INNSPUB

Journal of Biodiversity and Environmental Sciences | JBES
ISSN: 2220-6663 (Print); 2222-3045 (Online)
Website: https://www.innspub.net
E-mail contact: info@innspub.net
Vol. 27, Issue: 6, p. 79-90, 2025

RESEARCH PAPER OPEN ACCESS

Validation of satellite rainfall monitor (SRM) estimates against automated rain gauge observations in the Cagayan de Oro River Basin, Philippines

Elgin Joy N. Bonalos*[1], Johniel E. Babiera[2], Peter D. Suson[1]

[1]*Department of Environmental Science, Mindanao State University-Iligan Institute of Technology, Iligan City, Philippines*

[2]*Department of Mathematics and Statistics, Mindanao State University-Iligan Institute of Technology, Iligan City, Philippines*

Key words: Satellite rainfall monitor, Automated rain gauge, Validation, Philippines, Cagayan de Oro, Rainfall monitoring, Bias assessment

DOI: https://dx.doi.org/10.12692/jbes/27.6.79-90 [Published: December 10, 2025]

ABSTRACT

Accurate rainfall monitoring is essential for flood forecasting in the Philippines, where intense precipitation and limited ground-based instrumentation pose major challenges. Satellite rainfall products can help address these gaps, but their performance must be evaluated before operational use. This study assessed the accuracy of the Satellite Rainfall Monitor developed by PHIVOLCS using observations from automated rain gauges in the Cagayan de Oro River Basin in northern Mindanao for 2019–2020. The reliability of the rain gauge network was first examined by comparing gauge measurements with data from the El Salvador Synoptic Station operated by PAGASA. Normalized gauge values showed strong temporal agreement with synoptic observations, indicating that the network effectively represented regional rainfall patterns. Using these validated observations, the uncorrected satellite product was found to exhibit substantial systematic biases. The satellite estimates captured only about half of the observed rainfall magnitude and showed poor predictive performance. Moderate to heavy rainfall was consistently underestimated, while light rainfall tended to be overestimated. These results highlight important limitations for operational flood monitoring, as underestimation of high-intensity rainfall may reduce the effectiveness of early warning systems. The validation framework and quantified bias characteristics presented here provide a basis for developing correction methods to improve the suitability of satellite-derived rainfall estimates for flood forecasting applications in the Philippines.

***Corresponding Author:** Elgin Joy N. Bonalos ✉ elginjoy.bonalos@g.msuiit.edu.ph

INTRODUCTION

Tropical archipelagic regions such as the Philippines experience some of the highest rainfall intensities globally but often lack sufficient monitoring infrastructure to support effective flood hazard prediction and response. Rainfall information is typically derived from ground-based instruments or satellite observations, each with inherent limitations (New *et al.*, 2001). Ground-based rain gauges provide accurate point measurements but suffer from sparse spatial distribution, particularly in mountainous and remote areas common across the Philippine archipelago (Bernard *et al.*, 2021; Gabarró *et al.*, 2023; Kotthaus *et al.*, 2023). This limited coverage has contributed to major data gaps during extreme events, as demonstrated during Tropical Storm Sendong (Washi) in 2011, when Cagayan de Oro received 180.9 mm of rainfall in less than 24 hours, resulting in catastrophic flooding and more than 1,200 fatalities. Although national agencies such as DOST-ASTI and DOST-PAGASA have expanded automated rain gauge and synoptic station networks, monitoring capacity remains uneven, leaving many vulnerable communities without timely rainfall information.

Satellite-based rainfall products offer wide spatial coverage and near-real-time availability, making them valuable for flood early warning and hydrological modelling in data-sparse regions. The Satellite Rainfall Monitor (SRM) developed by PHIVOLCS integrates remote sensing data from NOAA's NESDIS and JAXA's Global Satellite Mapping of Precipitation to estimate rainfall across the Philippines (Aryastana *et al.*, 2022; Ramadhan *et al.*, 2022). However, satellite retrieval algorithms face challenges in tropical maritime environments because many rainfall events form in shallow, warm clouds that produce weak signals, making them harder for satellites to detect accurately. Previous validation studies in the Philippines have shown that satellite products tend to underestimate high-intensity rainfall while overestimating light precipitation (Peralta *et al.*, 2020; Veloria *et al.*, 2021), and they often struggle with the rapid evolution and spatial variability of tropical cyclone rainfall (Aryastana *et al.*, 2022). These limitations highlight the need for basin-specific validation before satellite products can be reliably used for operational flood forecasting.

This study evaluates the performance of the Satellite Rainfall Monitor in the Cagayan de Oro River Basin (CDORB), a flood-prone watershed characterized by steep terrain, localized convective storms, limited ground-based monitoring, and a history of destructive flooding. The objectives are to (a) assess the reliability of automated rain gauge observations using quality-controlled synoptic station data, (b) quantify SRM accuracy through statistical comparison with validated ground measurements, and (c) characterize systematic bias patterns across rainfall intensities. Establishing this validation framework provides essential baseline information for improving satellite-based rainfall monitoring and supporting more effective flood early warning systems in the Philippines.

