

An assessment of IMERG rainfall products over Bali at multiple time scale

I Wayan Andi Yuda^{1,4*}, *Rakhmat Prasetya*¹, *Abd. Rahman As-syakur*², *Takahiro Osawa*³, and *Masahiko Nagai*⁴

¹The Agency for Meteorology, Climatology, and Geophysics of the Republic of Indonesia (BMKG), Jalan Leli No. 9 BB Agung, Negara, Bali 82212, Indonesia

²Environmental Research Center (PPLH), Udayana University, PB Sudirman Street, Denpasar, Bali, Indonesia

³Center for Remote Sensing and Ocean Sciences (CRoSOS) Udayana University, Indonesia

⁴Graduate School of Sciences and Technology for Innovation, Yamaguchi University, Japan

Abstract. Evaluation of first five years of the Global Precipitation Measurement - Integrated Multi-satellite Retrievals for GPM (IMERG) final precipitation product was performed over Bali – Indonesia using surface observation data which derived from The Agency for Meteorology, Climatology, and Geophysics of the Republic of Indonesia (BMKG) as a reference. This study evaluated IMERG's performance in describing the temporal characteristics of rainfall variation over various time periods (including daily, monthly, and seasonal). The analysis concentrated on the period of April 2014 to April 2019. The results of statistical measurements consisted Probability of Detection (POD), linear correction coefficient (r), Mean Bias Error (MBE), and Root Mean Square Error (RMSE). In general, the results showed that IMERG rainfall estimation value was lower than rain gauges data. The statistical assesment indicated IMERG data was highly accurate on monthly to seasonal timescales. However, a moderate correlation was shown between the daily data comparison from IMERG to ground references. IMERG Performed better in wet season period (November -April) than in dry season period (May – Oktober). The probability of detection rain events on daily time scale was good. Overall, data from IMERG has the potential to be useful as a complement to rain gauge data in areas where rainfall observations are not available in the field.

1 Introduction

Rainfall is the most important meteorological elements in tropical area such as Indonesia's maritime continent and it observed by observers at weather observation stations every day. This is done considering that the data and information on rainfall are much needed in various kinds of human activities such as agriculture, plantations, fisheries, transportation (land, sea, air), and others [1]. Many locations in Indonesian region are very vulnerable to rainfall conditions. Rainfall Conditions above normal often cause floods and landslides. The opposite

* Corresponding author: andiyudaiwayan@gmail.com

condition, below normal rainfall causes dryness. Both conditions result in a decrease in food production. For areas that are vulnerable to these conditions, accurate observation and temporal knowledge about rainfall is important to understanding multi-scale interactions between weather, climate and environmental systems. Data and knowledge about rainfall can be utilized to escalate our capability to maintain water fount and to prepare an early prevention of extreme weather impacts such as landslides and floods [2]. However, there is a big challenge in rainfall observation due to its temporal variation and many forms of a geographical landscape that make measuring rainfall is difficult in the field. Another factor that makes it difficult to obtain rainfall data is the presence of uninhabited and remote areas making it difficult to apply a rain gauge network on the surface [3].

For now, the possibility of obtaining rainfall data needed in various scientific applications can be obtained from meteorological satellites. Rainfall products from satellites have become a very promising alternative as the most important source of rainfall data in hydrological, climatological and meteorological studies over the past few decades. Applications of satellite rainfall products are increasing rapidly because of the ability to monitor large areas, continuous measurement, the fact that they are free of charge and the availability of some products in near real time via the internet. Most importantly, satellite rainfall products could overcome the spatial limitations of point-based ground observations in mountain and ocean areas that are less accessible [4].

