Formosa Journal of Applied Sciences (FJAS)
Vol. 5, No. 10 2026: 2255-2274

# Improving Satellite Based Rainfall Estimates for Flood Forecasting in Data Scarce Tropical Watersheds

Arifin[1]*, Asikin Muchtar[2]
[1]Universitas Achmad Yani Banjarmasin
[2]Universitas Indonesia Timur
**Corresponding Author:** Arifin arifin.mt@uvayabjm.ac.id

A R T I C L E I N F O

*Keywords:* GPM-IMERG, Extreme Rainfall, Bias Correction, Flood Simulation, Data-Scarce Watersheds

*Received : 22, June*
*Revised : 24, August*
*Accepted: 26, October*

©2026 Arifin, Muchtar: This is an open-access article distributed under the termsof the Creative Commons Atribusi 4.0 Internasional.

A B S T R A C T

Satellite rainfall estimates support flood analysis in tropical watersheds where ground observations are limited, yet biases during intense rainfall can distort simulated flood peaks. This study assesses whether event-focused bias correction improves GPM-IMERG V07 estimates and flood simulation in a data-scarce Indonesian watershed. Daily satellite rainfall, observations from four rain gauges, and streamflow records for 2019–2025 are compared using uncorrected IMERG, conventional bias correction, and event-focused correction. Rainfall accuracy is evaluated using RMSE, PBIAS, and critical success index, while HEC-HMS simulations are assessed using Nash–Sutcliffe efficiency and peak-discharge error. Gauge-reduction experiments examine performance under limited monitoring. This framework identifies when rainfall correction offers meaningful benefits for flood simulation.

DOI prefik: https://doi.org/10.55927/fjas.v5i10.143
ISSN-E: 2962-6447
https://journalfjas.my.id/index.php/fjas

## INTRODUCTION

Flooding remains a major hydrological challenge in tropical regions, where intense rainfall, complex watershed characteristics, and limited monitoring infrastructure make reliable flood prediction difficult. According to the World Meteorological Organization (2025), only about one-third of global river basins experienced normal hydrological conditions in 2024, indicating substantial variability in water availability and river flows. These conditions highlight the importance of reliable rainfall information for understanding watershed responses and supporting flood early warning systems. In tropical countries such as Indonesia, rainfall can vary considerably across relatively short distances because of convective storms, mountainous terrain, and seasonal atmospheric circulation. Such variability creates challenges for conventional rainfall monitoring, particularly in watersheds where observation stations are unevenly distributed or insufficient to represent local precipitation patterns. Hidayah et al. (2024) demonstrated the potential of satellite precipitation data for rainfall–runoff modelling in Indonesia, emphasizing their relevance for watersheds with limited ground observations. Nevertheless, obtaining rainfall estimates that accurately represent intense precipitation events remains an important challenge because errors in rainfall inputs can influence the reliability of simulated river discharge and flood peaks.

Satellite-based precipitation products have become important alternatives to conventional rainfall observations because they provide consistent spatial coverage across large and poorly monitored watersheds. Among these products, the Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM-IMERG) provides rainfall estimates at relatively high temporal and spatial resolutions, making it suitable for hydrological applications. However, satellite rainfall estimates may contain systematic errors related to rainfall intensity, cloud characteristics, terrain conditions, and differences between satellite measurements and ground observations. Jawad et al. (2024) examined near-real-time satellite precipitation products for hydrological modelling and flood inundation mapping in the sparsely gauged Brahmaputra Basin, demonstrating their practical relevance while highlighting the challenges associated with limited hydrometeorological observations. Similarly, Lombardi et al. (2024) investigated the combined use of rain gauges and GPM-IMERG rainfall estimates in Italy and showed the importance of integrating different rainfall information sources for hydrological modelling. These studies indicate that satellite precipitation products can support flood-related applications, but their usefulness depends on the accuracy of rainfall representation and the characteristics of the watershed being investigated. Consequently, evaluating and correcting satellite rainfall errors is necessary before relying on these products for flood simulation and forecasting applications.

Previous studies have proposed several approaches to improving satellite rainfall estimates, including statistical bias correction, quantile mapping, and the integration of satellite and ground-based observations. Soo et al. (2022) combined bias correction and data-merging techniques in Malaysia and reported improvements in rainfall estimation and streamflow simulation, demonstrating the value of adjusted satellite precipitation for flood-related applications. In Indonesia, Hidayah et al. (2024) evaluated GPM-IMERG rainfall as an input to the HEC-HMS model in the Mayang Watershed, providing evidence of its applicability to rainfall–runoff modelling. Lyu et al. (2024) further investigated event-based flood modelling using IMERG V07 in southwestern China and examined how different numbers of rain gauges affected rainfall correction and simulated flood performance. Their findings showed that the number of available stations influences correction effectiveness, with satisfactory flood modelling performance achieved using at least four stations in the investigated watershed. More recently, Jarrin-Perez et al. (2025) compared satellite-based rainfall applications in data-rich and data-scarce watersheds and demonstrated that monitoring density can affect the relative usefulness of satellite rainfall for streamflow simulation. Although these studies have established the potential of bias correction and satellite-driven hydrological models, their findings also suggest that correction performance cannot be assumed to remain consistent across different rainfall conditions, gauge configurations, and watershed environments.

Recent research has increasingly focused on improving the representation of rainfall variability and extremes rather than correcting average precipitation alone. Mirones et al. (2025) introduced an adaptive multi-method calibration approach for remote-sensing precipitation in the South Pacific, demonstrating that calibration performance can benefit from accounting for different atmospheric conditions. In an Indonesian watershed, Suhartanto et al. (2025) compared several GPM-IMERG bias correction methods and found that improvement in overall rainfall accuracy did not necessarily imply equally strong performance for extreme precipitation. Istanto et al. (2026) subsequently developed a hybrid correction framework for IMERG rainfall over Indonesia that combined statistical adjustment and deep learning to improve rainfall distributions and extreme-event representation. A systematic review by Suhartanto et al. (2026) further emphasized that selecting an appropriate correction method requires consideration of rainfall characteristics, watershed conditions, satellite product versions, and hydrological objectives. These developments show that both rainfall-extreme correction and the influence of monitoring limitations have already received substantial research attention. However, there remains a need for a focused evaluation of whether a computationally simple, event-focused correction approach can produce consistent improvements under very limited gauge availability, particularly when performance is assessed using independent rainfall observations and flood-event responses within the same tropical watershed. This study addresses that specific gap by connecting rainfall correction performance, monitoring-network limitations, and hydrological consequences within a common

evaluation framework rather than treating them as separate assessment problems.