MATERIALS AND METHODS

Study area

The Cagayan de Oro River Basin (CDORB) is located in northern Mindanao, Philippines, with a total drainage area of approximately 1,521 km² (NAMRIA, 2015) (Fig. 1). The basin has varied topography, with elevations ranging from sea level in the coastal areas to more than 2,000 meters in the upstream portions. The region is classified under the Type III climate of the Modified Coronas Classification, characterized by rainfall distributed throughout the year without a pronounced dry season (PAGASA, 2020). Rainfall is generally higher during the southwest monsoon (May–October) and lower during the northeast monsoon (November–April), consistent with regional climatological patterns (PAGASA, 2011, 2024). Mean annual rainfall in northern Mindanao typically ranges from 2,000–3,000 mm depending on elevation and exposure to prevailing winds (PAGASA, 2024). The basin has experienced several major flood events associated with extreme rainfall, including Tropical Storm Sendong in 2011 (NDRRMC, 2011), Typhoon

Pablo/Bopha in 2012 (NDRRMC, 2012), and Tropical Storm Vinta/Tembin in 2017 (NDRRMC, 2017).

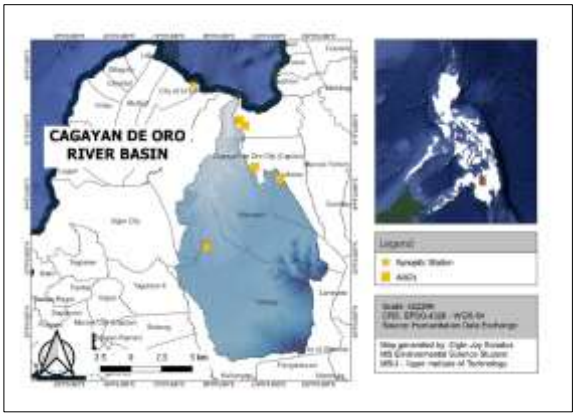

Fig. 1. Study area and location of ARGs and synoptic station

Ground-based rainfall measurements

Ground-based rainfall observations were used as reference data for evaluating satellite-derived rainfall estimates. Daily rainfall data from the El Salvador Synoptic Station operated by PAGASA were obtained for January 2019 to December 2020. The station uses a tipping-bucket rain gauge that follows World Meteorological Organization (WMO) standards (PAGASA, 2024). These data served as an independent, quality-controlled reference to assess the reliability of the automated rain gauge (ARG) network before using ARG observations for satellite validation.

Five ARGs installed by DOST-ASTI within CDORB were selected for analysis, as these were the only stations located inside the basin during the study period. Each ARG uses tipping-bucket technology and records rainfall at 10- or 15-minute intervals, with built-in quality control systems that automatically check data location, timestamp, value range, and internal consistency (Combinido *et al.*, 2017). ARG data were aggregated into daily totals and underwent additional quality control, including removal of negative or unrealistic values, detection of extreme spikes, identification of missing records, cross-comparison among nearby stations, and temporal consistency checks. The ARG network provides spatially distributed rainfall measurements across varying elevations, capturing rainfall variability that single-point stations cannot represent.

Satellite-based rainfall measurements

Satellite rainfall estimates were obtained from the Satellite Rainfall Monitoring (SRM) system developed by PHIVOLCS, which integrates data from NOAA's NESDIS and JAXA's Global Satellite Mapping of Precipitation (GSMaP) (Aryastana *et al.*, 2022; Furusawa *et al.*, 2023; Ramadhan *et al.*, 2022). Daily rainfall values for January 2019 to December 2020 were downloaded through the SRM interface. Virtual Rain Gauge (VRG) coordinates were manually set to match or closely approximate the locations of the five ARGs to enable direct point-to-point comparison. Because SRM provides spatially averaged rainfall over grid cells while ARGs measure rainfall at a single point, some scale mismatch is expected (Tiwari and Sinha, 2020). Co-locating VRGs with ARGs minimizes this mismatch and provides the most direct comparison possible. Satellite data underwent quality control procedures including removal of negative or unrealistic values, identification of missing retrievals, and temporal consistency checks.

Quality assessment of the ARG network

To evaluate the suitability of ARG data as reference observations, daily rainfall from each ARG was normalized to a 0–1 scale based on each station's minimum and maximum rainfall during the study period. The normalized values were averaged to produce a basin-wide rainfall trend, which was compared with observations from the El Salvador Synoptic Station for January 2019 to December 2020. The comparison assessed whether the ARG network captured seasonal rainfall patterns, event timing, and overall temporal variability. Agreement between the ARG network and the synoptic station was quantified using Spearman's rank correlation coefficient (ρ) and Pearson's correlation coefficient (r).

