

Correlation Analysis Between Measured Rain Data with Satellite at Rainfall Station in Merapi

Jazaul Ikhsan^{1*}, Amalia Kurnia Sari¹, Ani Hairani¹ and Alidina Nurul Hidayah²

¹ Department of Civil Engineering, Faculty of Engineering, Universitas Muhammadiyah Yogyakarta, Jalan Brawijaya, Kasihan, Bantul, Yogyakarta 55183, Indonesia

² Sabo Engineering Divison, Directorate of Water Resources, Ministry of Public Works, and Public Housing, Indonesia, Sopalana, Maguwoharjo, Depok, Sleman, Yogyakarta 55282, Indonesia

* Corresponding author email: jazaul.ikhsan.umy.ac.id



Keywords:

rainfall; correlation coefficient; satellite; GPM; PERSIANN.

Abstract

Manual and automatic rain gauges often need to be revised in measuring rainfall due to various constraints. Therefore, using rain data from satellites will be a promising alternative. The rain data used is measured hourly rainfall data >20mm in 2012, 2017, and 2022. In addition, rain data from the PERSIANN and GPM satellites were also used. The analysis was carried out using the correlation coefficient (r) method, which aims to find the correlation between measured rain data and satellite rain data. The results show that the PERSIANN satellite has the highest correlation value in rain duration in two years, while the GPM satellite has the highest total depth and intensity correlation value in two years. Therefore, it can be concluded that the GPM satellite has better accuracy than the PERSIANN satellite in monitoring rain.

INTRODUCTION

Indonesia has 127 active volcanoes, one of which is Mount Merapi located in the Special Region of Yogyakarta. The existence of volcanoes, especially active volcanoes, is one of the potential sources of natural disasters that needs to be watched out for (Badri & Yerizon, 2021; Fakhrudin & Elmada, 2022; Maharani et al., 2023). The disaster caused by a volcano is not only a primary disaster (when eruption occurs), but also a secondary disaster, for example, is a lava flood that occurs due to heavy rains (Munir, 2019).

In Yogyakarta and its surroundings, cold lava floods often occur due to the flow of water mixed with volcanic eruption materials. Cold lava floods carrying eruption materials cause siltation of the river so that the river can no longer withstand the flow, resulting in overflowing floods to various places and even to residential areas (Anonymous, 2019).

In flood calculations, the rainfall data required are the maximum daily rainfall rate, rainfall intensity, rainfall distribution patterns, and a network of rainfall stations capable of monitoring rainfall characteristics within the watershed with adequate rainfall recording periods (BSN, 2016). Rainfall data is sourced from rain gauge posts, divided into two types: manual and automatic rain gauges. However, this rainfall data is generally limited in space and time, so rainfall data often contains sampling errors (Anonymous, 2022).

One way that can be used to replace rain gauge data is to use satellite rain data (Ginting et al., 2019). Satellite rain data has many advantages compared to rain gauge data: accuracy, spatial coverage, timeliness, and cost efficiency (Vernimmen et al., 2012). However, there are still errors or differences in values between satellite rain data and rain gauge data (Mamenun et al., 2014). So, research is needed to determine the difference in value between satellite rain data and rain gauge posts.

Satellite Rainfall

The PERSIANN (Precipitation Estimation from Remotely Sensed Information Using Neural Networks) satellite has three products with different catchment area resolutions. The three products are PERSIANN (0.25°x0.25°) using Artificial Neural Networks (ANN) for precipitation estimation at various spatial and temporal scales, PERSIANN-CCS (0.04°x0.04°) using Cloud Classification System (CCS) for precipitation estimation and showing a tendency to underestimate rainfall in some areas and PERSIANN-CDR (0, 25°x0.25°) is a Climate Data Record (CDR) that provides global precipitation data with specific spatial and temporal resolution and is more accurate in capturing spatial and temporal patterns of extreme precipitation (Nguyen et al., 2018).

The GPM (Global Precipitation Measurement) satellite is an observation system designed to monitor rainfall globally across the Earth's surface. Unlike its predecessor, TRMM (Tropical Rainfall Measuring Mission), which only focused on observing rainfall in the tropics, GPM has a wider scope to observe rainfall around the world (Tan & Duan, 2017). In addition, GPM is also one of the rain-measuring satellites besides PERSIANN that is suitable for use in hilly or unevenly contoured areas. Although GPM has a lower resolution than PERSIANN, which is 0.1° x 0.1°, both satellites are considered to have the latest technology and high sensitivity in detecting rain data in an area. These advantages make GPM and PERSIANN the right choice for areas with diverse topographic conditions (Zhang et al., 2018).

Rainfall intensity

Rainfall intensity is the amount of rainfall in a unit of time, which is usually expressed in mm/hour, mm/day, mm/week, mm/month, mm/year and so on. The classification of rainfall intensity can be seen in table 1.

Table 1 Classification of rainfall intensity

Category	Rainfall (mm/day)
Cloudy	0
Light Rain	0.5-20
Medium Rain	20-50
Heavy Rain	50-100
Very Heavy Rain	100-150
Extreme Rain	>150

(Source : BMKG, 2024)

Correlation Coefficient

Correlation analysis is one of the statistical methods used to determine the strength of the relationship between two variables (Jarwanti et al., 2021). Correlation is expressed by a coefficient which shows the relative (linear) relationship between two variables. The correlation coefficient equation r is:

$$r = \frac{\sum_{i=1}^n (S_i - \bar{S})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (S_i - \bar{S})^2} \sqrt{\sum_{i=1}^n (O_i - \bar{O})^2}} \quad (1)$$

If the correlation coefficient value is greater, the stronger the relationship between the two, so that the estimated value pattern will be closer to the actual data pattern (Syaifullah, 2014). Categories of correlation coefficient values are shown in table 2.