One of the latest generation of weather observers, the Global Precipitation Measurement (GPM) is an international mission to set continuous rain and snow observations. The GPM Core Observatory Satellite was launched by NASA and the Japan Space Exploration Agency (JAXA) on the of February, 2014, bringing sophisticated instruments that became the new standard for measuring rainfall from space. The core observatory is supported by other collaborative satellites to complement the data density partially and temporally in the frame of the product algorithm called the Integrated Multi-Satellite Retrieval for GPM (IMERG). This product presents better spatial (0.1°) and temporal (30 minutes) resolution than TRMM and Multi-Satellite Precipitation Analysis (TMPA) products. In addition, IMERG coverage ($60^\circ\text{N} - 60^\circ\text{S}$) is also greater than that of TMPA products ($50^\circ\text{N} - 50^\circ\text{S}$). The method used to produce IMERG rainfall is provided by Huffman et al. (2017) [5], which is summarized briefly. The IMERG Core observatory is equipped with sensor package includes a dual-frequency rainfall radar and multi-scanning multichannel microwaves imager. IMERG is currently available from April 2014 to the present.

IMERG's performance in describing rainfall features in various regions (such as Singapore [4], United States of America [6], Netherlands [7], Tibetan Plateau [8], Nile River Valley [9], Indian Ocean [10], Taiwan [11], Brazil [12] has been examined. This shows that there is interest in the use of the relatively new satellite technology for monitoring rainfall in these places. However, IMERG's ability to describe variations in rainfall in Bali, Indonesia (ie, study area in this research) has not been evaluated in detail by existing studies. Therefore, this subject is examined here.

Bali Province, located at south part of Indonesia, is a dense area and is a world tourism destination. Bali island is a mountainous area and hills covering most of the center region and vast lowlands at the south and it surrounded by ocean. This complex topography causes rainfall fluctuations [14]. Besides local influence, Bali has a monsoonal type of rainfall with one peak in the rainy season triggered by North-west monsoon from December to February) and one peak in the dry season which corresponds to the South-east monsoon from June to August) [13, 14]. Large-scale climate event such as El Nino Southern Oscillation (ENSO) also influence variations in rainfall in this region [15, 16].

Rainfall in Bali is well documented by The Agency for Meteorology, Climatology, and Geophysics of the Republic of Indonesia (BMKG) and the distribution of the observation area includes the coast to the highlands. The main goal of this study is to assess the performance

of IMERG in describing variability in rainfall in Bali compared to the results of surface rainfall observations from the period April 2014 to April 2019 (5 years). This analysis focuses on climatological timescale in the form of seasonal, monthly, and daily. The results of this study are expected to increase our understanding of the ability of IMERG products to contribute to rainfall observation especially in Bali. If at any time there are disruptions on land related to human error, equipment damage, and / or natural disasters, then with the results of the assessment, IMERG can be considered as alternative data.

2 Study area

Geographically, Bali Province is located at south part of Indonesia Archipelago. This island with total territory of 5636.66 km² is the area of interest for this research. Bali island (Figure 1) is a mountainous small island surrounded by oceans. Bali also has extensive lowland areas in the south. This varied Balinese island relief became a distinct advantage in this study supported by a good distribution of rainfall observation. Local rainfall variation and Bali's climate pattern which is strongly influenced by the west and east monsoon winds becomes the biggest challenge for IMERG to capture it.

