

Research Article

Evaluation of Seven Gap-Filling Techniques for Daily Station-Based Rainfall Datasets in South Ethiopia

Alef Chasho^{1,2}, Bobe Bedadi,¹ Tesfaye Lemma,¹ Tamado Tana,³ Tilahun Hordofa,⁴ and Bisrat Elias⁵

¹*Africa Center of Excellence for Climate-SABC, Haramaya University, P.O. Box 138, Haramaya, Ethiopia*

²*Department of Environmental Science, Wolaita Sodo University, P.O. Box 138, Wolaita Sodo, Ethiopia*

³*Department of Crop Production, Faculty of Agriculture, University of Eswatini, P.O. M205, Luyengo, Eswatini*

⁴*Ethiopia Institute of Agricultural Research, Melkasa Research Center, P.O. Box. 436, Adama, Ethiopia*

⁵*Faculty of Meteorology and Hydrology, Arba Minch University, P.O. Box. 21, Arba Minch, Ethiopia*

Correspondence should be addressed to Alefu Chinasho; chinalefu11@gmail.com

Received 5 June 2021; Revised 21 July 2021; Accepted 13 August 2021; Published 19 August 2021

Academic Editor: Stefano Federico

Copyright © 2021 Alefu Chinasho et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Meteorological stations, mainly located in developing countries, have gigantic missing values in the climate dataset (rainfall and temperature). Ignoring the missing values from analyses has been used as a technique to manage it. However, it leads to partial and biased results in data analyses. Instead, filling the data gaps using the reference datasets is a better and widely used approach. Thus, this study was initiated to evaluate the seven gap-filling techniques in daily rainfall datasets in five meteorological stations of Wolaita Zone and the surroundings in South Ethiopia. The considered gap-filling techniques in this study were simple arithmetic means (SAM), normal ratio method (NRM), correlation coefficient weighing (CCW), inverse distance weighting (IDW), multiple linear regression (MLR), empirical quantile mapping (EQM), and empirical quantile mapping plus (EQM⁺). The techniques were preferred because of their computational simplicity and appreciable accuracies. Their performance was evaluated against mean absolute error (MAE), root mean square error (RMSE), skill scores (SS), and Pearson's correlation coefficients (*R*). The results indicated that MLR outperformed other techniques in all of the five meteorological stations. It showed the lowest RMSE and the highest SS and *R* in all stations. Four techniques (SAM, NRM, CCW, and IDW) showed similar performance and were second-ranked in all of the stations with little exceptions in time series. EQM⁺ improved (not substantial) the performance levels of gap-filling techniques in some stations. In general, MLR is suggested to fill in the missing values of the daily rainfall time series. However, the second-ranked techniques could also be used depending on the required time series (period) of each station. The techniques have better performance in stations located in higher altitudes. The authors expect a substantial contribution of this paper to the achievement of sustainable development goal thirteen (climate action) through the provision of gap-filling techniques with better accuracy.

1. Introduction

Rainfall (precipitation) is one of the key inputs in many disciplines such as climatology (climate variability and change), meteorology (weather conditions), irrigation engineering (irrigation scheduling), hydrology (water cycle), and environmental hazard assessment (floods). Despite its overriding uses, the rainfall dataset of meteorological stations has gigantic missing values, mainly in developing countries [1–3]. Data gaps in rainfall time series are

predominantly caused by the provisional absence of observers, equipment miscarriage, data archiving, and irregular calibration of devices [4, 5]. Ignoring the missing values from analyses has been used as a technique to manage it [6–8]. However, it leads to partial (coarse resolution) and biased results in data analyses [9–11]. Instead, filling the data gaps using reference datasets such as reanalysis products or estimates from the surrounding stations are better and widely used approaches [12–15]. Ample gap-filling techniques have been evaluated and suggested in the literature to

fill in the missing daily rainfall time series at different parts of the world. The majority of gap-filling techniques are spatial interpolation methods.

Among the spatial interpolation techniques, simple arithmetic mean (SAM) was indicated for its best performance and computational simplicity in some studies [16–18]. But, [19, 20] prioritized inverse distance weighting (IDW) over other spatial interpolation techniques. Out-performance was also reported for the normal ratio method (NRM) [16, 21, 22], correlation coefficient weighing (CCW) [23], and multiple linear regression (MLR) [24]. Nevertheless, Longman et al. [25] specified no statistical differences (similar performance) between five spatial interpolation techniques (normal ratio method, linear regression, inverse distance weighting, quantile mapping, and single best estimator) for large gaps. Machine learning processes such as the artificial neural network (ANN), Kernel approaches, and kriging are also suggested in some studies to fill the rainfall data gaps. The best performance of the machine learning process was stated for ANN [26, 27], ordinary kriging [28, 29], and Kernel approaches [30]. Besides, Grillakis et al. [31] indicated the acceptable performance of empirical quantile mapping in filling the discontinued daily rainfall data in the Mediterranean island of Crete.

Combining or modifying the previously existing gap-filling techniques is also reported for better performance than using the techniques separately. Teegavarapu et al. [32] indicated that the linear weight optimization method (LWOM) with a single best estimator (SBE) performed better than SBE only in Florida. Similarly, Kim and Pachepsky [33] concluded that the regression tree (RT) with ANN showed better performance than solely using RT or ANN in the Chesapeake Bay watershed of the USA. Furthermore, Khosravi et al. [34] presented better performance of the modified geographical coordinate (GC) method than the previously available methods in 24 station gauges in Iran. Martínez et al. [35] also showed that the generalization of the modified normal ratio with the inverse distance weighting and the generalization of modified correlation coefficient with the inverse distance weighting method outperformed NRM, NRM weighted with correlation, NRM modified with IDW, CCW, modified CCW, IDW, modified correlation coefficient with IDW, IDW weighing of NRM with correlation, and IDW-modified height. Similarly, Rahman et al. [36] indicated that the generalized linear model with gamma and Fourier series was outperformed over SAM, NRM, CCW, and IDW in estimating the missing daily rainfall series. The Gaussian mixture model-based KNN imputation showed better performance level than KNN only [37].

Filling the rainfall data gaps using the reanalysis products is also another widely used approach. For example, Cordeiro and Blanco [38] indicated that the Climate Hazards Group InfraRed Precipitation with Stations (CHIRPS) product outperformed the tropical rainfall measuring mission (TRMM) and Morphing Technique (CMORPH-CPC) in estimating daily rainfall time series in the Amazon region. Further, Tang et al. [14, 15] filled the data gaps in daily rainfall of North America (serially complete NA) and the

globe (Serially Complete Earth) dataset using the global historical climatology network daily (GHCND), a global surface summary of the day (GSOD), and Environment and Climate Change Canada (ECCC). In addition, Noh and Ahn [39] developed a new gridded rainfall dataset (K-Hydra) over the Korean peninsula to fill rainfall data gaps, which has comparable performance with global precipitation climatology project (GPCP), climate prediction center (CPC), tropical rainfall measuring mission (TRMM), and Asian precipitation highly resolved observational data integration towards evaluation (APHRODITE).

In Ethiopia, some studies evaluated and suggested different gap-filling techniques for daily rainfall time series. For instance, Boke [40] evaluated five spatial gap-filling techniques in ten meteorological stations in Ethiopia and suggested the nearest neighbor, inverse distance weighting average, and modified inverse distance weighting average for the country. Woldesenbet et al. [41] also tested four gap-filling techniques in 38 stations in the upper Blue Nile basin of Ethiopia in which CCW showed the best performance over NRM, modified NRM, and IDW. Similarly, Armanuos et al. [17] assessed twenty-one (21) gap-filling methods in 15 stations and suggested that NRM, MLR, IDW, CCW, and SAM fill in the missing rainfall data in Ethiopia. The reviewed literature indicates that the performances of gap-filling techniques vary between stations, considered evaluation criteria, statistical properties of data [17], and density and the geometrical organization of the station network [42]. Yet, to the authors' best knowledge, none of the reviewed literature and no related study covered the meteorological stations located in Wolaita Zone and the surroundings. Moreover, the applicability of gap-filling methods is limited by many factors including the required computational skill and the percentage of gaps in the data [43]. On the other hand, Ethiopia is a large country covering about 1,104,300 square kilometers [44] in which directly using any of the suggested techniques for the entire country is not representative and can lead to biased results. So, testing the gap-filling techniques at local levels is very important.

Thus, this study was initiated to evaluate the performances of seven gap-filling techniques to fill in the missing values of daily rainfall data in the meteorological stations of Wolaita Zone and the surroundings in South Ethiopia. The seven selected techniques were simple arithmetic mean (SAM), normal ratio method (NRM), inverse distance weighting (IDW), correlation coefficient weighing (CCW), multiple linear regression (MLR), empirical quantile mapping (EQM), and empirical quantile mapping plus (EQM⁺). The techniques were preferred among others due to their computational simplicity, wider application, and comparable performance with other techniques [45]. Performances of the techniques were tested against four evaluation criteria such as mean absolute error (MAE), root mean square error (RMSE), skill score (SS), and Pearson's correlation coefficients. As well, the performance consistency was evaluated on different time scales.

The authors of this paper expect momentous contributions of the paper to environmentalists, engineers, climatologists, agriculturalists, and natural resource

management experts facing rainfall data gaps. The techniques included in this study can be tracked on any other location and their performances can be compared with the findings of this work. Besides, the filled rainfall datasets of five meteorological stations (Areka, Bele, Boditi, Hosana, and Shone) are freely available based on requests. Moreover, our findings have a substantial contribution to sustainable development goal (SDG) thirteen (climate action) by providing the filled and summarized rainfall data freely so that the policymakers of the country can use it to understand the climate variability and change in the study area with reduced error level. So, it provides imperative information to take action on climate change adaptation and mitigation measures. The rest part of this paper is organized into four sections. Section 2 describes the materials and methods: study area and data description and methodology for gap-filling techniques and evaluation criteria. Section 3 presents the results of the gap-filling techniques of the missing daily precipitation data in five meteorological stations. Section 4 discusses and interprets the results. Finally, Section 5 concludes the findings of this study.

