MEDIA KOMUNIKASI TEKNIK SIPIL
Jurnal Ilmu dan Terapan Bidang Teknik Sipil
Website : mkts.sipil.undip.ac.id
e-ISSN : 2549 - 6778
doi: 10.14710/mkts.v31i1.72441

Comparison of GPM and ARR Rain Distribution Patterns in Design Flood Simulation

*Yunita Ayu Setiyowati, Donny Harisuseno, M. Amar Sajali
Faculty of Engineering, Brawijaya University Malang
*)yunita.wre@gmail.com

Received: 15 April 2025 Revised: 8 July 2025 Accepted: 4 August 2025

Abstract

This study evaluates the performance of Global Precipitation Measurement (GPM) satellite-based rainfall data in comparison to Automatic Rainfall Recorder (ARR) data in forming rainfall distribution patterns and assessing its impact on flood discharge simulation using the HEC-HMS model. Statistical validation was conducted using the Pearson Correlation Coefficient, the ratio of standard deviation of observations to RMSE (RSR), Percent Bias (PBIAS), and Mean Absolute Percentage Error (MAPE). The results show that GPM has a strong correlation with ARR (r = 0.875) and a low RSR value (RSR= 0.256), yet it exhibits a notable negative bias (PBIAS = –24.41%), indicating an underestimation of rainfall values. In contrast, simulations using ARR rainfall patterns produce peak discharges that closely match actual discharge records at the Jatigede Dam outlet, with an average deviation of less than 3% and a MAPE of 1.17%, categorized as very good. The GPM simulation produces peak discharges 13–16% higher than actual observations, with a MAPE of 14.53%, which still falls into the good category. These results suggest that while ARR provides higher accuracy, GPM remains a viable alternative source, especially in data-scarce areas, provided that appropriate calibration methods such as bias correction are applied. This study supports future research in satellite data calibration using machine learning and multivariate statistical approaches.

Keywords: *rainfall distribution, GPM, ARR, HEC-HMS, discharge, validation*

Abstrak

Penelitian ini mengevaluasi kinerja data curah hujan berbasis satelit Global Precipitation Measurement (GPM) dibandingkan dengan Automatic Rainfall Recorder (ARR) dalam membentuk pola distribusi hujan dan dampaknya terhadap simulasi debit banjir menggunakan model HEC-HMS. Validasi statistik dilakukan menggunakan Koefisien Korelasi Pearson, rasio RMSE terhadap deviasi standar observasi (RSR), Persentase Bias (PBIAS), dan Mean Absolute Percentage Error (MAPE). Hasil menunjukkan bahwa GPM memiliki korelasi tinggi terhadap ARR (r = 0,875) dan nilai RSR yang rendah (RSR = 0,256), tetapi menunjukkan bias negatif yang cukup signifikan (PBIAS = –24,41%), yang mengindikasikan kecenderungan meremehkan nilai curah hujan. Sebaliknya, simulasi debit dengan pola hujan ARR menghasilkan debit puncak yang sangat mendekati data aktual di outlet Bendungan Jatigede, dengan rata-rata deviasi kurang dari 3% dan MAPE sebesar 1,17% (akurasi sangat baik). Sementara itu, simulasi GPM menghasilkan debit puncak 13–16% lebih tinggi dengan MAPE sebesar 14,53% (masih tergolong akurasi baik). Hal ini menunjukkan bahwa meskipun ARR lebih akurat, GPM tetap dapat digunakan sebagai sumber alternatif, khususnya di wilayah minim data, asalkan dilakukan kalibrasi terlebih dahulu. Studi ini juga mendorong penelitian lanjutan dalam pengembangan metode kalibrasi satelit berbasis pembelajaran mesin dan pendekatan statistik multivariat.

Kata kunci: *distribusi hujan, GPM, ARR, HEC-HMS, debit, validasi*

Introduction

The temporal distribution of rainfall is an important aspect of hydrological analysis, particularly in flood discharge estimation and water control infrastructure design. Rainfall distribution patterns affect the shape and the peak of runoff discharge, so the accuracy of their selection is crucial in water resources planning. In Indonesia, empirical rain distribution patterns such as the Huff Rain

Distribution are still widely used, especially in conditions of limited high-intensity rainfall data and hourly resolution.

With the development of remote sensing technology, satellite rainfall data such as Global Precipitation Measurement (GPM) is an important alternative in filling the void of observational data. GPM data has the advantage of half-hourly temporal resolution and wide spatial coverage, making it a promising source for hydrological analysis in the tropics. Various studies have shown the potential of GPM in approximating actual rainfall values at the surface, although the accuracy of these data is still affected by factors such as topography and local rainfall intensity (Nurul 'aini et al., 2024);(Sharma et al., 2020).

However, most research on the validation of GPM data against surface rainfall data still focuses on daily to monthly scales (Narulita, 2016). Studies that specifically analyze GPM rainfall distribution on a sub-hourly time scale and compare it with design distribution patterns such as ARR are very limited. In fact, differences in temporal distribution patterns can significantly affect discharge calculations, so a comprehensive evaluation of GPM distribution patterns is needed if they are to be used for hydrological design analysis (Tunas & Tanga, 2000).

For example, research in the Cikapundung Upstream watershed shows that the use of rainfall distribution methods that do not match local characteristics can lead to considerable deviations in peak discharge estimates. (Christian et al., 2017). This finding reinforces the importance of selecting the right rainfall distribution pattern, as errors in this stage can directly impact the accuracy of hydraulic design and flood risk management. In addition, the study in Bekasi and Belitung Island emphasized that each region has a unique rainfall pattern, so a locally calibrated rainfall distribution approach is needed to make the hydrological analysis results more representative ((Ginting, 2022);(Narulita, 2016)).

