Comparison of bias correction techniques for rainfall over the Lower Mekong River Basin in Cambodia

Sophanny Phan[1] Supattra Visessri [1, 2, *]and Ty Sok[3]

[1]*Department of Water Resources Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, THAILAND*
[2] *Center of Excellence in Disaster and Risk Management Information Systems, Chulalongkorn University, Bangkok 10330, THAILAND*
[3] *Faculty of Hydrology and Water Resources Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, CAMBOODIA*
**Corresponding author; E-mail address: Supattra.vi@chula.ac.th*

Abstract

Rainfall is a key factor for water resources management and crucial for informed disaster management planning. The analysis of flood and drought conditions heavily relies on accurate and comprehensive rainfall datasets. According to the limitation of rainfall data scarcity in certain regions, using secondary datasets of rainfall data from satellites and reanalysis products could offer useful alternative in providing a wider coverage with acceptable resolution. Several bias correction techniques have been applied to further improve the accuracy of secondary rainfall data. This study aims to compare two techniques of bias correction which are quantile mapping (QM), and Bayesian theorem applied between observed rainfall and reanalysis rainfall. QM is a statistical transformation that attempts to match between model output and observation data using cumulative distribution function (CDF), while Bayesian theorem is a probabilistic framework that allows the integration of prior information with new data to generate an improved posterior distribution. The comparison of the two bias correction techniques is demonstrated using a case study in the Lower Mekong River Basin in Cambodia, which covers the area of 122,000 km^2 with available rainfall data from 1985-2022. The results illustrate Bayesian method performs better than QM method for all statistical performance indicators including R (0.90-1.00), NSE (0.81-0.99), and RMSE (6.00-37.50mm) over the regions of the study area. The findings from this study demonstrate the potential of using bias-corrected data in enhancing disaster planning and mitigation efforts, particularly for drought and flood in data scarce regions

.

Keywords: Bias correction, Quantile mapping, Bayesian theorem, Cambodia, Mekong River basin

1. Introduction

The analysis of flood and drought conditions heavily relies on accuracy and comprehensive rainfall datasets. The effectiveness of drought monitoring and flood warning systems through ground observation is constrained by their limited and inconsistent coverage [1, 2]. Satellite rainfall and reanalysis data offers possibility for hydrological assessment particularly for regions lacking rain gauge networks, though they are indirect measurements of rainfall and often available at a coarse scale [3, 4].

Rain gauges are essential for providing direct, local information on rainfall variability and serve as reference data for verifying both satellite-derived and reanalysis rainfall estimates. However, due to historical political instability and conflict in Cambodia, the observation network was abandoned, and gauging equipment destroyed, leading to a reduced number of operational rain gauges [5]. Consequently, rain gauge observations have predominantly been limited to populated and low-altitude areas until recently. These constraints have compelled previous studies to rely on rain gauge data from neighboring countries, satellite data, or reanalysis data like ERA5 [6, 7]. This reliance is typical in developing countries like Cambodia, where ground observations may not always be feasible due to financial constraints, even in the context of frequent droughts that cause significant damage [8, 9]. Cambodia is highly vulnerable to climate change-related to disaster, with its condition leading to have high frequency on drought. In 2015-2016, Cambodia experienced its worst drought due to El Niño phenomenon, leading to reduce rainfall and warm temperature. This drought event affected to over 2 million people, agriculture, and crop damage [10].

Given the limitations of ground-based observations, satellites and reanalysis data serve as vital secondary sources for rainfall data, necessitating the improvement of derived rainfall estimates through bias correction. Various methods are available to enhance the accuracy of these data, including linear scaling, quantile mapping (QM) [11], delta quantile mapping (DQM) [12], regression [13], and Bayesian methods [14]. While traditional bias correction methods such as Bayesian and QM have been widely applied, recent research highlights their limitations in handling extremes, non-stationarity, and multivariate dependencies. To address these challenges, machine learning techniques, copula-based, and hybrid correction method have been developed [15-17]. These emerging methods offer enhanced flexibility and abruptness, strengthening the contemporary relevance and effectiveness of bias correction practices. These approaches help bridge the gap caused by the sparse ground observation network, providing more reliable data for environmental and disaster management planning.

In this study, the Bayesian and QM methods were selected due to their widespread use, conceptual simplicity, strong foundation basis in probabilistic adjustment, and proven effectiveness. They have been found to improve the reliability of rainfall data in sparse rain gauge region [8, 11]. Bayesian approach provides a rational framework for personal beliefs in uncertainty situations, avoiding behavioral inconsistencies Bayesian theorem explains how to update the believable in the model based on new data or information [18, 19]. Quantile mapping (QM) method is a common approach for correcting bias by transferring function to match the cumulative distribution function (CDF) of gridded rainfall from any satellite or reanalysis product with to that of the observed rainfall [19-21].

The objective of this study is to compare two techniques of bias correction which are Bayesian theorem and QM approach applied between observed rainfall and reanalysis rainfall data.