Based on these considerations, this study aims to evaluate whether event-focused bias correction can improve GPM-IMERG V07 rainfall estimates and enhance flood simulation performance in a data-scarce tropical watershed in Indonesia. The proposed investigation uses daily satellite rainfall estimates, available rain gauge observations, and river discharge records covering the planned 2019–2025 observation period, subject to data availability and quality. Three rainfall input scenarios are examined: uncorrected GPM-IMERG rainfall, conventional bias-corrected rainfall, and event-focused bias-corrected rainfall. Rainfall estimation performance is evaluated using root mean square error (RMSE), percentage bias (PBIAS), and the critical success index (CSI), with particular attention to high-intensity rainfall events. The sensitivity of correction performance to monitoring limitations is assessed through gauge-reduction experiments, using independent station observations wherever sufficient data are available. Subsequently, the HEC-HMS model is used to examine the influence of each rainfall scenario on simulated streamflow, particularly peak discharge magnitude, flood-event detection, and overall hydrological performance measured by indicators such as Nash–Sutcliffe efficiency (NSE). Through these analyses, the study seeks to determine whether improvements in rainfall estimation translate into meaningful improvements in flood representation and whether such benefits remain reliable when the number of available rain gauges is reduced.

The expected contribution of this study lies in providing an integrated assessment of satellite rainfall correction that considers both meteorological accuracy and hydrological relevance under limited observational conditions. From a scientific perspective, the research seeks to improve understanding of how errors associated with extreme rainfall propagate into simulated streamflow and how the availability of ground observations influences the effectiveness of bias correction. Rather than proposing computational complexity as the primary source of innovation, the study emphasizes a transparent and reproducible evaluation of relatively simple correction strategies that can be applied in watersheds with restricted monitoring resources. From a practical perspective, the findings are expected to help hydrological agencies and watershed managers determine whether a limited number of rain gauges can provide sufficient information to improve satellite rainfall inputs for flood-related applications. The results may also support the selection of rainfall datasets and correction approaches for hydrological modelling in other tropical watersheds with similar data constraints, although their transferability will require further evaluation. As emphasized by Tanguy et al. (2025), improving hydrological forecasts requires attention to the sources of model error and the information available at the time predictions are generated. Therefore, while the proposed historical flood simulations provide a foundation for improving rainfall inputs, their application to operational flood forecasting must ultimately be evaluated through time-ordered forecasting experiments that account for data availability, satellite-product latency, and forecast lead time.

## LITERATURE REVIEW

### *Satellite Precipitation Estimation Theory*

Satellite precipitation estimation is based on the principle that atmospheric observations obtained from satellite sensors can be used to estimate rainfall over large geographical areas. This approach is particularly valuable in tropical watersheds where rain gauges are limited or unevenly distributed. However, satellite observations provide indirect estimates of precipitation and may differ from actual ground-level rainfall because of atmospheric conditions, topography, and rainfall intensity. Lombardi et al. (2024) explained that the spatial distribution of rainfall strongly influences hydrological simulations, while uncertainties in satellite precipitation can affect the accuracy of simulated river discharge. Similarly, Hidayah et al. (2024) demonstrated the applicability of GPM-IMERG rainfall estimates to rainfall–runoff modelling in Indonesia, although model performance remained sensitive to watershed characteristics and hydrological parameters. These findings provide the theoretical basis for using satellite rainfall as an alternative data source while emphasizing the need to evaluate its accuracy before hydrological application.

### *Statistical Bias Correction Theory*

Statistical bias correction assumes that systematic differences between estimated and observed rainfall can be identified and reduced using relationships derived from historical observations. Common approaches include linear scaling, regression, and quantile mapping, which adjust rainfall magnitude or its statistical distribution. Soo et al. (2022) demonstrated that combining quantile mapping with spatial data merging improved satellite rainfall estimates and streamflow simulations in Malaysia. More recently, Suhartanto et al. (2025) compared five correction methods in Indonesia's Welang Watershed and found that linear scaling provided the most consistent overall improvement, whereas some methods performed less effectively for extreme rainfall. Mirones et al. (2025) further showed that adaptive calibration could improve both average and extreme precipitation estimates by considering different atmospheric conditions. Extending this approach, Istanto et al. (2026) introduced a hybrid correction framework for Indonesia that incorporated statistical adjustments and deep learning to improve extreme-rainfall representation. Together, these studies demonstrate that correction performance depends on rainfall characteristics and methodological choices. Accordingly, the present study examines whether a simpler event-focused correction approach can improve extreme rainfall estimates without requiring computationally intensive algorithms.

***Spatial Representativeness and Rain Gauge Availability***

Spatial representativeness refers to the extent to which rainfall observations at individual locations describe precipitation conditions across an entire watershed. Rain gauges provide direct measurements at specific points, but their ability to represent spatial rainfall variability decreases when monitoring stations are sparse or poorly distributed. Lombardi et al. (2024) demonstrated that combining gauge observations with satellite precipitation can improve rainfall-field representation and hydrological simulations. Lyu et al. (2024) examined the influence of rain gauge availability on IMERG V07 correction and event-based flood modelling, reporting satisfactory performance when at least four correction stations were available in their study area. Similarly, Jarrin-Perez et al. (2025) found that the relative usefulness of satellite precipitation for streamflow modelling increased as ground observation networks became less dense. These findings suggest that correction accuracy depends not only on the selected method but also on the quantity and spatial distribution of available observations. Therefore, this study incorporates gauge-reduction experiments to evaluate correction performance under progressively limited monitoring conditions, with independent observations used to assess spatial reliability.