Other studies have also shown that when hourly rainfall data is not available, the use of empirical rainfall distribution patterns can be an alternative solution for flood discharge simulation. Studies in the Jurug watershed show that the Mononobe method is more suitable than the Alternating Block Method (ABM) in producing realistic peak discharges. (Pratiwi & Satria Negara, 2023). This confirms that the selection of a rainfall distribution method depends not only on the availability of data but also on the suitability of rainfall pattern characteristics to the conditions of the study area.

However, studies that not only compare the rainfall distribution from GPM and ARR but also test its impact directly on the simulated discharge results, which are then validated with field observation data, are still very limited in number. This research gap needs to be filled so that the use of GPM data can truly be optimized in water resources planning in Indonesia. By juxtaposing the three key elements of rain distribution from high-resolution GPM data, ARR-designed rain distribution, and actual discharge data from field recording, it is expected that a comprehensive picture of the accuracy and effectiveness of the satellite-based rain distribution approach can be obtained.

Based on this background, this study aims to analyze and compare rain distribution patterns from half-hourly GPM data and ARR empirical patterns to determine which method produces more accurate and reliable discharge estimations for hydrological design. Both types of distribution patterns will be used in hydrological simulations to calculate the planned discharge in the Jatigede Dam Catchment Area (DTA), and then the results are validated with actual recorded discharge data. Thus, the results of this research are expected to make a real contribution in enriching the reference of rain distribution techniques in Indonesia as well as strengthening the basis for decision-making in the management and mitigation of hydrometeorological disaster risk.

Methods

The methodology of this research was centered on a comparative analysis using hydrological simulations within the Jatigede Dam catchment. This analysis was designed to evaluate the performance of two distinct rainfall distribution patterns: one derived from GPM satellite data and the other representing the ARR empirical patterns commonly used in Indonesia. The selection of the Jatigede Dam catchment was based on the complexity of its hydrological characteristics and the availability of adequate rainfall and discharge recording data required for the analysis.

The main data used consists of two groups, namely satellite-based rainfall data and surface rainfall data. Satellite data was obtained from the Integrated Multi-satellitE Retrievals for GPM (IMERG) Final Run product managed by NASA. The products used were GPM_3IMERGDF_06 for daily resolution and GPM_3IMERGHH v07 for half-hourly resolution. These data are downloaded from the NASA Giovanni portal and include spatially distributed precipitation information at 0.1° × 0.1° resolution. (G. Huffman et al., 2020).

Meanwhile, observational rainfall data was collected from six Automatic Rainfall Recorder (ARR) stations managed by the Cimanuk-Cisadane River Basin Center (BBWS). The stations include ARR Cikajang, Bayongbong, Leuwingitis, Sadawangi, Darmaraja, and Jatigede. Each ARR station is associated with the nearest GPM grid based on location coordinates, namely Grids 5, 9, 10, 17, 22, and 27, so that the comparison between GPM data and ARR data is done in the same area spatially. The complete study location can be seen in Figure 1. This process ensures the validity of the spatial representation between the two types of data and avoids distortion of results due to location mismatches.

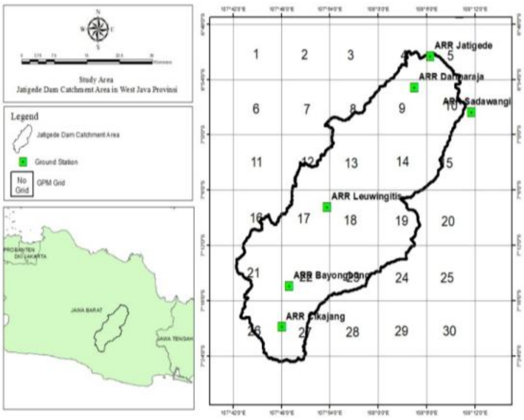

Figure 1. Study location

The data processing stage in this study begins with a data collection process that includes three main sources, namely GPM satellite rainfall data, ARR rain distribution patterns, and field recording discharge data. GPM data was downloaded in two time resolutions, namely daily and half-hourly, through the GPM IMERG Final Run product (GPM_3IMERGDF_06 and GPM_3IMERGHH v07). The six ARR stations used cover the Jatigede Dam catchment area, namely Cikajang, Bayongbong, Leuwingitis, Sadawangi, Darmaraja, and Jatigede.

After the extraction process, the GPM and ARR data were arranged in the form of hyetographs based on half-hourly time resolution, in order to visualize the temporal characteristics of rain distribution. Next, a comparison was made between the actual rain distribution pattern from GPM and the ARR design distribution pattern divided into four quartiles (Q1, Q2, Q3, and Q4). This step was carried out to evaluate the suitability of the shape and time of the rain peak between the two data sources.

To strengthen the analysis, rainfall data from both sources were then used as input for flood discharge simulation in the HEC-HMS model. This simulation aimed to observe how differences in temporal rainfall distribution patterns influence the resulting hydrograph characteristics, especially the timing and magnitude of peak discharge. The results of this modeling process serve as an important reference in assessing the potential impact of using satellite-based rainfall data for hydrological design, particularly in areas where ground station data is limited or sparse. This research process can be seen in Figure 2.