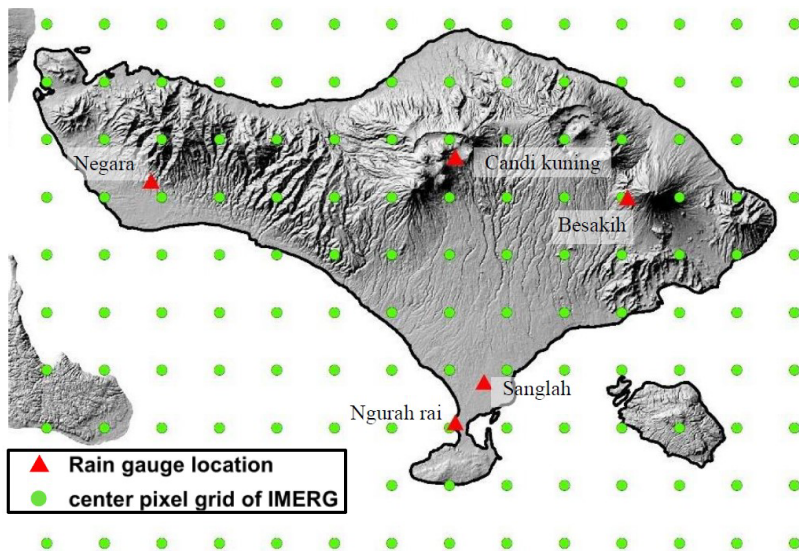


Fig. 1. Study Area and rain gauge locations

3 Data and Methods

3.1 Data

Daily in situ rainfall data derived from rain gauge measurement location (Fig. 1) from April 2014 – April 2019 provided by BMKG Climatology Station of Jembrana, Bali were used in this research. Daily rainfall from five surface observation location on Bali island were utilized as an actual data to validate satellite estimates. Months with unrecorded data were removed from calculations. Satellite rainfall products compared to reference measurement data sets have a daily and monthly time scales derived from IMERG level 3 data with the same period to the reference gauges. The data was downloaded from NASA Earth Data Homepage. The accessible website is <https://disc.gsfc.nasa.gov/datasets?page=1&keywords=imerg> .

The Integrated Multi-satellite Retrievals for GPM (IMERG) is a US integrated algorithm that provides multi-satellite rainfall products for the GPM team [5]. Precipitation is estimated from various satellites in the GPM constellation carrying associated passive satellite microwave sensors (PMW) computed using the 2017 version of Goddard Profiling Algorithm [17] then gridded, inter-calibrated to the GPM Combined Instrument product, and combined into 30 minutes and $0.1^{\circ} \times 0.1^{\circ}$ grids space. The system is run several times for each observation time to produce a quick precipitation estimation (IMERG “early run”) with data availability about 4 hours and successively providing better estimates as more data arrive (IMERG “late run”) with data latency about 14 hours. The Final step uses monthly gauge data to create research-level product (IMERG “final run”). This post-real-time form of satellite-gauge product with availability about 3.5 months after observation and gauge analysis has been carried out.

3.2 Validation

This study examined five years data period (April 2014 - April 2019). First, adjustments were made to the recording period of rain data between rain gauge and the IMERG because the rainfall data from the rain gauge was noted in the measurement date. Meanwhile, satellite data was noted on the date of rain events. Hence, an adjustment was made in the form of one day forward for daily rain gauge data. Furthermore, validation was done for daily satellite estimates products, then the same steps were applied to monthly rainfall. Accumulation of rainfall for monthly observations was calculated by combining daily values for one month. In this study, daily and monthly satellite data were obtained from level 3 IMERG data or "Final" products. Validation of monthly satellite products were also done in two different seasons in Bali island (rainy season and dry season) separately. The seasons separation reflects the monsoon cycle influence. The wet season usually lasts from November to April while the remaining time period from May to October is the dry season period [14, 18]. For all types of data per unit of time, there are two types of analysis carried out, namely point analysis and spatial average analysis.

Point analysis is done by comparing the measurement data and satellite data at the point of observation. Satellite data were extracted by its center point of pixels and compared to nearest rain gauge station. Data of extracted satellite and ground references were then juxtaposed to see their scatterplot distributions, statistical correspondence and error values. Then the procedure was repeated by combining the five reference data points so that they became their spatial averages. Spatial average analysis was an points average comparison for five years long-term daily time-series data as same as the long-term monthly time-series data. Monthly spatially averaged general pattern was also analyzed in this study.