2. Materials and Methods

2.1. Study Area and Data Description. Five meteorological stations located in two zones (Wolaita and Hadiya) of southern nations' nationalities and people's regional state of Ethiopia were included in this study (see Figure 1). From the five stations, two (Hosana and Shone) are located in Hadiya Zone and three stations (Areka, Bele, and Boditi) are located in Wolaita Zone. The five meteorological stations considered in this study are sufficient and comparable with the four stations [46, 47] and six stations [48] of similar studies. The stations are located from 6.92 to 7.57° (latitude) and from 37.5 to 37.95° (longitude) and in the altitudinal ranges of 1240–2397 meters above sea level (see Table 1). The observed daily rainfall and maximum and minimum temperature data of five stations for periods (1987–2017) were obtained from the National Meteorological Agency (NMA) of Ethiopia [49].

The stations have huge missing values up to 30.2% in daily rainfall, 29.4% in maximum temperature, and 19.4% in minimum temperature (see Table 1). Besides, Bele and Shone stations did not have the dataset for maximum and minimum temperature in the study period. So, the two stations were not considered in analyses of maximum and minimum temperature (see Table 1). The rainfall datasets of five stations have a bimodal pattern (two peaks in the year) even though the months of obtaining peak values slightly vary from station to station (see Figure 2, presented in bar charts). Two peak values of rainfall were observed in April and August in Areka and Hosana stations, April and July in Shone, May and August in Boditi, and May and July in Bele stations. The study area received an annual rainfall between 1,212 (in Hosana) and 1,561 mm (in Shone). Besides, the maximum temperature has a bimodal distribution pattern (see Figure 2: presented in lines). The mean monthly maximum temperature varies between 19.34°C (in Hosana) and 29.5°C (in Areka) (see Table 1). The minimum temperature of the area has a continuously decreasing trend

from February and March to December. It ranges between 8.65°C (in Hosana) and 15.4°C (in Areka) (see Table 1).

2.2. Methodology. The methodology of this work trailed the following processing steps. First, the data matrixes of five stations with complete data (excluding the years of data missing) were prepared. For the gap-filling techniques other than quantile mapping and quantile mapping plus, the datasets of five (all) stations were considered. In the empirical quantile mapping (EQM and EQM⁺), the datasets of three stations (one target and two with higher correlation coefficients) were used. The correlation between a target station and the surrounding stations is more important than proximity (physical distance) of the stations [25]. Then, the seven gap-filling techniques were cross-validated using four evaluation criteria, and the missing values were estimated using the best-performed technique. In the case when there is no data from neighboring stations, the method used by Ismail and Ibrahim [50], using the mean on the same day and month but at different years, was used to estimate the missing value on that particular date. The detailed methodology is described in the following paragraphs.

2.2.1. Simple Arithmetic Mean (SAM). It estimates the missing values in the target station from the surrounding stations by simply taking the average of surrounding stations' data in the same period of missing value [51]. This is the simplest technique of estimating missing values used when the missing value has less than 10% [52]. It is expressed as

$$V_o = \frac{\sum_{i=1}^n V_i}{n}, \quad (1)$$

where V_o is the estimated value of the missing data, V_i is the value of the same variable at the i^{th} nearest station, and n is the number of nearest weather stations considered for averaging.

2.2.2. Normal Ratio Method (NRM). It considers the correlation coefficients between the target station and the surrounding stations. It is recommended for filling in missing values if more than 10% of the data is missing [52]. It gives weight to the data of surrounding stations based on their correlation with the target station. It is expressed as follows:

$$V_o = \frac{\sum_{i=1}^n W_i V_i}{\sum_{i=1}^n W_i}, \quad (2)$$

where V_o is the estimated value, W_i depicts the weight of the i^{th} surrounding weather station, and V_i is the value of the same variable at the i^{th} station. The weight of the surrounding station is calculated using the following equation:

$$W_i = r_i^2 \left(\frac{n_i - 2}{1 - r_i^2} \right), \quad (3)$$

where W_i is the weight of the i^{th} station, r_i corresponds to the correlation coefficient between the target station and the i^{th} surrounding station, and n_i is the number of points used to calculate the correlation coefficient.

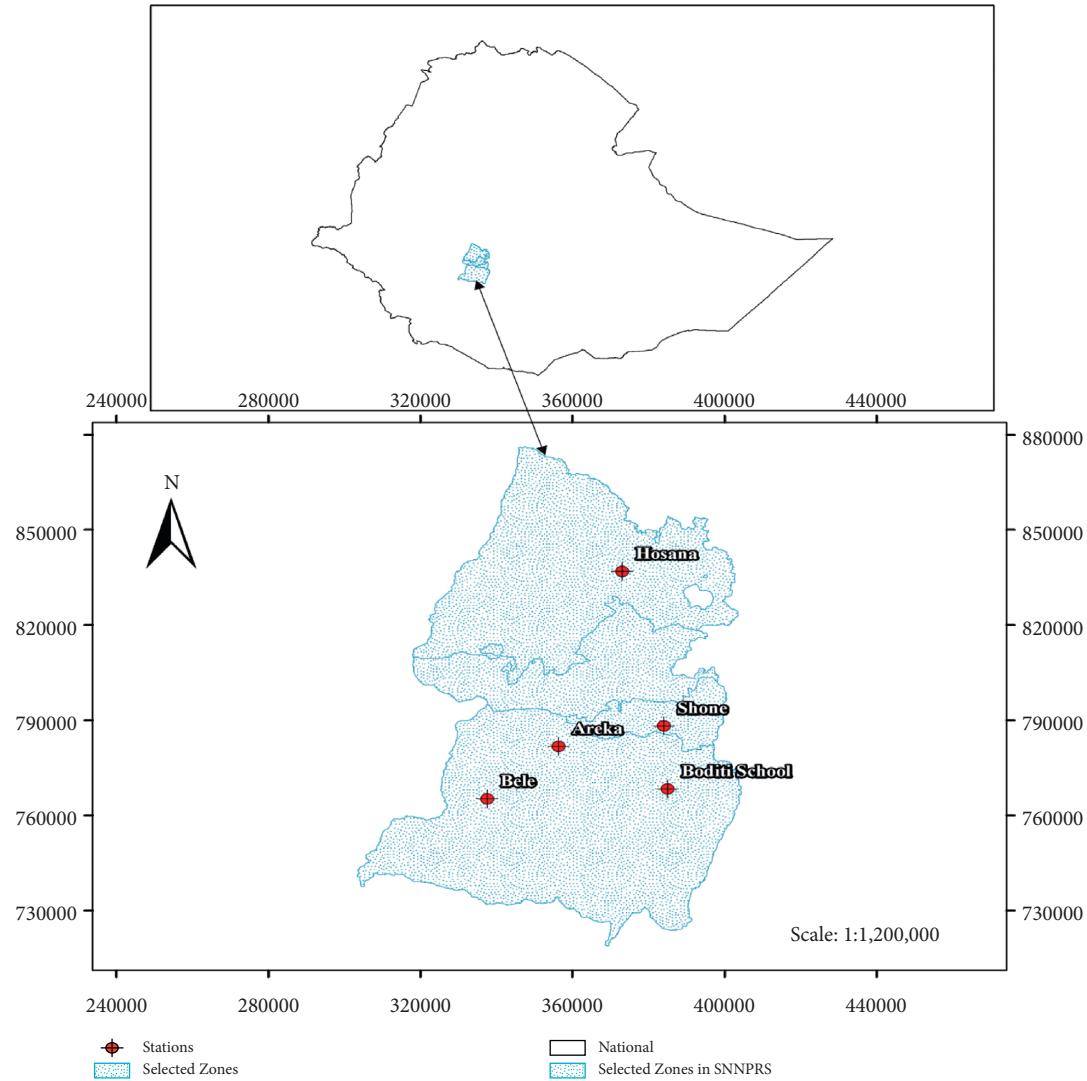


FIGURE 1: Location map of five meteorological stations in Wolaita Zone and the surroundings. SNNPRS is southern nations' nationalities and peoples' regional state of Ethiopia.

TABLE 1: Description of five meteorological stations in Wolaita Zone and the surroundings.

<i>Geographic information</i>	Stations				
	Areka	Bele	Boditi	Hosana	Shone
Latitude (degree)	7.07	6.92	6.95	7.57	7.13
Longitude (degree)	37.7	37.53	37.96	37.85	37.95
Altitude (meter)	1804	1240	2043	2307	1959
<i>Rainfall</i>					
Data coverage	1988/2017	1987–2018	1987–2018	1987–2018	1987–2018
Missing data (%)	30.2	21.5	4.5	2.7	8.4
Monthly total value	1536.84	1250.16	1264.66	1212.34	1560.58
Minimum value	29.54	28.74	34.72	11.92	37.06
Maximum value	209.1	174.44	182.08	165.66	235.94
Mean value	128.07	104.18	105.4	101.03	130.05
Standard deviation	68.7	61.13	53.94	55.1	68.63
<i>Maximum temperature</i>					
Data coverage	1992–2017	NA	1987–2018	1987–2018	NA
Missing data (%)	29.4	NA	4.6	14.8	NA
Minimum value	22.2	NA	21	19.34	NA
Maximum value	29.5	NA	28.4	25.6	NA

TABLE 1: Continued.

Geographic information	Areka	Bele	Stations		
			Boditi	Hosana	Shone
Mean value	25.7	NA	25	22.7	NA
Standard deviation	2.5	NA	2.5	2.14	NA
<i>Minimum temperature</i>					
Data coverage	1992–2017	NA	1987–2018	1987–2018	NA
Missing data (%)	19.36	NA	4.7	12.5	NA
Minimum value	12.95	NA	11.99	8.65	NA
Maximum value	15.44	NA	14.29	12.03	NA
Mean value	14.14	NA	13.28	10.81	NA
Standard deviation	0.72	NA	0.75	0.95	NA

NA is data not available.