Table 2. Correlation coefficient categories

Interval	Coefficient Level of Relationship
0.00 - 0.19	Very Weak
0.20 - 0.39	Weak
0.40 - 0.59	Medium
0.60 - 0.79	Strong
0.80 - 1.00	Very Strong

Source: Sugiyono (2013)

RESEARCH METHODS

The research was conducted using rainfall data that has been measured at Ngandong Station. Geographically, the station is located at coordinates 110° 24'11.10 (East) and 07° 35'46.10" (LS). The research location map can be seen in figure 1.

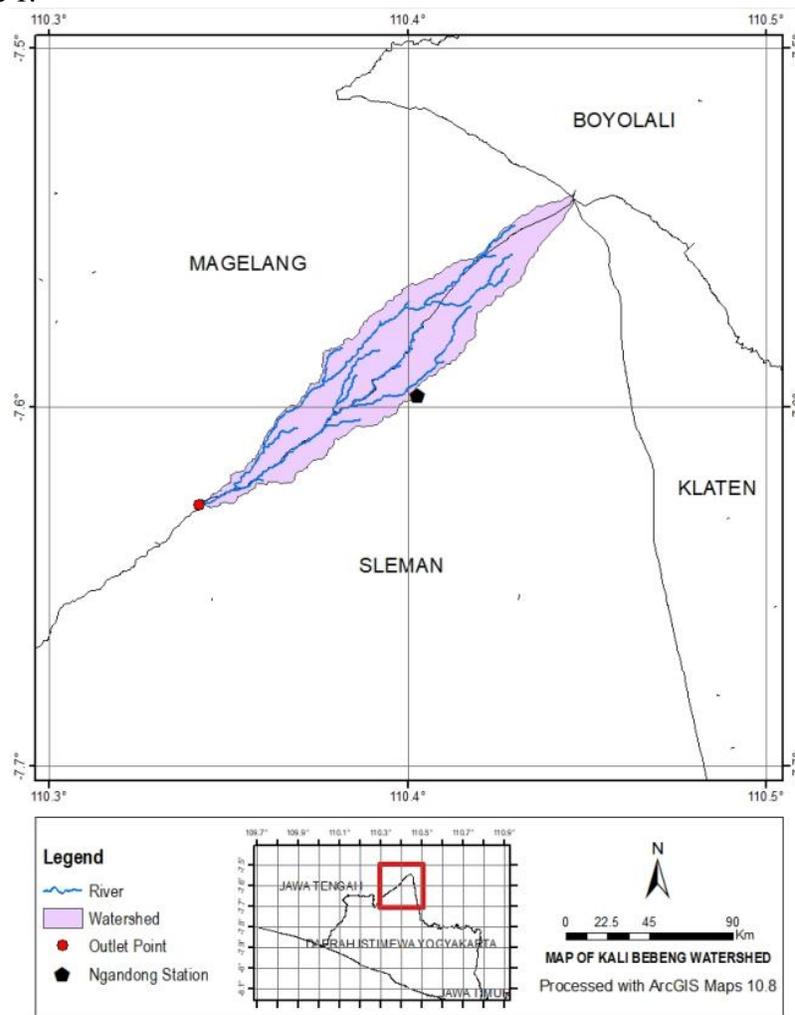


Figure 1. Map of Kali Bebeng Watershed

This study uses measured hourly rain data >20 mm per day at Ngandong Station, owned by the Sabo Engineering Center. In addition, this study also utilizes satellite rain data, namely PERSIANN (0.25° x 0.25°) accessed on the website: <http://chrdata.eng.uci.edu> and GPM satellite (0.1° x 0.1°) accessible on the website:

<https://journal.umy.ac.id/index.php/st/issue/view/1064>

<https://giovanni.gsfc.nasa.gov/giovanni/>. The use of measured and satellite rain data both use hourly rain data in the same period, namely 2012, 2017 and 2022.

Correlation analysis was conducted using measured rain data from Ngandong Station which was used as a reference in validating satellite rain data. The correlation analysis process between the two types of rain data is carried out by paying attention to the R^2 value on the trendline in the graph describing the relationship between measured rain data and satellite rain data. The analysis method uses the CORREL function available in Microsoft Excel. The parameters used in this study to perform correlation analysis are rain duration, total depth, and maximum intensity.

Flow Chart

The stages of research carried out can be described in figure 2. First, rain data collection of more than 20 mm per hour for ground rainfall data and satellite rainfall data. Furthermore, the two-rain data were mapped for rain data that corresponds to the time of occurrence. After that, the data is made a correlation and the results obtained are analyzed.