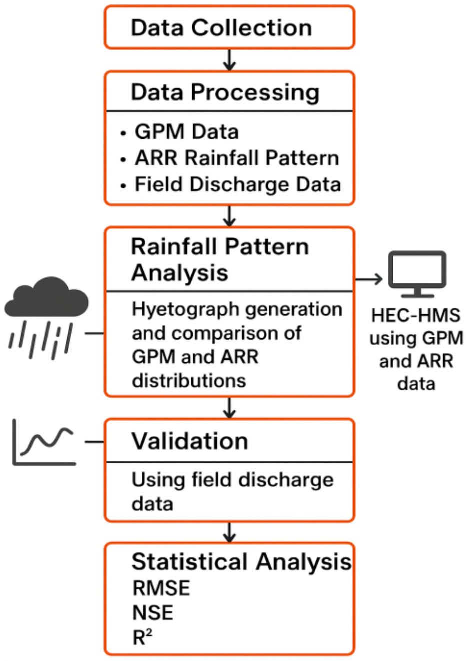

Figure 2. Research flow chart

Each rainfall event is analyzed by normalizing the duration of the event, by comparing the cumulative time to a certain interval to the total duration of rain. The rainfall data used in the analysis of rain distribution patterns, are rainstorm data with rain intensity > 50 mm/hour. (Huff, 1990). This step produces a cumulative time ratio, which represents the proportion of time to the total duration of rainfall events. Meanwhile, the rainfall depth at each interval is also normalized by dividing it by the total rainfall during one event, resulting in a cumulative rainfall ratio with a maximum value of 1.

Cumulative rainfall calculation is done using the following formula in Equation 1. (Harto, 1993).

$$P* = \frac{P_j \sum_{j=1}^{i}}{P_{total}} \quad (1)$$

Description P∗ is cumulative rainfall ratio (unitless, 0-1 scale), Pj is rain depth in period j (mm), $\sum_{j}^{i}$ is Accumulated rainfall in period I, Ptotal is total rainfall in one event

This normalization process aims to allow the comparison of distribution patterns between rainfall events of different durations and intensities. After the normalization process is performed, each rainfall event of similar duration is averaged to describe the general characteristic distribution.

The next step is the classification of rainfall distribution patterns based on the Huff method. (Huff, 1990), by grouping rainfall events into Quartile 1 to Quartile 4, depending on when the peak rainfall intensity occurs during the duration of the event. Quartile 1 indicates the peak occurs at the beginning of the duration, while Quartile 4 indicates the peak occurs at the end of the rainfall event. If the peak rainfall occurs in the first quarter of the total duration of the event, then the event is categorized as Quartile 1 (Q1). If the peak occurs in the second quarter, then it falls into Quartile 2 (Q2), and so on up to Quartile 4 (Q4).

This classification is done by identifying the time of occurrence of the highest rainfall intensity in an event, then comparing it with the standard distribution pattern of the Huff quartiles. After the grouping is done, for each quartile, the average value of the rainfall distribution of all events included in the category is calculated. So that the rainfall distribution of ARR data and GPM data is obtained.

The discharge simulation in this study was carried out using HEC-HMS software version 4.12, which is specifically designed to simulate the hydrological response of a watershed to rainfall inputs. The software allows modeling of various components of the hydrological system, such as the transformation of rainfall into surface runoff, in-channel flow propagation, and the incorporation of discharge from sub-watersheds. The HEC-HMS model was chosen for its ability to handle various rainfall scenarios and its flexibility in incorporating input data from various sources, including satellite data such as GPM.

In the initial stage of the simulation, the main parameters of the model were collected and determined. Parameters such as watershed area, time of concentration (Tc), land slope, main stream length, and Curve Number (CN) are obtained through spatial analysis using Geographic Information Systems (GIS), DEM (Digital Elevation Model) images, and references from relevant literature and secondary data. (Nageswara Rao, 2020). Curve Number is determined based on land use classification and soil type, which represents the infiltration capacity and surface runoff potential of an area. (Kincl et al., 2021). The parameters of the physical condition of the Jatigede watershed can be seen in Table 1, and for the HEC-HMS model used, it can be seen in Figure 3.

Simulations were conducted for three main scenarios, namely: (1) using rain distribution patterns based on ARR data, which is commonly used in Indonesia as a reference for hydrological design, (2) using actual rain distribution data from the GPM satellite with a resolution of 30 minutes, which is arranged in the form of a hyetograph to describe the temporal rain intensity, and (3) comparing the simulation results of the two scenarios with actual discharge data from field records at the Jatigede Dam outlet. This approach allows the evaluation of the performance of GPM data in generating discharge plans that are representative of real conditions.

Table 1. Parameter Model HEC-HMS

Land Cover	Soil Type	HSG	Area (km^2)	CN	CN Composite	S	Ia
Open land	Loam	B	10.40	82	72.99	94.02	9.40
Open land	Clay	D	5.20	89			
Dryland agriculture	Loam	B	403.03	71			
Dryland agriculture	Clay	D	134.34	81			
Mixed dryland farming	Loam	B	196.65	69			
Mixed dryland farming	Clay	D	98.33	80			
Sawah	Loam	B	170.35	72			
Primary dryland forest	Loam	B	4.84	79			
Secondary dryland forest	Loam	B	92.78	83			
Forest plantation	Loam	B	204.24	65			
Forest plantation	Clay	D	51.06	82			
Belukar	Loam	B	9.16	67			
Belukar	Clay Loam	D	0.19	83			
Plantation	Loam	B	8.30	73			
Settlement	Loam	B	90.32	75			
Water body	Clay	D	34.69	98			