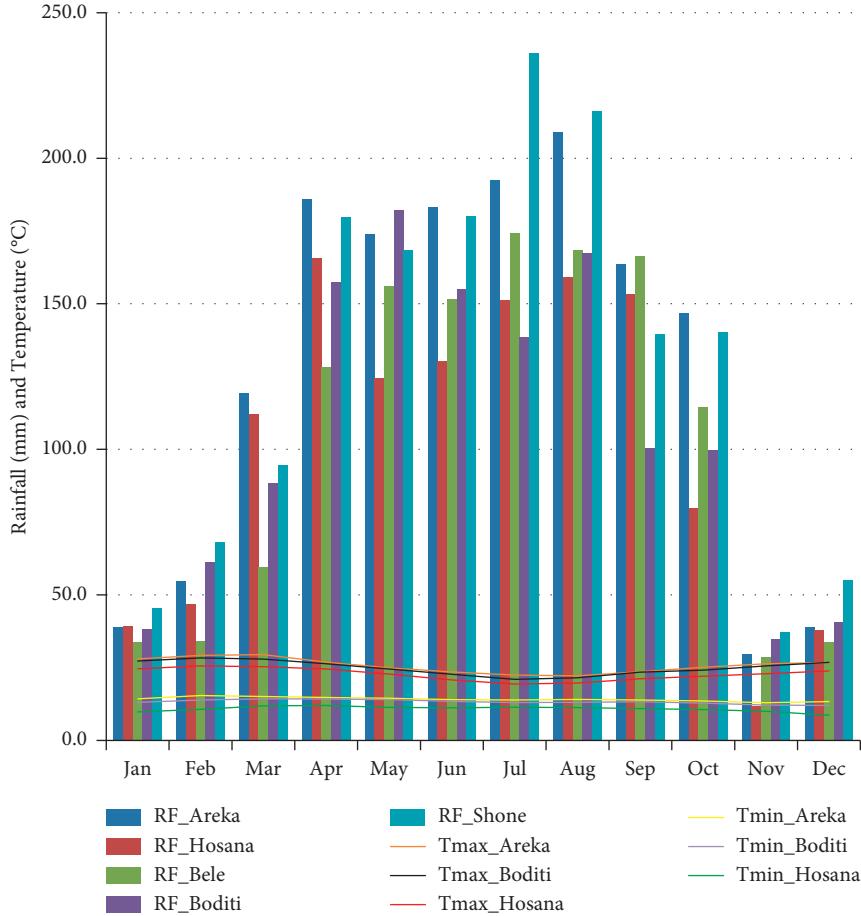


FIGURE 2: Mean monthly annual rainfall and temperature (maximum and minimum) patterns in Wolaita Zone and the surroundings. RF is the total rainfall (mm), Tmax is the mean maximum temperature (°C), and Tmin is the mean minimum temperature (°C).

2.2.3. Inverse Distance Weighting (IDW). It is the most commonly used technique for estimating the missing values of daily rainfall [53]. It assumes that the closer the surrounding stations to the target station, the better the estimation of missing values and the lower the error in estimation or the better the accuracy. It is calculated using the following equation

$$V_o = \frac{\sum_{i=1}^n V_i / d_i}{\sum_{i=1}^n (1/d_i)}, \quad (4)$$

where V_o is the estimate obtained for missing value, V_i is the observed value at the i^{th} station, d_i is the i^{th} surrounding station distance, and n is the number of stations used. The distance between the target station and the surrounding stations is calculated using the Pythagoras formula.

$$d_i = \sqrt{(X_a - X_i)^2 + (Y_a - Y_i)^2}, \quad (5)$$

where d_i is the distance between the target station and the surrounding i^{th} station, X_a and X_i are the longitudes, and Y_a

and Y_i are the latitudes of the target and the i^{th} surrounding stations, respectively. Then, the values in degrees are multiplied by 111 to convert them to kilometers.

2.2.4. Correlation Coefficient Weighing (CCW). In this approach, distance is replaced by Pearson's correlation coefficients [54]. It assures that the datasets of surrounding stations having a better positive correlation with that of the target station give better estimates of missing values in target stations than that of less correlated ones. Thus, Pearson's correlation coefficients between rainfall data of five meteorological stations were analyzed, and the missing values in target stations were determined using the following equation:

$$V_o = \frac{\sum_{i=1}^n r_i V_i}{\sum_{i=1}^n r_i}, \quad (6)$$

where V_o is the missing value of the target station, r_i is the correlation coefficient of the i^{th} surrounding station, and V_i is the value of the same variable in the i^{th} surrounding station.

2.2.5. Multiple Linear Regression (MLR). It was carried out by considering the linear significant relationship between the observed values of the target station and the surrounding stations [55]. The dataset of the target station was considered as a dependent variable and the surrounding stations' datasets were considered as independent variables. Accordingly, the multiple linear regressions were carried out for five stations. Then, the missing values in the target station were filled using the intercept and the coefficients of the variables were expressed as follows:

$$V_o = a_0 + \sum_{i=1}^n (a_i V_i), \quad (7)$$

where V_o is the estimated value, a_0, a_1, \dots, a_n are regression coefficients, and V_i is the value of the same parameter at the i^{th} weather station.

2.2.6. Empirical Quantile Mapping (EQM). It is a very common technique used to downscale the global circulation model (GCM) outputs of rainfall to the regional and local levels [56, 57]. It requires three elements (observed, historical, and projected datasets) for analysis. However, very few studies used this technique to estimate the missing daily rainfall of the target station [58, 59]. We used the observed data of the target station and that of two stations having a better correlation with the target station data and carried out EQM using R-software. QM is expressed as follows:

$$Qm(t) = F_o^{-1}[F_s[Q_s(t)]], \quad (8)$$

where $Qm(t)$ is the t^{th} estimated daily data at the target station, F_o^{-1} is the inverse cumulative distribution function (CDF) of the available data at the target station, $Q_s(t)$ is the t^{th} daily data at the neighboring station, and F_s is the CDF of the daily data at the neighboring station.

2.2.7. Empirical Quantile Mapping Plus (EQM⁺). In this study, we used the name "empirical quantile mapping plus (EQM⁺)" to refer to the empirical quantile mapping applied to the outputs (values estimated by all six techniques). The study [31] obtained a better result (reduced mean absolute error) after applying quantile mapping on the outputs generated by other techniques. Thus, we applied it to evaluate its performance on the outputs obtained by other techniques. First, the data matrix was made between the observed, an average of observed data, and the outputs of SAM, NRM, CCW, MLR, IDW, and EQM. Then, the data matrix was refed to R-software for the empirical quantile mapping process. Finally, the output was subjected to cross-validation analysis against preset criteria.

2.3. Performance Evaluation of Gap-Filling Techniques. Cross-validation was assessed by comparing the observed data and the data estimated by different gap-filling techniques. It was used to evaluate the quality (performance) of different gap-filling techniques based on four commonly used statistical validation (evaluation) criteria. The considered evaluation criteria are mean absolute error (MAE), root mean square error (RMSE), skill score (SS), and Pearson's correlation coefficients (R). Similar evaluation criteria were considered in identical studies conducted by [60–62]. The best-performing technique was selected based on cross-checking its performances under many criteria. The gap-filling techniques were evaluated as the best performing under three different conditions. The first condition (default) to decide the gap-filling techniques as best performing was when the technique has the lowest mean absolute error (MAE) and root mean square error (RMSE) and the highest skill score (SS) and correlation coefficient (R). Next, when the estimation technique fulfilled three evaluation criteria out of the four. The last condition was when the estimation technique fulfilled two evaluation criteria and showed at least equal performance level with other good performing techniques at least in one criterion. On the other hand, no gap-filling technique was decided as best performing under conditions other than these three. A similar procedure was followed to evaluate whether the application of empirical quantile mapping plus (EQM⁺) improved the performances of gap-filling techniques or not.

$$\begin{aligned}
 \text{MAE} &= \frac{\sum_{i=1}^n (V_{\text{obs},i} - V_{\text{est},i})}{n}, \\
 \text{RMSE} &= \sqrt{\frac{\sum_{i=1}^n (V_{\text{obs},i} - V_{\text{est},i})^2}{n}}, \\
 \text{SS}(\text{Vest, V}_{\text{obs}} \text{ and } \text{A}_{\text{Vobs}}) &= 1 - \frac{\text{MSE}(\text{Vest, V}_{\text{obs}})}{\text{MSE}(\text{Vavg, V}_{\text{obs}})}, \\
 \text{MSE}(\text{Vest, V}_{\text{obs}}) &= \frac{\sum_{i=1}^n (V_{\text{est},i} - V_{\text{obs},i})^2}{n}, \\
 \text{MSE}(\text{A}_{\text{Vobs}}, \text{V}_{\text{obs}}) &= \frac{\sum_{i=1}^n (\text{Vavg} - V_{\text{obs},i})^2}{n}, \\
 R &= \frac{\sum_{i=1}^n (V_{\text{obs},i} - \text{A}_{\text{Vobs}}) * (V_{\text{est},i} - \text{A}_{\text{Vest}})}{\sqrt{\sum_{i=1}^n (V_{\text{obs},i} - \text{A}_{\text{Vobs}})^2 * \sum_{i=1}^n (V_{\text{est},i} - \text{A}_{\text{Vest}})^2}}
 \end{aligned} \tag{9}$$

where $V_{\text{est},i}$ and $V_{\text{obs},i}$ are the estimated and observed i^{th} values, respectively, A_{Vobs} and A_{Vest} are the average of observed and estimated values, respectively, and n is the number of data points. MAE, RMSE, SS, and R are the mean absolute error, root mean squared error, skill score, and Pearson's correlation coefficient, respectively.

