National-scale Rainfall Bias Correction CMIP6 Projections Using Statistical Methods with Explicit SSP Comparison and Seasonal Regimes

A. D. S. Iresh
Department of Civil Engineering
Faculty of Engineering Technology
The Open University of Sri Lanka
Nugegoda, Sri Lanka
shahikairesh@gmail.com

B.C.L. Athapattu
Department of Civil Engineering
Faculty of Engineering Technology
The Open University of Sri Lanka
Nugegoda, Sri Lanka
bcliy@ou.ac.lk

W. C. D. K. Fernando
Department of Civil Engineering
Faculty of Engineering
General Sir John Kotelawala Defense University
Ratmalana, Sri Lanka
kumari@kdu.ac.lk

Jayantha Obeysekera
Sea Level Solution Center
Institute of Environment
Florida International University
Miami, USA
jobeysek@fiu.edu

***Abstract*— General Circulation Models (GCMs) play a vital role in forecasting future climate changes; however, their rainfall outputs frequently exhibit significant biases due to coarse resolution and oversimplified processes. Consequently, it is essential to apply bias correction methods to this data before utilizing it for climate change initiatives. This study introduces a context-aware framework that integrates multiple statistical bias correction methods specifically tailored to the climatic zones and rainfall characteristics of Sri Lanka. Rainfall data from 68 monitoring stations across the island were utilized alongside outputs from the CNRM-CM6-1 GCM. The study was conducted under moderate and high climate change scenarios, Shared Socioeconomic Pathways (SSP) (SSP2-4.5 and SSP5-8.5) from the CMIP6 program. Bias correction techniques, including Linear Scaling (LS), Empirical Quantile Mapping, Delta Change, Power Transformation, and Local Intensity Scaling, were employed to evaluate the performance of each method. The framework dynamically identifies the most effective method based on metrics such as RMSE, R^2, and Taylor diagrams. The results indicate that LS significantly enhances rainfall accuracy, particularly for monthly precipitation, outperforming other methods under SSP2-4.5. However, LS has limitations, particularly regarding its inability to account for topographic variation and spatial biases, highlighting the necessity for complementary approaches. Spatial analysis revealed regional disparities, especially in northern areas, attributed to local climate dynamics. Furthermore, the corrected data demonstrate improved accuracy and effectively capture seasonal trends, including intensified extremes projected under SSP5-8.5. This research underscores the importance of adaptive, regionally tailored correction strategies to enhance the reliability of rainfall data in climate impact assessments and water resource management.**

I. Introduction

A precise evaluation of future climate conditions is crucial for effective water resource management, agriculture, and disaster risk reduction. GCMs are essential tools for simulating climate processes and projecting future scenarios based on greenhouse gas emissions [1]. However, the outputs from GCMs often exhibit significant biases when compared to local climate data, particularly regarding rainfall, which varies significantly across different regions and periods [2]. Using GCM outputs directly for regional climate impact assessments can be misleading unless adjustments are made to account for these uncertainties. Consequently, bias correction is vital for aligning GCM-simulated rainfall with observed data, thereby enhancing the reliability of these outputs for local applications [3].

Biases can emerge from various factors, including the accuracy of physical process representations and local climate conditions that the model may inadequately capture. The primary aim of bias correction is to adjust model outputs to align more closely with observed data, such as rainfall, thereby rectifying systematic errors and enhancing the reliability of climate information [3]. Rainfall bias correction methods can be categorized into four main types: statistical bias correction, which is relatively straightforward; dynamic downscaling, which is more complex and requires greater computational resources; statistical downscaling methods, which establish statistical relationships between large-scale atmospheric variables from GCMs and local climate; and hybrid or Artificial Intelligence (AI) based methods, which combine the strengths of statistical and dynamic approaches or utilize AI to improve accuracy and efficiency [4], [5], [6].

Research indicates that dynamically downscaled Regional Climate Models (RCMs) generally demonstrate better performance in temperate climates compared to tropical regions [7]. This discrepancy is primarily attributed to the nature of tropical rainfall, which is predominantly convective and occurs on a sub-daily timescale. Consequently, dynamic downscaling may not be suitable for tropical countries like Sri Lanka, where bias correction is essential.

In contrast, statistical downscaling methods require long, high-quality observational records, which are often difficult to obtain [5]. Similarly, hybrid or AI-based methods are complex and require substantial computational power and extensive long-term data, which can be challenging to secure. Therefore, statistical bias correction methods may serve as a more viable alternative to other methods in situations where computational resources and data access are limited. Additionally, statistical bias correction methods are particularly appreciated for their simplicity and effectiveness in addressing these discrepancies.

A variety of statistical bias correction methods have been developed to address biases in GCM variables, including rainfall [3], [7], [8]. These techniques range from straightforward scaling methods to more complex approaches, such as distribution mapping, all designed to enhance the accuracy and reliability of statistical results. The effectiveness of these methods can vary depending on the climatic conditions of different regions. Therefore, it is crucial to evaluate the suitability of bias correction methods for regional applications by comparing their effectiveness [7].

The LS technique is one of the most widely adopted statistical bias correction methods for adjusting precipitation outputs generated by GCMs [9]. This method corrects systematic biases by aligning the long-term mean of simulated precipitation with that of observed data over a designated historical reference period [10]. A scaling factor, typically determined as the ratio of observed to modeled mean precipitation, is then applied to adjust future simulated values [11]. Due to its simplicity, ease of implementation, and ability to preserve the temporal structure of the original data, LS is particularly advantageous in hydrological modeling and climate impact assessments [12]. However, it is important to note that this method does not address biases in variability, distribution, or extreme events, which may limit its accuracy for specific applications [13].

The empirical quantile method (EQM) is a widely utilized statistical technique for correcting variable bias in GCMs and was introduced by Déqué [14]. The EQM aligns the quantiles of modeled climate variables with those observed during a specified historical period, thereby correcting the entire distribution, in contrast to the LS method. This non-parametric approach is particularly efficient for rainfall bias correction, as it adeptly adjusts for extreme rainfall events that GCMs often underestimate [15]. It can be applied to both daily and monthly timescales. However, the EQM method assumes stationarity, which may limit its validity in the context of climate change. Additionally, it may struggle to project unprecedented extremes that were not observed in the historical data and relies on long-term, high-quality datasets [16]. There is also a risk of overfitting to historical data if the method is not implemented with caution.