The relationship of IMERG products with rain gauge data is done using several statistical scores. The measurement of reliability of satellite estimates with ground references values utilizing the linear correlation coefficient (r), the root mean square error (RMSE), and the mean bias error (MBE). We also utilized other statistical tools to validate the fraction of observed events that are correctly diagnosed by satellite estimates. Categorical verification is carried out in this study using probability detection (POD). The equation is defined as follows [19, 20]:

$$r_{x_{sat}x_{ob}} = \frac{COV_{x_{sat}x_{ob}}}{S_{x_{sat}}S_{x_{ob}}} = \frac{\sum_{i=1}^n (x_{sat_i} - \bar{x}_{sat})(x_{ob_i} - \bar{x}_{ob})}{(\sum_{i=1}^n (x_{sat_i} - \bar{x}_{sat})^2 \sum_{i=1}^n (x_{ob_i} - \bar{x}_{ob})^2)^{1/2}} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_{sat_i} - x_{ob_i})^2}{n}} \quad (2)$$

$$MBE = \frac{\sum_{i=1}^n x_{sat_i} - x_{obs_i}}{n} \quad (3)$$

Where x_{sat_i} represented the IMERG rainfall values, x_{ob_i} reflected the surface observation rainfall, \bar{x}_{sat} and \bar{x}_{obs} were their respective mean and n deputize the number of data. The correlation coefficient (r) shows the degree of linear relationship between the rainfall estimated and observed distributions. The MBE providing a measure of the overestimation or underestimation of the gauge data by the satellite estimates by showing the systematic component of the error. RMSE is sensitive to extreme values because it involves the square of the difference from the observed value [21]. All errors are expressed as a percentage to facilitate comparison of daily, monthly and seasonal rainfall variations in a location. Therefore, the reliability of IMERG products can be seen in the different rainfall characters in each region.

The quantitative method was the other validation statistic applied in this study which based on the contingency tables shown in Table 1. The threshold of rainfall used to state rain/non rain events are 0.5 mm/d related to BMKG operational data. Categorical verification statistics measure the correspondence between predicted and true events that occur after being observed. This method is usually used in dichotomous (yes / no) forecasts. Description of these statistics has been given in many references [4, 22]. Most were based on a 2x2 contingency table of yes/no events, such as rain / no rain, shown in Table 1. "Hit" describes the forecast that rainfall will occur, and actually occur. "Miss" shows the forecast rainfall will not occur, but in reality it happens. "False alarm" represents the forecast that rainfall will occur, but it does not occur in the field. "Correct negative" indicate condition when rainfall is predicted not to occur, and true do not occur in reality. Using the results shown in Table 1, the following equations were used to calculate statistical parameters namely Probability of Detection (POD) and False Alarm Ratio (FAR).

$$POD = \frac{hits}{hits + misses} \quad (4)$$

$$FAR = \frac{false\ alarm}{hits + false\ alarm} \quad (5)$$

Where, POD explains the ability of the IMERG estimates are in detecting the rainfall occurrence. FAR shows ratio score when IMERG predicted rainfall but actually no rain in reality.

Table 1. Contingency table of yes or no rain event

		Rain Gauges	
		Yes	No
IMERG	Yes	<i>hits</i>	<i>false alarm</i>
	No	<i>misses</i>	<i>correct negatives</i>

4 Results

The first results of this study were obtained from the comparison of daily rain gauge data with IMERG data on a daily time scale. In general, rainfall from IMERG is lower than rainfall from rainfall measurement data. The rainfall average showed by the surface observation was 6.18 mm / d, while the IMERG average estimation was 5.83 mm / d. Then, the scatterplot of daily rain gauge data and IMERG data was shown by Figure 2. The figure showed the IMERG data have quite good relationship with the observed rainfall ($r = 0.49$), and the MBE score was -0.35 mm/d (-5.71%). Futhemore, the values of the RMSE was 14.13 mm/d (228.53%). POD of IMERG was close to 81% and FAR was generally about 43%.