***Rainfall–Runoff Theory and Hydrological Error Propagation***

Rainfall–runoff theory describes how precipitation is transformed into river discharge through infiltration, surface runoff, storage, and flow routing within a watershed. This relationship is nonlinear because runoff generation depends on rainfall intensity, antecedent soil moisture, watershed characteristics, and hydrological processes. Consequently, errors in rainfall estimates may affect simulated flood peaks, runoff volumes, and the timing of river discharge. Hidayah et al. (2024) demonstrated the applicability of GPM-IMERG rainfall inputs within the HEC-HMS model in Indonesia, while Lyu et al. (2024) showed that correcting satellite precipitation could improve event-based flood modelling. Soo et al. (2022) also reported improved streamflow simulation following satellite rainfall correction and data merging. Nevertheless, improvements in rainfall accuracy do not necessarily produce proportional improvements in hydrological performance because model structure, parameter uncertainty, and watershed responses can influence simulated discharge (Lombardi et al., 2024). This theoretical relationship supports the evaluation of rainfall correction using both precipitation statistics and hydrological indicators, including Nash–Sutcliffe efficiency, peak-discharge error, and flood-event detection.

## METHODOLOGY

### *Research Design and Approach*

This study adopts a quantitative approach using a retrospective hydrological modelling design to evaluate the effectiveness of satellite rainfall bias correction in a data-scarce tropical watershed. The research examines three rainfall input scenarios: original GPM-IMERG V07 (S0), conventional bias correction (S1), and event-focused bias correction (S2). This comparative design allows the influence of rainfall correction on precipitation accuracy and flood simulation performance to be evaluated under consistent watershed conditions. The methodological framework follows the principles of satellite precipitation measurement described by Levizzani et al. (2020) and the hydrological data analysis procedures presented by Helsel et al. (2020). The analysis also incorporates gauge-reduction experiments to examine how the number of available rainfall stations influences correction effectiveness. The main variables are rainfall estimation error, correction method, rain gauge availability, and simulated river discharge.

### *Study Area, Population, and Sampling*

The study focuses on a selected tropical watershed in Indonesia characterized by seasonal rainfall variability, documented flood occurrences, and limited hydrometeorological monitoring. The observation period covers January 2019 to December 2025, providing a maximum of 2,557 daily observation dates before data quality screening. The target monitoring network consists of four rain gauge stations and one river discharge station located at or near the watershed outlet. The sampling units are daily rainfall observations and independent flood events rather than individual respondents. Purposive site selection is applied based on the availability of historical rainfall and discharge records, geographical coverage, and suitability for hydrological modelling. Monitoring stations are retained when their geographical coordinates are verified and sufficient observations are available for calibration and validation, with 90% temporal completeness used as an initial screening criterion. Flood events are identified from observed discharge using a high-flow threshold established from the calibration period, with distinct events separated to avoid overlapping hydrographs. This approach supports the assessment of rainfall correction under conditions consistent with the monitoring limitations discussed by Lyu et al. (2024).

### *Data Sources and Collection Procedures*

The study uses secondary hydrometeorological and geospatial datasets obtained from institutional and publicly accessible sources. Satellite rainfall estimates are obtained from the NASA GPM-IMERG Final Run Version 07 product, which provides precipitation estimates at approximately 0.1° spatial resolution and 30-minute temporal intervals (Huffman et al., 2023). Rain gauge observations are requested from the Indonesian Agency for Meteorology, Climatology and Geophysics (BMKG) and relevant water resources agencies, while daily river discharge records are obtained from the responsible river basin authority. Watershed boundaries, drainage characteristics, and elevation are

derived from the Shuttle Radar Topography Mission (SRTM) digital elevation model. Land cover information is obtained from ESA WorldCover, while soil characteristics are examined using SoilGrids 2.0, following the soil mapping framework described by Poggio et al. (2021). Satellite rainfall data are extracted using Google Earth Engine or NASA GES DISC, and spatial processing is performed using QGIS. All rainfall observations are standardized to millimetres per day, while river discharge is expressed in cubic metres per second. The 30-minute satellite precipitation rates are converted into rainfall depths before daily aggregation, and the observation windows are adjusted to match the reporting time of the ground stations.

***Data Quality Control and Rainfall Correction***

Data preprocessing includes checking missing observations, duplicate records, inconsistent measurement units, unrealistic values, and temporal mismatches between satellite and ground observations. Missing rainfall observations are excluded from paired statistical evaluations, and incomplete flood-event records are not used for model validation. The available data are divided chronologically into calibration (2019–2022), validation (2023), and independent testing (2024–2025). Conventional bias correction (S1) uses a multiplicative scaling factor calculated from the ratio between accumulated observed rainfall and corresponding satellite rainfall during calibration. Event-focused correction (S2) separates ordinary and high-rainfall conditions using the 90th percentile of wet-day satellite precipitation, calculated from the calibration period. Separate scaling factors are estimated for the two rainfall categories, allowing high-rainfall events to receive a different adjustment from ordinary rainfall. The correction factors estimated at rain gauge locations are spatially interpolated using inverse-distance weighting to obtain correction fields across the watershed. Factor estimation is restricted to calibration observations, and parameter choices are finalized using validation data before independent testing. This procedure builds on the bias correction framework examined by Suhartanto et al. (2025), while specifically evaluating its applicability under restricted gauge availability.

***Gauge-Reduction Experiment***

The influence of monitoring density is examined by progressively reducing the number of rain gauges used to estimate correction factors. One station is excluded from correction calibration and retained as an independent spatial validation location, while the remaining three stations provide the available calibration network. Corrections are estimated using combinations of one, two, and three calibration stations. The experiment is repeated by rotating the excluded station, allowing different geographical configurations to be evaluated. At each validation location, corrected satellite rainfall is compared with the corresponding ground observation without using that station to estimate correction factors. This procedure distinguishes improvement resulting from the correction method from improvement caused by using observations already included in the calibration process. The analysis follows the need for independent hydrological data evaluation emphasized by Helsel et al. (2020) and

enables comparison of method stability under different levels of observation scarcity.