The Delta Change Approach (DCA) is a statistical downscaling and bias correction method used to correct GCM variables, dating back to the 1990s, particularly introduced by Barrow et al. [17] under the United Kingdom climate impact programme. The DCA is a simple and efficient method that alters observed historical time series using the mean change (deltas), which are projected by GCMs [18]. The resulting time series retains the observed temporal structure but reflects the expected climate change signal. Specifically, the DCA is suitable for assessing the impact of mean change. Similar to EQM, DCA also assumes stationarity and is not suitable for future extreme predictions, and is not suitable for nonlinear changes in climate behavior [16].

The power transformation (PT) is a statistical bias correction method specifically designed to correct nonlinear biases in GCMs' rainfall data. Utilizing the power exponent, PT adjusts the scale and shape of rainfall distribution, which LS, EQM, and DCA failed to do [19]. The PT method improves the representation of skewed distribution, especially for heavy rainfall. Additionally, parameters can be tailored to different seasons and locations, making them applicable to daily data. However, raising values to power can lead to units and interpretation issues. Sensitivity for zero values is a major drawback, as it results in undefined values of the power transformation [20]. Similarly, PT fails in correcting extreme precipitation [21].

Similar to the other four methods, the Local Intensity Scaling (LIS) method is a statistical bias correction and downscaling method designed to correct GCMs and Regional Climate Models (RCMs). However, unlike other statistical methods, LIS operates by differentiating between wet and dry conditions and scaling precipitation intensity locally [19]. Specifically, LIS was developed by Schmidli et al. [22] to perform bias correction for extreme events. The LIS corrects both frequency and intensity biases and effectively adjusts the high intensities. LIS performs correction locally in time and space, hence improving the accuracy for seasonal/monthly climate variability. However, LIS requires high-resolution observed data, which is challenging to obtain and does not fully correct the entire distribution. Also, it assumes stationarity [20].

The Coupled Model Intercomparison Project Phase 6 (CMIP6) provides a range of GCMs designed for analyzing the impacts of climate change. Among these, the CNRM-CM6 climate model represents a significant advancement in climate modeling. The GCM was developed by the Centre National de Recherches Météorologiques (CNRM), which operates under Météo-France, France's national meteorological service. It has been identified as the most suitable GCM for the context of Sri Lanka [23]. The model was created in partnership with the Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS), an institution specializing in advanced scientific computing [24]. The CNRM-CM6-1 model is a primary GCM that can be utilized in the Sri Lankan context. It plays a vital role within the CMIP6 framework, which underpins international climate research initiatives. This comprehensive Earth System Model (ESM) is designed to analyze and predict climate change and its associated variability across various scales. In the current study, statistical bias correction methods were employed to adjust biases in the CNRM-CM6-1 GCM data, leveraging observed data to assess their effectiveness.

II. Materials and Methods

A. Study Area

Sri Lanka is an island situated in the Indian Ocean, bordered by India to the north and surrounded by the Indian Ocean on the west and south. To the east, it borders the Bay of Bengal, extending from 5°55′ to 9°51′N and from 79°42′ to 81°53′ E [25]. Due to being an island and its tropical location, air temperatures across the country do not exhibit significant variation, except in the central mountain region [26], [27]. Therefore, the notable climatic differences are primarily linked to the amount of rainfall received. Sri Lanka is significantly influenced by the Indian Ocean monsoon system, resulting in a seasonal distribution of rainfall across the island throughout the year. The average annual rainfall in Sri Lanka ranges from less than 1,000 mm on the southeast coast to over 4,500 mm on the western slopes of the highlands [26], [27]. The two main seasonal monsoon patterns contribute a substantial portion of rainfall to the island: the southwest monsoon (SWM) from May to September and the northeast monsoon (NEM) from December to February [28]. Additionally, there are two inter-monsoon periods: the first inter-monsoon in March/April and the second inter-monsoon in October/November. The two inter-monsoons were influenced by tropical cyclones, depressions, and thunderstorms associated with the migration of the intertropical convergence zone [26].

B. Observed Climate Data

Daily rainfall data spanning from 1975 to 2014 were utilized, sourced from sixty-eight rain gauge stations managed by the Irrigation Department and the Department of Meteorology. Fig. 1 displays the rain gauges that were used in the study. The missing data were filled using the inverse distance method [29], [30], which was deemed suitable given that the percentage of missing values across all 68 stations was below 5%.

C. GCM Data and Climate Scenarios

The CNRM-CM6-1 GCM gridded rainfall data, featuring a nominal resolution of 50 km, served as the primary GCM for the analysis. Fig. 1 illustrates the GCM grid points over Sri Lanka. The CNRM-CM6-1 GCM was utilized under the SSP2-4.5 and SSP5-8.5 to assess the performance of bias correction. The dataset from the CNRM-CM6-1 model encompasses two time periods: the control period from 1975 to 2014 and the future period from 2015 to 2100, both of which were essential to the bias correction analysis.

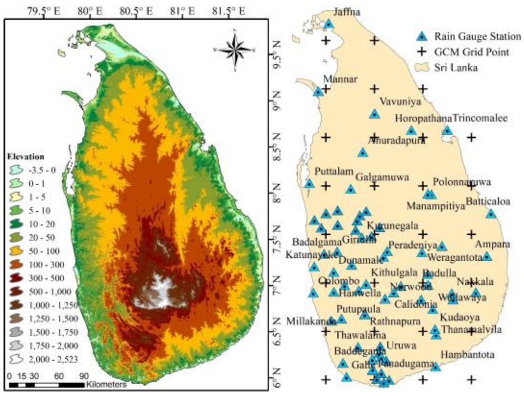

Figure 1. Elevation Change and Rain Gauge Observational Network with GCM Grid Point Locations.

D. Experimental Setup

Initially, time series data from 25 grid points located in Sri Lanka were extracted using the Climate Data Operator (CDO) [31] for both the control and future periods. The CDO is a suite of command-line tools developed by the Max Planck Institute for Meteorology, specifically for processing and analyzing climate and numerical weather prediction data, and it is widely utilized in climate science for handling large-scale climate datasets, particularly those in NetCDF, GRIB, and HDF5 formats [32]. Similarly, time series data (1975-2014) for the exact grid locations were developed using the observed 68 rain gauge stations through the Thiessen polygon weighted method. Twenty-five time series of data were developed to bias-correct the CNRM-CM6-1 model grid point data, as shown in Fig. 1.