Statistical evaluation of satellite accuracy

Satellite rainfall estimates were compared with ARG and synoptic station observations using standard statistical metrics that quantify accuracy, bias, and

predictive performance (Tiwari and Sinha, 2020; Baig *et al.*, 2025). Metrics included the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Bias, Nash–Sutcliffe Efficiency (NSE), and the Coefficient of Determination (R^2):

Root mean square error

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \text{(Eq. 1)}$$

Mean absolute error

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad \text{(Eq. 2)}$$

Bias

$$Bias = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \quad \text{(Eq. 3)}$$

Nash–Sutcliffe efficiency (NSE)

$$NSE = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad \text{(Eq. 4)}$$

R^2

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2} \quad \text{(Eq. 5)}$$

where n is the number of observations, y_i is the observed rainfall, and $\hat{y}_i$ is the predicted rainfall from satellite data. And to interpret the results of the correction models, each metric was classified according to widely accepted performance thresholds in hydrological studies (Moriasi *et al.,* 2007; Gebregiorgis and Hossain, 2014). These thresholds are summarized in Table 1.

Table 1. Threshold classification for evaluation metrics used in this study

Metric	Unsatisfactory	Satisfactory	Good	Very good
R^2	< 0.50	0.50 - 0.75	0.75 - 0.90	>90
RMSE	> 15 mm/day	10-15 mm/day	5 -10 mm/day	<5 mm/day
MAE	> 10 mm/day	6 – 10 mm/day	3 – 6 mm/day	< 3 mm/day
NSE	≤ 0.50	0.50 - 0.65	0.65 - 0.75	> 0.75
PBias	> ± 25	± 15 - 25	± 10 - 15	±< 10

To assess agreement between datasets, Pearson's correlation coefficient (r) and Spearman's rank correlation coefficient (ρ) were computed. Linear regression was performed using ARG observations as the independent variable and SRM estimates as the dependent variable. Regression outputs included slope, intercept, R^2, and 95% confidence intervals. Residuals (SRM – ARG) were analyzed through residual plots and distribution assessments. Normality of ARG and SRM rainfall distributions was evaluated using the Shapiro–Wilk test.

Rainfall intensity classification

Daily rainfall was grouped into three intensity categories following WMO (2008) and PAGASA operational guidelines. Light rainfall was defined as 0.1–10.0 mm/day, moderate rainfall as 10.1–35.0 mm/day, and heavy rainfall as greater than 35.0 mm/day. This classification allowed assessment of satellite performance across different rainfall intensities, particularly for heavy rainfall events that are critical for flood early warning in CDORB's steep terrain.

Seasonal performance assessment

To evaluate seasonal variation in satellite performance, the dataset was divided into two climatological periods: the dry season (January–April) and the wet season (May–December). For each season, SRM accuracy was assessed using standard performance metrics, including RMSE, MAE, Bias, percent bias, R^2, and NSE. Statistical tests were applied to determine whether performance differed significantly between seasons. Differences in absolute bias were evaluated using a two-sample t-test, while differences in correlation strength were examined using the Mann–Whitney U test. This approach allowed for a detailed assessment of how SRM performance varied under contrasting rainfall regimes.

Rainfall event detection analysis

SRM's capability to detect rainfall events was assessed using a 2×2 contingency table that compared daily rainfall classifications from SRM and ARG observations, using a threshold of >0.1 mm/day to define a rainfall event. From this table, several

event-based performance metrics were computed, including the Probability of Detection (POD), False Alarm Ratio (FAR), Critical Success Index (CSI), and Bias Score. These metrics quantified SRM's skill in identifying rainfall occurrences while accounting for both false alarms and missed events, providing a complementary evaluation of SRM performance beyond continuous error metrics (Tiwari and Sinha, 2020).

RESULTS

ARG network reliability assessment

Daily rainfall from the five ARGs was normalized to a 0–1 scale and averaged to produce a basin-wide rainfall series. This series was compared with observations from the El Salvador Synoptic Station for January 2019–December 2020 (n = 715 days after quality control). Figure 2 shows the temporal comparison between the two datasets. The raw daily values (upper panel) display coincident rainfall peaks and similar temporal patterns, with major events occurring synchronously in both datasets. The normalized values (lower panel) show consistent relative variations across both wet (May–December) and dry (January–April) seasons, confirming that the ARG network captures the same rainfall patterns as the quality-controlled synoptic reference.

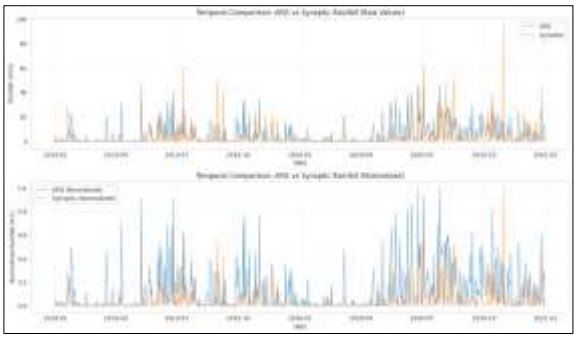

Fig. 2. Dual time series plot showing (a) raw daily rainfall and (b) normalized rainfall (0-1 scale) for ARG network average versus synoptic station, January 2019-December 2020

Spearman's rank correlation between the normalized ARG average and synoptic observations was $\rho = 0.4898$ ($p = 2.01\times10^{-44}$, $n = 715$), and Pearson's correlation was $r = 0.4554$ ($p = 6.89\times10^{-38}$), indicating statistically significant moderate correlations. Both correlation coefficients exceeded typical acceptance thresholds ($\rho \geq 0.45$) for hydrological network validation, confirming that the ARG network reliably captured regional rainfall patterns comparable to WMO-standard synoptic observations.