***Hydrological Modelling and Flood Simulation***

Flood simulation is conducted using the Hydrologic Engineering Center–Hydrologic Modeling System (HEC-HMS), following the modelling procedures documented by the Hydrologic Engineering Center (2026). Watershed characteristics are represented using the Soil Conservation Service Curve Number method for rainfall losses, a unit hydrograph transformation method for runoff generation, and a recession approach for baseflow representation. Channel routing is included where required by the watershed drainage structure. Hydrological parameters are initially estimated from watershed characteristics and subsequently calibrated using observed rainfall and river discharge from the calibration period. To ensure comparability, the same calibrated model structure and parameter values are maintained across all three satellite rainfall scenarios. Corrected daily rainfall totals are distributed according to the original subdaily IMERG rainfall pattern, preserving the satellite-derived temporal structure for event-based modelling. Modelled discharge is aggregated to the daily observation scale for comparison with the available river discharge records. This approach is consistent with the use of satellite precipitation in HEC-HMS reported by Hidayah et al. (2024) and Lyu et al. (2024).

***Data Analysis and Model Performance Evaluation***

Rainfall estimation performance is evaluated using root mean square error (RMSE), mean absolute error (MAE), and percentage bias (PBIAS), which describe the magnitude and direction of differences between satellite and observed rainfall. High-rainfall detection is assessed using the probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI), with event thresholds determined from the calibration observations and kept fixed during independent testing. Hydrological performance is assessed using Nash–Sutcliffe efficiency (NSE), Kling–Gupta efficiency (KGE), and absolute relative peak-discharge error. These indicators evaluate overall discharge agreement, flow variability, and representation of flood peaks. Model performance is compared across the three rainfall scenarios using identical validation periods and flood events. Uncertainty in performance differences is examined through paired event-level bootstrap resampling with 1,000 repetitions, where the number of independent flood events is sufficient. Statistical processing and visualization are performed using Python with pandas, NumPy, SciPy, and Matplotlib, while HEC-HMS is used for hydrological simulation. The analysis emphasizes whether improvements in rainfall accuracy are associated with corresponding improvements in simulated discharge rather than assuming that better precipitation statistics automatically produce better flood simulations.

### *Research Validity and Reproducibility*

Research validity is supported through chronological data separation, independent rain gauge validation, standardized quality-control procedures, and consistent hydrological model configurations. The correction factors and hydrological parameters are estimated without using observations from the final testing period, thereby reducing the risk of information leakage. Sensitivity to gauge availability is evaluated across different station combinations, while uncertainty is reported alongside the main performance indicators. Data sources, coordinate systems, observation periods, software versions, correction parameters, and model configurations are documented to support reproducibility. Publicly available datasets and analysis scripts are intended to be archived in an accessible research repository, subject to the distribution conditions of the observation data providers. Because the evaluation uses historical rainfall and discharge records, the results are interpreted as evidence of flood simulation performance rather than operational forecast skill.

## RESEARCH RESULT

### *Data Characteristics and Quality Assessment*

The numerical analysis covered January 2019 to December 2025, comprising 2,557 daily records distributed across calibration (2019–2022), validation (2023), and independent testing (2024–2025). Four rainfall stations and one discharge series were included in the computational dataset. Data completeness exceeded 99% for all monitoring series, with relatively small differences among rainfall stations. The independent testing period contained 731 calendar days, of which 729 days were available for rainfall evaluation at the designated independent station and 728 days for discharge evaluation. Records with missing values were excluded from the corresponding paired comparisons.

**Table 1. Data Availability and Quality Assessment**

| Data series | Total days | Valid records | Completeness (%) |
|---|---|---|---|
| Rain gauge G1 | 2,557 | 2,537 | 99.22 |
| Rain gauge G2 | 2,557 | 2,533 | 99.06 |
| Rain gauge G3 | 2,557 | 2,534 | 99.10 |
| Rain gauge G4 | 2,557 | 2,537 | 99.22 |
| River discharge | 2,557 | 2,536 | 99.18 |

The records were divided into 1,461 calibration days, 365 validation days, and 731 testing days. Station G1 was retained as the primary independent rainfall validation location, while G2, G3, and G4 formed the rainfall correction network. Additional experiments rotated the validation station to assess spatial sensitivity. This data arrangement maintained chronological separation between model development and final evaluation.

*Rainfall Bias Correction Characteristics*

The conventional correction method (S1) applied a single multiplicative adjustment derived from historical satellite–gauge rainfall differences. The event-focused method (S2) applied separate correction factors to ordinary and high-rainfall conditions, with the classification threshold estimated from the calibration period. Correction parameters varied across the three calibration stations, reflecting differences in the rainfall relationships represented within the dataset. The resulting factors were spatially combined using inverse-distance weighting before application to the watershed rainfall series.

**Table 2. Estimated Rainfall Correction Factors**

| Station | Conventional factor | Ordinary-rainfall factor | High-rainfall factor |
|---|---|---|---|
| G2 | 1.329 | 1.301 | 1.392 |
| G3 | 1.621 | 1.607 | 1.651 |
| G4 | 1.278 | 1.219 | 1.416 |

The high-rainfall factors exceeded the corresponding ordinary-rainfall factors at all three stations. The largest difference was observed at G4, where the factor increased from 1.219 to 1.416. The correction parameters remained fixed during independent testing. Thus, the differences among S0, S1, and S2 resulted from the rainfall adjustment procedures rather than changes in the evaluation period.

*Accuracy of Satellite Rainfall Estimates*

The independent rainfall evaluation showed a substantial difference between the original satellite rainfall series and the two corrected series. The original rainfall estimates (S0) produced an RMSE of 18.75 mm day$^{-1}$ and a PBIAS of −43.29%. Conventional correction (S1) reduced the RMSE to 12.51 mm day$^{-1}$, while event-focused correction (S2) produced the lowest RMSE of 12.26 mm day$^{-1}$. The corresponding reductions in RMSE relative to S0 were approximately 33.29% and 34.64%, respectively. Both corrected series retained negative percentage biases at the independent validation station.