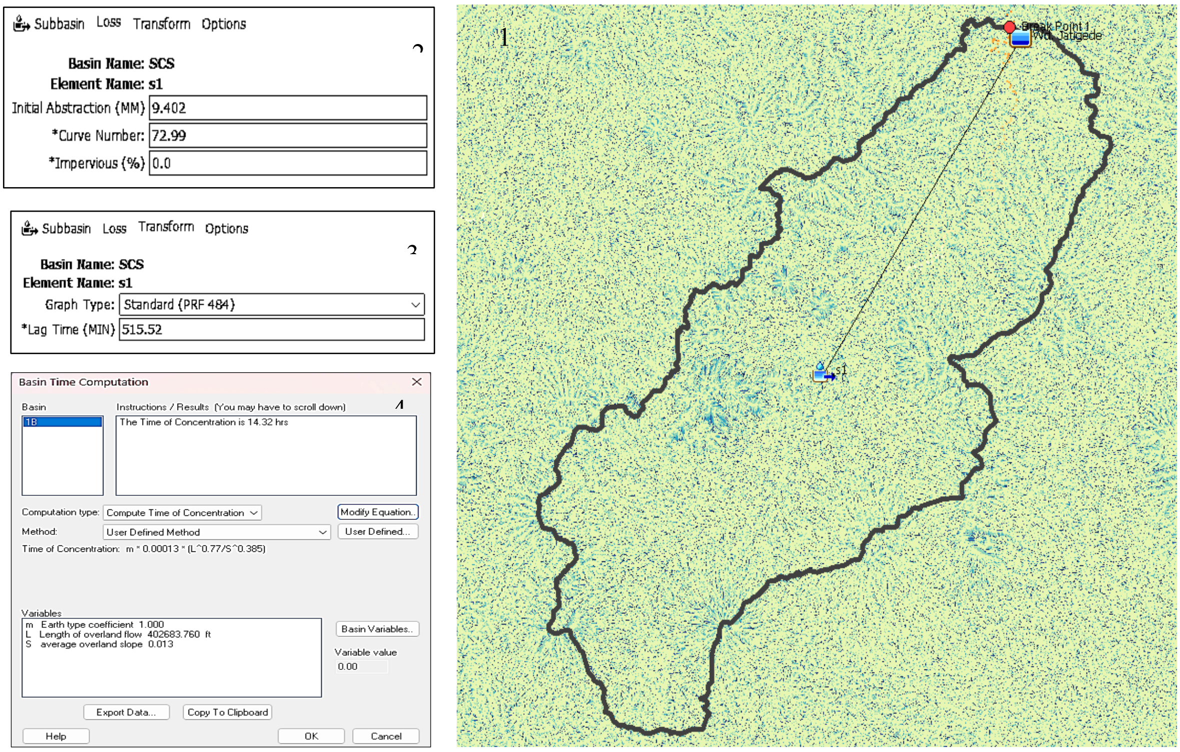

Figure 3. HEC-HMS Model 1) Single basin 2) Loss method 3) Transform method 4) TR-55

In evaluating the level of agreement between GPM satellite rainfall data and observations from the ARR rainfall station, three main statistical approaches commonly used in the validation of hydrological models and estimation data are used. These three methods aim to measure the closeness between the estimated results and the actual data, both in terms of linear relationship, error rate, and tendency of overestimation or underestimation.

First, the Pearson Correlation Coefficient (r) is used to measure how strong the linear relationship is between the two variables, namely, rainfall data from GPM and data from ARR stations. The value of this coefficient is in the range of -1 to 1, where values close to 1 indicate a very strong positive relationship, while values close to -1 indicate a strong negative relationship. If the value is close to zero, it can be concluded that there is no significant linear relationship between the two datasets.

The calculation of the r value is performed using the formula shown in Equation 2 (Daniel S. Wilks, 2011). Furthermore, an analysis using the RMSE-observations standard deviation ratio (RSR) is conducted to determine how large the mean square deviation is between the value predicted by the GPM data and the actual observed value of ARR. A smaller RSR value indicates a higher level of estimation accuracy. The formula to calculate RSR shown in Equation 3 and Equation 4 (D. N. Moriasi et al., 2007).

$$r = \frac{\Sigma_{i=1}^{N}(xi-\bar{x})(yi-\bar{y})}{\sqrt{\Sigma_{i=1}^{N}(xi-\bar{x})^2}\sqrt{\Sigma_{i=1}^{N}(yi-\bar{y})^2}} \tag{2}$$

Description r is correlation between GPM and ARR data, xi is ARR data in period I, yi is GPM data in period I, $\bar{x}$ is average ARR rainfall, $\bar{y}$ is average GPM rainfall, n is number of data.

$$RSR = \sqrt{\frac{1}{N}\Sigma_{i=1}^{N}(xi-yi)^2} \tag{3}$$

$$RSR = \frac{RMSE}{STDEVarr} \tag{4}$$

Description xi is ARR data in period I, yi is GPM data in period I, n is number of data

Percentage Bias (PBIAS), which gives an idea of the systematic tendency of the GPM in estimating rainfall. A positive bias value indicates that the GPM tends to overestimate the amount of rainfall compared to observational data, while a negative value indicates a tendency to underestimate. Ideally, a bias value close to zero indicates that the GPM estimate is very close to the observed value. The percentage bias is calculated by the formula that shown in Equation 5 (G. J. Huffman et al., 2020).

$$Bias = \frac{\Sigma_{i=1}^{N}(xi-yi)}{\Sigma_{i=1}^{N}xi} \; x \; 100\% \tag{5}$$

Description xi is ARR data in period I, yi is GPM data in period I, n is number of data.

To assess the level of performance of GPM satellite data in representing observed rainfall data from ARR stations, the results of statistical calculations such as Pearson correlation coefficients, RSR values, and percent bias need to be compared with predetermined evaluation criteria.