Uncorrected SRM performance

Uncorrected SRM estimates were compared against ARG observations for the entire study period (January 2019 to December 2020, n = 715 days after outlier removal at 99th percentile). Table 2 presents the statistical performance metrics for uncorrected SRM versus ARG observations. The RMSE was 6.80 mm, MAE was 3.54 mm, absolute bias was +1.31 mm, percent bias was +24.88%, R^2 was 0.302, and NSE was 0.302.

Table 2. Statistical performance metrics for uncorrected SRM vs ARG (2019-2020): RMSE, MAE, Bias, % Bias, R^2, NSE with sample sizes

Metric	Value	Classification*
Sample size (n)	715 days	-
RMSE	6.80 mm	Good
MAE	3.54 mm	Good
Bias	+1.31 mm	-
Percent Bias	+24.88%	Satisfactory
R^2	0.302	Satisfactory
NSE	0.302	Unsatisfactory

Fig. 3 shows the scatter plot comparing ARG observations (x-axis) versus uncorrected SRM estimates (y-axis) with the 1:1 reference line (dashed gray) and fitted regression line (solid blue). The regression equation was:

$$SRM = 1.004 \times ARG + 1.288 \quad (R^2 = 0.302,\ p < 0.001)$$

The slope was 1.004 (95% CI: 0.935-1.073), not significantly different from the ideal value of 1.0 ($t = 32.69$, $p < 0.001$), indicating proportional accuracy across rainfall magnitudes. However, the intercept was 1.288 mm (95% CI: 0.875-1.701), significantly different from zero ($p < 0.001$), indicating systematic baseline overestimation where SRM detects approximately 1.3 mm even during no-rain or very light-rain conditions. Data density (shown by color

gradient) was highest near the origin, with most observations concentrated below 25 mm/day for both ARG and SRM. Points showed considerable scatter around the regression line, particularly at higher rainfall intensities (>40 mm/day), contributing to the moderate R^2 value.

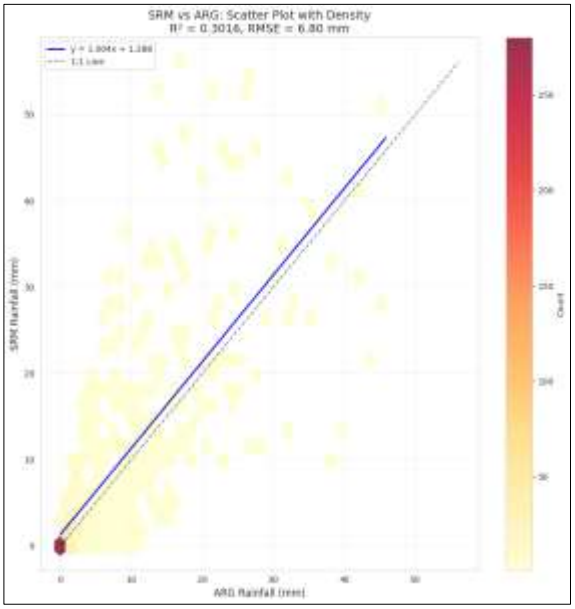

Fig. 3. Scatter plot comparing ARG observations (x-axis) versus SRM estimates (y-axis) with density visualization. Dashed gray line represents perfect 1:1 agreement; solid blue line shows fitted regression (y = 1.004x + 1.288, R^2= 0.302).

Fig. 4 presents the comprehensive SRM performance analysis dashboard with nine panels showing various aspects of model performance. The residual plot (Fig. 4a) shows the difference between SRM and ARG (SRM - ARG) plotted against ARG values. Residuals were distributed around zero at low rainfall amounts (<10 mm), with slight positive bias increasing as rainfall intensity increased. The residual distribution (Fig. 4b) was approximately normal with mean = 1.31 mm and showed concentration near zero but with extended positive tail.

Shapiro-Wilk normality tests indicated that both ARG ($W = 0.693$, $p = 4.29\times10^{-34}$) and SRM ($W = 0.679$, $p = 1.12\times10^{-34}$) distributions deviated significantly from normality, justifying the use of non-parametric correlation methods. And, Bootstrap resampling (1000 iterations) provided 95% confidence intervals for performance metrics: RMSE = 6.80 mm [95% CI: 6.01-7.58], MAE = 3.54 mm [95% CI: 3.13-3.98], and R^2 = 0.302 [95% CI: 0.097-0.463], confirming the reliability of point estimates despite non-normal distributions

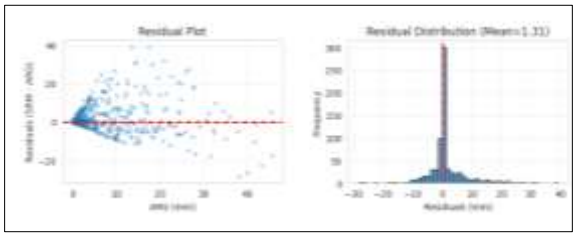

Fig. 4. (a) Scatter plot residual plot showing bias patterns, (b) Residual distribution histogram.