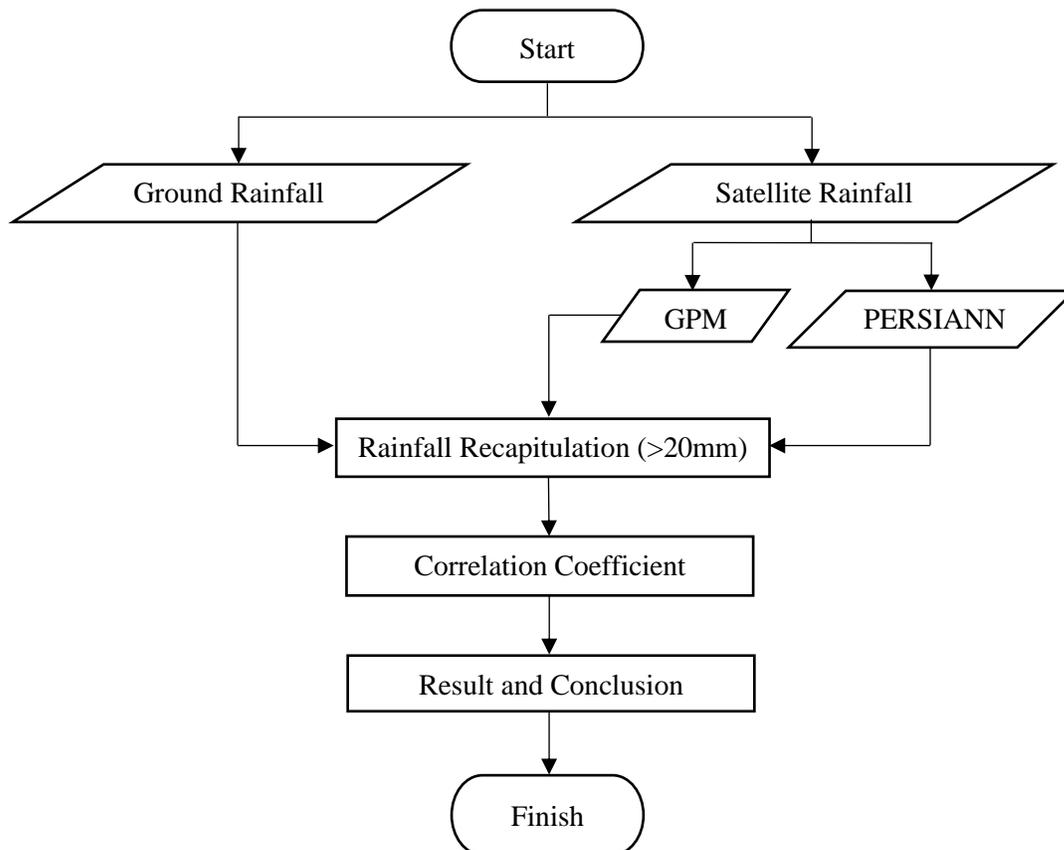


Figure 2. Research flow chart

RESULTS AND DISCUSSION

Recapitulation of Rain Data

The rainfall data used in this study came from Ngandong Station in 2012, 2017, and 2022, obtained from the Balai Teknik Sabo. Of all the rainfall data, only those with a total depth of more than 20mm in a day were used.

There were 47 rainfall events in 2012, 58 rainfall events in 2017, and 68 rainfall events in 2022 that had more than 20 mm of rainfall in one day which will be shown in figure 3.

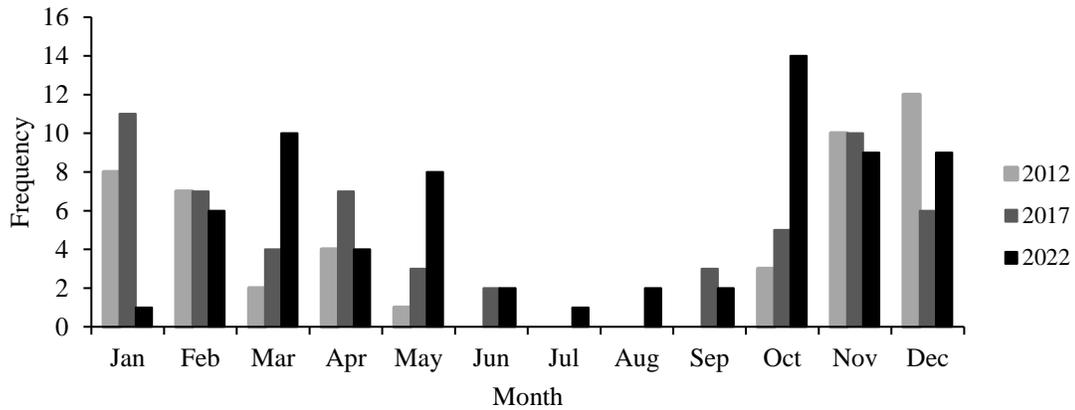


Figure 3. Rainfall intensity in 2012, 2017 and 2022

From figure 3, the results show that the incidence of rain above 20 mm per day, from 2012 to 2017, has increased. This shows that high-intensity rain events are increasingly frequent because of climate change. These results are consistent with research that has been conducted by several experts that climate change increases the rainfall intensity (Aristizabal et al., 2022; Chapman et al., 2021; Otto et al., 2023; and Tradowsky et al., 2023).