Initially, bias correction was applied to the GCM grid point data series using observed data from the control period of 1975-2014. Five statistical bias correction methods were employed to estimate bias correction parameters for the CNRM-CM6-1 model. Subsequently, the GCM data for the future period was adjusted using these bias correction parameters. The Climate Model Data for Hydrological Modeling (CMhyd) was employed for the bias correction process. CMhyd is a Python-based tool designed to extract and rectify biases in data from both global and regional climate models [33]. The CMhyd program utilizes various Python packages, including NetCDF, NumPy, SciPy, and PyQt. CMhyd was explicitly developed to effectively leverage observed data corresponding to the locations of gauges used in the basin study or the specific region, utilizing NetCDF files or datasets that align with the original observed rain gauge data. Fig. 2 shows the step-by-step methodology adopted for bias correction. The effectiveness of bias correction methods is evaluated by comparing the adjusted outputs from GCMs with observed precipitation data.

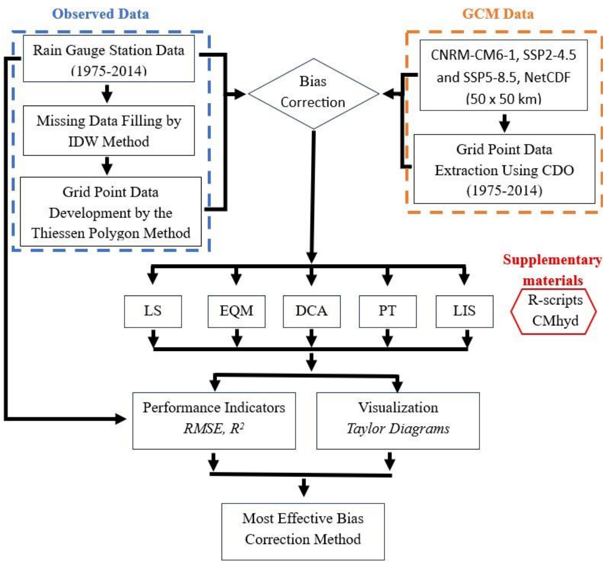

Figure 2. Flowchart Illustrating the Step-by-Step Methodology for Analyzing Bias Correction Performance.

This comparison utilizes statistical metrics such as Root Mean Square Error (*RMSE*), Coefficient of Determination (R^2), and Taylor diagrams [34]. This evaluation process helps identify the most effective technique for aligning model-generated data with actual observations, particularly concerning extreme precipitation events, which are crucial for water resource management and disaster preparedness. Following the implementation of five bias correction methods, a comprehensive analysis revealed the most accurate and reliable approach for correcting future precipitation projections. The corrected data is subsequently plotted for visual comparison alongside observed and adjusted outputs, ensuring a precise representation of precipitation events. The *RMSE* and R^2 were estimated as follows.

$$RMSE = \sqrt{\frac{1}{p}\sum_{i=1}^{p}(y_i - \check{y}_i)^2} \tag{1}$$

where, y_i is the actual value, $\check{y}_i$ is the predicted value, and p is the number of data points. The *RMSE* metric gives greater weight to more significant errors due to the squaring of differences. Consequently, a lower *RMSE* showcases a closer alignment between model predictions and actual observations; however, it is also sensitive to outliers, which can skew the results owing to the squaring mechanism.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \check{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{2}$$

where, $\bar{y}$ is the mean of the actual value. The R^2 value ranges from 0 to 1, with a value of 1 indicating a perfect fit. If R^2 equals 1, the model explains 100% of the variability; conversely, an R^2 of 0 indicates that the model explains none of the variability, and if $R^2 < 0$, the model performs worse than simply predicting the mean.

D. Bias Correction Methods

1) Linear Scaling

The LS bias correction method is a straightforward and effective approach to mitigating biases in precipitation data [35]. This method adjusts the observed precipitation by applying a scaling factor based on the ratio of mean observed to mean modeled precipitation derived from a GCM, thereby aligning their averages [36].

The LS method is designed to ensure that the corrected monthly precipitation values match the observed values. This correction is performed monthly, using factors that account for the discrepancies between raw GCM output and actual data. By applying a multiplier and an additive term, the LS method addresses systematic biases in GCM outputs, resulting in improved accuracy for precipitation forecasts and enhanced utilization of GCM data for hydrological and climate-related applications.

$$P_{cor,m}(t) = P_{mod,m}(t) \text{ X } \frac{\bar{P}_{obs}}{\bar{P}_{mod}} \tag{3}$$

where P_{cor} is the corrected precipitation at m^{th} month and time (day) t, $\bar{P}_{obs}$ is the mean observed precipitation, and P_{mod} is the mean modeled precipitation at the m^{th} month and time (day) t.

2) Empirical Quantile Mapping

The process begins with the calculation of the Cumulative Distribution Function (CDF) for both observed and modeled precipitation. Subsequently, quantiles were identified to facilitate the mapping of model values to their corresponding quantiles in the observed data. This adjustment significantly enhances the accuracy of parameters such as the mean, standard deviation, and quantiles, ultimately resulting in a more realistic representation of precipitation. EQM employs the empirical cumulative distribution function (ECDF) and its inverse for these adjustments [35], [36].

$$P_{cor,m}(t) = ECDF^{-1}{}_{obs,m}(ECDF_{mod,m,t}(P_{mod,m,t}) \tag{4}$$

3) Delta Change Approach

The DCA is a methodology that employs outputs from GCMs for catchment-scale analysis and hydrological modeling by creating change factors. Instead of providing absolute values, it modifies observed climate data, such as temperature and precipitation, using these factors to reflect projected climate variations resulting from greenhouse gas emissions and other influences. This process facilitates the estimation of future climate scenarios and the assessment of their potential impacts, thereby providing a clearer understanding of the effects of climate change through alterations in current conditions [18].