2. Methodology

2.1 Study Area

Cambodia is located in Southeast Asia, between 10°N- 15°N latitudes and 102°E-108°E longitudes. It is consisted of expansive river basins and plains bordering Loas, Thailand, Vietnam, and Gulf of Thailand [22]. There are five main river basins for the entire Cambodia including Tonle Sap, 3S (Se San, Se Kong, Sre Pok), Upper Mekong, Cambodia Mekong Delta, and Coastal Zone.

This study is conducted in the Cambodia lower Mekong River basin which covers both the Tonle Sap River basin and Mekong delta River basin with a total area of 122, 000 km^2. The Tonle Sap River basin covers the Tonle Sap Lake ('Great Lake'), and 11 major tributaries with an approximated area of 86,000 km^2 [23]. The Mekong Delta River basin covers an area about 36,000 km^2 in southern part of Cambodia with the entire mainstream of Mekong Delta River basin is about 297 km long [24].

Fig. 1 shows the topography of Cambodia lower Mekong River basin. This study area originated from a high elevation in mountainous southwest, ranging approximately from 1800 meters to a low elevation of mean sea level in the lowland area. There are a total of 221 observed rainfall stations with 159 gridded rainfalls.

Cambodia's average annual rainfall is 1,836 mm over the current climatology (1985-2022). The level of annual rainfall occurs based on each region in Cambodia. For instance, the southwestern coastal and highland received annual rainfall of more than 2,000 mm, while annual rainfall in eastern mountains and plains ranged from 1,500 mm to 1,900. In the central plain of the country received amount of annual rainfall between 1,300 mm to 1,800 mm, less than others two regions [25].

Cambodia experiences a moist tropical monsoon climate, with subtropical condition at higher elevation in average annual temperature about 27.41°C [25]. As the Intertropical Convergence Zone (ITCZ) migrates south, characterized Cambodia within two seasons which are dry season (November to April) and wet season (May to October). The minimum average temperature varies from (24°C to 35°C) in the warmest month in April, while the coolest month of December, the range of maximum average temperature is 20°C to 30°C [25].