Duration of Rain Correlation

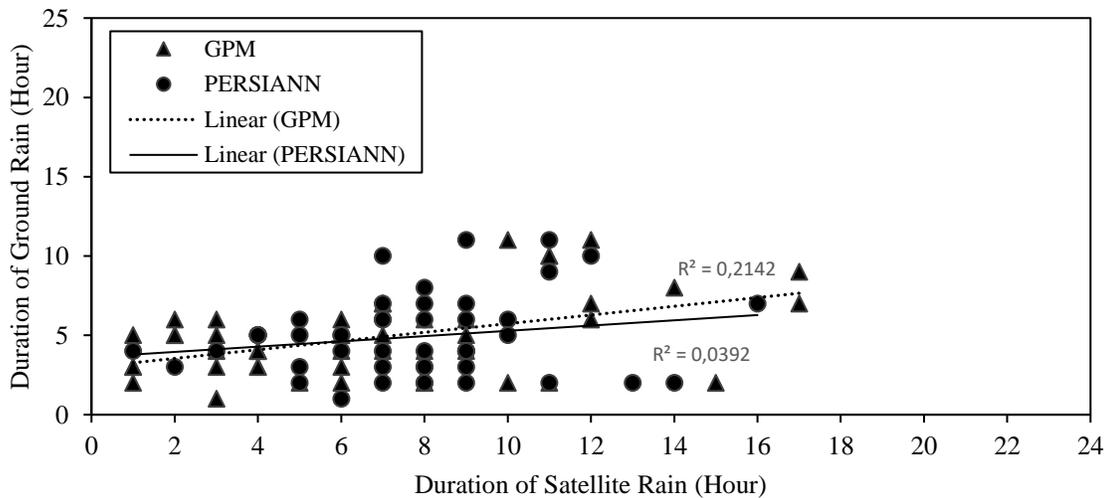


Figure 4. Duration of rain correlation in 2012

Based on figure 4, the correlation of rain duration between PERSIANN satellite rain data and ground rain data has an R-value of 0.20, indicating a weak relationship. In contrast, the correlation on the GPM satellite has an R-value of 0.46, indicating a medium relationship. From the results in 2012 shown on the GPM satellite, the correlation is better than the PERSIANN satellite when viewed from the parameters of the rain event duration.

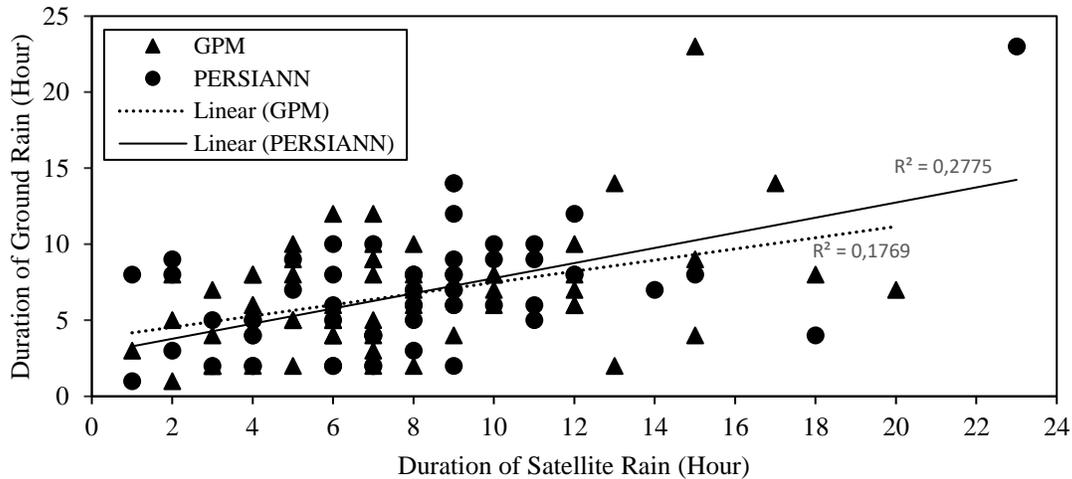


Figure 5. Duration of Rain Correlation in 2017

Based on figure 5, the correlation of rain duration between PERSIANN satellite rain data and ground rain data has an R-value of 0.53, higher than the GPM satellite with an R-value of 0.42. both are in the medium relationship category. Based on the results shown in Figure 5, data from the PERSIANN satellite is more accurate than data from the GPM satellite for parameters of rain duration in 2017.

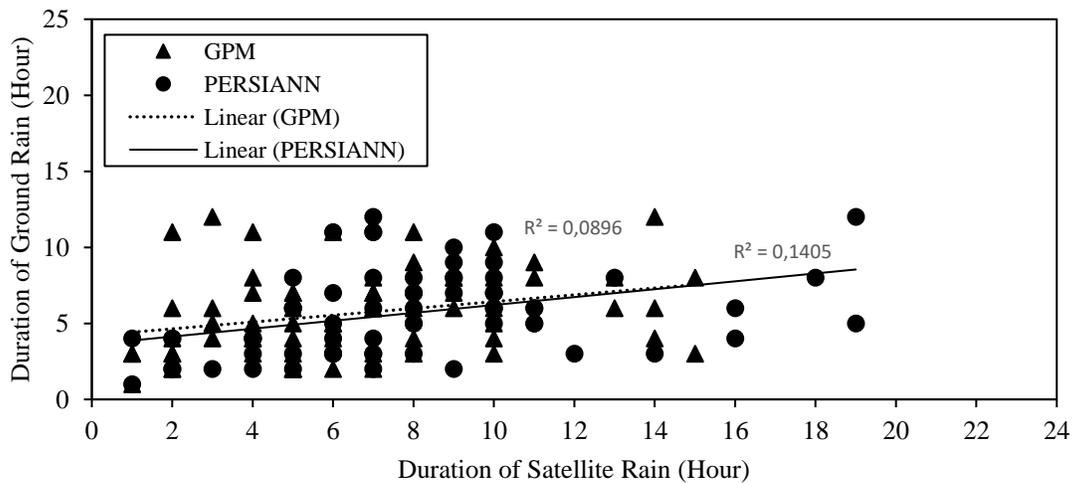


Figure 6. Duration of rain correlation in 2022

Based on figure 6, the correlation of rain duration between PERSIANN satellite rain data and rain data on the ground has an R-value of 0.37, higher than the GPM satellite with an R-value of 0.30. both are in the weak relationship category. Based on the results shown in Figure 6, data from the PERSIANN satellite is more accurate than data from the GPM satellite for parameters of rain duration in 2022. The results shown in Figure 4 to Figure 6 show that no pattern can be drawn regarding the best data trend between the PERSIANN satellite and GPM satellite data. However, on average, the correlation of data from the PERSIANN and GPM satellites is 0.36 and 0.39, respectively. Based on the correlation analysis of rain duration in the three years, the PERSIANN satellite has a weak correlation value, while the GPM satellite has a medium correlation value. The conclusion that can be obtained is that data from the GPM satellite is more accurate than data from the PERSIANN satellite. These results are in line with those that have been done by previous researchers (Adib et al., 2018; Islam et al., 2020; Kesarwani et al., 2023; and Le et al., 2020).

