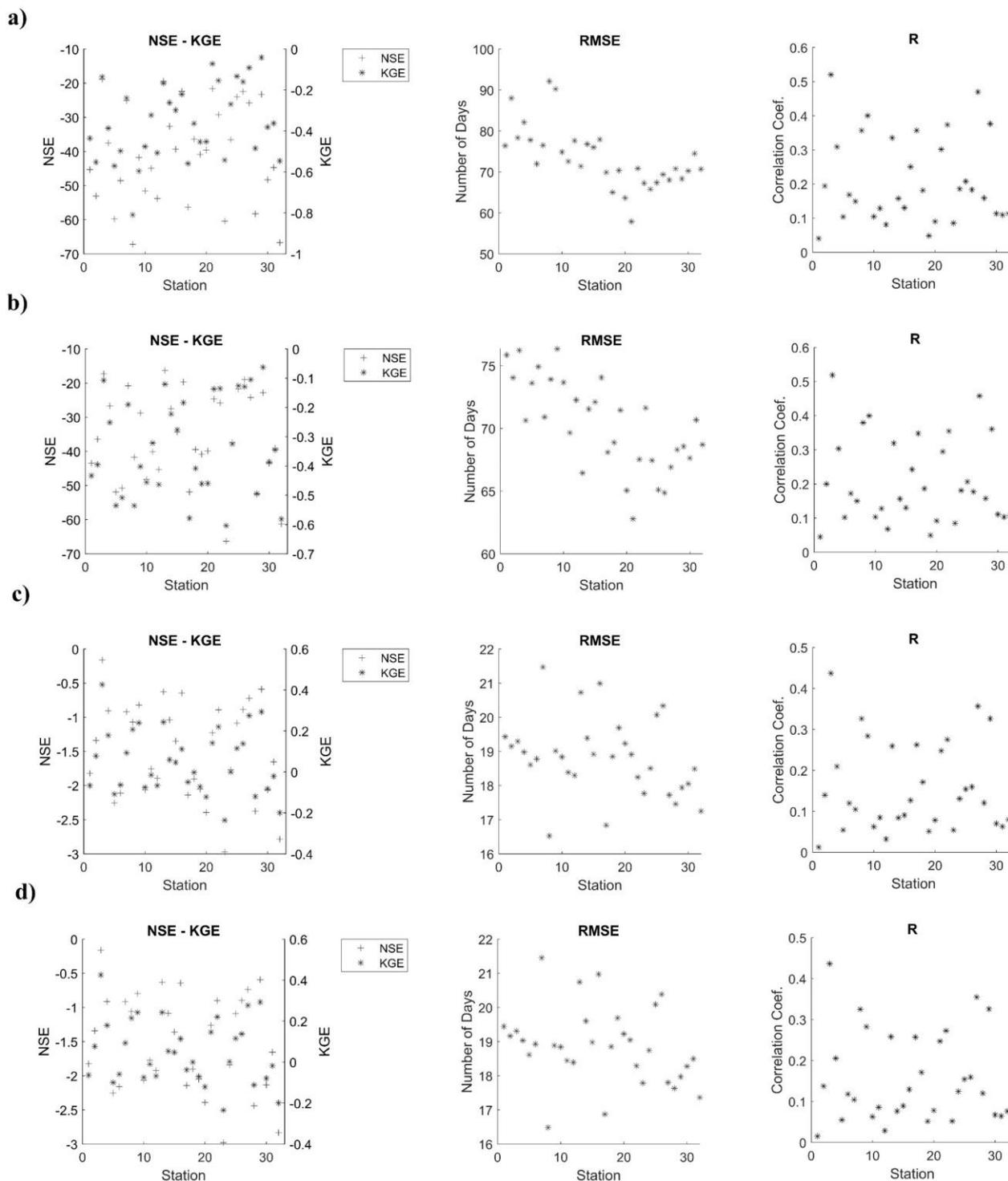



**Figure 2 Average error metrics for the number of rainy days by hydrologic year
. a) Without DownScaling – (WDS), b) DM, c) QM, and d) RT**