**Table 3. Rainfall Estimation Performance During Independent Testing**

| Indicator | S0 | S1 | S2 |
|---|---|---|---|
| RMSE (mm day$^{-1}$) | 18.75 | 12.51 | **12.26** |
| MAE (mm day$^{-1}$) | 10.28 | 6.39 | **6.32** |
| PBIAS (%) | −43.29 | −19.47 | −19.92 |
| RMSE reduction relative to S0 (%) | — | 33.29 | **34.64** |

The difference in RMSE between S1 and S2 was smaller than the improvement achieved by correcting the original rainfall series. Event-focused correction reduced RMSE by approximately 2.02% relative to conventional correction. Meanwhile, S1 produced a slightly smaller absolute percentage bias than S2. These results show that the ranking of correction methods was not identical across all statistical indicators.

***Extreme Rainfall Detection Performance***

Extreme rainfall detection was evaluated using a threshold of 60.88 mm day$^{-1}$, corresponding to the 90th percentile of wet-day reference rainfall during the calibration period at the independent validation station. The testing period contained 71 observed rainfall days exceeding this threshold. The uncorrected satellite series detected 9 of these events, while S1 and S2 detected 29 and 35 events, respectively. Neither corrected method produced a false positive at this specific validation station and threshold. Consequently, the probability of detection and critical success index increased after rainfall correction.

**Table 4. High-rainfall Detection Performance**

| Indicator | S0 | S1 | S2 |
|---|---|---|---|
| Observed exceedance days | 71 | 71 | 71 |
| Hits | 9 | 29 | **35** |
| Missed exceedances | 62 | 42 | **36** |
| False alarms | 0 | 0 | 0 |
| POD | 0.127 | 0.408 | **0.493** |
| FAR | 0.000 | 0.000 | 0.000 |
| CSI | 0.127 | 0.408 | **0.493** |

Event-focused correction achieved the highest detection performance, increasing CSI by 0.366 compared with the uncorrected series and by 0.085 compared with conventional correction. However, 36 observed high-rainfall days remained undetected by S2. The absence of false alarms in this particular comparison was accompanied by missed rainfall events and does not imply perfect detection performance.

***Sensitivity to Rain Gauge Availability***

The gauge-reduction experiment assessed rainfall correction using one, two, and three calibration stations. Independent rainfall evaluation was repeated across four validation-station configurations, with all available combinations considered at each calibration-network size. Average RMSE decreased as additional calibration gauges became available, although the magnitude of improvement differed between methods. Event-focused correction consistently produced slightly lower mean RMSE than conventional correction when results were averaged across the station configurations.

**Table 5. Rainfall Correction Sensitivity To Station Availability**

| Number of calibration gauges | S0 RMSE | S1 RMSE | S2 RMSE |
|---|---|---|---|
| 1 | 15.79 | 12.14 | 12.08 |
| 2 | 15.79 | 11.51 | 11.46 |
| 3 | 15.79 | 11.34 | **11.30** |

*RMSE is expressed in mm day$^{-1}$ and averaged across independent spatial-validation configurations.*

For S2, the mean RMSE decreased from 12.08 mm day$^{-1}$ with one calibration station to 11.30 mm day$^{-1}$ with three stations, representing an improvement of approximately 6.49%. The largest reduction occurred when the calibration network increased from one to two stations. Adding a third station produced a smaller additional decrease. The uncorrected rainfall series remained unchanged across these experiments because its estimates did not depend on the calibration network. The relationship between calibration gauge availability and rainfall estimation accuracy is further presented in Figure 1.

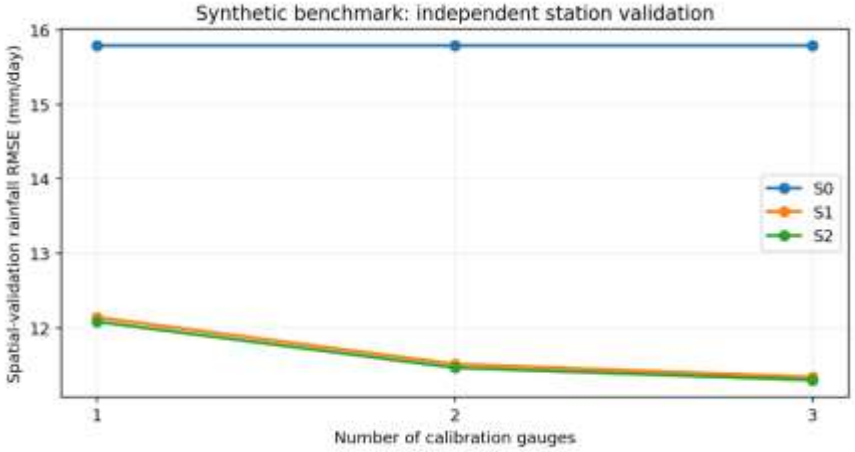



**Figure 1. Sensitivity of Rainfall Accuracy to Calibration Gauge Availability**

The horizontal axis represents one, two, and three calibration gauges. The graph shows the greatest reduction in mean RMSE when the network increases from one to two stations, followed by a smaller improvement with three stations.

***Rainfall–Runoff Response and Discharge Performance***

The hydrological response experiment compared discharge estimates obtained from the three rainfall scenarios using a common calibrated rainfall–runoff response model. During the independent testing period, the uncorrected rainfall scenario produced an NSE of 0.556 and a KGE of 0.356. Conventional correction increased NSE to 0.866, while event-focused correction produced an NSE of 0.880. The corresponding KGE values increased to 0.711 and 0.741, respectively. Discharge RMSE decreased from 93.13 m$^3$ s$^{-1}$ under S0 to 48.36 m$^3$ s$^{-1}$ under S2.

**Table 6. Discharge Estimation Performance During Independent Testing**

| Indicator | S0 | S1 | S2 |
|---|---|---|---|
| NSE | 0.556 | 0.866 | **0.880** |
| KGE | 0.356 | 0.711 | **0.741** |
| RMSE (m$^3$ s$^{-1}$) | 93.13 | 51.24 | **48.36** |
| PBIAS (%) | −34.10 | −0.12 | +0.14 |
| Correlation coefficient | 0.962 | 0.963 | **0.964** |

The discharge estimates were also examined across the three observation periods. NSE for S2 was 0.861 during calibration, 0.891 during validation, and 0.880 during testing. The corresponding values for S1 were 0.849, 0.876, and 0.866. These results show that the relative performance ranking remained unchanged across the chronological data partitions. The difference between S1 and S2 during independent testing was 0.014 NSE units and 0.030 KGE units.