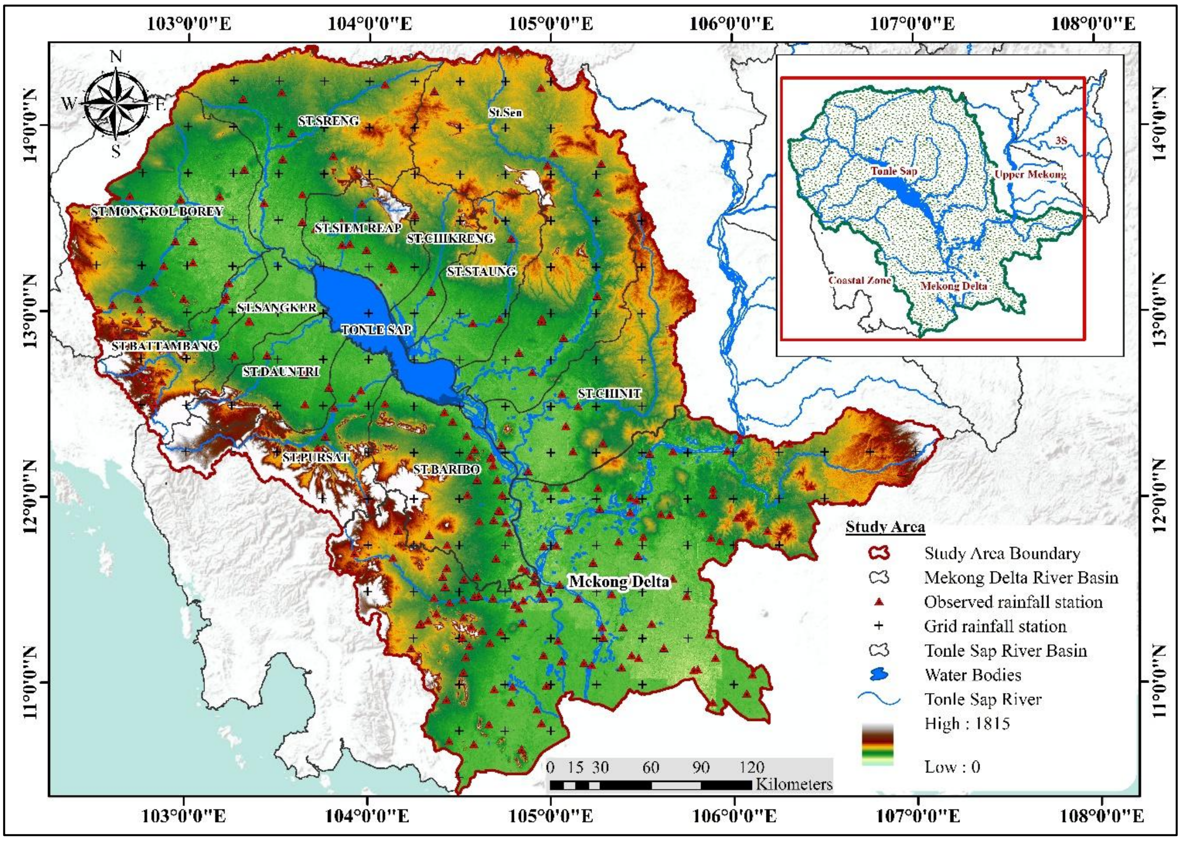

Fig. 1 Description of study area map

2.2 Dataset

In this study, there are two types of meteorological data as illustrated in Table 1 which are observed rainfall data and gridded rainfall data. Observed and gridded rainfall are obtained from MOWRAM and ERA-5 reanalysis product, respectively. Observed rainfall data has 221 rainfall stations in 1985 to 2022 with 38 years at the daily time step. Gridded rainfalls cover the same period as observed rainfall with their spatial resolution at 25km x 25km which is available 159 grids cover the study area at daily time step. The ERA5 reanalysis dataset with spatial resolution of 25kmx25km was considered for the lower Mekong River Basin due to this study area's relatively flat topography and modest regional climate variability. Additionally, the 25km^2 resolution adequately capture the synoptic-scale weather system such as monsoon rainfall pattern, that govern the majority of the study area's climate and hydrological processes.

Table 1 Data availability for rainfall over Cambodia lower Mekong River Basin

Datasets	Sources	Resolution	
		Temporal	Spatial
Observed Rainfall	Ministry of Water Resources and Meteorology (MOWRAM)	Daily	Point (221 stations)
Reanalysis rainfall	ERA-5 Reanalysis	Daily	25km x 25km (159 Grids)

2.3 Data processing

Bias correction on rainfall dataset allows the researchers to compare observed rainfall from MOWRAM, represented as point data, and reanalysis rainfall from ERA-5, which is provided as grid data. The observed rainfall is transformed from point measurements to areal gridded rainfall using the simple average method. This approach facilitates the comparison and subsequent adjustment of discrepancies between observed and reanalysis data. The simple average is simplest and widely used method of computing the average rainfall over the area. This suitable method stems from its appropriateness in the context of this study area's predominantly flat topography, where rain gauges area uniformly distributed, and rainfall patterns display remarkable regularity. Consequently, the results obtained from this method could be expected to provide sufficient satisfactory. However, this method has limited accuracy because it depends on rainfall distribution and the size of the study area. It also requires a sufficient number of observation stations to ensure reliable and consistent results [26]. Simple average method allows the observed rainfall to be converted into grid rainfall where the grid size of the spatial resolution of both datasets are matched. These two datasets can be used to compare and preprocess on bias correction between observed and reanalysis rainfall data. Double mass curve analysis is commonly used to confirm this relative consistency [27]. Double Mass Curve method is applied in this study to verify data consistency [28].

In terms of temporal resolution, both daily observed and reanalysis rainfall data are converted into monthly rainfall, as the further study are expected to focus on drought analysis. Additionally, performance at daily timesteps tend to show low accuracy compared to monthly timesteps. Therefore, using monthly rainfall data is deemed more suitable for achieving better performance and for accurately assessing drought analysis in future studies.

2.4 Bias correction method

2.4.1 Bayesian bias correction

The Bayesian approach provides a probabilistic framework which allows for integration of prior information with new data to improve the estimation between ground observations and satellite data. The the normal distribution may not perfectly capture the intricacies of rainfall data, its use in initial Bayesian bias correction efforts can be justified by its simplicity, mathematical properties, and the practicality of establishing a baseline model from which more refined models can be developed. This approach allows for an iterative improvement in model sophistication as more about the data's behavior is learned. All the probability of posterior, prior, and likelihood function are assumed to follow normal distribution. It is because normal distribution is a simple distribution characterized by mean and variance. It can be used to estimate variance of error and normalization weighting factor. When assuming a normal distribution for the errors in a model, it allows for a straightforward estimation of the variance. This variance quantifies the expected magnitude of deviations or errors around the mean, which is crucial in assessing the accuracy and reliability of the model predictions. Knowing the error variance helps in understanding the confidence levels and uncertainty in the predictions. Normalization often involves adjusting data from different sources or scales to a common scale, typically to ensure that each dataset contributes proportionately to the final results. When assuming a normal distribution, the normalization can use the variance or standard deviation as a weighting factor, which helps in minimizing the influence of outliers and scaling the data effectively. This method ensures that all data points are treated uniformly in terms of their relative importance or influence on the model. This assumption of normal distribution is robust in reprocessing algorithms for hydrological component, especially for rainfall data. However, it is noted that the Bayesian method could be sensitive to the choice of prior distributions [17]. and may struggle to adequately update the prior beliefs, leading to persistent uncertainties or skewed corrections especially in regions with sparse or unevenly distributed ground observed [8]. In this study, Bayesian method is applied to correct bias between observed rainfall data and ERA-5 reanalysis data over a period of 38 years.

The fundamental concept of Bayesian method was described by [18] and summarized as shown in Eq.(1):

$$P(S|O) = \frac{P(O|S) \times P(S)}{P(O)} \tag{1}$$

where, O refers to observed rainfall, S refers to reanalysis rainfall.

- $P(S|O)$ is the posterior probability of reanalysis rainfall estimates being corrected given observed rainfall data.
- $P(O|S)$ is the likelihood of observed rainfall data given reanalysis rainfall estimates.
- $P(S)$ is the prior probability of the rainfall reanalysis rainfall estimates being corrected.
- $P(O)$ is the probability of observed rainfall data under all possibles reanalysis rainfall estimates.