Total Depth Correlation

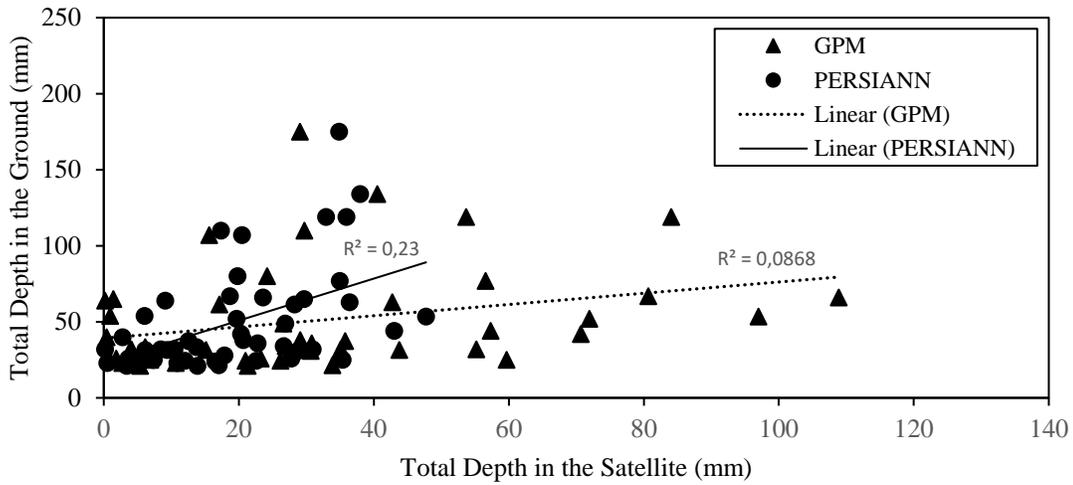


Figure 7. Total depth correlation in 2012

Based on figure 7, the total depth correlation between PERSIANN satellite rain data and rain data on the ground has an R-value of 0.48, indicating a moderate relationship. In contrast, the GPM satellite has an R-value of 0.29, indicating a weak relationship.

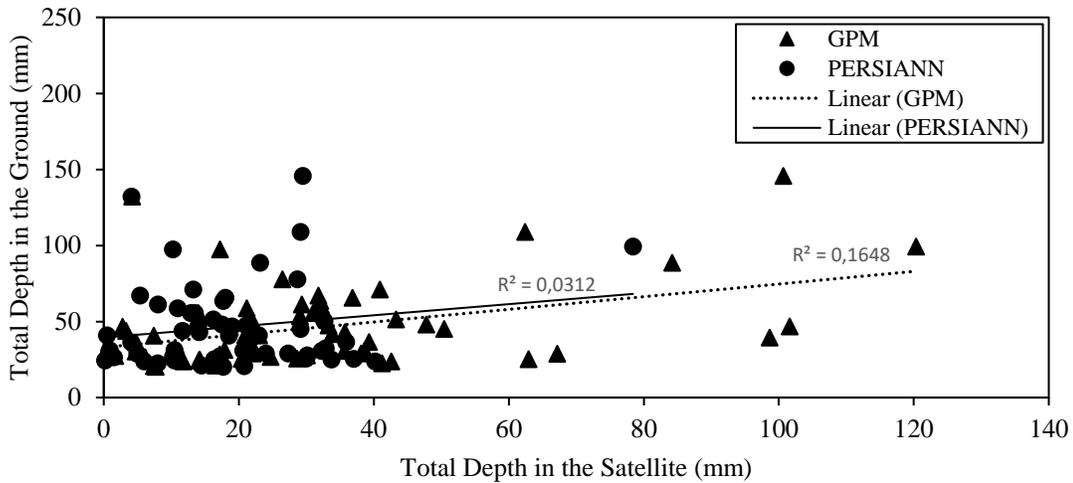


Figure 8. Total depth correlation in 2017

Based on figure 8, the total depth correlation between PERSIANN satellite rain data and rain data on the ground has an R-value of 0.18, indicating a weak relationship. In contrast, the correlation on the GPM satellite has an R-value of 0.41, indicating a moderate relationship.