3. Results

3.1. Performances of Gap-Filling Techniques in the Areka Meteorological Station. The multiple linear regression (MLR) gap-filling technique outperformed other techniques under the root mean square error (RMSE), skill scores (SS), and Pearson's correlation coefficient (R) evaluation criteria in Areka station (see Table 2). It showed the lowest RMSE (7.75 mm) and the highest SS (0.16) and R (0.29). Its performance of estimating observed rainfall was appreciable up to the year 2010; however, it became poor after 2010 as indicated by its largest MAE value (see Table 2 and Figure 3(a)). Except for the empirical quantile mapping (EQM), the rest of the gap-filling methods showed comparable (similar) and second-ranked performance. It is only EQM that showed the negative (poor value) of skill scores ($SS = -0.31$). The performance levels of gap-filling techniques became poor after empirical quantile mapping plus: all the techniques overestimated the observed values (see Table 2 and Figure 3(b)). Despite its poor performance after 2010, no other gap-filling techniques (including EQM^+) have comparable performance in Areka station. So, MLR gap-filling technique can be used to fill rainfall data gaps in Areka station.

3.2. Performances of Gap-Filling Techniques in Bele Station. As in Areka station, MLR outperformed other techniques by considering three evaluation criteria in Bele station (RMSE, SS, and R) (see Table 3). The lowest RMSE (7.43 mm) and the highest SS (0.11) and R (0.33) values were observed in MLR. Except for the periods 1988–1991 and 2004–2009, MLR

(yellow line) was outperformed in estimating the observed annual total rainfall (red line) of Bele station in all the years (1988–2017) (see Table 3 and Figure 4(a)). Besides, no other techniques obtained the positive value of the skill score except MLR. Except for EQM, the rest of the gap-filling techniques showed similar performance (no significant variation in their statistics) and ranked second. Similarly, they (except for EQM) showed comparable performance with MLR in estimating the annual total rainfall of Bele station up to the year 2002 (see Figure 4(a)). Thus, these techniques could be used if the data required are up to the year 2002 only; otherwise, MLR is suggested to fill rainfall data gaps in Bele station. But, the performance levels of all gap-filling techniques became poorer after the application of empirical quantile mapping plus (see Figure 4(b)). All of the techniques overestimated the annual total rainfall of Bele station after EQM^+ . Thus, using EQM^+ is not advisable to fill the rainfall data gaps in Bele station.

3.3. Performances of Gap-Filling Techniques in Boditi-School Station. Under the three evaluation criteria (RMSE, SS, and R), MLR showed the best performance in Boditi-School station. The lowest RMSE (7.01 mm) and the highest SS (0.2) and R (0.45) were observed in MLR (see Table 4). The yellow line of MLR showed a consistent performance in estimating the observed annual total rainfall (OBS: red line) throughout the estimation period except for (2006–2007 and 2014–2015) in Boditi-School station (see Figure 5(a)). The remaining techniques overestimated the observed annual total rainfall between 1992 and 2000 (see Figure 5(a)). The poorest performance was observed in EQM (the only technique with the negative values of SS). Except for MLR and EQM, the performance levels of all techniques were improved after applying EQM^+ up to the year 2005 and became poorer after 2005 (see Figure 5(b)). Due to the limited improvement in the performance levels of gap-filling techniques, EQM^+ is not suggested for use in Boditi-School station.

TABLE 2: Performance evaluation results of gap-filling techniques in Areka station before and after empirical quantile mapping plus (EQM⁺).

Performance evaluation criteria	SAM	NRM	Gap-filling techniques			
			CCW	IDW	MLR	EQM
<i>Before empirical quantile mapping plus</i>						
MAE (mm)	4.16	4.14	4.15	4.21	4.67	5.40
RMSE (mm)	7.96	7.98	7.96	8.04	7.75	9.70
SS (unitless)	0.12	0.11	0.12	0.10	0.16	-0.31
R (unitless)	0.29	0.29	0.29	0.29	0.29	0.22
<i>After empirical quantile mapping plus (EQM⁺)</i>						
MAE (mm)	5.01	4.98	5.00	5.06	5.74	6.29
RMSE (mm)	8.11	8.12	8.10	8.21	7.99	10.08
SS (unitless)	0.08	0.08	0.09	0.06	0.11	-0.42
R (unitless)	0.29	0.29	0.29	0.30	0.27	0.22

MAE, RMSE, SS, and R are mean absolute error, root mean square error, skill score, and Person's correlation coefficients, respectively.

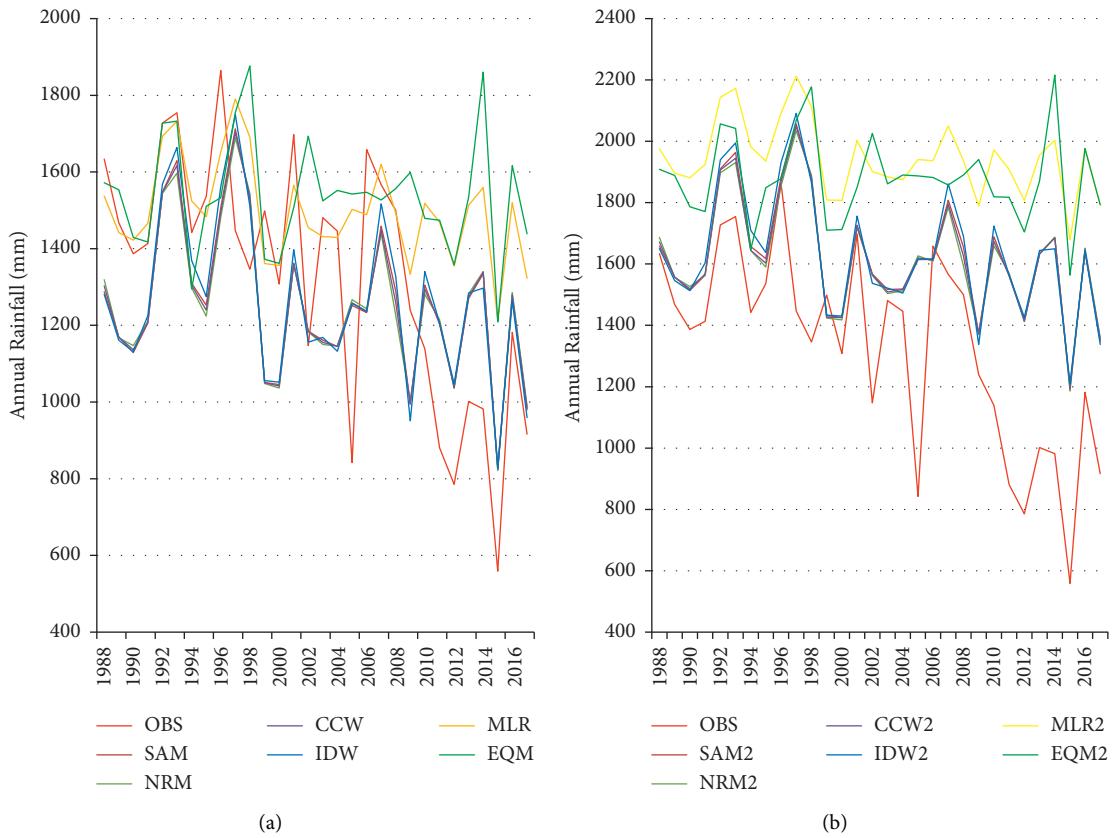


FIGURE 3: Performance consistency of gap-filling techniques in estimating the observed annual total rainfall in millimeter (OBS) before (a) and after (b) quantile mapping plus (EQM⁺) in Areka station. The performance of gap-filling methods after EQM⁺ is expressed with an extension of number 2 (e.g., SAM to SAM2).

3.4. Performances of Gap-Filling Techniques in Hosana Station. The outperformance of MLR was observed under all (MAE, RMSE, SS, and R) evaluation criteria in Hosana station (see Table 5). MLR has the lowest MAE (3.37 mm) and RMSE (5.69 mm) and the highest SS (0.37) and R (0.69). Its performance in estimating the observed daily rainfall data was consistent with time up to 2010 and decreased after 2010 (see Figure 6(a)). However, it showed poorer performance after empirical quantile mapping plus (see Figure 6(b)). Except for EQM, the rest of the gap-filling techniques showed similar performance under all of the

considered evaluation criteria (see Table 5). The performance levels of four techniques (SAM2, NRM2, CCW2, and IDW2) were improved up to the year 2001 (see Figure 6(b)). The poorest performance was observed in EQM (the only technique with a negative SS value) both before and after EQM⁺ (underestimated the observed rainfall). EQM⁺ did not improve the values of a Person's correlation coefficient (R) in all techniques. Due to the absence of weighty improvements in performance levels of gap-filling techniques after EQM⁺, its application is not suggested in Hosana station.

TABLE 3: Performance evaluation results of gap-filling techniques in Bele station before and after empirical quantile mapping plus.

Performance evaluation criteria	SAM	NRM	Gap-filling techniques			
			CCW	IDW	MLR	EQM
<i>Before empirical quantile mapping plus</i>						
MAE (mm)	4.38	4.37	4.37	4.47	4.47	5.21
RMSE (mm)	7.91	7.95	7.92	8.13	7.43	10.18
SS (unitless)	-0.01	-0.02	-0.02	-0.07	0.11	-0.68
R (unitless)	0.33	0.33	0.33	0.32	0.33	0.23
<i>After empirical quantile mapping plus (EQM⁺)</i>						
MAE (mm)	4.51	4.48	4.49	4.59	4.75	5.33
RMSE (mm)	8.03	8.06	8.03	8.25	7.44	10.25
SS (unitless)	-0.04	-0.05	-0.04	-0.10	0.11	-0.70
R (unitless)	0.33	0.33	0.33	0.32	0.33	0.24

MAE, RMSE, SS, and R, are mean absolute error, root mean square error, skill score, and Person's correlation coefficients, respectively.