2.4.2 Quantile mapping (QM)

QM is a statistical transformation that attempts to find a function that maps the model output to a new distribution such that the resulting distribution matches that of observations [29, 30]. QM algorithms use empirical quantile mapping to align the cumulative distribution function (CDF) of model output with observations. In general, QM applies to all possible meteorological parameters, while applications based on CDF may become problematic for parameters that do not fit theoretical functions [31]. Therefore, empirical quantile mapping (EQM) is used instead of assuming parametric distributions due to QM-based statistical bias correction has been widely used, and its performance was found to be satisfactory in comparison to the other methods [30, 31]. The limitation of QM is their assumption of bias stationarity. They presume that biases observed in historical data will remain unchanged in the future. This assumption may not hold true under changing climate conditions, potentially leading to inaccurate projection [32]. Furthermore, while QM effectively adjust the statistical distribution of variables, it often fails to preserve important temporal characteristics, such as seasonal patterns or variability, which are critical for capturing the full dynamic of climate change [33]. In this study, EQM is applied to correct biases between observed rainfall data and ERA-5 reanalysis data derived rainfall over a period of 38 years.

Mathematically, EQM is formulated based on the concept [34] as shown in Eq.(2):

$$X_{corrected} = F_{obs}^{-1}\left(F_{sat}(x_{sat})\right) \tag{2}$$

where,

- $X_{corrected}$ is bias-corrected reanalysis rainfall estimates
- F_{obs}^{-1} is the inverse of cdf or quantile function of observed rainfall
- x_{sat} is reanalysis rainfall estimates
- F_{sat} is cdf or quantile function of reanalysis rainfall estimates

3. Results and Discussion

3.1 Bayesian bias correction

Fig. 2 illustrates the effectiveness of Bayesian bias correction between observed and reanalysis rainfall over the Lower Mekong River Basin in Cambodia from 1985 to 2022. The blue dots represent the likelihood distribution of the observed rainfall, while the red dots and green diamonds refer to the prior distribution of reanalysis rainfall and the posterior distribution of corrected rainfall from reanalysis data, respectively. Fig. 2 shows the quantitative relationship between monthly rainfall on the x-axis ranging from 0 to 300mm and cdf on the y-axis ranging from 0 to 1. The posterior distribution shown in Fig. 2 closely follows the observed rainfall distribution, demonstrating that Bayesian method effectively reduces the biases in reanalysis data. For example, at the same 20th percentile, the reanalysis data overestimates monthly rainfall at approximately 110 mm, compared to the corrected value of around 100 mm. As expected, the posterior distribution is updated from reanalysis rainfall to closely match with corrected rainfall. The cumulative distribution function (cdf) curves show that the bias correction significantly improves the alignment of corrected rainfall with the observed rainfall. The bias-corrected results based on Bayesian technique agree very well with observed rainfall, indicating that this bias correction technique effectively reduces discrepancies and increases accuracy in rainfall data. Good performance of the Bayesian approach obtained from this study support previous research in which the Bayesian approach is recommended to adjust variability in rainfall estimates [8].

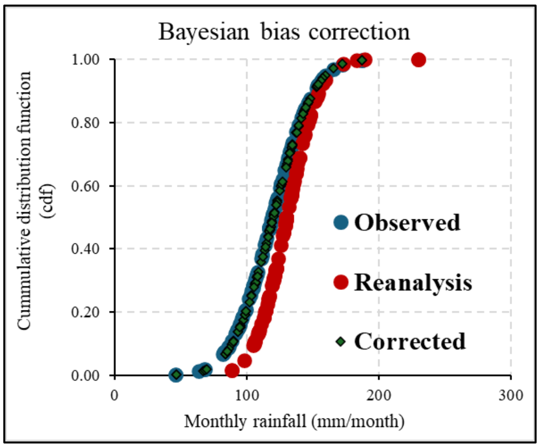

Fig. 2 Average corrected rainfall all grids using Bayesian approach

3.2 Quantile mapping bias correction

Quantile mapping bias correction is applied to monthly average rainfall data for all grids by using cumulative distribution function (cdf). In Fig. 3 with monthly rainfall on the x-axis (0-400 mm) and cdf on the y-axis (0-1) demonstrates the improvement in bias correction. The observed, reanalysis, and corrected rainfall are represented by blue dots, red dots, and green diamonds, respectively. At the same cdf percentile, the monthly rainfall of reanalysis data and corrected data have different rainfall amount. This shift indicates that EQM method effectively reduces the overestimation of corrected rainfall in the reanalysis data, aligning the observed rainfall more closely with corrected rainfall. This suggests that the empirical quantile mapping (EQM) method effectively enhances the accuracy of reanalysis rainfall by adjusting them to better much with observed rainfall. The EQM algorithm adjusted the total rainfall amount after correcting the reanalysis data, refining the various statistical characteristics and the narrowest ranges of variability, achieving the best match with the ensemble mean [34, 35].