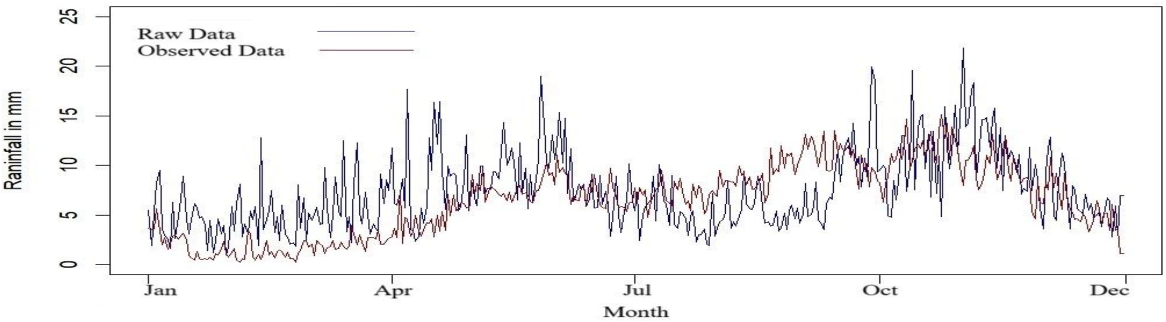

Figure 3. Annual Average Daily Time Series Plot of the Dandeniya Tank Rain Gauge, and Closest Grid Point Raw Data of CNRM-CM6-1, Under SSP2-4.5.

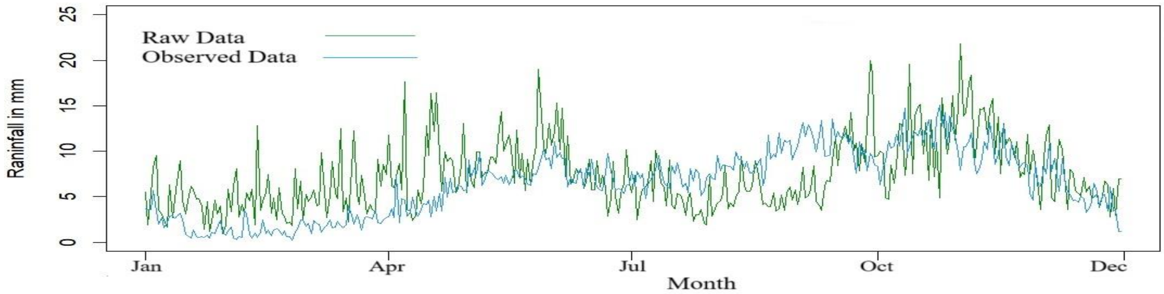

Figure 4. Annual Average Daily Time Series Plot of the Dandeniya Tank Rain Gauge and Closest Grid Point Raw Data from CNRM-CM6-1, Under SSP5-8.5.

$$P_{cor,m}(t) = P_{obs}(t) \text{ X } \left[\frac{P_{future,\ mod}(t)}{P_{historical,\ mod}(t)}\right] \quad (5)$$

Where, P_{obs} is the observed precipitation at time *t*, $P_{future,\ mod}$ is the GCM-modeled precipitation for future periods at time *t*, $P_{historical,\ mod}$ is the GCM-modeled precipitation for historical periods at time *t*. This method is often used in climate projections because relative changes are more important than absolute values.

4) Power Transformation

The polynomial or PT technique is utilized to address nonlinear biases in precipitation data, particularly those that affect the extremes of the distribution. This method involves applying a polynomial or power function to the modeled precipitation data to better align its distribution with the observed data. By fitting a mathematical function, the transformed values can be adjusted to correct discrepancies, particularly in the tail ends of the distribution, where extreme precipitation events typically occur. Additionally, PT addresses biases in variance and employs an exponential form to further refine the standard deviation (SD) of the rainfall series [36].

$$P_{cor,m}(t) = \left[P_{obs,m}(t) - \mu(P_{mod,m})\right] \text{X} \left[\frac{\sigma(P_{obs,m})}{\sigma(P_{mod,m})}\right] + \mu(P_{obs,m}) (6)$$

where σ represent the SD operator and μ represent the expectation operator.

5) Local Intensity Scaling

The LIS method addresses the correction of wet-day frequencies and intensity values in precipitation data, specifically targeting the challenge posed by an overabundance of drizzle days accompanied by low rainfall amounts. The method consists of two primary steps: First, a wet-day threshold for each month ($P_{thres,m}$) is established based on modeled precipitation to ensure that the frequency of wet days aligns with observed data. Second, a scaling factor (S_m) is computed to adjust rainfall intensity, ensuring that the mean rainfall of the corrected dataset aligns with the observed mean. This approach effectively enhances the accuracy of precipitation patterns [36].

$$S_m = \frac{\mu(P_{obs,m}(t)|P_{obs,m}(t)>0)}{\mu(P_{mod,m}(t)|P_{mod,m}(t)>P_{thres,m})} \quad (7)$$

$$P_{corr,m}(t) = \begin{cases} 0, & P_{mod,m}(t) < P_{thres,m} \\ P_{mod,m}(t) \cdot S_{m,} & otherwise \end{cases} \quad (8)$$

III. RESULTS AND DISCUSSION

A. Pre-Bias Correction Analysis

The analysis produced comparative time series plots for GCM grid points alongside observed rainfall data. The GCM grid point nearest to the Dandeniya Tank rain gauge was selected as a representative example. Under the SSP2_4.5 scenario, the GCM data initially overestimated rainfall during the first half of the year, while predictions were notably lower in the latter half; however, forecasts for November and December demonstrated greater accuracy (Fig. 3).

A similar pattern emerged in the SSP5-8.5 scenario, characterized by increased instances of intense rainfall and shifts in seasonality (Fig. 4). Despite some systematic biases, the GCM data effectively represented the overall seasonal cycle, although discrepancies in both magnitude and timing were evident. Both scenarios exhibited extreme values, highlighting the necessity for bias correction in future projections, especially under higher emission pathways.

B. Post-Bias Correction Analysis

The analysis focused on bias-corrected data from the CNRM-CM6-1 model, in conjunction with observed data spanning the period from 1975 to 2014. It employed five correction methods pertinent to Sri Lanka under the carbon emission scenarios SSP2_4.5 and SSP5_8.5. Table 1 presents a comparison between the observed and bias-corrected data under SSP2_4.5, while Fig. 5 showcases box plots of the *RMSE* for each method. A lower *RMSE* indicates superior performance, and the characteristics of the box plots reveal the consistency and variability among the methods, with outliers highlighting areas of suboptimal performance. A tightly clustered box indicates consistency, whereas a wider box reveals variability, emphasizing each method's sensitivity to precipitation patterns. Outliers highlight instances of poor performance.