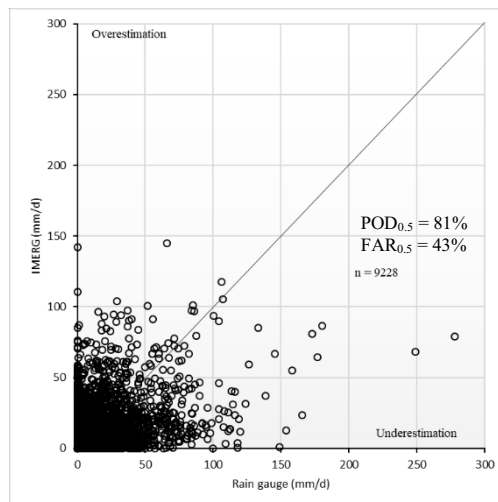


Fig 2. Scatter plots of daily IMERG rainfall and daily rain gauge data from 2014 to 2019

A comparison of five-year period spatial average of daily rainfall measured by IMERG and gauges was shown by figure 3. Both data has a very similar fluctuation of the long-term period which indicated by the graph in the figure, moreover the spatially average data from IMERG is quite capable of detecting high rainfall value by the rain gauge. Correlation coefficient value from this data was very good ($r = 0.72$), MBE was -0.37 mm/d (-6.0%), and RMSE was 7.34 mm/d (117.79%). When comparisons were made individually for each seasons, the result of correlation value was higher in the rainy season than in the dry season. The correlation coefficient on wet period was 0.70, while in the dry period, the score was only 0.59. MBE score for wet period and dry period were -0.61 mm/d (-6.13%) and -0.14 mm/d (-5.50%) successively. The values of RMSE were 8.72 mm/d (88.31%) and 5.60 mm/d (221.55%) respectively for wet season and dry season calculation period.

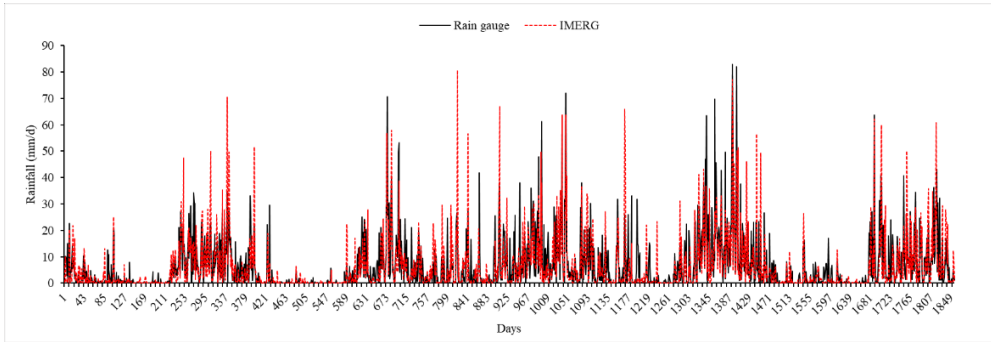


Fig 3. Spatial average daily rainfall of IMERG and surface observation stations 2014 – 2019.

The second result of this study was a statistical assesment between time series of monthly rainfall data and monthly IMERG data shown in Figure 4. The graph showed that pattern of monthly rainfall was significantly similar between two datasets. The data showed very good compatibility with ground references, resulting in high correlation for this product. The IMERG correlation was 0.96. To see its main climate patterns, monthly time series data were then averaged into annual patterns for 2014 to 2019 shown in Figure 5. The graph showed the same pattern as forming a U shape. This pattern was identical to the monsoon effect characteristic where there is a one peak of the rainy season and one peak of the dry season. The relationship between the annual pattern of rain gauges and IMERG was very strong ($r = 0.98$). This result indicated that IMERG data can be used as a source of data for remote areas that are difficult to reach or locations that do not yet have rain observation network infrastructure.