These criteria aim to provide a quantitative classification of the quality of rainfall estimates, making it easier to interpret how well satellite data is able to describe actual conditions in the field. This performance classification is divided into four levels, very good, good, satisfactory, and unsatisfactory, with the score ranges for each indicator detailed in Table 2.

In this study, the evaluation of the accuracy of discharge calculations was carried out using Mean Absolute Percentage Error (MAPE), which is a general metric for measuring how much the prediction error is relative to the actual value in percentage form. MAPE is calculated by the formula (Nabillah & Ranggadara, 2020):

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{At-Ft}{At}\right| x100\% \quad (5)$$

At is the actual value,Ft is the predicted value, and$n$ is the amount of data. The interpretation of MAPE values is generally categorized as follows: MAPE < 10% indicates excellent accuracy, 10-20% good, 20-50% fair, and >50% poor. The use of MAPE provides an intuitive picture of model performance, especially in the context of hydrology and river discharge forecasting, as it allows direct comparison between simulation results and observed data in easily understood units.

Results and Discussion

GPM and ARR rain distribution pattern analysis

The results of the analysis of rain distribution from two data sources, namely Automatic Rainfall Recorder (ARR) and Global Precipitation Measurement (GPM), show significant differences in their patterns and temporal characteristics. The ARR data shows that rainfall events with a duration of 5 hours are the most frequent in the study area (78 events).

Therefore, the entire analysis of the ARR rain distribution pattern is focused on rain events with a duration of 5 hours, as shown in Figure 4. Meanwhile, the GPM data shows that the dominant duration of the most frequently detected extreme rain events (storm events) is 9 hours (79 events), so the rain distribution pattern from GPM is analyzed based on events with this duration, as shown in Figure 5.

Table 2. Proximity Assessment

Performance assessment	Correlation coefficient	RSR	Relative bias (%)
Very good	0,75 –1,00	0,00 – 0,49	< ±10
Good	0,50 – 0,74	50 – 0,60	±10 – ±15
Simply	0,25 – 0,49	60 – 0,69	±16– ±25
Less	0,00 – 0,24	> 0,70	> ±25

Sources: (Krisnayanti et al., 2020); (Cabrera, 2009); (D. N. Moriasi et al., 2007)

Figure 4(1) shows the 5-hour distribution of ARR rainfall classified by the Huff method into four quartiles and shows modest variations between events. In Figure 4(2), the average ARR rain distribution tends to be closest to Huff Quartile 1, where the peak rain intensity occurs early in the event and decreases gradually. This reflects the early heavy rainfall characteristic of tropical regions with localized convective systems. The uniformity of the curve shape indicates that the ARR data provides more structured and consistent results, as it comes from direct recording at a single location point. (Da Silva et al., 2021)

In contrast, in Figure 5(1), the rain distribution from the GPM data over the 9-hour duration shows greater variation between events. This distribution suggests that the GPM captures highly variable temporal dynamics of rain, with peak rain positions that are not always at the beginning or end of the duration. Figure 5(2) shows that the average GPM rain distribution pattern falls between Huff Quartile 1 and Quartile 2. Visually, the shape resembles Quartile 1 because the rain peak tends to be at the beginning of the event, but the slope of the curve is more gentle.

This suggests that the rain intensity in the GPM increases more slowly and not as sharply as reflected in Huff Quartile 1. This characteristic is potentially due to the satellite estimation method based on cloud microphysics algorithms and passive sensors used in GPM, introducing uncertainties in the temporal resolution as well as the positional accuracy of the rain peak ((Da Silva et al., 2021);(Prakash et al., 2018);(Tang et al., 2016)).

The difference in distribution shape between ARR and GPM also indicates that the data acquisition method greatly influences the result of rainfall representation.

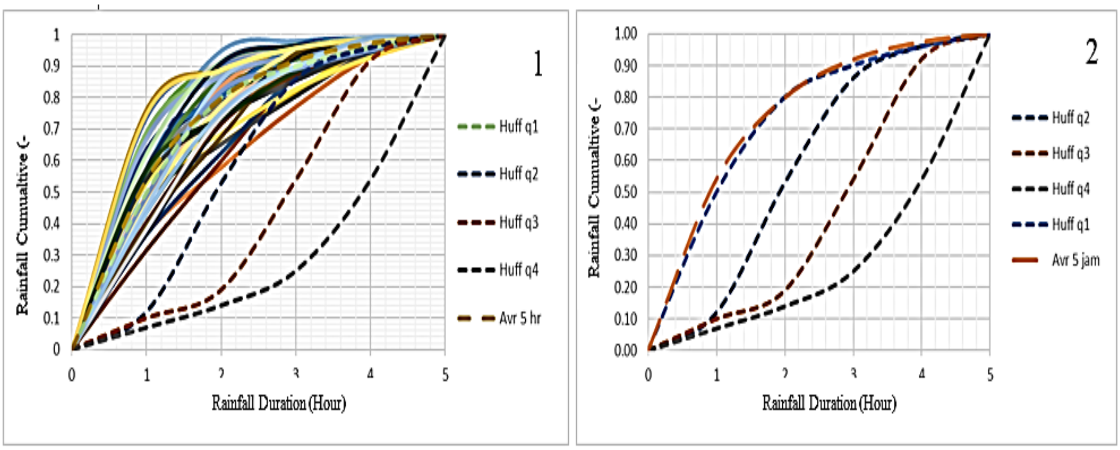

Figure 4 1) Comparison of ARR Rain Distribution Pattern with Huff (5 hours
2) Comparison of Average Rain Distribution Pattern of ARR with Huff (5 hours)