TABLE 1. The Results for the Bias Correction Statistical Performance Analysis for Daily Data.

Bias Correction Method	*RMSE*	R^2
Linear Scaling	18.88	0.0059
Delta Change	19.63	0.0052
local Intensity Scaling	19.52	0.0046
Power Transformation	20.73	0.0032
Empirical Quantile Mapping	22.47	0.0025

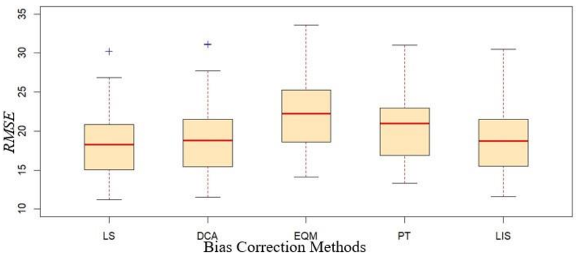

Figure 5. Boxplot showing the Performance of Five Bias Correction Methods in terms of *RMSE*.

Fig. 6 features Taylor diagrams that demonstrate the efficacy of bias correction methods at the GCM grid point near the Ampara rain gauge station, serving as a representative example, as similar results were observed in the other Taylor diagrams. Notably, the LS method achieved the highest correlation, with an average of 0.2 for daily data and 0.6 for monthly data. In the annual series, the LIS, EQM, and PT methods demonstrated strong correlations near 0.95, whereas other methods were generally around 0.3.

Although all methods exhibited improvements with monthly mean values, the LS method consistently surpassed the others in terms of standard deviation and correlation, despite underestimating variability at higher aggregation levels. The EQM and PT methods, although computationally intensive, produced more sophisticated results, while the more straightforward LS method encountered challenges related to nonlinear biases.

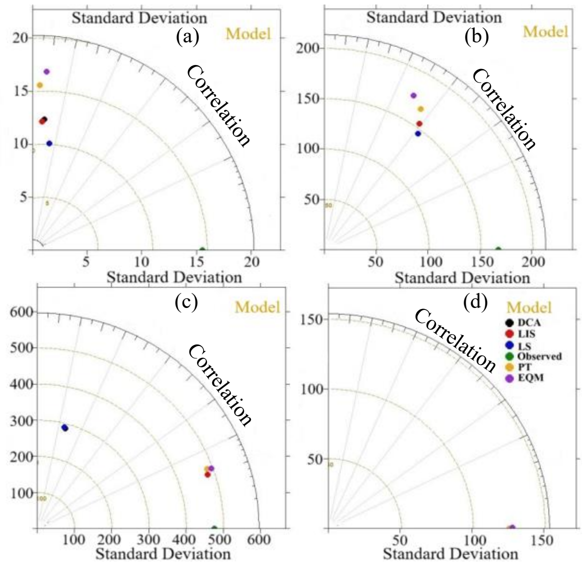

Figure 6. Taylor Diagram Developed for the Comparison of Bias Correction Methods for: (a) Daily Series, (b) Monthly Series, (c) Annual Series, and (d) Mean Monthly Data.

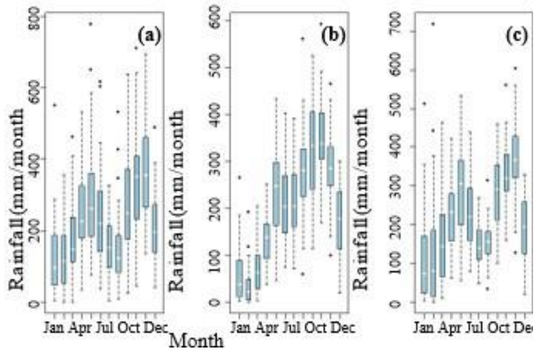

Figure 7. The Average Monthly Rainfall Records Box Plots of All Rain Gauges: (a) Observed Data, (b) Raw GCM (CNRM-CM6-1, SSP2-4.5) Data, and (c) Bias-corrected GCM Data by LS.

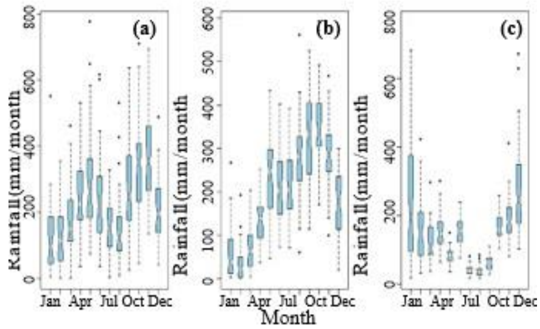

Figure 8. The Average Monthly Rainfall Records Box Plots of All Rain Gauges: (a) Observed Data, (b) Raw GCM (CNRM-CM6-1, SSP5-8.5) Data, and (c) Bias-corrected GCM Data by the LS Method.

Regional factors, including climate and topography, have a significant impact on performance, particularly in the context of water resource management and extreme weather modeling, where the EQM method particularly shines.

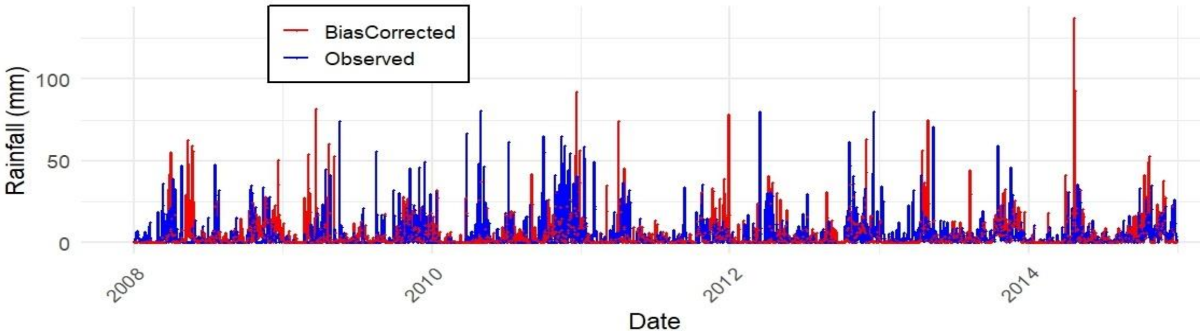

Figure 9. Verification Time Series Plot for the Calibration Period (Historical Data) of the Daisy Valley Rain Gauge.