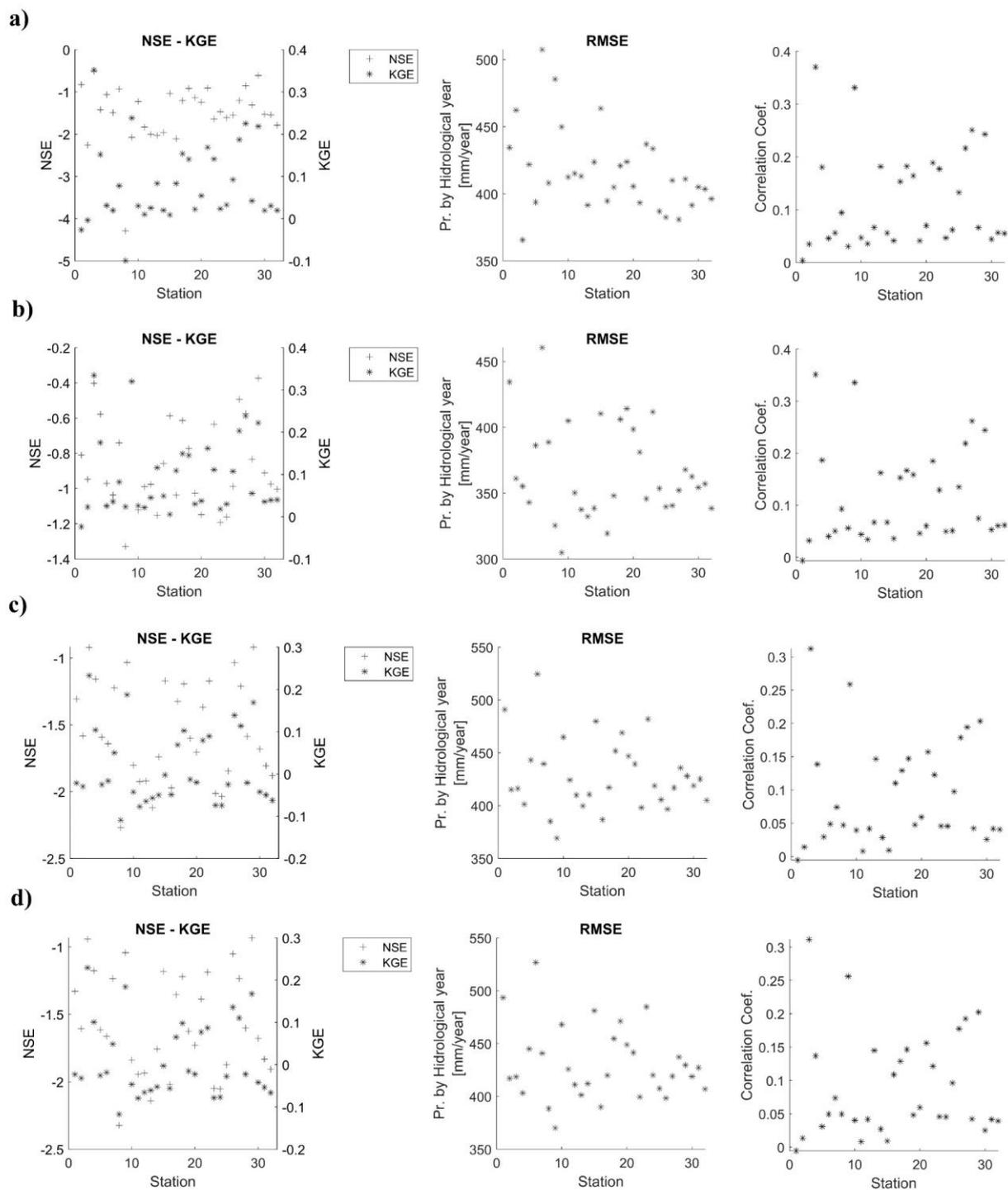



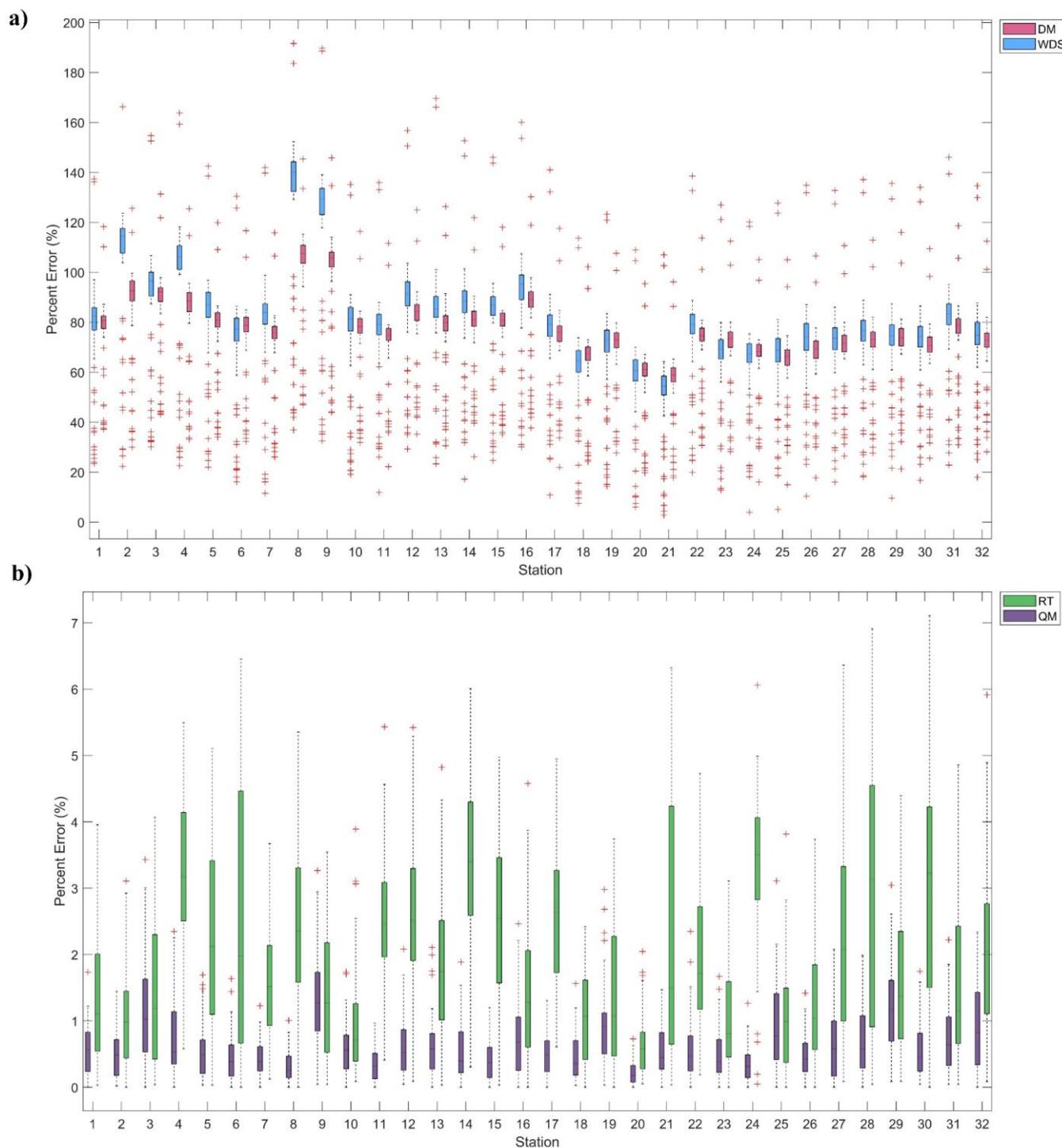
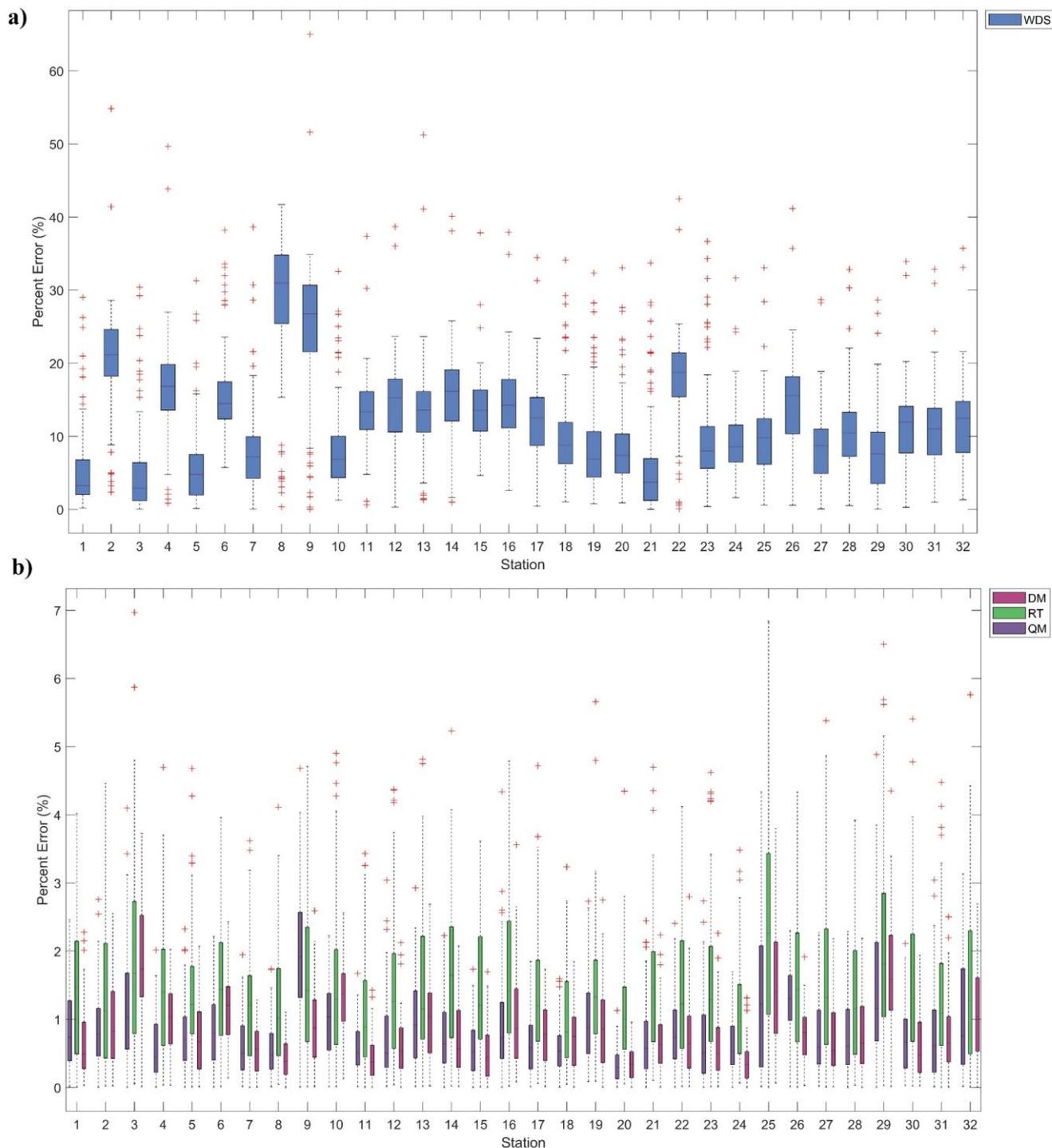
Figure 3 Average Error Metrics for Total Precipitation by hydrologic year.
a) WDS, b) DM, c) QM, and d) RT