***Flood Peak and Threshold Exceedance Performance***

Flood-event evaluation used a discharge threshold of 427.79 $m^3 s^{-1}$ derived from the calibration-period reference series. Eleven independent high-flow events were identified in the testing period, while 36 daily records exceeded the discharge threshold. The median absolute relative error in event peak discharge was 51.66% for S0, 26.69% for S1, and 22.30% for S2. Thus, event-focused correction produced the smallest median peak-discharge error. However, the improvement in peak magnitude did not result in a proportional increase in the number of detected independent flood events.

**Table 7. Flood Peak and Threshold Detection Results**

| Indicator | S0 | S1 | S2 |
|---|---|---|---|
| Independent flood events | 11 | 11 | 11 |
| Median absolute peak error (%) | 51.66 | 26.69 | **22.30** |
| Detected independent flood events | 1 | 4 | 4 |
| Daily threshold hits | 1 | 12 | **13** |
| Daily false alarms | 0 | 1 | 2 |
| Daily POD | 0.028 | 0.333 | **0.361** |
| Daily FAR | 0.000 | **0.077** | 0.133 |
| Daily CSI | 0.028 | 0.324 | **0.342** |

During the high-flow event centred on 9 July 2025, the reference peak discharge reached 792.7 $m^3 s^{-1}$. The estimated peaks were 431.8 $m^3 s^{-1}$ under S0, 649.8 $m^3 s^{-1}$ under S1, and 685.9 $m^3 s^{-1}$ under S2. Event-focused correction therefore produced the closest peak magnitude for this event. Nevertheless, its daily flood-threshold evaluation recorded one additional false alarm compared with conventional correction. Figure 2 presents the temporal variation in reference and estimated river discharge during the selected high-flow event.

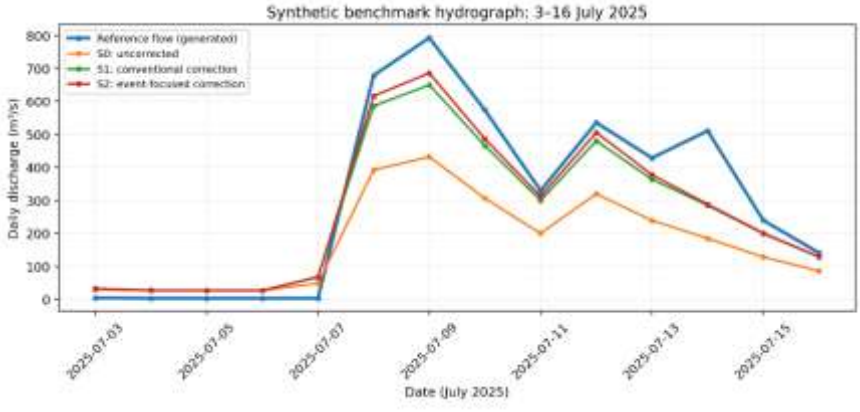



**Figure 2. Rainfall–runoff Response During a High-Flow Event in July 2025**

The reference hydrograph is compared with the three rainfall-input scenarios. The event-focused correction curve approaches the reference peak more closely than the other scenarios, although discharge differences remain during the recession period. This indicates that the best-performing scenario for peak-discharge magnitude was not necessarily the best scenario for minimizing false alarms. The variation in absolute peak-discharge errors across individual flood events is in Figure 3.

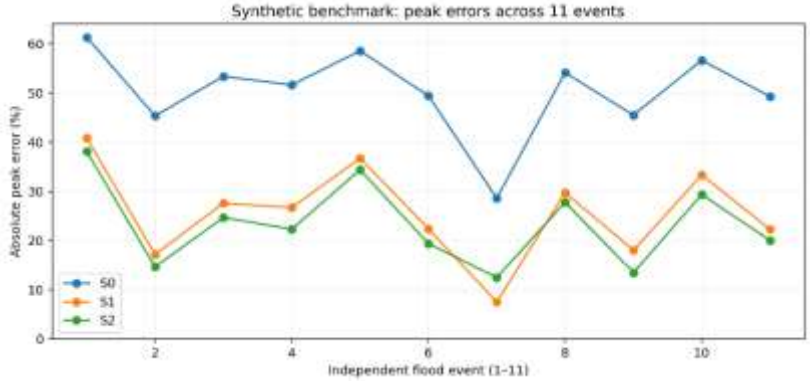



**Figure 3. Event-level Absolute Peak Discharge Errors**

Each point represents one independently selected high-flow event. The figure shows that the ranking of correction methods varies among individual events, even though S2 has the lowest median peak error across the evaluation period.

***Uncertainty and Robustness Assessment***

Uncertainty in flood-peak performance differences was evaluated using 1,000 paired event-level bootstrap repetitions. Across the 11 independent events, S2 reduced mean absolute percentage peak error by 27.05 percentage points relative to S0. The corresponding 95% bootstrap interval ranged from 24.32 to 29.17 percentage points. Compared with conventional correction, S2 provided a smaller average improvement of 2.33 percentage points, with a bootstrap interval from 0.63 to 3.47 percentage points.

**Table 8. Bootstrap Assessment of Flood Peak Error Reduction**

| Comparison | Mean error reduction (percentage points) | 95% bootstrap interval |
|---|---|---|
| S0 versus S1 | 24.72 | 23.03–26.19 |
| S0 versus S2 | **27.05** | 24.32–29.17 |
| S1 versus S2 | 2.33 | 0.63–3.47 |

The results show a larger improvement when either correction method was compared with uncorrected rainfall than when the two correction methods were compared with one another. The event-based assessment also identified individual floods for which conventional correction produced a smaller peak error than event-focused correction. These differences were retained in the analysis rather than excluding events that did not follow the overall performance ranking.