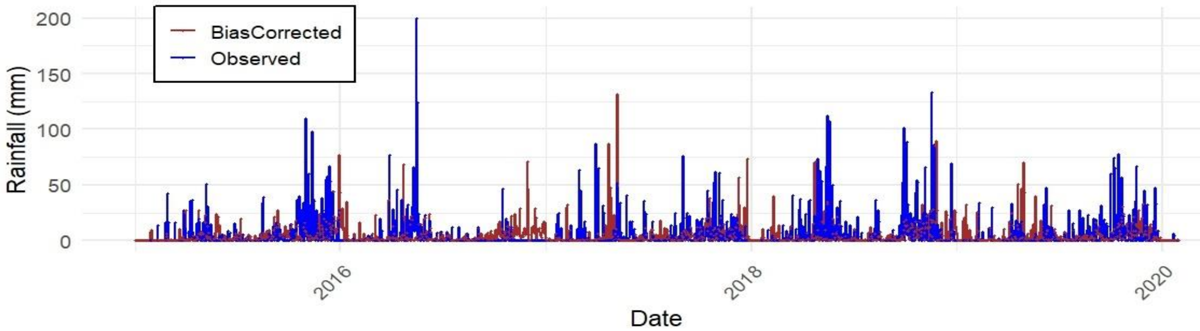

Figure 10. Verification Time Series Plot for the Validation Period (Future Data) of the Daisy Valley Rain Gauge.

Taylor diagrams reveal that the LS method achieved the highest daily correlation at an average of 0.2, which improved to about 0.6 for monthly data. In the annual correlations, LIS, EQM, and PT approached 0.95, while the daily series correlations for LIS and PT remained relatively low, ranging from 0.06 to 0.08.

The Taylor diagram and statistical analysis demonstrate that the LS method surpasses other approaches in terms of statistical measures, standard deviation, and correlation. The LS bias correction method proves to be robust, displaying a low median *RMSE* and moderate performance variability (R^2), particularly for monthly data; however, its effectiveness may fluctuate depending on duration and the comprehensiveness of the data. While more advanced methods, such as EQM and PT, offer accuracy, they necessitate careful consideration of their complexity. Although the LS method is relatively straightforward, it can be influenced by nonlinear biases, with contextual factors such as climate and topography further impacting its performance. The bias correction approach effectively aligned rainfall distribution metrics with observed data, thereby reducing outliers and the interquartile range. Performance comparisons indicate that the SSP2-4.5 scenario significantly outperformed the SSP5-8.5 scenario, highlighting varying effectiveness across different climate scenarios.

In the Concise presentation, only the bias-corrected historical data from the Daisy Valley rain gauge station for the years 2008 to 2014 is included as calibration data, while the bias-corrected future data from 2015 to 2020 serves as validation, as shown in Figures 9 and 10. Figure 9 displays a time series plot of the historical data collected from 2008 to 2014, which is designated as the calibration period for the Daisy Valley rain gauge station. In contrast, Figure 10 illustrates a time series plot that corresponds to the validation phase of the Daisy Valley rain gauge station for the future period.

To assess the efficacy of bias correction, *RMSE* indicated a reduction in uncertainty for both SSPs. Spatial maps were created to illustrate monthly LS method parameters. Fig. 11 underscores significant biases in the northern and northeastern regions, which were corrected except for November and December, when biases shifted to the southeast and southwest.

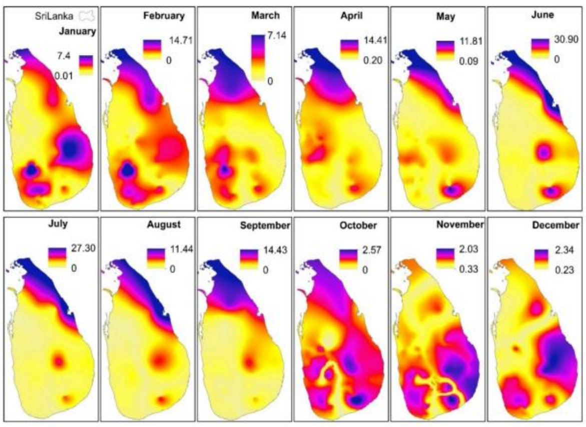

Figure 11. Spatial Maps of Monthly Variation of the LS Bias Correction Method Parameter (Monthly Factor) from January to December.

Fig. 12 compares observed and bias-corrected average daily rainfall data for SSP2-4.5 and SSP5-8.5, showing that while the corrected values slightly overestimate the observations, the low variability in the error suggests an effective correction. The clustering of scatter points further supports the appropriateness of the method, confirming the value of bias correction for Sri Lanka.

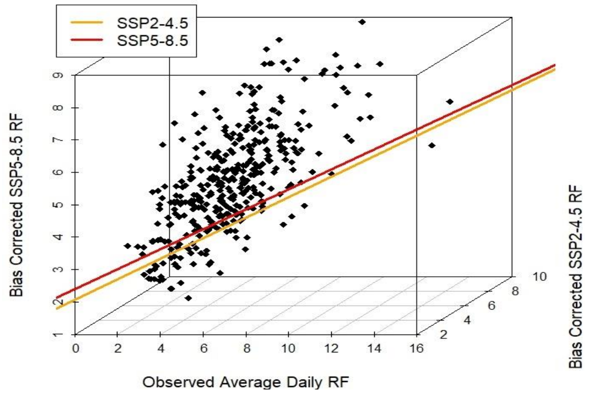

Figure 12. Spatial Graphical Representation Average Daily Rainfall of Observed, Bias-Corrected SSP2-4.5, and Bias-Corrected SSP5-8.5.