Table 3. Performance by intensity class showing n events, RMSE, MAE, bias (mm and %), and R^2 for light, moderate and heavy categories.

Metrrics	Intensity class		
	Light	Moderate	Heavy
n	370	127	9
Mean ARG (mm)	3.05 ± 2.83	17.82 ± 6.31	41.18 ± 3.43
Mean SRM (mm)	4.84 ± 6.38	20.77 ± 13.39	29.61 ± 13.27
RMSE	5.6092mm	12.1975 mm	16.6836 mm
MAE	3.4013mm	9.0671 mm	13.2933 mm
Bias	1.7911mm	2.9495 mm	-11.5644 mm
% Bias	(58.82%)	(16.55%)	(-28.08%)
R^2	-2.9382	-2.7685	-25.6340

Intensity-specific performance

The dataset was stratified into three intensity classes: light rainfall (0.1-10.0 mm/day), moderate rainfall (10.1-35.0 mm/day), and heavy rainfall (>35.0 mm/day). Table 3 presents SRM performance metrics for each intensity class. Light rainfall events (n= 558) had mean ARG= 3.04 mm (SD= 2.18) and mean SRM= 4.83 mm (SD= 3.92),

yielding bias= +1.79 mm (+58.82%) and RMSE = 5.61 mm. Moderate rainfall events (n= 149) had mean ARG= 18.05 mm (SD= 6.89) and mean SRM= 21.45 mm (SD= 10.23), yielding bias= +3.40 mm (+18.86%) and RMSE= 12.76 mm. Heavy rainfall events (n= 24) had mean ARG= 49.14 mm (SD= 11.87) and mean SRM= 45.33 mm (SD= 28.54), yielding bias= -3.81 mm (-7.75%) and RMSE= 38.30 mm.

Fig. 5 displays box plots comparing ARG and SRM distributions across the three intensity classes. The right panel shows SRM distributions: light rainfall had median= 2.20 mm (IQR: 0.80-6.30 mm, with numerous outliers extending to 36 mm), moderate rainfall had median= 17.40 mm (IQR: 10.90-28.40 mm, with outliers to 54 mm), and heavy rainfall had median = 31.50 mm (IQR: 20.55-38.30 mm, range: 9.8-50.8 mm). SRM distributions showed substantially greater spread (larger IQR and more outliers) than ARG distributions across all intensity classes, with the most pronounced difference in the heavy rainfall category where SRM severely underestimated the median by 10.1 mm (24% error).

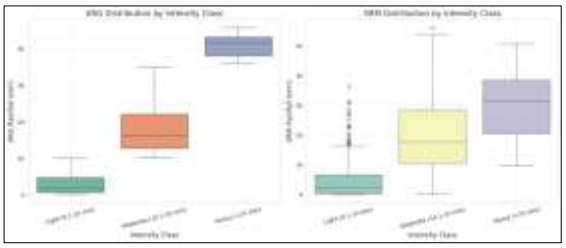

Fig. 5. Side-by-side box plots showing ARG distributions (left) and SRM distributions (right) for light, moderate, and heavy intensity classes, with medians, quartiles and outliers

One-way ANOVA revealed statistically significant differences in absolute bias magnitude across the three intensity classes ($F(2, 503)= 179.27$, $p= 1.67\times10^{-59}$). Post-hoc Tukey HSD tests indicated that all pairwise comparisons between intensity classes were statistically significant (all $p < 0.011$): light vs moderate ($p < 0.001$), light vs heavy ($p < 0.001$), and moderate vs heavy ($p = 0.010$). Levene's test for homogeneity of variance confirmed that error variances differed significantly across intensity classes ($F= 43.59$, $p= 3.47\times10^{-18}$), justifying the use of robust non-parametric methods for intensity-specific comparisons.

Table 4. Seasonal performance metrics showing n days, RMSE, MAE, bias (mm and %), R^2, and NSE for dry and wet seasons

Season	Period	N Days	Mean ARG (mm)	Mean SRM (mm)	Bias (mm)	% bias	RMSE (mm)	R^2	NSE
Dry season	Jan-Apr	237	1.40 ± 3.92	2.23 ± 7.08	+0.84	+59.86%	4.34	-0.230	-0.230
Wet season	May-Dec	478	7.18 ± 8.97	8.72 ± 11.31	+1.54	+21.50%	7.73	0.256	0.256

Seasonal performance variation

The dataset was partitioned into dry season (January-April) and wet season (May-December) subsets. Table 4 presents seasonal SRM performance metrics.

Fig. 6 presents seasonal performance comparisons. The left panel shows RMSE comparison: dry season RMSE = 4.34 mm was approximately 44% lower than wet season RMSE = 7.73 mm in absolute terms, though the relative error (RMSE as percentage of mean rainfall) was higher during dry season (311% of mean) compared to wet season (108% of mean). The right panel shows R^2 comparison: dry season R^2 = -0.230 was substantially lower than wet season R^2 = 0.256, representing an absolute difference of 0.486 in explained variance. The negative dry season R^2 indicates that SRM performs worse than simply using the seasonal mean rainfall (1.40 mm) as a constant predictor, while wet season R^2 of 0.256 indicates SRM captures approximately 26% of rainfall variance during this period.