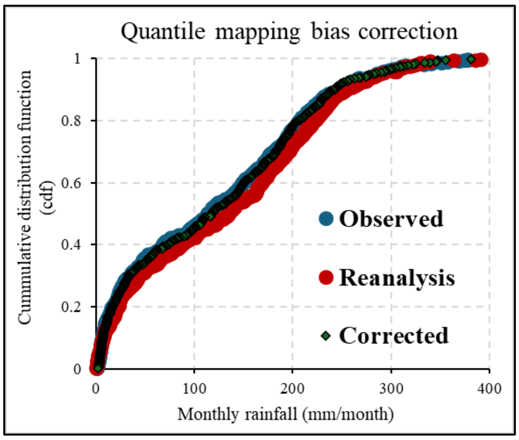

Fig. 3 Average corrected rainfall all grids using QM approach

3.3 Temporal performance between Bayesian and QM

Table 2 presents the seasonal statistics performance of the Bayesian and QM bias correction techniques. In overall, both methods perform better during the dry season compared to the wet season across all statistics performance (R, NSE, and RMSE). The Bayesian method consistently outperforms QM in both dry and wet seasons, achieving higher R and NSE values and significantly lower RMSE. During the dry season, Bayesian shows near-perfect performance (R = 0.99, NSE = 0.98 and RMSE = 1.21 mm), while in the wet season, it maintains strong accuracy (R = 0.99, NSE = 0.97, and RMSE = 4.06 mm), clearly outperforming QM, especially in reducing biases. These results highlight that the Bayesian method is more robust across both seasons, offering high reliability for both dry and wet periods, whereas QM tends to experience a notable performance decline, especially during the wet season when rainfall variability is higher.

Table 2 Seasonal performance in both Bayesian and QM method

	Dry (Nov-Apr)		Wet (May-Oct)	
Statistical performance	Bayesian	QM	Bayesian	QM
R	0.99	**0.96**	0.99	0.98
NSE	0.98	0.94	0.97	**0.82**
RMSE	1.21	8.82	4.06	**15.47**

3.4 Spatial performance between Bayesian and QM

Fig. 4 illustrates the comparison performance between two bias corrections techniques which are Bayesian and EQM method over Cambodia lower Mekong River Basin. There are 3 subregions including Northern Tonle Sap, Southern Tonle Sap, and Mekong Delta. These maps evaluate the performance based on three statistical performances including correlation coefficient (R), Nash-Sutcliffe efficiency coefficient (NSE), and root mean square error (RMSE). R explains how well the corrected rainfall correlates with observed rainfall. NSE describes how well the corrected rainfalls matches the observed rainfall. RMSE is an indicator of the average magnitude of errors in the corrected rainfall.

The colors coding demonstrates the performance of bias correction using Bayesian and EQM methods across different regions in Cambodia lower Mekong River basin, with dark blue indicating strong performance and light blue indicating weak performance for R and NSE. However, RMSE is a special case among the other two, where dark blue represents the weak performance and light blue indicates high performance. Thus, weak performance of RMSE signifies high error, whereas strong performance indicates low error in the region.

3.4.1. Correlation coefficient (R)

The Bayesian method for bias correction consistently shows the best performance with very high R value ranges from 0.9 to 1, showing that the corrected rainfall strongly correlates with observed rainfall mostly in all grids as shown in Fig. 4(a). However, a few grids located in Northern Tonle Sap shows weak correlation, where in Fig. 4(b) the QM exhibits slightly lower performance with values ranging from 0.8 to 0.98. Strong performance of R revealed in serval grids in Mekong Delta while most grids in the region shows moderate performance. Only a few grids are observed weak R values, located in Northern Tonle Sap and Southern Tonle Sap. Thus, weak performance of R occurs in the same region for both Bayesian and QM method, due to a few grids situated in front of mountain with steep slope, which can cause error between ground observed rainfall and rainfall from reanalysis product.

3.4.2. Nash-Sutcliffe efficiency (NSE)

NSE assesses the accuracy of observed rainfall correction using Bayesian method. Fig. 4(c) illustrates that most grids across all regions predominantly feature dark blue, representing high NSE values ranging from 0.81 to 0.99. This color indicates that the corrected rainfall closely aligns with the observed rainfall, ensuring high reliability and effectiveness within all regions. In contrast, a few grids particularly in Northern Tonle Sap denote lower NSE values, meaning that less accurate predictions and suggesting that rainfall from reanalysis product may not fully capture the rainfall dynamics in those grids. Following the EQM method as shown in Fig. 4(d), NSE performance varies from 0.60 to 0.99, signaling a range from weak to strong performance. The most robust NSE results appear in the Mekong Delta, while only one grid received weak performance of NSE in Southern Tonle Sap. This low score could be due to the scrubland impacting the accuracy of both rain gauge and grid rainfall data. Meanwhile, grids in all regions generally show a moderate performance of NSE, reflecting a varied effectiveness across the region.