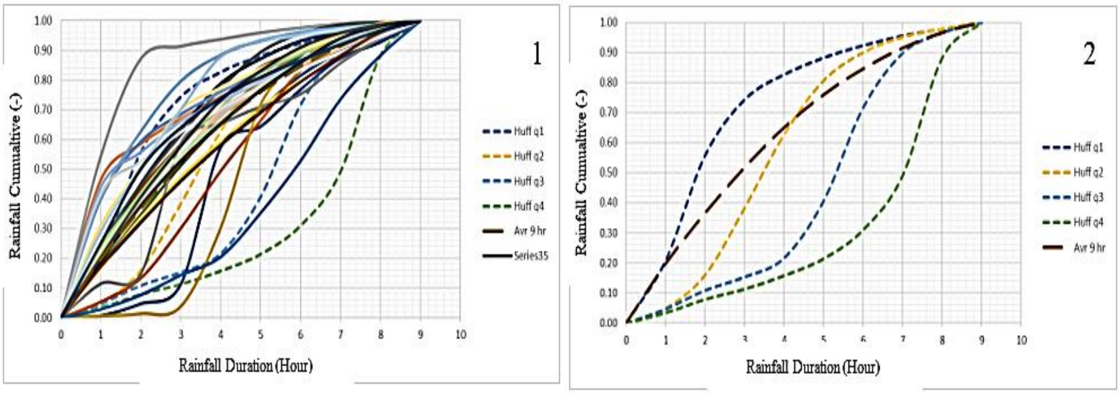

Figure 5. 1) *Comparison of GPM Rain Distribution Pattern with Huff (9 hours)*
2) Comparison of Average Rain Distribution Pattern of GPM with Huff (9 hours)

Point-based ARR data provides a more consistent distribution structure, while GPM as satellite-based spatial data shows a more diffuse pattern. Nevertheless, the GPM distribution is still able to show a general pattern similarity with the 1st Quartile of the Huff method, which strengthens its potential as an alternative rain modeling for areas that lack direct recording data.

Validation of GPM rain distribution pattern against ARR data

To evaluate the closeness between the GPM and ARR rainfall distribution data, a statistical analysis based on interpolation and time scale adjustment was performed. The ARR data originally available only up to the 5th hour was linearly interpolated and extrapolated up to the 9th hour, following the accumulated rain distribution approach, which generally peaks at mid-period and remains constant thereafter. (Dunkerley, 2022).

The validation results showed a Pearson Correlation Coefficient (r) value of 0.875, which statistically falls into the "Excellent" category (0.75-1.00). The RSR value of 0.256 was also classified as "Excellent" (0.00-0.49), indicating a low absolute deviation between the two curves. However, the PBIAS value of -24.41% indicates a significant negative bias, which is classified as "Fair" (±16% - ±25%) based on the model performance assessment criteria. The detailed validation results can be seen in Figure 6.

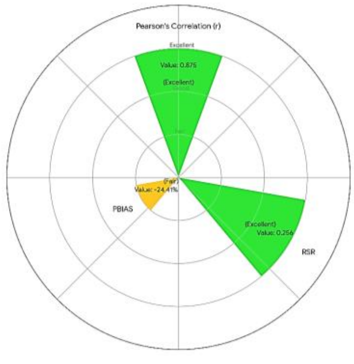

Figure 6. Polar Bar Chart Correlation, RSR, and PBIAS

Although the statistical values perform well numerically, visual interpretation of the accumulated rainfall distribution graphs shows a striking difference in the shape of the curves. The GPM distribution is consistently below the ARR, indicating that the GPM systematically underestimates the cumulative rainfall amount compared to the ARR observational data, especially in the early to mid rainy period. This calls into question the validity of statistical indicators such as correlation and RSR in describing absolute value matches. High correlations in this context reflect similarity in time trend patterns rather than a quantitative match to rainfall totals, so their use needs to be balanced with bias indicators such as PBIAS as well as visual interpretation. (Chai & Draxler, 2014).

Thus, although GPM performs well in capturing the temporal pattern of rainfall distribution, special attention needs to be paid to its accuracy in estimating actual rainfall volume, especially for hydrological applications such as flood discharge calculation or reservoir operation. (D. N. Moriasi et al., 2007). Although the RSR and correlation values show good results, the GPM distribution still shows quantitatively significant deviations. Therefore, a calibration process between GPM rainfall data (half-hourly resolution) and hourly ARR data is needed to make the estimation more representative. This also opens up further research opportunities in developing spatial-temporal correction and calibration methods for satellite data, especially in the context of tropical regions with high rainfall dynamics.

Simulation using GPM and ARR distribution patterns

Hydrological simulations in this study were conducted using HEC-HMS software version 4.12 with rainfall data input from two main sources, namely Global Precipitation Measurement (GPM) satellite data and local observation data from ARR developed in the form of Huff distribution patterns for Quartile 1 to Quartile 4. The GPM data, which has a temporal resolution of half an hour, was first converted to hourly form to match the model input format. The simulation includes three flood return period scenarios, namely Q5, Q10, and Q20 years, to get an idea of the performance of each rainfall distribution approach on the design flood discharge results.