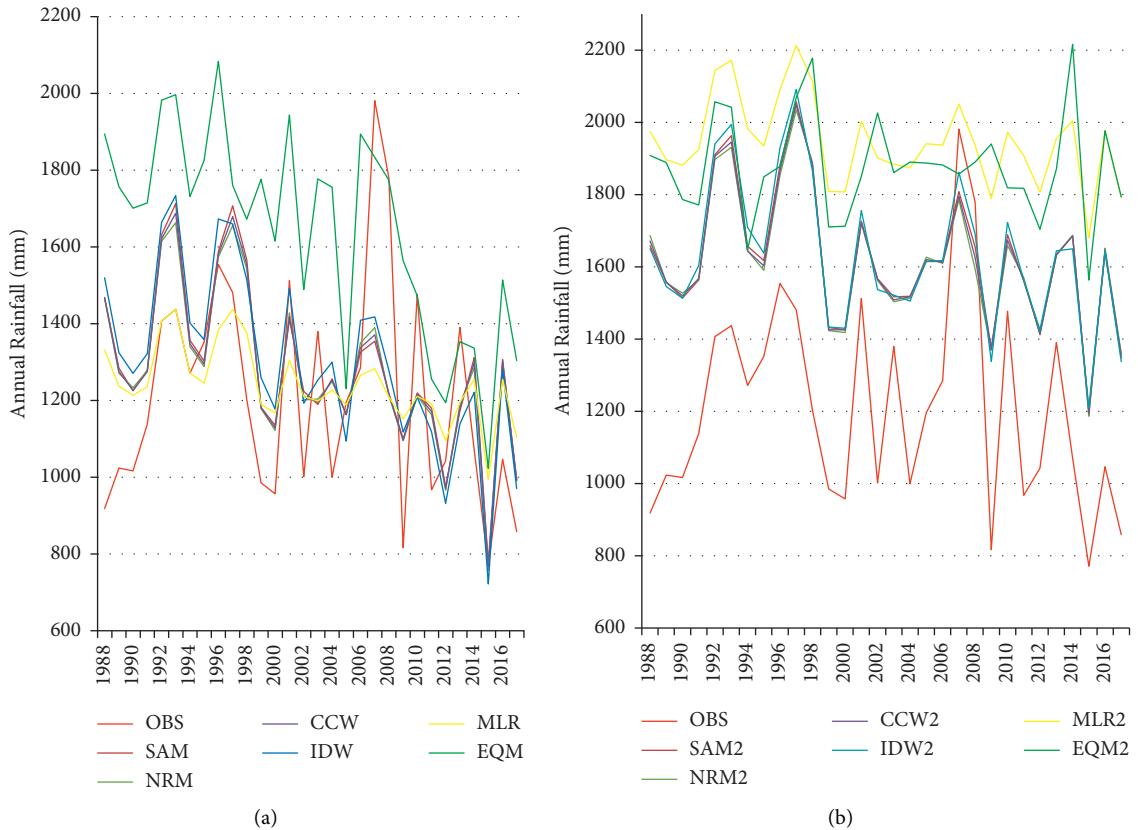


FIGURE 4: Performance consistency of gap-filling techniques in estimating the observed annual total rainfall in millimeter (OBS) before (a) and after (b) quantile mapping plus (EQM⁺) in Bele station. The performance of gap-filling methods after EQM⁺ is expressed with an extension of number 2 (e.g., SAM to SAM2).

3.5. Performances of Gap-Filling Techniques in Shone Station. MLR also outperformed other techniques in estimating the annual total rainfall of Shone station (see Table 6). However, it showed poor performance between 1992 and 1998 (see Figure 7(a)), which is expressed in its highest value of MAE (4.4 mm). Except for EQM, the rest of the gap-filling techniques showed similar performance in Shone station. The performance levels of gap-filling techniques were not improved after the application of empirical quantile mapping plus (Figure 7(b)). All of the considered techniques failed to catch the highest and lowest rainfall values of

observed data between 1992 and 1998 in Shone station (see Figures 7(a) and 7(b)). When compared with other stations, all of the gap-filling techniques showed the poorest performance in Shone station.

4. Discussion

Evaluating the performances of gap-filling techniques at station levels has paramount benefits. MLR outperformed other techniques in all of the considered (five) meteorological stations in Wolaita Zone and the surroundings. The

TABLE 4: The performance evaluation result of gap-filling techniques in Boditi-School station before and after empirical quantile mapping plus.

Performance evaluation criteria	SAM	NRM	Gap-filling techniques			
			CCW	IDW	MLR	EQM
<i>Before empirical quantile mapping plus</i>						
MAE (mm)	3.82	3.84	3.83	3.89	3.83	4.11
RMSE (mm)	7.22	7.23	7.21	7.33	7.01	8.65
SS (unitless)	0.15	0.15	0.15	0.13	0.20	-0.22
R (unitless)	0.44	0.45	0.45	0.43	0.45	0.31
<i>After empirical quantile mapping plus (EQM⁺)</i>						
MAE (mm)	3.39	3.39	3.38	3.44	3.24	3.84
RMSE (mm)	7.14	7.13	7.13	7.23	7.17	8.48
SS (unitless)	0.17	0.17	0.17	0.15	0.16	-0.17
R (unitless)	0.44	0.44	0.44	0.43	0.45	0.30

MAE, RMSE, SS, and R are mean absolute error, root mean square error, skill score, and Person's correlation coefficients, respectively.

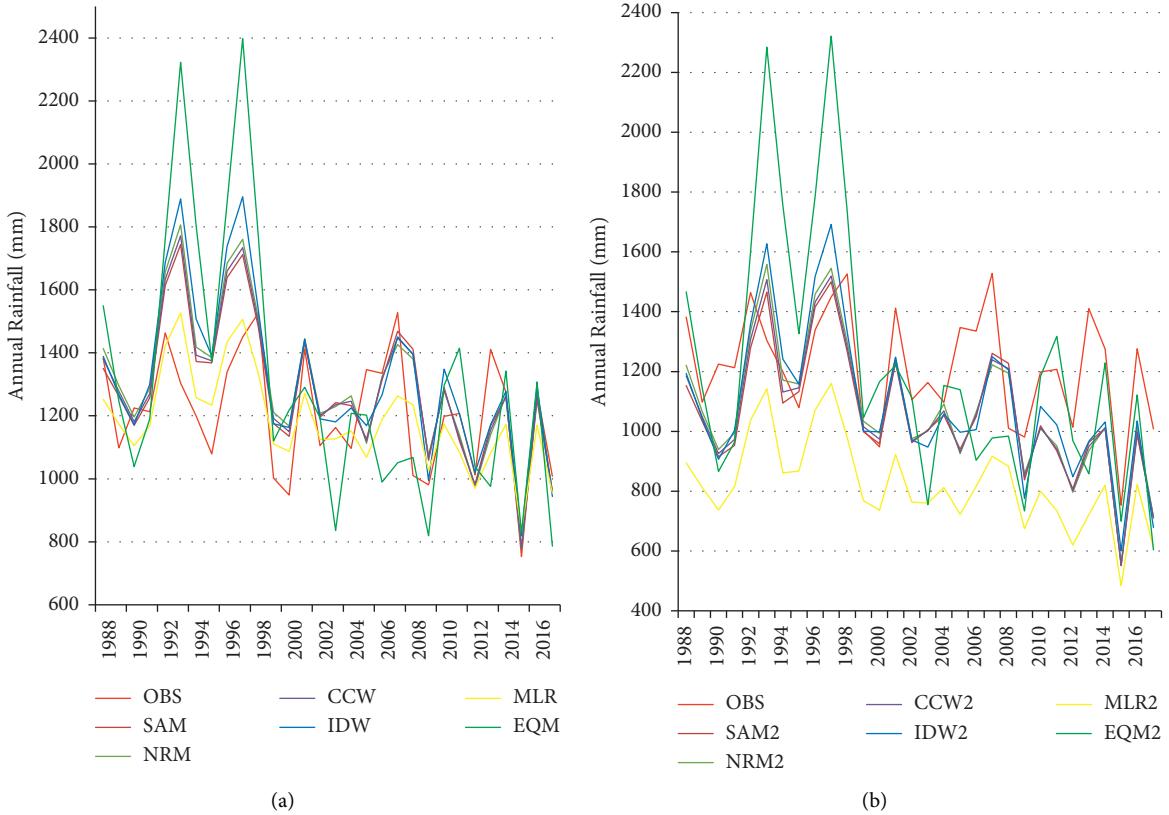


FIGURE 5: Performance consistency of gap-filling techniques in estimating the observed annual total rainfall in millimeter (OBS) before (a) and after (b) quantile mapping plus (EQM⁺) in Boditi-School station. The performance of gap-filling methods after EQM⁺ is expressed with an extension of number 2 (e.g., SAM to SAM2).

outperformance of MLR in estimating the observed rainfall data was also observed in meteorological stations of southern Iran [63] and South Central Chile [64]. In the study area, MLR estimates the daily observed rainfall with the mean absolute error of 4.67 mm in Areka, 4.47 mm in Bele, 3.83 mm in Boditi-School, 3.37 mm in Hosana, and 4.4 mm in Shone stations. In some years in the considered period (1988–2017), MLR showed poorer performances like failure to catch extreme values. The inconsistency of the MLR performance is associated with the conditional (inconsistent) process of

rainfall [65, 66], yet MLR considers the linear relationships between the observed and estimated values. Besides, MLR increased the number of rainy days (i.e., no zero was obtained in estimates of MLR) which limited its performance. The poor performance of MLR in estimating the observed daily rainfall dataset was observed in China under MAE and RMSE evaluation criteria [67]. Teegavarapu [68] also appreciated the performance of MLR in estimating the daily rainfall of Kentucky in the USA but did not suggest using it due to its negative correlation coefficient.