The results of the bias correction analysis, as detailed in Table 2, reveal varying degrees of agreement between the bias-corrected rainfall data for the SSP2-4.5 and SSP5-8.5 scenarios and the observed historical rainfall values. For daily average rainfall, both scenarios tend to overestimate relative to observations (3.4 mm/day), with SSP2-4.5 reporting 5.0 mm/day and SSP5-8.5 at 5.4 mm/day, resulting in positive deviations of approximately 47% and 59%, respectively. Although these differences are moderate, the overestimation may impact subsequent hydrological modeling by inflating simulated runoff. The mean monthly rainfall for SSP2-4.5 is 219.5 mm, which closely aligns with the observed value of 228.1 mm, exhibiting a slight negative bias of about 3.8%. In contrast, SSP5-8.5 significantly underestimates monthly totals at 143.1 mm, approximately 37% lower than the observed values. This indicates that the bias correction for SSP5-8.5 may require further refinement to capture monthly-scale variability accurately.

Upon examining seasonal means, the NEM mean rainfall for both SSP2-4.5 (1020.0 mm) and SSP5-8.5 (1068.3 mm) closely aligns with the observed total of 1049.9 mm, with deviations within ±3%. Conversely, the SWM rainfall displays mixed results: SSP2-4.5 (972.5 mm) shows a slight overestimation (+1.3%), while SSP5-8.5 (882.3 mm) significantly underestimates (-8%) relative to the observed value of 959.6 mm.

Error statistics reveal that the *RMSE* for daily values is slightly more favorable in SSP2-4.5 (3.4 mm) than in SSP5-8.5 (3.8 mm), indicating a marginally better fit to observed daily variability. Moreover, for monthly *RMSE*, SSP2-4.5 performs significantly better (4.0 mm) compared to SSP5-8.5 (39.3 mm). This reinforces the conclusion that the bias-corrected data from SSP2-4.5 provides a more consistent alignment with historical observations, particularly at the monthly scale.

Overall, the bias correction approach enhances alignment with observations for both scenarios; however, SSP2-4.5 demonstrates greater reliability across daily, monthly, and seasonal timescales. The relatively high monthly *RMSE* and significant underestimation in mean monthly rainfall under SSP5-8.5 indicate a need for further calibration or alternative bias correction methods to improve model accuracy for this high-emission scenario.

TABLE 2. Bias Correction Analysis Results for SSP2-4.5 and SSP5-8.5

Evaluation Index	Observed	Bias-corrected Data	
		SSP2-4.5	SSP5-8.5
Daily Average Rainfall	3.4	5	5.4
Mean Monthly Rainfall	228.1	219.5	143.1
NEM Mean Rainfall	1049.9	1020	1068.3
SWM Mean Rainfall	959.6	972.5	882.3
RMSE for Daily	-	3.4	3.8
RMSE for Monthly	-	4	39.3

IV. CONCLUSION

Bias correction of GCM rainfall data is crucial for enhancing climate projections in areas such as water management, agriculture, and disaster preparedness. This study revealed the limitations of raw GCM outputs, which often overestimate or underestimate rainfall and struggle to capture localized extremes. Among the five evaluated bias correction methods, LS proved most effective, showing the lowest median *RMSE* and consistent performance across various rainfall regimes. Post-correction analyses indicated better alignment with observed data, particularly under the moderate SSP2-4.5 scenario.

However, the LS method has limitations, such as its inability to correct for variance and spatial biases related to topography, indicating a need for complementary approaches. The spatial analysis highlighted regional disparities, especially in northern areas, underlining local climate dynamics. Despite these challenges, the corrected data showed improved accuracy and better captured seasonal trends, particularly intensified extremes under SSP5-8.5.

These results underscore the importance of bias correction in climate studies and position the LS method as an effective tool for enhancing the reliability of GCMs in Sri Lanka. Future research should focus on hybrid methods for addressing non-linear biases and validating techniques across different geographical contexts for better climate projections.

ACKNOWLEDGMENT

The authors express their gratitude to the Director and staff of the Hydrology Division of the Irrigation Department of Sri Lanka for providing the essential data required for this study.

REFERENCES

[1] *IPCC*. 2023. Accessed: Apr. 04, 2023. [Online]. Available: https://dlib.hust.edu.vn/handle/HUST/21737

[2] J. Chen, F. P. Brissette, and P. Lucas-Picher, "Assessing the limits of bias-correcting climate model outputs for climate change impact studies," *J Geophys Res*, vol. 120, no. 3, pp. 1123–1136, Feb. 2015.

[3] X. Li and Z. Li, "Assessment of bias correction methods for high resolution daily precipitation projections with CMIP6 models: A Canadian case study," *J Hydrol Reg Stud*, vol. 58, Apr. 2025.

[4] A. Menapace *et al.*, "Review of bias correction methods for climate model outputs in hydrology," *J Hydrol (Amst)*, vol. 660, p. 133213, Oct. 2025.

[5] Y. Sun, K. Deng, K. Ren, J. Liu, C. Deng, and Y. Jin, "Deep learning in statistical downscaling for deriving high spatial resolution gridded meteorological data: A systematic review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 208, pp. 14–38, Feb. 2024.

[6] F. Wang, D. Tian, and M. Carroll, "Customized deep learning for precipitation bias correction and downscaling," *Geosci Model Dev*, vol. 16, no. 2, pp. 535–556, Jan. 2023.

[7] J. Chen, F. P. Brissette, D. Chaumont, and M. Braun, "Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America," *Water Resour Res*, vol. 49, no. 7, pp. 4187–4205, Jul. 2013.

[8] R. Rahimi, H. Tavakol-Davani, and M. Nasseri, "An Uncertainty-Based Regional Comparative Analysis on the Performance of Different Bias Correction Methods in Statistical Downscaling of Precipitation," *Water Resources Management*, vol. 35, no. 8, pp. 2503–2518, Jun. 2021.

[9] R. Mahmood, S. Jia, N. K. Tripathi, and S. Shrestha, "Precipitation extended linear scaling method for correcting GCM precipitation and its evaluation and implication in the transboundary Jhelum River basin," *Atmosphere (Basel)*, vol. 9, no. 5, Apr. 2018.

[10] P. Dhawan *et al.*, "A comprehensive comparison of bias correction methods in climate model simulations: Application on ERA5-Land across different temporal resolutions," *Heliyon*, vol. 10, no. 23, Dec. 2024.