3.4.3. Root means square error (RMSE)

RMSE is depicted with an inverse color scheme compared to the other two statistical performance, where dark blue signifies weak performance and light blue indicates strong performance. When the Bayesian method is applied, regions depicted in dark blue, such as Mekong Delta show RMSE values as high as 37.5, suggesting significant error between corrected and observed rainfall as shown in Fig. 4(e). Conversely, Fig. 4(f), the area in light blue, potentially in Northern and Southern Tonle Sap, demonstrate low RMSE values around 6, indicating that the corrected rainfalls are both accurate and reliable. The EQM method is used to correct rainfall, assessed through RMSE value that range significantly. A value of 16.1 indicates weak performance but reflects low error, suggesting minimal deviation from corrected rainfall. On the other hand, a high RMSE at 65.8 indicates strong performance but it is characterized by substantial error, emphasizing significant deviation in Mekong Delta. EQM exhibits a consistently robust performance across the study area, especially in Mekong delta, which reflects a moderate performance according to the RMSE result. This result highlights that the most significant error is found in Mekong Delta due to the complex topography of the region. Therefore, the comparison between these two methods of bias correction, indicating that all the performance of Bayesian method has high potential and performance well than QM method. This suggests that the Bayesian method is more effective in correcting bias observed rainfall in these specific regions.

3.4.4. Discussion on the performance statistics

In overall, lower value for all statistics performance mostly occurs in Northern part and the highest performance is located in the Southern part and Mekong delta in both bias correction techniques. The southern part performs well under both bias correction techniques, as this region has a higher density of rainfall stations, which leads to greater accuracy in the corrected rainfall. The medium range of these statistics are performed in central part of the study area. According to this region mostly covers with the mountainous area which caused error to rainfall pattern. Bayesian method is suitable for reduction of systematic error, especially in high elevation regions. However, it is sensitive to sparse rain gauge distribution in regions of deep convection system due to cirrus effects that cause overestimation while the limitation of QM method may struggle with non-stationarity, and it has random error caused by day-to-day precipitation variations in sparse rain gauge in mountainous area. It is suggested to have the reference data of a higher or equal spatial scale for higher rainfall spatial representations [8]. Therefore, the potential directions for further research could be incorporating multiple satellite data sources and hybrid corrections techniques to further improvement. The key finding of corrected rainfall datasets can be used as input into hydrological models such as SWAT, HEC-HMS to improve the accuracy of runoff simulation, flood forecasting, and drought prediction. Moreover, it can elaborate on integrating into enhancing disaster planning and mitigation efforts, particularly regarding drought and flood management.