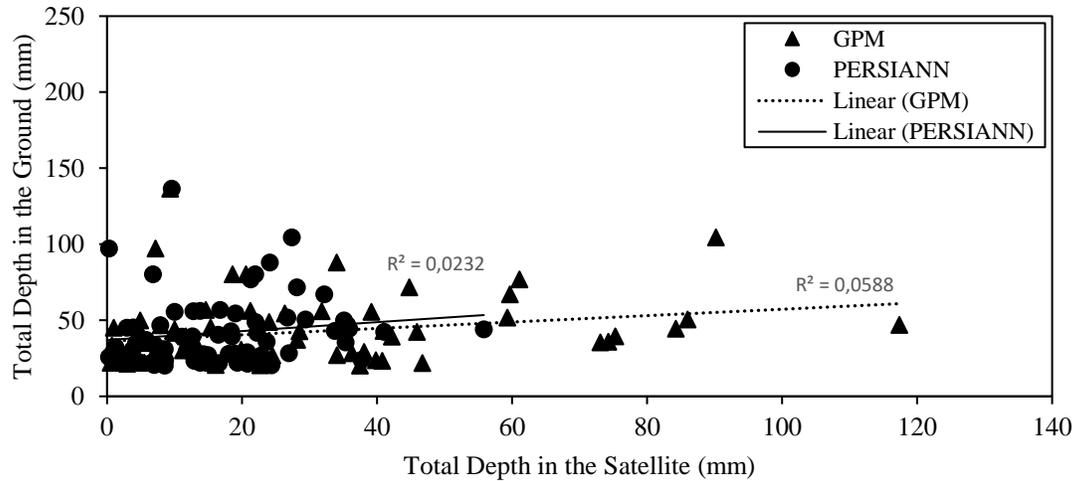


Figure 9. Total depth correlation in 2022

Based on figure 9, the total depth correlation between PERSIANN satellite rain data and rain data on the ground has an R-value of 0.15, indicating a very weak relationship. In contrast, the GPM satellite has an R-value of 0.24 indicating a weak relationship. Based on the correlation analysis of the total depth of rain in the three years, the PERSIANN satellite has a correlation value that tends to be very weak, while the GPM satellite has a correlation value that tends to be weak. However, in general, these results show that the correlation quality of GPM satellite data is better than data from the PERSIANN satellite. This result is in line with the results obtained by previous researchers, such as Islam et al. (2020); Kesarwani et al. (2023) and Le et al. (2020).

Intensity Maximum Correlation

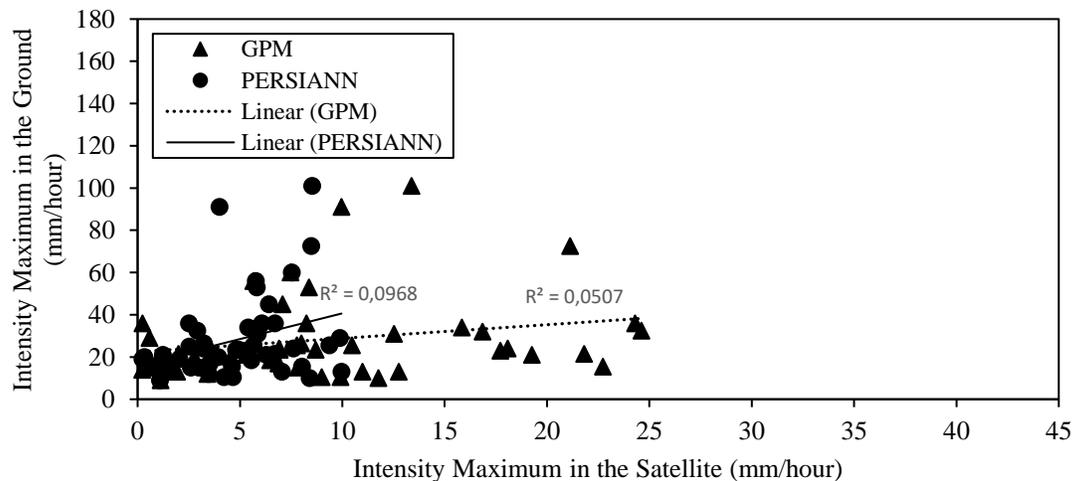


Figure 10. Intensity maximum correlation in 2012

Based on figure 10, the maximum intensity correlation between PERSIANN satellite rain data and rain data on the ground has an R-value of 0.31, higher than the GPM satellite with an R-value of 0.23. both fall into the weak relationship category.

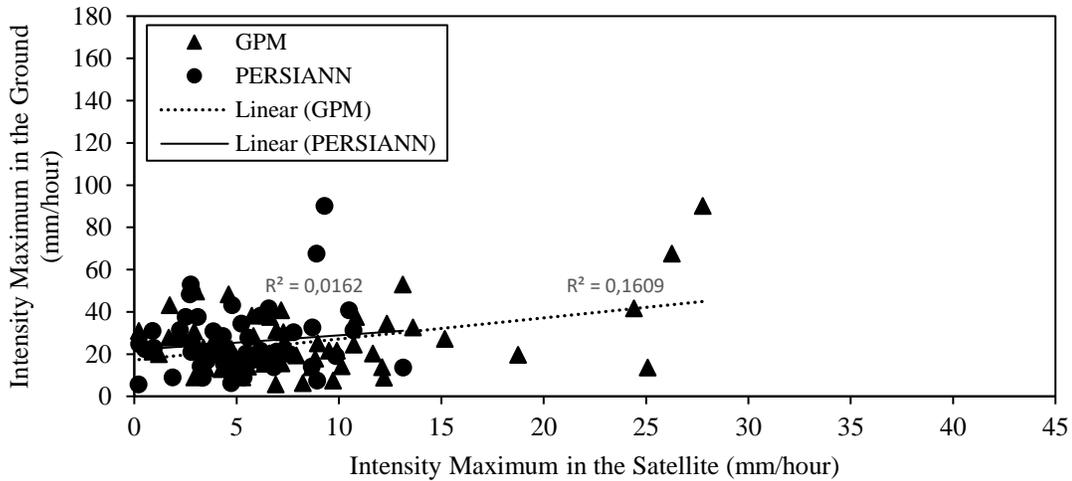


Figure 11. Intensity maximum correlation in 2017

Based on figure 11, the maximum intensity correlation between PERSIANN satellite rain data and rain data on the ground has an R-value of 0.13, indicating a very weak relationship. In contrast, the correlation on the GPM satellite has an R-value of 0.40, indicating a moderate relationship.