TABLE 5: Performance evaluation results of gap-filling techniques in Hosana station before and after empirical quantile mapping plus.

Performance evaluation criteria	SAM	NRM	Gap-filling techniques			
			CCW	IDW	MLR	EQM
<i>Before empirical quantile mapping plus</i>						
MAE (mm)	3.85	3.84	3.84	3.86	3.37	3.89
RMSE (mm)	7.14	7.14	7.14	7.16	5.69	8.37
SS (unitless)	0.01	0.01	0.01	0.00	0.37	-0.36
R (unitless)	0.38	0.38	0.38	0.38	0.69	0.28
<i>After empirical quantile mapping plus (EQM⁺)</i>						
MAE (mm)	3.48	3.47	3.47	3.49	3.16	3.64
RMSE (mm)	6.86	6.86	6.86	6.87	6.00	8.00
SS (unitless)	0.08	0.08	0.08	0.08	0.30	-0.25
R (unitless)	0.37	0.37	0.37	0.37	0.69	0.26

MAE, RMSE, SS, and R are mean absolute error, root mean square error, skill score, and Person's correlation coefficients, respectively.

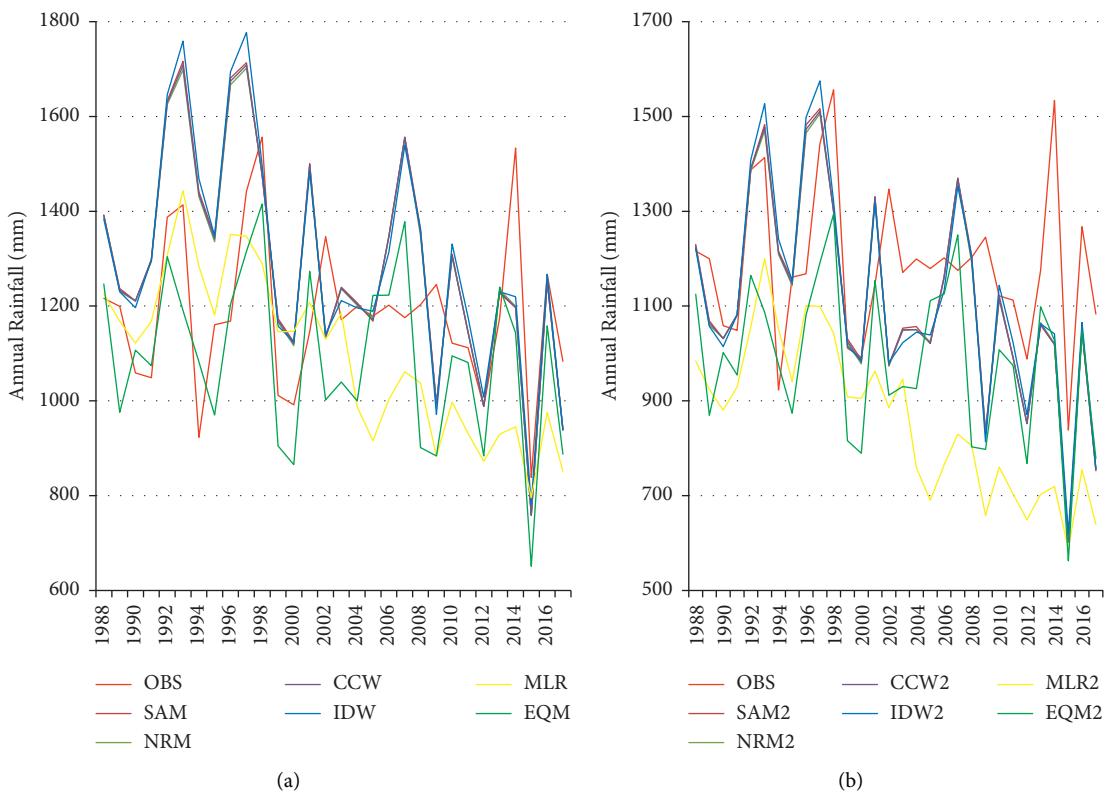


FIGURE 6: Performance consistency of gap-filling techniques in estimating the observed annual total rainfall in millimeter (OBS) before (a) and after (b) quantile mapping plus (EQM⁺) in Hosana station. The performance of gap-filling methods after EQM⁺ is expressed with an extension of number 2 (e.g., SAM to SAM2).

With little variations between stations, four of the gap-filling techniques (SAM, NRM, CCW, and IDW) ranked second in estimating the observed rainfall in the Wolaita Zone and the surroundings. Empirical quantile mapping showed the poorest performance in all of the five stations. It is the only technique that showed the negative skill scores values in all stations. Empirical quantile mapping plus (EQM⁺) improved the performance levels of some gap-filling techniques, though the improvement is not substantial. However, it did not improve the performance level of MLR in any of the five stations. There was a statistically nonsignificant difference in the relative performance of gap-

filling techniques between stations. This is because all stations have a similar rainfall pattern of the bimodal (peak and valley). Besides, their statistics of variation between the observed and mean values (standard deviation) between stations are comparable (53.94–68.7 mm) (see Table 1).

To determine the relationships between the altitude of meteorological stations and performance levels of the techniques, Pearson's correlation was analyzed. The results indicated that there was a strong negative correlation between altitude and MAE (-0.91) and between altitude and RMSE (-0.72). Besides, a strong positive correlation was obtained between altitude and SS (0.82) and between altitude

TABLE 6: Performance evaluation results of gap-filling techniques in Shone station before and after empirical quantile mapping plus.

Performance evaluation criteria	SAM	NRM	Gap-filling techniques			
			CCW	IDW	MLR	EQM
<i>Before empirical quantile mapping plus</i>						
MAE (mm)	4.0	4.0	4.0	4.0	4.4	4.3
RMSE (mm)	7.3	7.3	7.3	7.4	6.9	8.5
SS (unitless)	0.1	0.1	0.1	0.0	0.2	-0.3
R (unitless)	0.4	0.4	0.4	0.4	0.4	0.3
<i>After empirical quantile mapping plus (EQM⁺)</i>						
MAE (mm)	4.0	4.0	4.0	4.0	4.6	4.2
RMSE (mm)	7.2	7.2	7.1	7.2	6.9	8.1
SS (unitless)	0.1	0.1	0.1	0.1	0.2	-0.2
R (unitless)	0.4	0.4	0.4	0.4	0.4	0.3

MAE, RMSE, SS, and R, are mean absolute error, root mean square error, skill score, and Pearson's correlation coefficients, respectively.

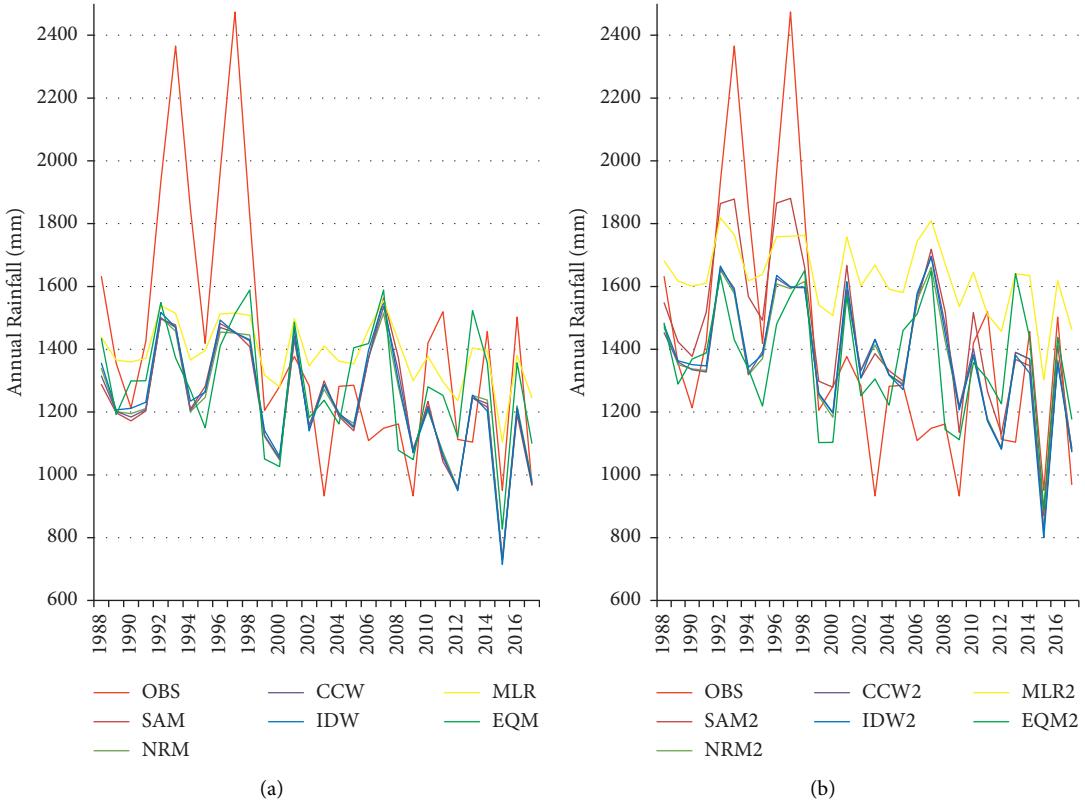


FIGURE 7: Performance consistency of gap-filling techniques in estimating the observed annual total rainfall in millimeter (OBS) before (a) and after (b) quantile mapping plus (EQM⁺) in the Shone station. The performance of gap-filling methods after EQM⁺ is expressed with an extension of number 2 (e.g., SAM to SAM2).

and R (0.75). This showed that the gap-filling techniques have better performance in meteorological stations located at a higher altitude than those at lower altitudes. This is because the meteorological stations located in higher elevations obtain more rainfall in a better pattern (better data statistics) and those located in lower altitudes received erratic rainfall. The preference to use any of the tested gap-filling techniques has to consider the performances of each technique in each station. In general, MLR without EQM⁺ can be suggested for use in the study area. It is in agreement with the study [69] indicating that the precision of MLR in estimating the observed rainfall is less affected by the

increase of failure (percent of missing data) and has better performance when the distance between stations is short.