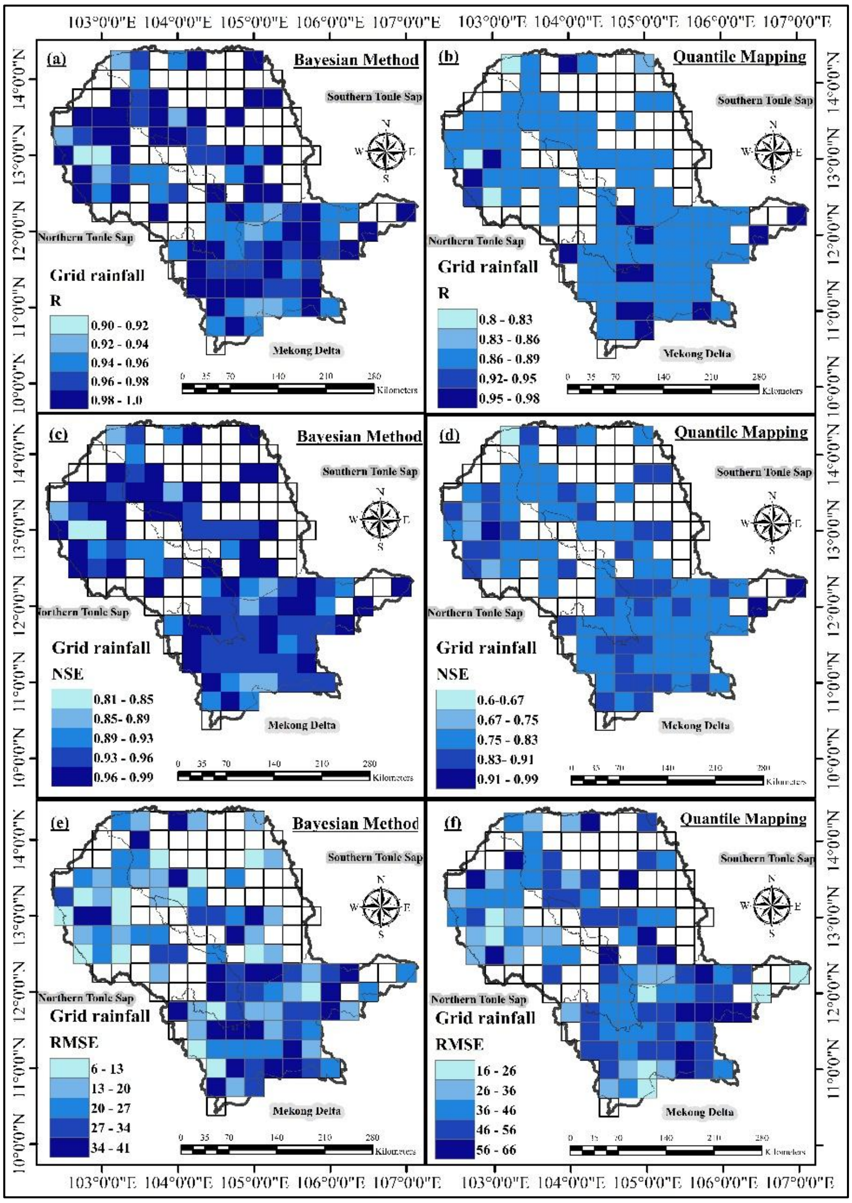

Fig. 4 Spatial performance patterns over Cambodia lower Mekong River Basin

4. Conclusions

The effectiveness of drought monitoring and flood warning system through ground observations is limited by their inconsistent and inadequate coverage. This study examines the comparative accuracy between two bias correction techniques which are Bayesian and QM approach. The observed rainfall is having 221 available rainfall stations and reanalysis rainfall consists of 159 gridded rainfalls over the study area. According to the corrected rainfall from reanalysis data, both Bayesian and QM methods adjusts to closely match between observed and reanalysis rainfall. Although the performance of these two techniques is similar; however, the Bayesian method achieves higher accuracy than the QM method. This is evidenced by 3 statistical performances used to test including R, NSE, and RMSE. The Bayesian method's performance shows R (0.9-1) and NSE (0.81-0.99), both slightly higher than performance of QM method. Particularly notable is the RMSE performance of Bayesian method is RMSE (6-37.5) indicate low error compared to RMSE (16-65.8) of QM indicates high error. Therefore, the Bayesian method provides more suitable for correcting between observed and reanalysis data in this study area. Better performance was observed in regions of the study area with a higher density of rain gauge stations, particularly in low attitude zones, where it greatly aids in identifying and predicting flash drought and floods early-warning systems. To further enhance rainfall accuracy, it is recommended that the government prioritize the installation of additional rain gauge stations, especially in mountainous and high attitude areas. Increasing the density of ground observed in these regions would importantly improve the capture of spatial rainfall variability, leading to more accurate forecasting and better support for disaster preparedness and climate resilience efforts. In addition to installing rain gauges, the two bias correction methods used in this study should be considered by agencies such as MOWRAM and NGOs, as they offer complementary strengths that can enhance climate risk management efforts. By combining these two techniques, agencies can create more reliable early warning systems and strengthen climate resilience, particularly by integrating bias correction into operational workflows and collaborating with research institutes.

Acknowledgement

The author would like to thank the Ministry of Water Resources and Meteorological (MOWRAM) for providing observed rainfall and the reviewers for constructive comments.

References

[1] H. Guo *et al.*, "Meteorological drought analysis in the Lower Mekong Basin using satellite-based long-term CHIRPS product," *Sustainability,* vol. 9, no. 6, p. 901, 2017.

[2] R. Sun, H. Yuan, X. Liu, and X. Jiang, "Evaluation of the latest satellite–gauge precipitation products and their hydrologic applications over the Huaihe River basin," *Journal of Hydrology,* vol. 536, pp. 302-319, 2016.

[3] M. Saber and K. K. Yilmaz, "Evaluation and bias correction of satellite-based rainfall estimates for modelling flash floods over the Mediterranean region: Application to Karpuz River Basin, Turkey," *Water,* vol. 10, no. 5, p. 657, 2018.

[4] H. Vergara *et al.*, "Effects of resolution of satellite-based rainfall estimates on hydrologic modeling skill at different scales," *Journal of Hydrometeorology,* vol. 15, no. 2, pp. 593-613, 2014.

[5] (2024). *Weather Observation and Forecasting in Cambodia*

[6] T. Satomura, K. Yamamoto, B. Sysouphanthavong, and S. Phonevilay, "Diurnal variation of radar echo area in the middle of Indochina," *Weather Records. Vol. 2,* vol. 89, pp. 299-305, 2011.

[7] H. G. Takahashi, "Orographic low-level clouds of Southeast Asia during the cold surges of the winter monsoon," *Atmospheric research,* vol. 131, pp. 22-33, 2013.

[8] M. W. Kimani, J. C. Hoedjes, and Z. Su, "Bayesian bias correction of satellite rainfall estimates for climate studies," *Remote sensing,* vol. 10, no. 7, p. 1074, 2018.

[9] N. Chhinh and A. Millington, "Drought monitoring for rice production in Cambodia," *Climate,* vol. 3, no. 4, pp. 792-811, 2015.

[10] UNDRR, *Disaster risk reduction in Cambodia: Status report 2019*. 2019.

[11] N. Y. Nguyen, T. N. Anh, H. D. Nguyen, and D. K. Dang, "Quantile mapping technique for enhancing satellite-derived precipitation data in hydrological modelling: a case study of the Lam River Basin, Vietnam," *Journal of Hydroinformatics,* vol. 26, no. 8, pp. 2026-2044, 2024.

[12] A. J. Cannon, S. R. Sobie, and T. Q. Murdock, "Bias correction of GCM precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes?," *Journal of Climate,* vol. 28, no. 17, pp. 6938-6959, 2015.