The time windows and the totality of the available historical records obtained through the application of DownScaling techniques were considered in the estimation of TP and RD at the multiannual level. For example, if the station under analysis had records from 1970 to 2014, the total rainfall and the number of rainy days estimated at the multiannual level were associated with this period. The number of years considered ranged from 24 to 69 years.

As shown in Figure 4, for the number of rainy days, it is observed that the percentage errors of the DM are similar to those obtained when no downscaling technique is applied. However, it is evident that the application of DM generates a greater dispersion of the percentage errors when compared to those generated by QM and RT. It can be seen that, in general terms, the percentage errors associated with the estimation of the number of rainy days at the multi-year level are greater than 50% when DM is used as a downscaling technique.

Thus, Figure 4 shows that the estimate of the RD at the multi-year level from the rainfall series obtained by applying the QM and RT generate percentage errors lower than 7%. In this sense, it is concluded that for the study region, the application of QM is appropriate in cases where it is intended to estimate the number of rainy days at the multi-year level from precipitation simulations of CMIP6, since to difference of RT, presents a smaller dispersion of the percentage errors.

In cases where no DownScaling techniques is applied, the estimated of total multi-year precipitation shows smaller percentage errors than those obtained in the number of RD, being below 40%. On the other hand, a smaller number of outliers is also observed, which indicates a smaller dispersion of the percentage errors (See Figure 5).

It was observed that estimating the total multi-year precipitation from the series obtained by applying DM, QM, and RT generates percentage errors lower than 7% in most cases. However, the DM presents a smaller dispersion of percentage errors, followed by QM and RT. Thus, it is recommended to employ DM over QM and RT in cases where it is desired to estimate changes in total precipitation at the multi-year level from the MCGs daily precipitation simulations of CMIP6.

255   **Figure 4 Percentage errors of the number of rainy days at the multiannual level for station.**
a) Without DownScaling (WDS) and applying DM like DownScaling technique, b) With the application of QM and RT like
DownScaling techniques.

260   **Figure 5 Percent errors of total precipitation at the multiannual level for each station. a) Without DownScaling (WDS) b) With the application of DM, QM and RT like DownScaling techniques.**

3.1 Frequency Analysis

265 Considering that 238 frequency analyses were performed per station, i.e. one for each model and for each DownScaling technique evaluated, we proceeded to estimate the mean values of NSE, KGE and RMSE to facilitate the analysis and interpretation of the results presented in Figure 6.

The results of stations and models where the null hypotheses of homogeneity and stationarity of the Spearman and Mann-Whitney tests could not be accepted were disregarded. A total of 5 stations and 60 simulations of 25 stations were disregarded (See appendix 1).

270 In the case of the DM, a low correlation is perceived between the quantiles estimated from the historical series and those estimated with the Downscaling series, which are expressed in low NSE, KGE, high RMSE and percentage errors. On the other hand, Figure 7 shows that the performance of the DM was similar to that obtained when no DownScaling technique was applied. In this sense, it is not recommended to use of this method as a downscaling technique in studies that seek to identify changes in the frequency of occurrence of daily rainfall from the simulated rainfall for CMIP6 since the percentage errors are 275 high and the efficiency metrics do not show a good performance.