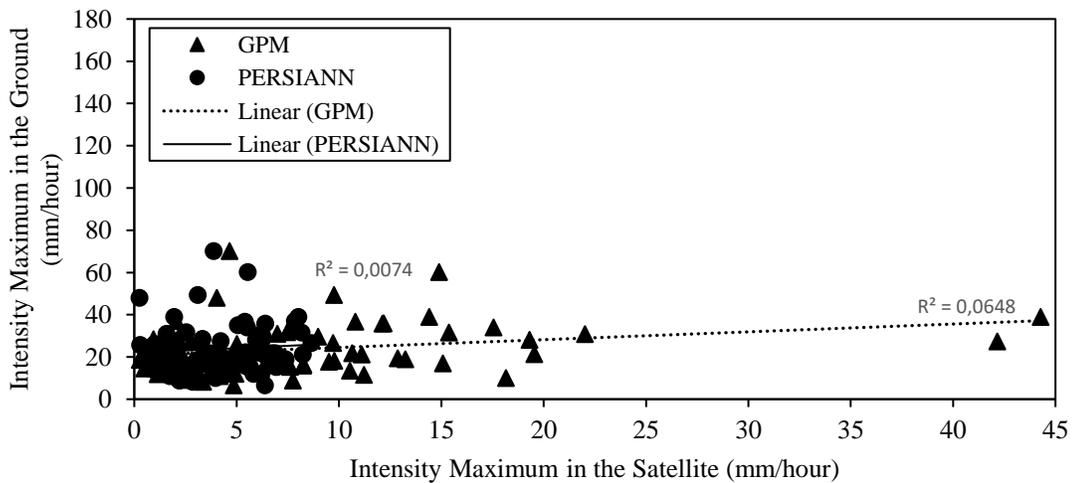


Figure 12. Intensity maximum correlation in 2022

Based on figure 12, the maximum intensity correlation between PERSIANN satellite rain data and rain data on the ground has an R-value of 0.09, indicating a very weak relationship. In contrast, the correlation on the GPM satellite has an R-value of 0.25, indicating a weak relationship. Based on the correlation analysis of intensity maximum in the three years, the PERSIANN satellite has a correlation value that tends to be very weak, while the GPM satellite has a correlation value that tends to be weak. Based on a review of the Intensity Maximum, the results obtained show that there is a tendency for data from GPM satellites to be better than data from PERSIANN satellites. The results obtained are in accordance with the results obtained from previous studies (Islam et al., 2020; Kesarwani et al., 2023; and Le et al., 2020).

CONCLUSION

The test results show that the PERSIANN satellite has a correlation value of rain duration that tends to be weak in 2012 with an R-value of 0.20 and in 2022 with an R-value of 0.37, while the GPM satellite has a correlation value that tends to be moderate in 2012 with an R-value of 0.46 and 2017 with an R-value of 0.42. In the total depth parameter, the PERSIANN satellite has a correlation value that tends to be very weak in 2017 with an R-value of 0.18 and in 2022 with an R-value of 0.15, while the GPM satellite has a correlation value that tends to be weak in 2012 with an R-value of 0.29 and 2022 with an R-value of 0.24. In the maximum intensity parameter, the PERSIANN satellite has a correlation value that tends to be very weak in 2017 with an R-value of 0.13 and in 2022 with an R-value of 0.09, while the GPM satellite has a correlation value that tends to be weak in 2012 with an R-value of 0.23 and 2022 with an R-value of 0.25. Overall, the GPM satellite has better accuracy than the PERSIANN satellite.

REFERENCES

- Adib Azka, M., Kadar Dzikiro, T., Kusuma Wardani, U., & Fadlan, A. (2018). Uji akurasi data model estimasi curah hujan satelit TRMM, GSMAP, dan GPM selama periode siklon tropis Cempaka dan Dahlia di wilayah Jawa validation of TRMM, GSMAP, and GPM Modeling Data Accuracy during tropical cyclone event in Java region. *Seminar Nasional Penginderaan Jauh, July 2018*, 983–991.
- Anonim. (2019). *Banjir*. Badan Penanggulangan Bencana Daerah (BPBD) Daerah Intimewa Yogyakarta (DIY). <http://bpbdd.jogjaprovo.go.id/banjir> (Accessed 24 February 2024)
- Anonim. (2022). *Modul 1 Analisis Curah Hujan*. Direktorat Jenderal Sumber Daya Air. Jakarta.
- Aristizabal, E., Garcia, E. F., Marin, R. J., Gomez, F., & JuanGuzman-Martinez. (2022). Rainfall-intensity effect on landslide hazard assessment due to climate change in north-western Colombian Andes. *Revista Facultad de Ingenieria*, 103. <https://doi.org/10.17533/udea.redin.20201215>
- Badri, R., & Yerizon. (2021). Development of the learning instruction based on problem based learning models oriented with mitigation of mount eruption and lava floods on the mathematical reasoning ability of class VIII students of SMP / MTs. *Journal of Physics: Conference Series*, 1742(1). <https://doi.org/10.1088/1742-6596/1742/1/012001>
- BMKG (METEOROLOGICAL, CLIMATOLOGICAL, AND GEOPHYSICAL AGENCY). (2024). Probabilistik Curah Hujan 20 mm. <https://www.bmkg.go.id/cuaca/probabilistik-curah-hujan.bmkg?mm=20&hour=24&gen=yeyoybj0smcr31fx1l> (accessed 24 April 2024)
- BSN. (2016). *SNI 2415:2016. Tata Cara Perhitungan Debit Banjir Rencana*. 1–4.
- Chapman, S., Birch, C. E., Galdos, M. V., Pope, E., Davie, J., Bradshaw, C., Eze, S., & Marsham, J. H. (2021). Assessing the impact of climate change on soil erosion in East Africa using a convection-permitting climate model. *Environmental Research Letters*, 16(8). <https://doi.org/10.1088/1748-9326/ac10e1>
- Fakhrudin, I., & Elmada, M. A. G. (2022). Local wisdom as a part of disaster communication: a study on the local storytelling in disaster mitigation. *ETNOSIA: Jurnal Etnografi Indonesia*. <https://doi.org/10.31947/etnosia.v7i2.22145>
- Ginting, J. M., Sujono, J., & Jayadi, R. (2019). Analisis Hubungan Data Hujan Satelit dengan Hujan Terukur ARR Kalibawang. *Prociding Konferensi Nasional Pascasarjana Teknik Sipil (KNPTS) X 2019, November*, 89–102.
- Jarwanti, D. P., Suhartanto, E., & Fidari, J. S. (2021). Validasi Data Curah Hujan Satelit TRMM (Tropical Rainfall Measuring Mission) dengan Data Pos Penakar Hujan di DAS Grindulu, Kabupaten Pacitan, Jawa Timur. *Jurnal Teknologi Dan Rekayasa Sumber Daya Air*, 1(2), 772–785. <https://doi.org/10.21776/ub.jtresda.2021.001.02.36>