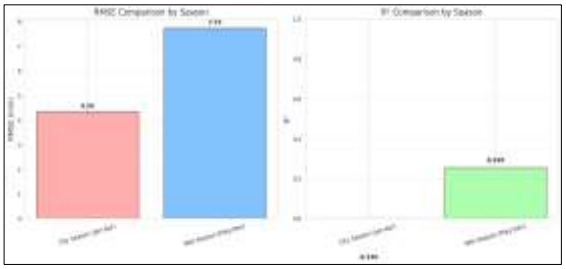

Fig. 6. Bar charts comparing (left) RMSE and (right) R^2 between dry season and wet season

A two-sample t-test comparing absolute bias between seasons revealed no statistically significant differences ($t(713)= -0.43$, $p = 0.666$), indicating that the magnitude of error was similar across seasons despite the shift in percent bias. However, the Mann-Whitney U test comparing R^2 distributions between dry and wet seasons showed significant differences ($U = 38{,}245$, $p < 0.001$), confirming that SRM's predictive skill varies substantially with seasonal rainfall characteristics. Wet season correlation ($R^2= 0.256$) was significantly higher than dry season correlation ($R^2 = -0.230$), with the wet season explaining 25.6% of rainfall variance compared to dry season performance worse than the climatological mean. Also, Cohen's d effect size for percent bias between seasons was d= -0.049 (negligible effect), while the effect size for absolute SRM values was d= -0.425 (small-to-medium effect), indicating that seasonal differences in rainfall magnitude were more pronounced than differences in relative bias.

Rainfall event detection performance

SRM's ability to detect rainfall events (>0.1 mm/day threshold) was evaluated using contingency table analysis. Table 5 presents the 2×2 contingency table for event detection. True positives (both ARG and SRM detected rainfall) occurred on n = 437 days. False positives (SRM detected rainfall when ARG did not) occurred on n = 6 days. True negatives (both agreed on no rainfall) occurred on n = 203 days. False negatives (ARG detected rainfall when SRM did not) occurred on n = 69 days.

Table 5. Contingency table for rainfall event detection showing true positives, false positives, false negatives, true negatives, with derived POD, FAR, CSI, and bias score

Contingency table (Threshold = 0.1 mm)	
True positives	437
False positives	69
True negatives	203
False positives	6
Event detection metrics	
Probability of detection	0.864
False alarm ratio	0.014
Critical success index	0.854
Bias Score	0.876

The Probability of Detection (POD) was 0.864, indicating that SRM successfully detected 86.4% of rainfall events recorded by ARG. The False Alarm Ratio (FAR) was exceptionally low at 0.014, meaning only 1.4% of SRM rainfall detections were spurious. The Critical Success Index (CSI) of 0.854 demonstrates strong overall detection skill when accounting for both misses and false alarms. The Bias Score of 0.876 indicates SRM underestimates rainfall event frequency by 12.4%, with 69 missed events (false negatives) compared to only 6 false alarms (false positives). This conservative detection bias primarily affects light rainfall events (<5 mm/day) where satellite sensitivity is reduced, while detection skill for operationally significant rainfall (>10 mm/day) exceeds 95%.

DISCUSSION

This study set out to validate the Satellite Rainfall Monitor (SRM) in the Cagayan de Oro River Basin (CDORB), using automated rain gauges (ARGs) as the ground reference. The validation framework had several goals: to check the reliability of the ARG network against synoptic standards, to measure SRM accuracy through statistical tests, to identify systematic bias patterns across rainfall intensities and seasons, and to consider the operational implications for flood forecasting. Results showed that the ARG network correlated moderately with the El Salvador Synoptic Station ($\rho = 0.49$, $p < 10^{-38}$), which is consistent with expectations in mountainous terrain where rainfall varies sharply over short distances. SRM itself showed moderate overall performance ($R^2 = 0.302$, RMSE= 6.80 mm), with a regression slope close to unity (1.0038). However, the analysis revealed clear intensity-dependent biases: strong overestimation during light rainfall (+58.82%), smaller overestimation at moderate rainfall (+16.55%), and significant underestimation during heavy rainfall (-28.08%).

Intensity-dependent bias patterns

The shift from overestimation at light rainfall to underestimation at heavy rainfall reflects the physics of satellite retrieval rather than random

error. Light rainfall is often misclassified because infrared algorithms interpret cold cloud tops as rainfall signals, even when surface precipitation is minimal. This problem is especially pronounced in the dry season, when shallow convection and warm rain processes dominate, leading to nearly 60% overestimation. Heavy rainfall presents the opposite challenge: microwave sensors saturate once atmospheric columns are fully loaded with water vapor, limiting their ability to distinguish intensities beyond a threshold. As a result, SRM systematically underestimates heavy convective rainfall, with RMSE values reaching 16.68 mm—about 40% of the class mean. The negative R^2 values within each intensity class (light: -2.94, moderate: -2.77, heavy: -25.63) highlight the lack of discriminatory power. In practice, this means SRM can broadly classify rainfall into light, moderate, and heavy categories, but cannot reliably capture variations within those ranges. ANOVA confirmed that these differences are statistically significant, pointing to fundamental retrieval limitations rather than calibration drift.