In contrast, in the case of QM and RT, there is a high correlation between the quantiles estimated with the historical series and those estimated with the DownScaling series, which are expressed in NSE and KGE close to 1, and low RMSE and percentage errors.

When comparing the QM with the RT, one realizes that RT guarantee lower RMSE, percentage errors and higher NSE and 280 KGE for the study area. It is recommended to employ the RT over the quantile mapping in studies that aim to look for changes in the frequency of occurrence of daily precipitation from the simulations of CMIP6.

Unlike the DM, Figure 7 shows for QM and RT the percentage errors do not vary significantly as the return time increases, and for all return times the DM performance was similar to that obtained when no DownScaling technique was applied.

285

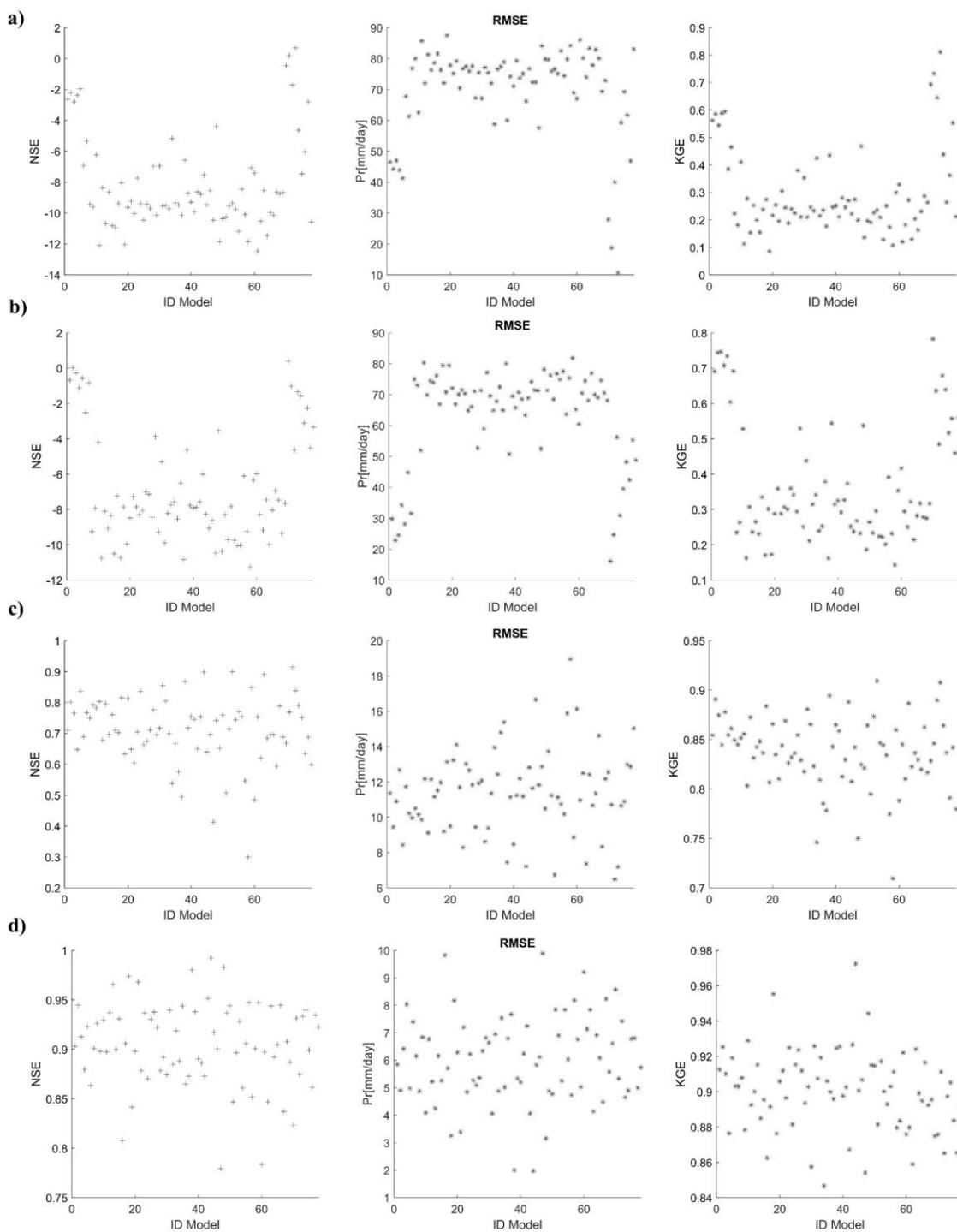



Figure 6 Average of frequency analysis error metrics without application of DownScaling Techniques -WDS (a), and applying DM (b) QM (c) and RT (d) like DownScaling Techniques.