5. Conclusion

The performances of seven gap-filling techniques in estimating the daily observed rainfall data in five meteorological stations of Wolaita Zone and the surrounding were tested in this study. The techniques were tested against the four widely used evaluation criteria such as mean absolute error (MAE), root mean square error (RMSE), skill score (SS), and Pearson's correlation coefficients (R). MLR fulfilled three of

the evaluation criteria (RMSE, SS, and R) in all of the five meteorological stations in Wolaita Zone and the surroundings. It showed the lowest RMSE and the highest SS and R over all of the considered techniques. With some exceptional cases, its performance was consistent with the analysis period (1988–2017). None of the remaining methods has comparable performance with MLR. However, its performance becomes poorer after the empirical quantile mapping plus applied. Except for EQM, the rest of the gap-filling techniques (SAM, NRM, CCW, and IDW) did not show a statistically significant difference in their performance levels in all of the considered stations. Thus, prioritizing any of these four techniques depends on the intended analysis period, the purpose of end-use, and individual station. EQM showed the poorest performance under all of the evaluation criteria in all of the stations. Besides, its application on the outputs obtained by other techniques (EQM⁺) did not bring noteworthy improvements in the performance levels of gap-filling techniques. So, it is suggested to use the MLR technique to fill in the missing values of daily rainfall in the stations of Wolaita Zone and the surroundings. In addition, using any of the second-ranked techniques depending on their performance range (period and station) is acceptable to be used in the study area.

Data Availability

The data of meteorological stations are available in the National Meteorological Agency (NMA) of Ethiopia, which is free for the study purpose and charged for other purposes. An official letter is needed to access the data of meteorological stations. Important datasets used for this study are included within the paper.

Disclosure

This paper is part of a Ph.D. dissertation work of a corresponding author supervised by co-authors.

Conflicts of Interest

The authors declare that there are no conflicts of interest concerning the publication of this article.

Acknowledgments

The authors would like to acknowledge the National Meteorological Agency of Ethiopia for freely providing the data of the stations. This work was sponsored by the Africa Center of Excellence for Climate-Smart Agriculture and Biodiversity Conservation of Haramaya University (World Bank) and Wolaita Sodo University, Ethiopia.

References

- [1] M. T. Dastorani, A. Moghadamnia, J. Piri, and M. Rico-Ramirez, “Application of ANN and ANFIS models for reconstructing missing flow data,” *Environmental Monitoring and Assessment*, vol. 166, no. 1–4, pp. 421–434, 2010.
- [2] L. M. Castro, J. Gironás, and B. Fernández, “Spatial estimation of daily precipitation in regions with complex relief and scarce data using terrain orientation,” *Journal of Hydrology*, vol. 517, pp. 481–492, 2014.
- [3] F. Oriani, S. Stisen, M. C. Demirel, and G. Mariethoz, “Missing data imputation for multisite rainfall networks: a comparison between geostatistical interpolation and pattern-based estimation on different terrain types,” *Journal of Hydrometeorology*, vol. 21, no. 10, pp. 2325–2341, 2020.
- [4] G. A. A. Saeed, Z. L Chuan, R. Zakaria, W. W. N. Yusoff, and M. Z. Salleh, “Determination of the best single imputation algorithm for missing rainfall data treatment,” *Journal of Quality Measurement and Analysis (JQMA)*, vol. 12, no. 1–2, pp. 79–87, 2016.
- [5] S. Beguería, M. Tomas-Burguera, R. Serrano-Notivolí, D. Peña-Angulo, S. M. Vicente-Serrano, and J.-C. González-Hidalgo, “Gap filling of monthly temperature data and its effect on climatic variability and trends,” *Journal of Climate*, vol. 32, no. 22, pp. 7797–7821, 2019.
- [6] A. R. P. Hernández, R. C. Balling, and L. R. B. Martínez, “Comparative analysis of indices of extreme rainfall events: variations and trends from southern México,” *Atmósfera*, vol. 22, pp. 219–228, 2009.
- [7] A. Mekasha, K. Tesfaye, and A. J. Duncan, “Trends in daily observed temperature and precipitation extremes over three Ethiopian eco-environments,” *International Journal of Climatology*, vol. 34, no. 6, pp. 1990–1999, 2014.
- [8] S. H. Gebrechorkos, S. Hülsmann, and C. Bernhofer, “Statistically downscaled climate dataset for east Africa,” *Scientific Data*, vol. 6, no. 31, p. 31, 2019.
- [9] O. Harel and X.-H. Zhou, “Multiple imputation: review of theory, implementation and software,” *Statistics in Medicine*, vol. 26, no. 16, pp. 3057–3077, 2007.
- [10] A. J. Newman, M. P. Clark, J. Craig et al., “Gridded ensemble precipitation and temperature estimates for the contiguous United States,” *Journal of Hydrometeorology*, vol. 16, no. 6, pp. 2481–2500, 2015.
- [11] A. Chaudhry, W. Li, A. Basri, and F. Patenaude, “A method for improving imputation and prediction accuracy of highly seasonal univariate data with large periods of missingness,” *Wireless Communications and Mobile Computing*, vol. 2019, Article ID 4039758, 13 pages, 2019.
- [12] V. Singh and Q. Xiaosheng, “Data assimilation for constructing long-term gridded daily rainfall time series over southeast Asia,” *Climate Dynamics*, vol. 53, no. 5–6, pp. 3289–3313, 2019.
- [13] H. Aguilera, C. Guardiola-Albert, and C. Serrano-Hidalgo, “Estimating extremely large amounts of missing precipitation data,” *Journal of Hydroinformatics*, vol. 22, no. 3, pp. 578–592, 2020.
- [14] G. Tang, M. P. Clark, A. J. Newman et al., “SCDNA: a serially complete precipitation and temperature dataset for north America from 1979 to 2018,” *Earth System Science Data*, vol. 12, no. 4, pp. 2381–2409, 2020.
- [15] G. Tang, M. P. Clark, and S. M. Papalexiou, “SC-earth: a station-based serially complete earth dataset from 1950 to 2019,” *Journal of Climate*, vol. 34, no. 1, pp. 1–47, 2021.
- [16] R. P. De Silva, N. D. K. Dayawansa, and M. D. Ratnasiri, “A comparison of methods used in estimating missing rainfall data,” *Journal of Agricultural Sciences*, vol. 3, no. 2, pp. 101–108, 2007.
- [17] A. M. Armanuos, N. Al-Ansari, and Z. M. Yaseen, “Cross assessment of twenty-one different techniques for missing