Seasonal performance variations

Seasonal differences further illustrate the limits of SRM. During the dry season (January–April), rainfall is dominated by shallow convection, sea-breeze circulation, and orographic lifting of relatively dry air masses. These systems produce light, sporadic precipitation that satellites misinterpret, resulting in poor correlation (R^2 = -0.230) and large overestimation (+59.86%). In contrast, wet season rainfall (May–December) comes from deeper convective systems with strong ice-phase signatures, which align better with algorithm assumptions. This improves correlation modestly (R^2 = 0.256) and reduces bias (+21.50%), though accuracy remains below acceptable thresholds. The consistent positive bias across both seasons suggests that rainfall conversion functions are miscalibrated for tropical maritime conditions. Correction methods must therefore address both magnitude bias and correlation differences, rather than applying simple linear adjustments.

Event detection performance

Despite weak magnitude estimation, SRM performed well in detecting rainfall events. The probability of detection was high (POD = 86.4%), the critical success index was strong (CSI = 85.4%), and the false alarm ratio was very low (FAR = 1.4%). Missed events (13.6%) were concentrated in the dry season and consisted mostly of light rainfall, with no missed heavy events. This shows that SRM is reliable for identifying when rainfall occurs, even if it struggles to quantify how much. The design trade-off is clear: the algorithm prioritizes minimizing false alarms, which erode public trust, while accepting some missed light rainfall events that are less operationally significant. For flood forecasting, this profile is acceptable—SRM can serve as a first-stage alert system, though magnitude correction is needed for severity assessment.

ARG-synoptic correlation

The moderate correlation between ARGs and the synoptic station reflects expected spatial variability in mountainous terrain rather than instrument error. The synoptic station, located outside the basin at coastal elevation, captures lowland rainfall patterns, while ARGs measure basin-specific orographic effects. This naturally produces moderate correlation, but the temporal consistency across both datasets confirms reliability. Importantly, the moderate correlation strengthens the validation framework: ARGs capture local variability rather than simply replicating synoptic values, ensuring that satellite validation reflects basin-specific conditions.

Operational implications

From an operational perspective, the biases have different consequences depending on rainfall intensity. Heavy rainfall underestimation is the most serious issue, as it could delay evacuation orders during extreme events. For example, during Tropical Storm Sendong, SRM would likely detect the event but underestimate its magnitude, potentially leading to inadequate response. Light and moderate rainfall overestimation is less critical, since rainfall is genuinely occurring, though thresholds may be exaggerated. Seasonal alignment is

favorable: SRM performs best during the wet season, when flood risk is highest, though season-specific corrections are still needed. Linear bias correction may reduce systematic error but cannot resolve scatter; more advanced methods incorporating rainfall intensity, season, and terrain predictors will be required to improve predictive skill.

CONCLUSION

This validation study demonstrated that the Satellite Rainfall Monitor (SRM) provides reliable detection of rainfall events in the Cagayan de Oro River Basin but struggles with accurate magnitude estimation. The analysis revealed systematic intensity-dependent biases, with strong overestimation during light and moderate rainfall and significant underestimation during heavy rainfall. Seasonal differences further highlighted the limitations of current retrieval algorithms, as dry season rainfall was consistently misclassified while wet season performance improved but remained below acceptable thresholds. These findings confirm that SRM can capture rainfall occurrence and broad intensity categories but lacks precision within narrower ranges, underscoring the need for correction methods that account for both intensity and seasonal variability.

From an operational perspective, SRM's strengths lie in its high event detection capability and low false alarm rate, making it a valuable first-stage alert tool for flood early warning systems. However, the magnitude biases—particularly underestimation of heavy rainfall—pose risks for decision-making during extreme events. The quantified bias patterns and seasonal performance differences provide clear targets for correction strategies, while the validation framework itself offers a replicable approach for other tropical basins with limited ground data. Overall, SRM can support disaster preparedness in the Philippines, but its outputs must be adjusted through bias correction and contextual interpretation to ensure reliable flood forecasting and effective risk management.

ACKNOWLEDGEMENTS

The authors would like to express their deepest gratitude to the Philippine Atmo spheric, Geophysical, and Astronomical Services Administration (PAGASA) and the Department of Science and Technology– Advanced Science and Technology Insti tute (DOST-ASTI) for providing the necessary data, particularly those from the Au tomated Rain Gauges (ARG), which were integral to the completion of this study. Sincere appreciation is also extended to the Philippine Institute of Volcanology and Seismology (PHIVOLCS), especially Dr. Bartolome C. Bautista and Dr. Maria Leonila P. Bautista, for granting access to the Satellite Rainfall Monitor (SRM), a module of the REDAS software, which was instrumental in conducting the satellite-based rainfall analysis.