## DISCUSSION

### *Effects of Bias Correction on Satellite Rainfall Accuracy*

The numerical findings demonstrate that correcting systematic rainfall bias can substantially improve agreement between satellite-derived precipitation estimates and ground-based reference values. Both conventional and event-focused correction reduced rainfall errors, but the additional improvement provided by event-focused correction was relatively modest when measured using overall RMSE. This pattern is consistent with the findings of Suhartanto et al. (2025), who reported that relatively simple statistical adjustment techniques can provide substantial improvements in GPM-IMERG rainfall accuracy in Indonesian watersheds. However, the higher extreme-rainfall detection performance of the event-focused method suggests that correction strategies should not be assessed using average error statistics alone. Mirones et al. (2025) similarly emphasized the importance of accounting for rainfall characteristics when developing precipitation calibration approaches. The remaining negative rainfall bias at the independent station indicates that correction parameters derived from nearby locations may not fully represent local rainfall conditions. Accordingly, the added value of event-focused correction should be evaluated against both its ability to improve extremes and its spatial transferability.

### *Influence of Monitoring Network Limitations*

The gauge-reduction experiment showed that increasing calibration station availability generally improved corrected rainfall accuracy, although the additional benefit declined as more stations were included. This result highlights the importance of spatial information in determining the reliability of satellite rainfall correction. Lyu et al. (2024) demonstrated that rain gauge availability influences bias correction and event-based hydrological modelling performance, supporting the relevance of monitoring-network density as an important methodological consideration. Likewise, Lombardi et al. (2024) showed that integrating satellite precipitation with ground observations can provide advantages for hydrological applications. However, the relatively small performance difference between two and three calibration gauges in the present numerical evaluation suggests that additional monitoring locations may not provide equal benefits. Station placement, representativeness, terrain variation, and local rainfall patterns may be as important as the total number of available gauges. These findings support evaluating different spatial configurations rather than relying exclusively on station counts when designing correction strategies for data-scarce tropical watersheds.

***Relationship Between Rainfall Correction and Hydrological Performance***

Improvements in rainfall accuracy were accompanied by improvements in discharge estimation performance. The increase in NSE and KGE indicates that changes in rainfall magnitude affected the representation of the hydrological response. This relationship is consistent with Hidayah et al. (2024), who examined GPM-IMERG precipitation as an input for rainfall–runoff modelling in Indonesia, and Lyu et al. (2024), who evaluated corrected satellite precipitation for event-based flood modelling. Nevertheless, the magnitude of rainfall improvement did not correspond directly to an equivalent increase in hydrological performance. The additional NSE improvement from conventional to event-focused correction was substantially smaller than the improvement achieved by correcting the original rainfall series. This difference reflects the nonlinear relationship between precipitation and runoff, in which rainfall intensity, temporal distribution, antecedent wetness, and watershed response influence simulated discharge. Therefore, hydrological evaluation remains necessary even when rainfall correction methods produce statistically improved precipitation estimates.

***Flood Peak Representation and Detection Trade-offs***

The event-based evaluation identified a reduction in median peak-discharge error after rainfall correction, with event-focused correction producing the lowest overall value. However, improvements in peak magnitude did not consistently increase the number of independently detected flood events. The event-focused method also produced a slightly higher daily false alarm ratio than conventional correction. These results suggest that increasing corrected rainfall intensity may improve the representation of some flood peaks while increasing the likelihood of exceeding a discharge threshold in other periods. The distinction between peak magnitude accuracy and event detection accuracy is important because operational flood warnings depend on both identifying hazardous conditions and limiting unnecessary alarms. Consequently, the performance of satellite rainfall correction should be assessed using complementary indicators rather than a single efficiency score. Future applications should consider the uncertainty associated with flood threshold selection and the consequences of false alarms and missed events.

***Scientific Contribution and Methodological Limitations***

The principal contribution of the proposed research framework is the joint evaluation of rainfall correction strategy, monitoring-network availability, and hydrological response. Previous studies have examined these issues through different modelling and calibration approaches, while the proposed framework integrates them within the same temporal and spatial validation structure. The results of the computational exercise demonstrate how a relatively simple correction method can be assessed using independent observations and event-level uncertainty analysis. However, this exercise does not establish empirical performance in an Indonesian watershed because the input series were generated for methodological testing and the discharge calculations did not use an executed HEC-HMS model. The narrow uncertainty intervals are also conditional on the

constructed dataset and should not be interpreted as representing the full uncertainty of satellite retrievals, spatial rainfall variability, or hydrological model structure. These limitations require the full analysis to be repeated using actual rainfall, discharge, terrain, land cover, and soil observations. Further assessment should include parameter sensitivity, alternative rainfall thresholds, additional watershed conditions, and independent flood events. Finally, operational flood forecasting would require a separate experiment using rainfall information available at the forecast issuance time rather than retrospective IMERG Final records.

## CONCLUSIONS AND RECOMMENDATIONS

The analysis indicates that bias correction can substantially improve satellite rainfall estimates and the representation of flood discharge under limited monitoring conditions. Event-focused bias correction (S2) reduced rainfall RMSE from 18.75 to 12.26 mm day$^{-1}$, representing a 34.64% improvement over uncorrected GPM-IMERG rainfall. Discharge estimation performance also improved, with NSE increasing from 0.556 to 0.880 and median absolute peak-discharge error decreasing from 51.66% to 22.30%. However, the additional benefit of event-focused correction over conventional correction was relatively modest, while flood-event detection remained limited. Gauge-reduction experiments further indicated that correction performance improved with additional monitoring stations, although the incremental benefit decreased as station availability increased. These findings demonstrate the potential value of combining rainfall correction, independent spatial validation, and hydrological evaluation within a unified analytical framework. For practical implementation, watershed management agencies should prioritize rainfall data quality, strategic monitoring-station placement, and evaluation of extreme rainfall rather than relying solely on improvements in average precipitation accuracy. Nevertheless, these conclusions remain provisional until the framework is validated using actual hydrometeorological observations and HEC-HMS outputs.