[11] C. Teutschbein and J. Seibert, "Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods," *J Hydrol (Amst)*, vol. 456–457, pp. 12–29, Aug. 2012.

[12] G. Lenderink, A. Buishand, and W. Van Deursen, "Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach," 2007.

[13] S. Hagemann, C. Chen, J. O. Haerter, J. Heinke, D. Gerten, and C. Piani, "Impact of a statistical bias correction on the projected hydrological changes obtained from three GCMs and two hydrology models," *J Hydrometeorol*, vol. 12, no. 4, pp. 556–578, Aug. 2011.

[14] M. Déqué, "Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values," *Glob Planet Change*, vol. 57, no. 1–2, pp. 16–26, May 2007.

[15] M. Mendez, L. A. Calvo-Valverde, J. A. Hidalgo-Madriz, and J. A. Araya-Obando, "Assessing the Effect of Bias Correction Methods on the Development of Intensity–Duration–Frequency Curves Based on Projections from the CORDEX Central America GCM-RCM Multimodel-Ensemble," *Water (Switzerland)*, vol. 16, no. 23, Dec. 2024.

[16] F. Lehner, I. Nadeem, and H. Formayer, "Evaluating skills and issues of quantile-based bias adjustment for climate change scenarios," *Adv Stat Climatol Meteorol Oceanogr*, vol. 9, no. 1, pp. 29–44, Apr. 2023.

[17] E. Barrow, M. Hulmel, and M. Semenov, "Effect of using different methods in the construction of climate change scenarios: examples from Europe," 1996.

[18] A. Ruiter, "Delta-change approach for CMIP5 GCMs," 2012.

[19] T. M. Jakob, A. Gobiet, and A. Leuprecht, "Empirical-statistical downscaling and error correction of daily precipitation from regional climate models," *International Journal of Climatology*, vol. 31, no. 10, pp. 1530–1544, Aug. 2011.

[20] M. Mendez, B. Maathuis, D. Hein-Griggs, and L. F. Alvarado-Gamboa, "Performance evaluation of bias correction methods for climate change monthly precipitation projections over Costa Rica," *Water (Switzerland)*, vol. 12, no. 2, Feb. 2020.

[21] P. S. Smitha, B. Narasimhan, K. P. Sudheer, and H. Annamalai, "An improved bias correction method of daily rainfall data using a sliding window technique for climate change impact assessment," *J Hydrol (Amst)*, vol. 556, pp. 100–118, Jan. 2018.

[22] J. Schmidli, C. Frei, and P. L. Vidale, "Downscaling from GCM precipitation: A benchmark for dynamical and statistical downscaling methods," *International Journal of Climatology*, vol. 26, no. 5, pp. 679–689, Apr. 2006.

[23] C. Jayaminda, L. Gunawaradana, L. Rajapakshe, and J. Bamunawala, "Isohyetal maps for the assessment of Future Extreme Precipitation: A Case Study in the Kalu River Basin," in *National Conference on Disaster Risk Reduction and management in Sri Lanka; Science, policy and Practice*, 2023, pp. 3–4.

[24] R. Séférian *et al.*, "Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate," *J Adv Model Earth Syst*, vol. 11, no. 12, pp. 4182–4227, Dec. 2019.

[25] A. D. S. Iresh, J. Obeysekera, B. C. L. Athapattu, and W. C. D. K. Fernando, "Rainfall Intensity Duration Frequency Curves Using Cluster and Regional Frequency Analyses," *J Hydrol Eng*, vol. 29, no. 5, Oct. 2024.

[26] T. P. Burt and K. D. N. Weerasinghe, "Rainfall distributions in Sri Lanka in time and space: An analysis based on daily rainfall data," *Climate*, vol. 2, no. 4, pp. 242–263, Dec. 2014.

[27] A. D. S. Iresh, A. G. N. S. Marasinghe, A. M. T. S. H. Wedanda, G. P. Wickramasekara, M. D. J. P. Wickramasooriya, and M. T. C. Premathilake, "Development of a Hydrological Model for Kala Oya Basin Using SWAT Model," *Journal of the Institution of Engineers, Sri Lanka*, vol. 54, no. 01, pp. 57–65, 2021.

[28] A. D. S. Iresh, B. C. L. Athapattu, W. C. D. K. Fernando, and J. Obeysekera, "Analysis of Long-Term Seasonal Rainfall Trends in Sri Lanka after Eliminating the Serial Correlation Effect," *Engineer: Journal of the Institution of Engineers, Sri Lanka*, vol. 58, no. 1, pp. 71–84, Mar. 2025.

[29] 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 – Sri Lanka*, vol. 3, no. 2, 2007.

[30] A. D. S. Iresh, "Screening of Annual Rainfall Time-Series Data in Kala Oya Basin: Case Study in Sri Lanka," *Engineer*, vol. 53, no. 2, pp. 69–75, 2020.

[31] U. Schulzweida, "CDO User Guide," Oct. 2023, *Zenodo*: 2.3.0. Accessed: Aug. 05, 2025.

[32] F. Kaspar, D. Wetterdienst, S. Profile, and U. Schulzweida, "'Climate data operators' as a user-friendly processing tool for CM SAF's satellite-derived climate monitoring products," in *In Conference: EUMETSAT Meteorological Satellite Conference 2010*, 2010, pp. 20–24.

[33] H. Rathjens, K. Bieger, R. Srinivasan, I. Chaubey, and J. G. Arnold, "CMhyd User Manual Documentation for preparing simulated climate change data for hydrologic impact studies," May 2016.

[34] K. E. Taylor, "Summarizing multiple aspects of model performance in a single diagram," *Journal of Geophysical Research Atmospheres*, vol. 106, no. D7, pp. 7183–7192, Apr. 2001.

[35] M. Shrestha, S. C. Acharya, and P. K. Shrestha, "Bias correction of climate models for hydrological modelling – are simple methods still useful?" *Meteorological Applications*, vol. 24, no. 3, pp. 531–539, Jul. 2017.

[36] G. H. Fang, J. Yang, Y. N. Chen, and C. Zammit, "Comparing bias correction methods in downscaling meteorological variables for hydrologic impact study in an arid area in China," *Hydrology and Earth System Science*, vol. 11, pp. 12659–12696, Nov. 2014.