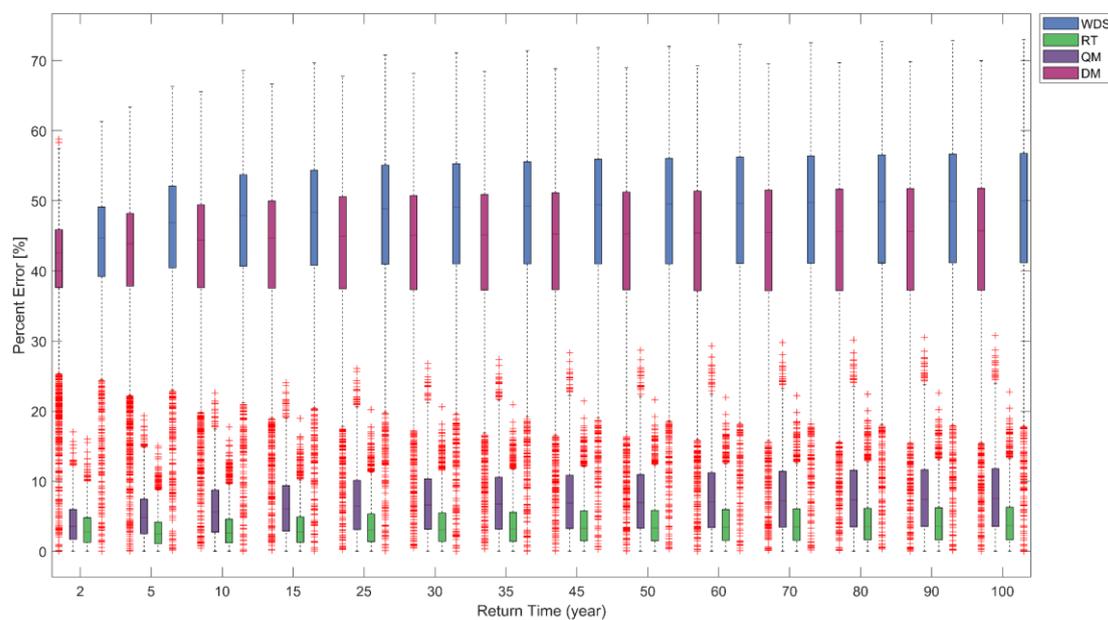



Figure 7 Percent errors from the frequency analysis for the different feedbacks evaluated without DownScaling application -WSD and applying the DownScaling techniques RT, QM and DM.

4. Conclusions

The estimation of total precipitation and the number of rainy days per hydrologic year from the daily precipitation series simulated by the MCGs participants of CMIP6, does not present a temporal correlation with the estimations undertaken from observed series, in these sense, different magnitudes were observed. It was observed, the application of DM, QM and RT as DownScaling techniques is not able to solve that drawback of CMIP6 models.

Nevertheless, the analysis of each precipitation characteristic separately, i.e. total precipitation and the number of rainy days, at the multi-year level from the rainfall simulated by the MCGs shows values close to the historical data in the case of the application of the DownScaling techniques QM and RT. In these two cases, the percentage of errors is lower than 10%, with smaller errors and relative dispersion with QM. In this sense, we recommend the application of QM over RT in studies that aim to identify the possible changes in total precipitation and the number of rainy days at the multi-year level in the study region from the daily precipitation projected by the models participating in the SSP1-2.6 and SSP5-8.5 emission scenarios.

Developing frequency analysis from the daily precipitation simulated by the MCGs allows obtaining quantiles close to those estimated with historical records when QM and RT are applied. However, it was observed that in the cases where RT was applied, the results were more accurate in terms of errors and dispersion. Thus, it is recommended the application of RT over

QM to estimate the possible changes in the frequency of occurrence of daily precipitation from the precipitation simulated by the MCGs participating in the scenarios SSP1-2.6 and SSP5-8.5, in the Metropolitan Region of Belo Horizonte- Brazil.

Except for the estimation of total precipitation at the multi-year level, the application of the DM as a DownScaling technique generates results close to those obtained when no DownScaling technique is applied. Because of this scarce performance, the
310 application of QM and RT over DM in the study region is recommended to identify the possible changes in the frequency of occurrence of daily precipitation, number of rainy days, and total precipitation per hydrologic year by the MCGs participating in the scenarios SSP1-2.6 and SSP5-8.5.

QM and RT performed better in developing frequency analyses. This performance may be associated with the statistical formulation of the process, which implicitly associates the high magnitudes of simulated precipitation with the highest
315 magnitudes of recorded precipitation.

Taking account, the results, it is recommended that the studies that aim to evaluate the changes RFA in the Region of study for the SSP1-2.6 and SSP5-8.5 scenarios, employ the RT like Downscaling technique, since it has shown a greater representativeness respect to the DM and QM.

References

320 Adachi, S. A., & Tomita, H. (2020). Methodology of the Constraint Condition in Dynamical Downscaling for Regional Climate Evaluation: A Review. *Journal of Geophysical Research: Atmospheres*, *125*(11). https://doi.org/10.1029/2019JD032166

Cannon, A. J., Sobie, S. R., & Murdock, T. Q. (2015). Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *Journal of Climate*, *28*(17), 6938–6959.
325 https://doi.org/10.1175/JCLI-D-14-00754.1

Eekhout, J. P. C., Hunink, J. E., Terink, W., & de Vente, J. (2018). Why increased extreme precipitation under climate change negatively affects water security. *Hydrology and Earth System Sciences*, *22*(11), 5935–5946. https://doi.org/10.5194/hess-22-5935-2018

Enayati, M., Bozorg-Haddad, O., Bazrafshan, J., Hejabi, S., & Chu, X. (2021). Bias correction capabilities of quantile mapping
330 methods for rainfall and temperature variables. *Journal of Water and Climate Change*, *12*(2), 401–419. https://doi.org/10.2166/wcc.2020.261

Fadhel, S., Rico-Ramirez, M. A., & Han, D. (2017). Uncertainty of Intensity–Duration–Frequency (IDF) curves due to varied climate baseline periods. *Journal of Hydrology*, *547*, 600–612. https://doi.org/10.1016/j.jhydrol.2017.02.013

Ghasemi Tousi, E., O'Brien, W., Doulabian, S., & Shadmehri Toosi, A. (2021). Climate changes impact on stormwater infrastructure design in Tucson Arizona. *Sustainable Cities and Society*, *72*, 103014. https://doi.org/10.1016/j.scs.2021.103014

Golkar Hamzee Yazd, H., Salehnia, N., Kolsoumi, S., & Hoogenboom, G. (2019). Prediction of climate variables by comparing the knearest neighbor method and MIROC5 outputs in an arid environment. *Climate Research*, *77*(2), 99–114. https://doi.org/10.3354/cr01545

Hashmi, M. Z., Shamseldin, A. Y., & Melville, B. W. (2011). Statistical downscaling of watershed precipitation using Gene Expression Programming (GEP). *Environmental Modelling & Software*, *26*(12), 1639–1646. https://doi.org/10.1016/j.envsoft.2011.07.007

Hassanzadeh, E., Nazemi, A., & Elshorbagy, A. (2014). Quantile-Based Downscaling of Precipitation Using Genetic Programming: Application to IDF Curves in Saskatoon. *Journal of Hydrologic Engineering*, *19*(5), 943–955. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000854

Heo, J.-H., Ahn, H., Shin, J.-Y., Kjeldsen, T. R., & Jeong, C. (2019). Probability Distributions for a Quantile Mapping Technique for a Bias Correction of Precipitation Data: A Case Study to Precipitation Data Under Climate Change. *Water*, *11*(7), 1475. https://doi.org/10.3390/w11071475