The simulation results show that the peak discharge generated from the GPM rainfall distribution pattern for the three flood return periods is 825.76 m3/s for Q5, 855.30 m3/s for Q10, and 931.25 m3/s for Q20. When compared with the actual discharge recording data at the Jatigede Dam outlet, which are 715.57 m3/s for Q5, 762.14 m3/s for Q10, and 803.00 m3/s for Q20, respectively, there is a significant difference. In contrast, simulations using the ARR rainfall distribution pattern show results that are much closer to actual conditions, with peak discharges of 734.81 m3/s for Q5, 768.20 m3/s for Q10, and 802.96 m(3)/s Q20. Details of the peak discharge and discharge difference can be seen in Figure 7.

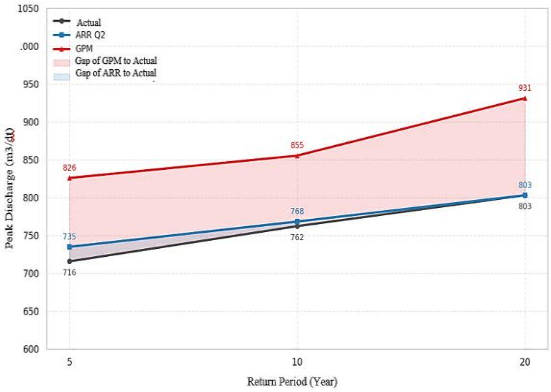

Figure 7. Graph of actual design flood discharge, ARR, and GPM

This difference indicates that the design flood peak discharge formed from the ARR rainfall distribution pattern is more representative of the actual conditions recorded in the field. This indicates that although GPM is quite capable of representing the shape of the temporal distribution of rain, the intensity and total volume of rain still show significant bias against reality. This finding is in line with previous studies by (Yuan et al., 2017) and (Liu et al., 2020) That highlighted the advantages of GPM in spatial and temporal coverage but suggested the need for bias correction before use in quantitative simulations.

By considering these results, it can be concluded that the rain distribution pattern developed from the ARR data is more accurate than the GPM in forming the design flood discharge. Based on the Mean Absolute Percentage Error (MAPE) calculation results, the peak discharge generated by the ARR model shows an error value of 1.17%, which is classified as very good accuracy, while the GPM model produces a MAPE of 14.53%, which is still in the good accuracy category.

This shows that although the ARR model is more representative of actual conditions, the rain distribution pattern from the GPM can still be used, but with the note that the calibration process needs to be done first. Without adjustment, the peak discharge from the GPM shows a considerable difference from the recorded data, making it less appropriate to use directly in planning. The calibration process can be done through bias correction approaches, such as quantile mapping or multivariate regression methods, which have proven effective in improving the accuracy of satellite rainfall estimates, as shown by (FNU Misnawati, 2022) . With the right calibration process, GPM data has great potential to be used as an alternative source of rainfall data in hydrological planning in areas that lack observational data while bridging the spatial limitations of conventional data.

Conclusion

This study concludes that the ARR empirical rainfall distribution pattern is significantly more accurate for simulating design flood discharge in the Jatigede Dam catchment compared to the pattern derived from uncalibrated GPM satellite data. The ARR pattern produced simulations with high precision (MAPE of 1.17%), closely matching observed data. In contrast, the GPM-derived pattern, despite showing a statistically good correlation, resulted in a considerable overestimation of peak discharge (MAPE of 14.53%). A key finding is that GPM tends to produce a temporal distribution that, when modeled, creates higher discharge volumes, highlighting a critical discrepancy for hydrological design.

Therefore, while GPM data is a valuable alternative for regions with limited ground observation, its direct use for hydrological design is not recommended without prior adjustment. Calibration is essential to correct the overestimation bias, for which methods like quantile mapping are highly recommended. Future research should be directed towards developing more robust calibration techniques, such as machine learning models, and integrating multiple remote sensing data sources to enhance the reliability of satellite-based rainfall for diverse hydrological applications.

Acknowledgment

With great respect, the authors would like to express their appreciation to the Education Fund Management Institution (LPDP) of the Ministry of Finance of the Republic of Indonesia for the scholarship support that has been provided. Thanks are also due to the Department of Irrigation Engineering, Brawijaya University, which has provided facilities and support during the process of conducting this research.

References

Cabrera. (2009). Calibración de Modelos Hidrológicos. *Imefen.Uni.Edu.Pe*, *1*, 1–7. http://www.imefen.uni.edu.pe/Temas_interes/modhidro_2.pdf

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Christian, K., Yudianto, D., & Rusli, S. R. (2017). Analisis Pola Distribusi Hujan Terhadap Perhitungan Debit Banjir DAS Cikapundung Hulu. *Jurnal Teknik Sumber Daya Air*, *3*(3), 153–160.

D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, & T. L. Veith. (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

Da Silva, N. A., Webber, B. G. M., Matthews, A. J., Feist, M. M., Stein, T. H. M., Holloway, C. E., & Abdullah, M. F. A. B. (2021). Validation of GPM IMERG extreme precipitation in the Peninsular Malaysia and Philippines by station and radar data. *IOP Conference Series: Earth and Environmental Science*, *893*(1). https://doi.org/10.1088/1755-1315/893/1/012020

Daniel S. Wilks. (2011). *Statistical Methods in the Atmospheric Sciences*.

Dunkerley, D. (2022). Regional Rainfall Regimes Affect the Sensitivity of the Huff Quartile Classification to the Method of Event Delineation. *Water (Switzerland)*, *14*(7). https://doi.org/10.3390/w14071047

FNU Misnawati. (2022). Bias Correction of Rain Estimation Using Gpm-Imerg Over Java Island. *Megasains*, *13*(2). https://doi.org/https://doi.org/10.46824/megasains.v13i2.113

Ginting, S. (2022). Pengembangan Kurva Distribusi Hujan Sintetis Di Kota Bekasi, Jawa Barat. *Jurnal Sumber Daya Air*, *18*(1), 13–25. https://doi.org/10.32679/jsda.v18i1.708

Harto, S. (1993). *Analisis Hidrologi*. Penerbit PT Gramedia Pustaka Utama.