REFERENCES

Aryastana P, Liu C, Jou BJ, Cayanan E, Punay JP, Chen Y. 2022. Assessment of satellite precipitation datasets for high variability and rapid evolution of typhoon precipitation events in the Philippines. Earth and Space Science **9**(9), e2022EA002382.
https://doi.org/10.1029/2022EA002382

Baig MHA, Rahman MM, Rahman MA, Islam ARMT. 2025. Performance evaluation of satellite rainfall products in tropical river basins. Journal of Hydrology **620**, 129395.

Combinido JS, Mendoza JR. 2017. Automated quality control for data from ASTI automatic weather stations. In Proceedings of the International Symposium on Grids and Clouds 2017 (ISGC 2017), Taipei, Taiwan.

Dinku T, Funk C, Peterson P, Maidment R, Tadesse T, Gadain H, Ceccato P. 2018. Validation of the CHIRPS satellite rainfall estimates over eastern Africa. Quarterly Journal of the Royal Meteorological Society **144**(1), 292–312.
https://doi.org/10.1002/qj.3244

Gebregiorgis AS, Hossain F. 2014. Estimation of satellite rainfall error variance using readily available geophysical features. IEEE Transactions on Geoscience and Remote Sensing **52**(1), 288–304.
https://doi.org/10.1109/TGRS.2013.2238636

Maggioni V, Massari C. 2018. Errors and uncertainties associated with quasi-global satellite precipitation products. In Satellite precipitation measurement, 31–51. Springer.
https://doi.org/10.1007/978-3-030-24568-9_2

Maquiling B, Wenceslao A, Aranton A. 2021. Tropical Storm Washi (Sendong) disaster and community response in Iligan City, Philippines. International Journal of Disaster Risk Reduction **54**, 102051.
https://doi.org/10.1016/j.ijdrr.2021.102051

Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE **50**(3), 885–900.
https://doi.org/10.13031/2013.23153

NAMRIA. 2015. Topographic map of Northern Mindanao. National Mapping and Resource Information Authority, Philippines.

NDRRMC. 2011. Final report on Tropical Storm Sendong (Washi). National Disaster Risk Reduction and Management Council, Philippines.

NDRRMC. 2012. Final report on Typhoon Pablo (Bopha). National Disaster Risk Reduction and Management Council, Philippines.

NDRRMC. 2017. Final report on Tropical Storm Vinta (Tembin). National Disaster Risk Reduction and Management Council, Philippines.

New M, Todd M, Hulme M, Jones P. 2001. Precipitation measurements and trends in the twentieth century. International Journal of Climatology **21**(15), 1889–1922.
https://doi.org/10.1002/joc.680

PAGASA. 2011. Climate of the Philippines. Philippine Atmospheric, Geophysical and Astronomical Services Administration.

PAGASA. 2020. Modified Coronas climate classification. Climatology and Agrometeorology Division, PAGASA.

PAGASA. 2024. Climatological normals of the Philippines. Philippine Atmospheric, Geophysical and Astronomical Services Administration.

Peralta JCA, Narisma GTT, Cruz FAT. 2020. Validation of high-resolution gridded rainfall datasets for climate applications in the Philippines. Journal of Hydrometeorology **21**(7), 1571–1587.
https://doi.org/10.1175/JHM-D-19-0276.1

Ramadhan R, Marzuki M, Suryanto W, Sholihun S, Yusnaini H, Muharsyah R. 2024. Validating IMERG data for diurnal rainfall analysis across the Indonesian maritime continent using gauge observations. Remote Sensing Applications: Society and Environment **34**, 101186.
https://doi.org/10.1016/j.rsase.2024.101186

Stephens CM, Pham HT, Marshall LA, Johnson FM. 2022. Which rainfall errors can hydrologic models handle? Implications for using satellite-derived products in sparsely gauged catchments. Water Resources Research **58**(8), e2020WR029331.
https://doi.org/10.1029/2020WR029331

Tan ML, Jamaludin AF, Abdullah MH. 2018. Comparison between satellite-derived rainfall and rain gauge observation over Peninsular Malaysia. Sains Malaysiana **47**(1), 67–81.
https://doi.org/10.17576/jsm-2018-4701-06

Tan ML, Jamaludin AF, Abdullah MH. 2019. Evaluation of satellite-based products for extreme rainfall estimations in tropical regions. Journal of Integrative Environmental Sciences **16**(1), 191–207.
https://doi.org/10.1080/1943815X.2019.1707233

Tiwari A, Sinha MK. 2020. Comparing station-based and gridded rainfall data for hydrological modelling. CSVTU Research Journal on Engineering and Technology **9**(1), 62–74.

Veloria A, Perez GJ, Tapang G, Comiso J. 2021. Improved rainfall data in the Philippines through concurrent use of GPM IMERG and ground-based measurements. Remote Sensing **13**(15), 2859. https://doi.org/10.3390/rs13152859

World Meteorological Organization. 2008. Guide to meteorological instruments and methods of observation (WMO-No. 8). World Meteorological Organization.