HOSKING, J. R. M. e WALLIS, J. R Regional Frequency Analysis - An Approach Based on L-Moments, 224 p. Cambridge University Press, Cambridge, Reino Unido, 1997

Hutengs, C., & Vohland, M. (2016). Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sensing of Environment*, *178*, 127–141. https://doi.org/10.1016/j.rse.2016.03.006

Im, J., Park, S., Rhee, J., Baik, J., & Choi, M. (2016). Downscaling of AMSR-E soil moisture with MODIS products using machine learning approaches. *Environmental Earth Sciences*, *75*(15), 1120. https://doi.org/10.1007/s12665-016-5917-6

IPCC, I. P. on C. C. (2014). *Climate Change 2014 Mitigation of Climate Change: Working Group III Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. https://www.ipcc.ch/site/assets/uploads/2018/02/ipcc_wg3_ar5_frontmatter.pdf

Jimenez, D. A. (2022). *Avaliação das alterações nas frequências de ocorrência das precipitações diárias máximas para a região Metropolitana de Belo Horizonte considerando diferentes cenários de climas futuros.* [Universidade Federal de Minas Gerais - UFMG]. https://repositorio.ufmg.br/handle/1843/46268

Khalid, I. A., & Sitanggang, I. S. (2022). *Machine Learning-Based Spatial Downscaling on Precipitation Satellite Data in Riau Province, Indonesia*. 10.

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. *Journal of Geophysical Research*, *115*(D10), D10101. https://doi.org/10.1029/2009JD012882

Liu, W., Bailey, R. T., Andersen, H. E., Jeppesen, E., Nielsen, A., Peng, K., Molina-Navarro, E., Park, S., Thodsen, H., & Trolle, D. (2020). Quantifying the effects of climate change on hydrological regime and stream biota in a groundwater-dominated catchment: A modelling approach combining SWAT-MODFLOW with flow-biota empirical models. *Science of The Total Environment*, *745*, 140933. https://doi.org/10.1016/j.scitotenv.2020.140933

Loh, W. (2011). Classification and regression trees. *WIREs Data Mining and Knowledge Discovery*, *1*(1), 14–23. https://doi.org/10.1002/widm.8

Mahla, P., Lohani, A. K., Chandola, V. K., Thakur, A., Mishra, C. D., & Singh, A. (2019). Downscaling Of Precipitation Using Statistical Downscaling Model and Multiple Linear Regression Over Rajasthan State. *Current World Environment*, *14*(1), 68–98. https://doi.org/10.12944/CWE.14.1.09

Mann, H. B., & Whitney, D. R. (1947). On the test of wheter one of random variables is stochastically larger than the other. *Annals of Mathematical Statistics*.

Naghettini, M., & Pinto, É. J. (2007). *Hidrologia Estatística*. CPRM.

Nashwan, M. S., & Shahid, S. (2022). Future precipitation changes in Egypt under the 1.5 and 2.0 °C global warming goals
380   using CMIP6 multimodel ensemble. *Atmospheric Research*, *265*, 105908. https://doi.org/10.1016/j.atmosres.2021.105908

NERC, N. E. R. C. (1975). *Flood Studies Report*.

Niazkar, M., Goodarzi, M. R., Fatehifar, A., & Abedi, M. J. (2022). Machine learning-based downscaling: Application of multi-gene genetic programming for downscaling daily temperature at Dogonbadan, Iran, under CMIP6 scenarios.
385   *Theoretical and Applied Climatology*. https://doi.org/10.1007/s00704-022-04274-3

Norris, J., Chen, G., & Li, C. (2020). Dynamic Amplification of Subtropical Extreme Precipitation in a Warming Climate. *Geophysical Research Letters*, *47*(14). https://doi.org/10.1029/2020GL087200

Nover, D. M., Witt, J. W., Butcher, J. B., Johnson, T. E., & Weaver, C. P. (2016). The Effects of Downscaling Method on the Variability of Simulated Watershed Response to Climate Change in Five U.S. Basins. *Earth Interactions*, *20*(11), 1–
390   27. https://doi.org/10.1175/EI-D-15-0024.1

Olsson, J., Arheimer, B., Borris, M., Donnelly, C., Foster, K., Nikulin, G., Persson, M., Perttu, A.-M., Uvo, C., Viklander, M., & Yang, W. (2016). Hydrological Climate Change Impact Assessment at Small and Large Scales: Key Messages from Recent Progress in Sweden. *Climate*, *4*(3), 39. https://doi.org/10.3390/cli4030039

Ostad-Ali-Askari, K., Ghorbanizadeh Kharazi, H., Shayannejad, M., & Zareian, M. J. (2020). Effect of Climate Change on
395   Precipitation Patterns in an Arid Region Using GCM Models: Case Study of Isfahan-Borkhar Plain. *Natural Hazards Review*, *21*(2), 04020006. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000367

Ozbuldu, M., & Irvem, A. (2021). Evaluating the effect of the statistical downscaling method on monthly precipitation estimates of global climate models. *Global NEST Journal*, *23*(2), 232–240. https://doi.org/10.30955/gnj.003458

Pouteau, R., Rambal, S., Ratte, J.-P., Gogé, F., Joffre, R., & Winkel, T. (2011). Downscaling MODIS-derived maps using GIS
400   and boosted regression trees: The case of frost occurrence over the arid Andean highlands of Bolivia. *Remote Sensing of Environment*, *115*(1), 117–129. https://doi.org/10.1016/j.rse.2010.08.011

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., Kc, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., … Tavoni,

M. (2016). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, *42*, 153–168. https://doi.org/10.1016/j.gloenvcha.2016.05.009

Roca, V., B., Beltrán, S. M., & Gómez, H. R. (2019). Cambio climático y salud. *Revista Clínica Española*, *219*(5), 260–265. https://doi.org/10.1016/j.rce.2019.01.004

Sachindra, D. A., Ahmed, K., Rashid, Md. M., Shahid, S., & Perera, B. J. C. (2018). Statistical downscaling of precipitation using machine learning techniques. *Atmospheric Research*, *212*, 240–258. https://doi.org/10.1016/j.atmosres.2018.05.022