- Maharani, S. A., Romadhona*, M., & Masnuna, M. (2023). Perancangan picture storybook berbasis augmented reality (AR) tentang edukasi sigap bencana alam untuk anak usia 9-12 tahun. *SYNAKARYA Visual Communication Design Student Journal*, 4(2). <https://doi.org/10.33005/synakarya.v4i2.89>
- Mamenun, Pawitan, H., & Sophaheluwakan, A. (2014). Validasi dan koreksi data satelit TRMM pada tiga pola hujan di Indonesia (Validation and correction of TRMM satellite data on three rainfall patterns in Indonesia). *Jurnal Meteorologi Dan Geofisika*, 15(1), 13–23. <http://202.90.199.54/jmg/index.php/jmg/article/view/169/155>
- Munir, M. D. (2019). Bangunan sabodam, fungsi dan potensinya sebagai bagian dari geowisata gunung api Merapi. *Jurnal Lingkungan Dan Bencana Geologi*, 10(2), 15–26. <https://doi.org/10.34126/jlbg.v10i2.202>
- Nguyen, P., Ombadi, M., Sorooshian, S., Hsu, K., AghaKouchak, A., Braithwaite, D., Ashouri, H., & Rose Thorstensen, A. (2018). The PERSIANN family of global satellite precipitation data: A review and evaluation of products. *Hydrology and Earth System Sciences*, 22(11), 5801–5816. <https://doi.org/10.5194/hess-22-5801-2018>
- Otto, F. E. L., Zachariah, M., Saeed, F., Siddiqi, A., Kamil, S., Mushtaq, H., Arulalan, T., AchutaRao, K., Chaithra, S. T., Barnes, C., Philip, S., Kew, S., Vautard, R., Koren, G., Pinto, I., Wolski, P., Vahlberg, M., Singh, R., Arrighi, J., ... Clarke, B. (2023). Climate change increased extreme monsoon rainfall, flooding highly vulnerable communities in Pakistan. *Environmental Research: Climate*, 2(2). <https://doi.org/10.1088/2752-5295/acbfd5>
- Sorooshian, S. (2020). *UCI CHRS Data Portal - PERSIANN*. Accesed (12 Febuary 2024), <http://chrsdata.eng.uci.edu>
- Sugiyono. (2013). *Statika untuk Penelitian*. Bandung: Alfabeta.
- Syaifulallah, M. D. (2014). Validasi data TRMM terhadap data curah hujan aktual di tiga DAS di Indonesia. *Jurnal Meteorologi Dan Geofisika*, 15(2), 109–118. <https://doi.org/10.31172/jmg.v15i2.180>
- Tan, M. L., & Duan, Z. (2017). Assessment of GPM and TRMM precipitation products over Singapore. *Remote Sensing*, 9(7), 1–16. <https://doi.org/10.3390/rs9070720>
- Tradowsky, J. S., Philip, S. Y., Kreienkamp, F., Kew, S. F., Lorenz, P., Arrighi, J., Bettmann, T., Caluwaerts, S., Chan, S. C., De Cruz, L., de Vries, H., Demuth, N., Ferrone, A., Fischer, E. M., Fowler, H. J., Goergen, K., Heinrich, D., Henrichs, Y., Kaspar, F., ... Wanders, N. (2023). Attribution of the heavy rainfall events leading to severe flooding in Western Europe during July 2021. *Climatic Change*, 176(7). <https://doi.org/10.1007/s10584-023-03502-7>
- Vernimmen, R. R. E., Hooijer, A., Mamenun, Aldrian, E., & Van Dijk, A. I. J. M. (2012). Evaluation and bias correction of satellite rainfall data for drought monitoring in Indonesia. *Hydrology and Earth System Sciences*, 16(1), 133–146. <https://doi.org/10.5194/hess-16-133-2012>
- Zhang, C., Chen, X., Shao, H., Chen, S., Liu, T., Chen, C., Ding, Q., & Du, H. (2018). Evaluation and intercomparison of high-resolution satellite precipitation estimates-GPM, TRMM, and CMORPH in the Tianshan Mountain Area. *Remote Sensing*, 10(10). <https://doi.org/10.3390/rs10101543>