Huff, F. (1990). Time Distributions of Heavy Rainstorms in Illinois Time Distributions of Heavy Rainstorms in Illinois. *Water Survey Campaign*, 22.

Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E., Sorooshian, S., Tan, J., & Xie, P. (2020). NASA GPM Integrated Multi-satellitE Retrievals for GPM (IMERG) Algorithm Theoretical Basis Document (ATBD) Version 06. *Nasa/Gsfc*, *January*, 29. https://pmm.nasa.gov/sites/default/files/imce/times_allsat.jpg%0Ahttps://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K.-L., Joyce, R. J., Kidd, C., Nelkin, E. J., Sorooshian, S., Stocker, E. F., Tan, J., Wolff, D. B., & Xie, P. (2020). *Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG) BT - Satellite Precipitation Measurement: Volume 1* (V.

Levizzani, C. Kidd, D. B. Kirschbaum, C. D. Kummerow, K. Nakamura, & F. J. Turk (eds.); pp. 343–353). Springer International Publishing. https://doi.org/10.1007/978-3-030-24568-9_19

Kincl, D., Kabelka, D., Vopravil, J., & Heřmanovská, D. (2021). Estimating the curve number for conventional and soil conservation technologies using a rainfall simulator. *Soil and Water Research*, *16*(2), 95–102. https://doi.org/10.17221/114/2020-SWR

Krisnayanti, D. S., Welkis, D. F. B., Hepy, F. M., & Legono, D. (2020). Evaluasi Kesesuaian Data Tropical Rainfall Measuring Mission (TRMM) dengan Data Pos Hujan Pada Das Temef di Kabupaten Timor Tengah Selatan. *Jurnal Sumber Daya Air*, *16*(1), 51–62. https://doi.org/10.32679/jsda.v16i1.646

Liu, J., Du, J., Yang, Y., & Wang, Y. (2020). Evaluating extreme precipitation estimations based on the GPM IMERG products over the Yangtze River Basin, China. *Geomatics, Natural Hazards and Risk*, *11*(1), 601–618. https://doi.org/10.1080/19475705.2020.1734103

Nabillah, I., & Ranggadara, I. (2020). Mean Absolute Percentage Error untuk Evaluasi Hasil Prediksi Komoditas Laut. *JOINS (Journal of Information System)*, *5*(2), 250–255. https://doi.org/10.33633/joins.v5i2.3900

Nageswara Rao, K. (2020). Analysis of surface runoff potential in ungauged basin using basin parameters and SCS-CN method. *Applied Water Science*, *10*(1), 1–16. https://doi.org/10.1007/s13201-019-1129-z

Narulita, I. (2016). Distribusi Spasial Dan Temporal Curah Hujan Di Das Cerucuk, Pulau Belitung. *Jurnal RISET Geologi Dan Pertambangan*, *26*(2), 141. https://doi.org/10.14203/risetgeotam2016.v26.194

Nurul 'aini, U., Bagus Adityawan, M., Prasetyo, A., Prabandini, F., Nugroho, R. W., Fathur, E., Madjodjo, A., Bandung, I. T., Ganesa, J., 10, N., Siliwangi, L., Coblong, K., Bandung, K., Barat, J., Teknik, B., Kementerian, P., Gilimanuk -Singaraja Km, J., Gerokgak, K., & Buleleng -Bali, K. (2024). Evaluasi Kesesuaian Data Satelit Global Precipitation Measurement (GPM) terhadap Stasiun Curah Hujan Disekitar Kawasan Inti Pusat Pemerintahan (KIPP) di Kabupaten Penajam Paser Utara. *Journal on Education*, *06*(02), 13438–13448.

Prakash, S., Mitra, A. K., AghaKouchak, A., Liu, Z., Norouzi, H., & Pai, D. S. (2018). A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *Journal of Hydrology*, *556*(February 2014), 865–876. https://doi.org/10.1016/j.jhydrol.2016.01.029

Pratiwi, D. W., & Satria Negara, A. (2023). Perbandingan Pola Distribusi Hujan Terukur dan Metode Empiris dalam Perhitungan Debit Banjir Rencana DAS Jurug. *Jurnal Teknik Sumber Daya Air*, *3*(1), 29–42. https://doi.org/10.56860/jtsda.v3i1.55

Sharma, S., Chen, Y., Zhou, X., Yang, K., Li, X., Niu, X., Hu, X., & Khadka, N. (2020). Evaluation of GPM-Era satellite precipitation products on the southern slopes of the central Himalayas against rain gauge data. *Remote Sensing*, *12*(11). https://doi.org/10.3390/rs12111836

Tang, G., Ma, Y., Long, D., Zhong, L., & Hong, Y. (2016). Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales. *Journal of Hydrology*, *533*, 152–167. https://doi.org/https://doi.org/10.1016/j.jhydrol.2015.12.008

Tunas, G., & Tanga, A. (2000). *Pengaruh Pola Distribusi Hujan Terhadap Penyimpangan Debit*.

Yuan, F., Zhang, L., Wah Win, K. W., Ren, L., Zhao, C., Zhu, Y., Jiang, S., & Liu, Y. (2017). Assessment of GPM and TRMM multi-satellite precipitation products in streamflow simulations in a data sparse mountainous watershed in Myanmar. *Remote Sensing*, *9*(3). https://doi.org/10.3390/rs9030302