Sachindra, D. A., Ahmed, K., Shahid, S., & Perera, B. J. C. (2018). Cautionary note on the use of genetic programming in statistical downscaling. *International Journal of Climatology*, *38*(8), 3449–3465. https://doi.org/10.1002/joc.5508

Sachindra, D. A., Huang, F., Barton, A., & Perera, B. J. C. (2014). Statistical downscaling of general circulation model outputs to precipitation—part 2: Bias-correction and future projections. *International Journal of Climatology*, *34*(11), 3282–3303. https://doi.org/10.1002/joc.3915

Salehnia, N., Hosseini, F., Farid, A., Kolsoumi, S., Zarrin, A., & Hasheminia, M. (2019). Comparing the Performance of Dynamical and Statistical Downscaling on Historical Run Precipitation Data over a Semi-Arid Region. *Asia-Pacific Journal of Atmospheric Sciences*, *55*(4), 737–749. https://doi.org/10.1007/s13143-019-00112-1

Salehnia, N., Salehnia, N., Saradari Torshizi, A., & Kolsoumi, S. (2020). Rainfed wheat (Triticum aestivum L.) yield prediction using economical, meteorological, and drought indicators through pooled panel data and statistical downscaling. *Ecological Indicators*, *111*, 105991. https://doi.org/10.1016/j.ecolind.2019.105991

Semenov, M., & Stratonovitch, P. (2010). Use of multi-model ensembles from global climate models for assessment of climate change impacts. *Climate Research*, *41*, 1–14. https://doi.org/10.3354/cr00836

Shahabul Alam, Md., & Elshorbagy, A. (2015). Quantification of the climate change-induced variations in Intensity–Duration–Frequency curves in the Canadian Prairies. *Journal of Hydrology*, *527*, 990–1005. https://doi.org/10.1016/j.jhydrol.2015.05.059

Tabari, H., Paz, S. M., Buekenhout, D., & Willems, P. (2021). Comparison of statistical downscaling methods for climate change impact analysis on precipitation-driven drought. *Hydrology and Earth System Sciences*, *25*(6), 3493–3517.

430   https://doi.org/10.5194/hess-25-3493-2021

Ullah, A., Salehnia, N., Kolsoumi, S., Ahmad, A., & Khaliq, T. (2018). Prediction of effective climate change indicators using statistical downscaling approach and impact assessment on pearl millet (Pennisetum glaucum L.) yield through Genetic Algorithm in Punjab, Pakistan—ScienceDirect. *Ecological Indicators*, 569–576.

Wang, G., Zhang, Q., Yu, H., Shen, Z., & Sun, P. (2020). Double increase in precipitation extremes across China in a 1.5

435   °C/2.0 °C warmer climate. *Science of The Total Environment*, *746*, 140807. https://doi.org/10.1016/j.scitotenv.2020.140807

Wang, L., Ranasinghe, R., Maskey, S., van Gelder, P. H. A. J. M., & Vrijling, J. K. (2016). Comparison of empirical statistical methods for downscaling daily climate projections from CMIP5 GCM's: A case study of the Huai River Basin, China: COMPARING EMPIRICAL STAISTITCAL DOWNSCALING METHODS. *International Journal of Climatology*,

440   *36*(1), 145–164. https://doi.org/10.1002/joc.4334

Waters, D., Watt, W. E., Marsalek, J., & Anderson, B. C. (2003). Adaptation of a Storm Drainage System to Accommodate Increased Rainfall Resulting from Climate Change. *Journal of Environmental Planning and Management*, *46*(5), 755–770. https://doi.org/10.1080/0964056032000138472

Weschenfelder, A. B., Klering, E. V., Alves, R. de C. M., & Pinto, É. J. de A. (2019). Geração de Curvas IDFs para Cenários

445   Projetados na Cidade de Porto Alegre/RS. *Revista Brasileira de Meteorologia*, *34*(2), 201–216. https://doi.org/10.1590/0102-77863340026

Worku, G., Teferi, E., Bantider, A., & Dile, Y. T. (2021). Modelling hydrological processes under climate change scenarios in the Jemma sub-basin of upper Blue Nile Basin, Ethiopia. *Climate Risk Management*, *31*, 100272. https://doi.org/10.1016/j.crm.2021.100272

450 Xu, C. (1999). From GCM's to river flow: A review of downscaling methods and hydrologic modelling approaches. *Progress in Physical Geography: Earth and Environment*, *23*(2), 229–249. https://doi.org/10.1177/030913339902300204

Zhang, Z., & Li, J. (2020). Big climate data. Em *Big Data Mining for Climate Change* (p. 1–18). Elsevier. https://doi.org/10.1016/B978-0-12-818703-6.00006-4

455

Appendix 1

Simulations disregarded per station

Model	Ensamble	Station
CESM2_R4	r4i1f1p1	2044024
CESM2_WACCM_R2	r2i1f1p1	1943004, 1943009, 1943022, 1943024, 1943042, 1943049, 1943055, 1944027, 2043002, 2043004, 2044008
CESM2_WACCM_R3	r3i1f1p1	1943009, 1943022, 1943024, 1943055, 1944007, 1944062, 2043042, 2044024
CMCC_CM2_SR5_R1	r1i1f1p1	1944007, 1944055, 2044008
CMCC_ESM2_R1	r1i1f1p1	2044024, 2044043, 2044053
EC_Earth3_R1	r1i1f1p1	1944004, 1944027, 2044016
EC_Earth3_R101	r101i1p1f1	1943009, 1943010, 1943022, 1943024, 2043002, 2043042, 2044016
EC_Earth3_R102	r102i1p1f1	1943004, 1943049
EC_Earth3_R103	r103i1p1f1	1943010
EC_Earth3_R104	r104i1p1f1	1944007, 1944026, 1944055, 1944062, 2043002, 2043004, 2043043, 2044008, 2044012, 2044024, 2044043, 2044053, 2044054
EC_Earth3_R105	r105i1p1f1	1944026, 2044024, 2044043
EC_Earth3_R106	r106i1p1f1	1943004, 1943009, 1943022, 1943024, 1943042, 1943049, 1943055, 1944026, 1944027, 1944055, 2043002, 2043004, 2043042, 2044012, 2044016, 2044024, 2044043, 2044054
EC_Earth3_R109	r109i1p1f1	1943004, 1943009, 1943010, 1943022, 1943024, 1943042, 1943049, 1943055, 1944004, 1944027, 2043002, 2043004, 2043042, 2044008, 2044016
EC_Earth3_R11	r11i1f1p1	1944007, 1944026, 1944055, 2043043, 2044008, 2044012, 2044024, 2044043, 2044053, 2044054
EC_Earth3_R110	r110i1p1f1	2044043
EC_Earth3_R111	r111i1p1f1	1943042, 1943049
EC_Earth3_R112	r112i1p1f1	1943004, 1943009, 1943024, 1944026, 1944062, 2043043, 2044012, 2044024, 2044043
EC_Earth3_R114	r114i1p1f1	1943049, 2043042
EC_Earth3_R115	r115i1p1f1	1943010, 1943022, 1944055, 2043004, 2043042, 2044008
EC_Earth3_R116	r116i1p1f1	1943042, 1943049
EC_Earth3_R118	r118i1p1f1	1943004, 1943022, 1943042, 1943049, 1944004, 2043002, 2044016
EC_Earth3_R119	r119i1p1f1	1944007, 1944026, 2043043, 2044012, 2044053, 2044054
EC_Earth3_R122	r122i1p1f1	1944062