[13] S. Chen, L. Xiong, Q. Ma, J.-S. Kim, J. Chen, and C.-Y. Xu, "Improving daily spatial precipitation estimates by merging gauge observation with multiple satellite-based precipitation products based on the geographically weighted ridge regression method," *Journal of Hydrology,* vol. 589, p. 125156, 2020.

[14] Y. Ma *et al.*, "Comprehensive evaluation of ensemble multi-satellite precipitation dataset using the dynamic bayesian model averaging scheme over the Tibetan Plateau," *Journal of hydrology,* vol. 556, pp. 634-644, 2018.

[15] M. Holthuijzen, B. Beckage, P. J. Clemins, D. Higdon, and J. M. Winter, "Robust bias-correction of precipitation extremes using a novel hybrid empirical quantile-mapping method: Advantages of a linear correction for extremes," *Theoretical and Applied Climatology,* vol. 149, no. 1, pp. 863-882, 2022.

[16] Q. Guo, J. Chen, X. Zhang, M. Shen, H. Chen, and S. Guo, "A new two-stage multivariate quantile mapping

method for bias correcting climate model outputs," *Climate Dynamics,* vol. 53, pp. 3603-3623, 2019.

[17] J. Carter, E. A. Chacón-Montalván, and A. Leeson, "Bayesian hierarchical model for bias-correcting climate models," *Geoscientific Model Development,* vol. 17, no. 14, pp. 5733-5757, 2024.

[18] D. J. C. Mackay, *Bayesian methods for adaptive models*. California Institute of Technology, 1992.

[19] X. Li, H. Wu, N. Nanding, S. Chen, Y. Hu, and L. Li, "Statistical bias correction of precipitation forecasts based on quantile mapping on the sub-seasonal to seasonal scale," *Remote Sensing,* vol. 15, no. 7, p. 1743, 2023.

[20] R. L. Wilby, L. E. Hay, and G. H. Leavesley, "A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River basin, Colorado," *Journal of Hydrology,* vol. 225, no. 1-2, pp. 67-91, 1999.

[21] M. Minville, F. Brissette, and R. Leconte, "Uncertainty of the impact of climate change on the hydrology of a nordic watershed," *Journal of hydrology,* vol. 358, no. 1-2, pp. 70-83, 2008.

[22] C. Phoeurn and S. Ly, "Assessment of satellite rainfall estimates as a pre-analysis for water environment analytical tools: A case study for Tonle Sap Lake in Cambodia," *Engineering Journal,* vol. 22, no. 1, pp. 229-241, 2018.

[23] C. Oeurng, T. A. Cochrane, S. Chung, M. G. Kondolf, T. Piman, and M. E. Arias, "Assessing climate change impacts on river flows in the Tonle Sap Lake Basin, Cambodia," *Water,* vol. 11, no. 3, p. 618, 2019.

[24] Z. Xu *et al.*, "Morphological characteristics of Cambodia Mekong Delta and Tonle Sap Lake and its response to river-lake water exchange pattern," *Journal of Water Resource and Protection,* vol. 12, no. 04, p. 275, 2020.

[25] Cambodia, *The World Bank Group: Climate Risk Profile*. 2024.

[26] I. G. N. P. Dharmayasa, P. I. D. Putri, I. P. Sugiana, R. Jindal, K. Surakit, and R. Thongdara, "Comparing Areal Rainfall Estimation Methods in the Ayung Watershed, Bali, Indonesia: A Comprehensive Analysis," in *2024 10th International Conference on Smart Computing and Communication (ICSCC)*, 2024: IEEE, pp. 212-217.

[27] R. K. Linsley Jr, M. A. Kohler, and J. L. Paulhus, *Hydrology for engineers: SI metric edition*. McGraw-Hill, 1988.

[28] J. M. Bernardo and A. F. Smith, *Bayesian theory*. John Wiley & Sons, 2009.

[29] V. Mishra, U. Bhatia, and A. D. Tiwari, "Bias-corrected climate projections for South Asia from coupled model intercomparison project-6," *Scientific data,* vol. 7, no. 1, p. 338, 2020.

[30] H. Tiwari and B. K. Pandey, "Non-parametric characterization of long-term rainfall time series," *Meteorology and Atmospheric Physics,* vol. 131, pp. 627-637, 2019.

[31] M. J. Themeßl, 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, 2011.

[32] M. B. Switanek *et al.*, "Scaled distribution mapping: a bias correction method that preserves raw climate model projected changes," *Hydrology and Earth System Sciences,* vol. 21, no. 6, pp. 2649-2666, 2017.

[33] X. Wu, Z. Liu, and Q. Duan, "A combined wavelet analysis–quantile mapping (WA-QM) method for bias correction: capturing the intra-annual temporal patterns in climate model precipitation simulations and projections," *Environmental Research Letters,* 2025.

[34] C. Piani *et al.*, "Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models," *Journal of hydrology,* vol. 395, no. 3-4, pp. 199-215, 2010.

[35] C. Teutschbein and J. Seibert, "Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods," *Journal of hydrology,* vol. 456, pp. 12-29, 2012.