- rainfall data estimation," *Atmosphere*, vol. 11, no. 389, pp. 1–34, 2020.
- [18] A. Aieb, K. Madani, M. Scarpa, B. Bonacorso, and K. Lefsih, "A new approach for processing climate missing databases applied to daily rainfall data in soummam watershed, Algeria," *Heliyon*, vol. 5, p. e1247, 2019.
- [19] K. N. Dirksa, J. E. Hayb, C. D. Stowa, and D. Harris, "High-resolution studies of rainfall on Norfolk island-part IV: observations of fractional time raining," *Journal of Hydrology*, vol. 208, pp. 156–176, 1998.
- [20] M. E. Moeletsi, Z. P. Shabalala, G. De Nysschen, and S. Walker, "Evaluation of an inverse distance weighting method for patching daily and dekad rainfall over the free state province, south Africa," *WaterSA*, vol. 42, no. 3, pp. 466–474, 2016.
- [21] J. El Kasri, A. Lahmili, O. Latifa, L. Bahi, and H. Soussi, "Comparison of the relevance and the performance of filling in gaps techniques in climate datasets," *International Journal of Civil Engineering & Technology*, vol. 9, no. 5, pp. 13–21, 2018.
- [22] H. P. G. M. Caldera, V. R. P. C. Piyathisse, and K. D. W. Nandalal, "A comparison of methods of estimating missing daily rainfall data," *Engineer: Journal of the Institution of Engineers, Sri Lanka*, vol. 49, no. 4, pp. 1–8, 2016.
- [23] R. S. V. Teegavarapu and V. Chandramouli, "Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records," *Journal of Hydrology*, vol. 312, no. 1–4, pp. 191–206, 2005.
- [24] N. Kanda, H. S. Negi, M. S. Rishi, and M. S. Shekhar, "Performance of various techniques in estimating missing climatological data over snowbound mountainous areas of Karakoram himalaya," *Meteorological Applications*, vol. 25, no. 3, pp. 337–349, 2018.
- [25] R. J. Longman, A. J. Newman, T. W. Giambelluca, and M. Lucas, "Characterizing the uncertainty and assessing the value of gap-filled daily rainfall data in Hawaii," *Journal of Applied Meteorology and Climatology*, vol. 59, no. 7, pp. 1261–1276, 2020.
- [26] T. R. Nkuna and J. O. Odiyo, "Filling of missing rainfall data in luvuhu river catchment using artificial neural networks," *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 36, no. 14–15, pp. 830–835, 2011.
- [27] P. Coulibaly and N. D. Evora, "Comparison of neural network methods for infilling missing daily weather records," *Journal of Hydrology*, vol. 341, no. 2, pp. 27–41, 2007.
- [28] A. Mair and A. Fares, "Comparison of rainfall interpolation methods in a mountainous region of a tropical island," *Journal of Hydrologic Engineering*, vol. 16, no. 4, pp. 371–383, 2011.
- [29] A. G. Frazier, T. W. Giambelluca, H. F. Diaz, and H. L. Needham, "Comparison of geostatistical approaches to spatially interpolate month-year rainfall for the Hawaiian islands," *International Journal of Climatology*, vol. 36, no. 3, pp. 1459–1470, 2016.
- [30] H. Lee and K. Kang, "Gap-filling of missing precipitation data using kernel estimations for hydrologic modeling," *Advances in Meteorology*, vol. 2015, Article ID 935868, 12 pages, 2015.
- [31] M. G. Grillakis, C. Polykretis, S. Manoudakis, K. D. Seiradakis, and D. D. Alexakis, "A quantile mapping method to fill in discontinued daily precipitation time series," *Water*, vol. 12, no. 8, p. 2304, 2020.
- [32] R. S. V. Teegavarapu, A. Aly, C. S. Pathak, J. Ahlquist, H. Fuelberg, and J. Hood, "Infilling missing precipitation records using variants of spatial interpolation and data-driven methods: use of optimal weighting parameters and nearest neighbour-based corrections," *International Journal of Climatology*, vol. 38, no. 2, pp. 776–793, 2018.
- [33] J.-W. Kim and Y. A. Pachepsky, "Reconstructing missing daily precipitation data using regression trees and artificial neural networks for SWAT streamflow simulation," *Journal of Hydrology*, vol. 394, no. 3–4, pp. 305–314, 2010.
- [34] G. Khosravi, A. R. Nafarzadegan, A. Nohegar, H. Fathizadeh, and A. Malekian, "A modified distance-weighted approach for filling annual precipitation gaps: application to different climates of Iran," *Theoretical and Applied Climatology*, vol. 119, no. 1–2, pp. 33–42, 2014.
- [35] J. L. M. Martínez, F. A. H. Rangel, I. S. Domínguez, A. R. Morua, and J. H. Hernández, "Analysis of a new spatial gap-filling weighting method to estimate missing data applied to rainfall records," *Atmósfera*, vol. 32, no. 3, pp. 237–259, 2019.
- [36] N. A. Rahman, S. M. Deni, and N. M. Ramli, "Generalized linear model for estimation of missing daily rainfall data," *AIP Conference Proceedings*, vol. 1830, 2017.
- [37] P. C. Chiu, A. Selamat, and O. Krejcar, "Infilling missing rainfall and runoff data for Sarawak, Malaysia using gaussian mixture model based K-nearest neighbor imputation," in *Advances and Trends In Artificial Intelligence, from Theory to Practice*, F. WotawaG. Friedrich et al., Eds., vol. 11606, Cham, Germany, Springer, 2019.
- [38] A. L. M. Cordeiro and C. J. C. Blanco, "Assessment of satellite products for filling rainfall data gaps in the amazon region," *Natural Resource Modeling*, vol. 34, no. 2, pp. 1–21, 2021.
- [39] G. H. Noh and K. H. Ahn, "New gridded rainfall dataset over the Korean peninsula: gap infilling, reconstruction, and validation," *International Journal of Climatology*, vol. 2021, pp. 1–18, 2021.
- [40] A. S. Boke, "Comparative evaluation of spatial interpolation methods for estimation of missing meteorological variables over Ethiopia," *Journal of Water Resource and Protection*, vol. 9, no. 8, pp. 945–959, 2017.
- [41] T. A. Woldesenbet, N. A. Elagib, L. Ribbe, and J. Heinrich, "Gap filling and homogenization of climatological datasets in the headwater region of the upper Blue Nile basin, Ethiopia," *International Journal of Climatology*, vol. 37, no. 4, pp. 2122–2140, 2017.
- [42] M. Garcia, C. D. P. Lidard, and D. C. Goodrich, "Spatial gap-filling of precipitation in a dense gauge network for monsoon storm events in the southwestern United States," *Water Resources Research*, vol. 44, no. 5, pp. 1–14, 2008.
- [43] J. J. Miró, V. Caselles, and M. J. Estrela, "Multiple imputation of rainfall missing data in the iberian mediterranean context," *Atmospheric Research*, vol. 197, pp. 313–330, 2017.
- [44] World Bank, "The World Factbook," 2021, <https://www.cia.gov/the-world-factbook/countries/ethiopia/>.
- [45] F.-W. Chen and C.-W. Liu, "Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan," *Paddy and Water Environment*, vol. 10, no. 3, pp. 209–222, 2012.
- [46] X. Li, L. Li, X. Wang, and F. Jiang, "Reconstruction of hydrometeorological time series and its uncertainties for the Kaidu river basin using multiple data sources," *Theoretical and Applied Climatology*, vol. 113, no. 1–2, pp. 45–62, 2013.
- [47] W. Sanusi, W. Z. Wan Zin, U. Mulbar, M. Danial, and S. Side, "Comparison of the methods to estimate missing values in monthly precipitation data," *International Journal of Advanced Science, Engineering and Information Technology*, vol. 7, no. 6, pp. 2168–2174, 2017.

- [48] E. M. Mokhele, Z. P. Shabalala, G. D. Nysschen, and S. Walker, "Evaluation of an inverse distance weighting method for patching daily and decadal rainfall over the free state province, south Africa," *Water*, vol. 42, no. 3, pp. 466–474, 2016.
- [49] NMA, "National meteorological agency of federal democratic republic of Ethiopia, Addis Ababa, Ethiopia," 2019, <http://www.ethiomet.gov.et/>.
- [50] W. N. W. Ismail and W. Z. W. Ibrahim, "Estimation of rainfall and streamflow missing data for Terengganu, Malaysia by using gap-filling technique methods," *Malaysian Journal of Fundamental and Applied Sciences*, vol. 13, no. 3, pp. 213–217, 2017.
- [51] M. Hasanpour Kashani and Y. Dinپashoh, "Evaluation of efficiency of different estimation methods for missing climatological data," *Stochastic Environmental Research and Risk Assessment*, vol. 26, no. 1, pp. 59–71, 2012.
- [52] W. Y. Tang, A. H. M. Kassim, and S. H. Abubakar, "Comparative studies of various missing data treatment techniques—Malaysian experience," *Atmospheric Research*, vol. 42, no. 1–4, pp. 247–262, 1996.
- [53] B. Ahrens, "Distance in spatial interpolation of daily rain gauge data," *Hydrology And Earth System Sciences*, vol. 10, no. 2, pp. 197–208, 2006.
- [54] N. F. A. Radi, R. Zakaria, and M. A. Z. Azman, "Estimation of missing rainfall data using spatial gap-filling and imputation methods," *AIP Conference Proceedings*, vol. 1643, pp. 42–48, 2015.
- [55] L. Muluken, "Techniques of filling missing values of daily and monthly rainfall data : a review," *SF Journal of Environmental and Earth Science*, vol. 3, no. 1, p. 1036, 2020.
- [56] L. Gudmundsson, J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen, "Technical note: downscaling RCM precipitation to the station scale using statistical transformations—a comparison of methods," *Hydrology and Earth System Sciences*, vol. 16, no. 9, pp. 3383–3390, 2012.
- [57] J. Ringard, F. Seyler, and L. Linguet, "A quantile mapping bias correction technique based on hydroclimatic classification of the Guiana shield," *Sensors*, vol. 17, no. 6, pp. 1–17, 2017.
- [58] C. Simolo, M. Brunetti, M. Maugeri, and T. Nanni, "Improving estimation of missing values in daily precipitation series by a probability density function-preserving approach," *International Journal of Climatology*, vol. 30, no. 10, pp. 1564–1576, 2010.
- [59] U. Devi, M. S. Shekhar, G. P. Singh, N. N. Rao, and U. S. Bhatt, "Methodological application of quantile mapping to generate rainfall data over northwest Himalaya," *International Journal of Climatology*, vol. 2019, pp. 1–11, 2019.
- [60] A. Di Piazza, F. L. Conti, L. V. Noto, F. Viola, and G. La Loggia, "Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy," *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, no. 3, pp. 396–408, 2011.
- [61] P. D. Wagner, P. Fiener, F. Wilken, S. Kumar, and K. Schneider, "Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions," *Journal of Hydrology*, vol. 464–465, pp. 388–400, 2012.
- [62] A. A. Abbas, N. A. Ali, G. Abozeid, and H. I. Mohamed, "Comparison of different methods for estimating missing monthly rainfall data," in *Proceedings of the Twenty-Second International Water Technology Conference (IWTC)*, Ismailia, Egypt, September 2019.
- [63] M.-T. Sattari, A. Rezazadeh-Joudi, and A. Kusiak, "Assessment of different methods for estimation of missing data in precipitation studies," *Hydrology Research*, vol. 48, no. 4, pp. 1032–1044, 2017.
- [64] A. Barrios, G. Trincado, and R. Garreaud, "Alternative approaches for estimating missing climate data: application to monthly rainfall records in south-central Chile," *Forest Ecosystems*, vol. 5, p. 28, 2018.
- [65] R. L. Wilby, C. W. Dawson, and E. M. Barrow, "Statistical downscaling model (SDSM 4.2): a decision support tool for the assessment of regional climate change impacts," *Environmental Modelling and Software*, vol. 17, no. 2, pp. 145–157, 2007.
- [66] G. Tardivo and A. Berti, "Comparison of four methods to fill the gaps in daily precipitation data collected by a dense weather network," *Science Journal of Environmental Engineering Research*, vol. 2013, no. 265, pp. 1–9, 2013.
- [67] T. Chen, L. Ren, F. Yuan et al., "Comparison of spatial gap-filling schemes for rainfall data and application in hydrological modeling," *Water*, vol. 9, pp. 1–18, 2017.
- [68] R. S. V. Teegavarapu, "Spatial interpolation using nonlinear mathematical programming models for estimation of missing precipitation records," *Hydrological Sciences Journal*, vol. 57, no. 3, pp. 383–406, 2012.
- [69] T. M. Ventura, G. S. S. Guarienti, R. M. G. Ventura, C. E. Guarienti, and J. M. Figueiredo, "Analysis of gap filling in rainfall data with statistical methods," in *Proceedings of the Simpósio Internacional de Climatologia VISIC*, Natal, Brazil, October 2015.