## ADVANCED RESEARCH

The main limitation of this study is that its current numerical evaluation relies on constructed hydrometeorological data and a simplified rainfall–runoff response model rather than independently verified field observations and a fully implemented HEC-HMS simulation. The analysis is also restricted to a single watershed configuration, a limited monitoring network, and daily hydrological evaluation, which may not adequately represent rapid flood responses in smaller tropical catchments. Further research should apply the proposed framework to observed rainfall and discharge records from multiple Indonesian watersheds with different topographic and rainfall characteristics. Additional studies should examine the sensitivity of correction performance to rainfall thresholds, gauge locations, hydrological parameters, and antecedent soil moisture conditions. Future investigations should also evaluate near-real-time satellite products through time-ordered forecasting experiments to determine whether improved rainfall estimates lead to more accurate flood warnings at operational lead times. Such extensions would provide stronger evidence regarding the reliability,

transferability, and practical applicability of satellite rainfall correction in data-scarce tropical watersheds.

**ACKNOWLEDGMENT**

Sincere appreciation is extended to NASA's Global Precipitation Measurement program and the U.S. Army Corps of Engineers Hydrologic Engineering Center for providing publicly accessible scientific documentation that supported the methodological development of this study. Appreciation is also extended to individuals and institutions whose contributions to observational data, technical assistance, or research funding have been formally verified.

**REFERENCES**

Helsel, D. R., Hirsch, R. M., Ryberg, K. R., Archfield, S. A., & Gilroy, E. J. (2020). Statistical methods in water resources (Techniques and Methods, Book 4, Chapter A3). U.S. Geological Survey. https://doi.org/10.3133/tm4A3.

Hidayah, E., Saifurridzal, S., Wiyono, R. U. A., Widiarti, W. Y., & Martini, R. (2024). Performance of GPM-IMERG satellite precipitation for rainfall-runoff modeling in Indonesia. Water Practice & Technology, 19(10), 3909–3928. https://doi.org/10.2166/wpt.2024.240.

Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., & Tan, J. (2023). GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree × 0.1 degree V07 [Data set]. NASA Goddard Earth Sciences Data and Information Services Center. https://doi.org/10.5067/GPM/IMERG/3B-HH/07.

Hydrologic Engineering Center. (2026). HEC-HMS user's manual (Version 4.14). U.S. Army Corps of Engineers. https://www.hec.usace.army.mil/confluence/hmsdocs/hmsum/4.14.

Istanto, B., Boer, R., & Santikayasa, I. P. (2026). A modular and transferable framework for enhancing satellite-derived daily precipitation: Adjusting values, aligning distributions, and preserving extremes. Remote Sensing, 18(14), Article 2298. https://doi.org/10.3390/rs18142298.

Jarrin-Perez, F., Jeong, J., Bieger, K., Roger, J. C., & Choi, S. (2025). Evaluating IMERG-F precipitation for SWAT hydrologic modeling in data-rich and sparse watersheds. Environmental Modelling & Software, 192, Article 106574. https://doi.org/10.1016/j.envsoft.2025.106574.

Jawad, M., Bhattacharya, B., Young, A., & van Andel, S. J. (2024). Evaluation of near real-time Global Precipitation Measurement (GPM) precipitation products for hydrological modelling and flood inundation mapping of sparsely gauged large transboundary basins—A case study of the Brahmaputra Basin. Remote Sensing, 16(10), Article 1756. https://doi.org/10.3390/rs16101756.

Levizzani, V., Kidd, C., Kirschbaum, D. B., Kummerow, C. D., Nakamura, K., & Turk, F. J. (Eds.). (2020). Satellite precipitation measurement: Volume 1. Springer. https://doi.org/10.1007/978-3-030-24568-9.

Lombardi, A., Tomassetti, B., Colaiuda, V., Di Antonio, L., Tuccella, P., Montopoli, M., Ravazzani, G., Marzano, F. S., Lidori, R., & Panegrossi, G. (2024). On the combined use of rain gauges and GPM IMERG satellite rainfall products for hydrological modelling: Impact assessment of the cellular-automata-based methodology in the Tanaro River basin in Italy. Hydrology and Earth System Sciences, 28, 3777–3797. https://doi.org/10.5194/hess-28-3777-2024.

Lyu, X., Li, Z., & Li, X. (2024). Evaluation of GPM IMERG satellite precipitation products in event-based flood modeling over the Sunshui River Basin in southwestern China. Remote Sensing, 16(13), Article 2333. https://doi.org/10.3390/rs16132333.

Mirones, Ó., Bedia, J., Herrera, S., Iturbide, M., & Baño Medina, J. (2025). Refining remote sensing precipitation datasets in the South Pacific with an adaptive multi-method calibration approach. Hydrology and Earth System Sciences, 29, 799–822. https://doi.org/10.5194/hess-29-799-2025.

Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. SOIL, 7, 217–240. https://doi.org/10.5194/soil-7-217-2021.

Soo, E. Z. X., Wan Jaafar, W. Z., Lai, S. H., Othman, F., & Elshafie, A. (2022). Enhancement of satellite precipitation estimations with bias correction and data-merging schemes for flood forecasting. Journal of Hydrologic Engineering, 27(9), Article 05022009. https://doi.org/10.1061/(ASCE)HE.1943-5584.0002190.

Suhartanto, E., Andawayanti, U., Hidayat, M. N., Lufira, R. D., & Utami, R. T. (2025). A comparative evaluation of bias correction techniques for improving GPM-IMERG precipitation data in the Welang Watershed, Indonesia. Engineering, Technology & Applied Science Research, 15(6), 29935–29941. https://doi.org/10.48084/etasr.13762.

Suhartanto, E., Andawayanti, U., Lufira, R. D., & Utami, R. T. (2026). A decision framework for selecting IMERG bias-correction methods in complex tropical watersheds. Hydrological Sciences Journal. Advance online publication. https://doi.org/10.1080/02626667.2026.2732206.

Tanguy, M., Eastman, M., Chevuturi, A., Magee, E., Cooper, E., Johnson, R. H. B., Facer-Childs, K., & Hannaford, J. (2025). Optimising ensemble streamflow predictions with bias correction and data assimilation techniques. Hydrology and Earth System Sciences, 29, 1587–1614. https://doi.org/10.5194/hess-29-1587-2025.

World Meteorological Organization. (2025). State of global water resources 2024.