Model	Ensamble	Station
EC_Earth3_R123	r123i1p1f1	1943009, 1943042
EC_Earth3_R124	r124i1p1f1	1943010, 1943022, 1943042, 1943055, 2043002, 2043004, 2043042, 2043043, 2044008, 2044043
EC_Earth3_R125	r125i1p1f1	1943004, 1943009, 1943010, 1943022, 1943024, 1943042, 1943049, 1943055, 1944004, 1944027, 2043002, 2043004, 2043042, 2043043, 2044008, 2044016
EC_Earth3_R127	r127i1p1f1	1944062, 2044043
EC_Earth3_R128	r128i1p1f1	1943004, 1944007, 1944062, 2043043, 2044012, 2044024, 2044043, 2044053, 2044054
EC_Earth3_R129	r129i1p1f1	1943004, 1943009, 1943010, 1943022, 1943024, 1943042, 1943049, 1943055, 2043002, 2043004, 2043042
EC_Earth3_R130	r130i1p1f1	1943010, 1943022, 1943055, 2043002, 2043004, 2043042
EC_Earth3_R131	r131i1p1f1	1943024
EC_Earth3_R133	r133i1p1f1	1943010, 1943022, 1944004, 1944027, 2044008
EC_Earth3_R134	r134i1p1f1	1943009, 1943010, 1943022, 1943024, 1943042, 1943049, 1943055, 1944004, 1944007, 1944026, 1944027, 1944055, 1944062, 2043002, 2043004, 2043042, 2043043, 2044008, 2044012, 2044016, 2044024, 2044043, 2044053, 2044054
EC_Earth3_R136	r136i1p1f1	2043004
EC_Earth3_R137	r137i1p1f1	1943004, 1943009, 1943010, 1943022, 1943042, 1943049, 1943055, 2043002, 2043004, 2043042
EC_Earth3_R138	r138i1p1f1	1943004, 1943009, 1943010, 1943022, 1943024, 1943042, 1943049, 1943055, 1944004, 1944027, 2043002, 2043004, 2043042, 2044008, 2044016, 2044043
EC_Earth3_R141	r141i1p1f1	1943049
EC_Earth3_R142	r142i1p1f1	1943042
EC_Earth3_R143	r143i1p1f1	1943024
EC_Earth3_R144	r144i1p1f1	1943009, 1943022, 1943055, 1944062, 2043002, 2043042, 2044043
EC_Earth3_R145	r145i1p1f1	1943004
EC_Earth3_R146	r146i1p1f1	1943010, 1944055, 1944062, 2043043, 2044024, 2044043
EC_Earth3_R147	r147i1p1f1	1943049, 2044043
EC_Earth3_R149	r149i1p1f1	1943022, 1943042, 1944004, 1944027, 2044008, 2044016
EC_Earth3_R15	r15i1p1f1	1943004, 1943022, 1943042, 1943049, 1943055, 1944055, 2043042, 2044043

Model	Ensamble	Station
EC_Earth3_R150	r150i1p1f1	01943009, 01943024, 01944007, 01944026, 01944055, 01944062, 02043043, 02044008, 02044012, 02044053, 02044054
EC_Earth3_R4	r4i1f1p1	01943009, 01943024
EC_Earth3_Veg_R1	r1i1f1p1	01943022, 01943042, 01943049, 01943055, 01944026, 02043002, 02043043, 02044024, 02044043, 02044054
EC_Earth3_Veg_R2	r2i1f1p1	01944026, 01944055, 02044024, 02044043, 02044054
EC_Earth3_Veg_R3	r3i1f1p1	01944004, 01944027, 02044008, 02044016, 02044024
EC_Earth3_Veg_R4	r4i1f1p1	01943009, 01943010, 01943022, 01943024, 01943055, 02043002, 02043004, 02043042
EC_Earth3_Veg_R6	r6i1f1p1	02043004
GFDL_CM4_R1	r1i1f1p1	02043002
GFDL_ESM4_R1	r1i1f1p1	01943024
INM_CM5_0_R1	r1i1f1p1	01943010, 01943022, 01943042, 01943055, 01944007, 01944026, 01944027, 01944055, 01944062, 02043002, 02043004, 02043042, 02043043, 02044008, 02044012, 02044016, 02044024, 02044043, 02044053, 02044054
MPI_ESM1_2_HR_R1	r1i1f1p1	01944026, 01944055, 02044024, 02044043
MPI_ESM1_2_HR_R2	r2i1f1p1	01943004, 01943024, 01943042
MRI_ESM2_0_R1	r1i1f1p1	01943010, 01943022, 01943049, 01943055, 02043002, 02043042, 02044008
NorESM2_MM_R1	r1i1f1p1	01944062
TaiESM1_R1	r1i1f1p1	01943004, 01943009, 01943024, 01943042, 01943049, 01944004, 02044016

460 **Simulations disregarded in the frequency analysis**

ID	Name	Station Code
1	Mineração Morro Velho	01943000
3	Sabará	01943006
4	Vespasiano	01944009
25	Fazenda Vista Alegre	02044019
26	Calambau	02044020
27	Alto da Boa Vista	02044021
30	Jardim	02044052