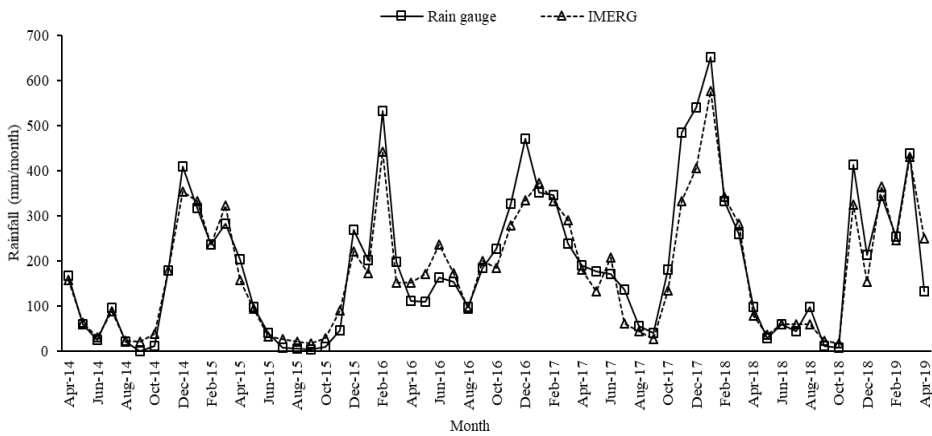


Fig. 4. Time series of monthly rainfall averaged from five obervation points for IMERG and Rain gauge data for the period 2014 - 2019

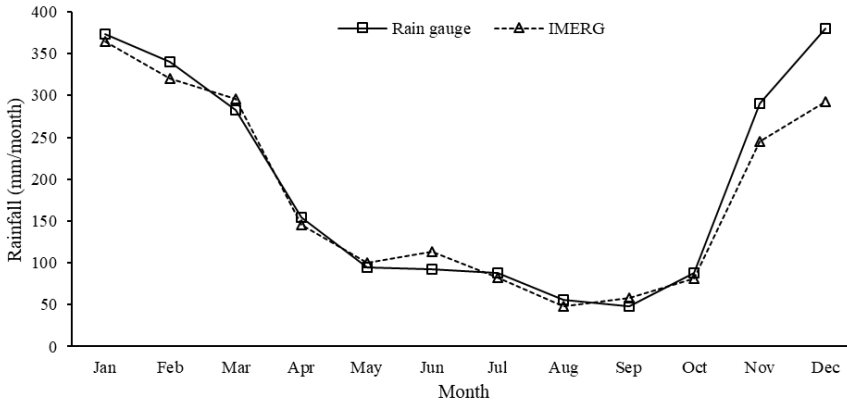


Fig. 5. Annual pattern of monthly average rainfall from five observation points of IMERG and Rain Gauge data for five years.

Figure 6 showed the comparison of monthly error statistical values for IMERG products. The correlation between IMERG and observational data was high ($r > 0.90$) except in April, May, July. The values of correlation coefficient were 0.75, 0.77, 0.77, respectively. More stable correlation coefficient was visible during wet season compared to the dry season. As a matter of fact, the levels of correlation were mostly higher in the wet season than dry season (see figure 6(a)). The values of MBE were varies from -23% to 24% showed the unstable pattern. In November (beginning of the rainy season), the value was negative. Then the value moved away from zero in December and stabilized near zero until April. Unsteady MBE values were spotted during the dry period, positive values occurred in May, June and September. However, negative values of MBE during dry season counted on July, August, and October (see figure 6(b)). The RMSE value was quite narrow, in the range of 10% to 42%. As well as, the RMSE value was relatively stable. The RMSE values during wet seasons were lower than dry seasons period (see figure 6(c)). From these results it can be seen that the random error of the estimated product was less than 50% of the amount of rainfall measured, thus it can be said that the estimate is quite accurate and reliable.

Scatter plots of all monthly data used from 5 observation points for comparison of IMERG and rain gauge were shown in Figure 7. The number of data used in the plot was 302 data. Generally, the IMERG monthly rainfall estimation was lower than actual observed rainfall. Statistically, monthly averaged rainfall from the rain gauges and IMERG were 188.34 mm/month and 178.97 mm/month respectively. Monthly data that proves good agreement between satellite estimations and ground references, results in a high correlation coefficient. The monthly IMERG correlation with reference data (shown in figure 8 (a)) was generally high ($r = 0.87$), but many data have underestimated distributions. This underestimated condition was proven by negative MBE errors. The MBE error from IMERG was -5% (see figure 8 (b)). Another statistical value from IMERG compared to rain gauge data is the RMSE with 48% percentage (see figure 8 (c)).

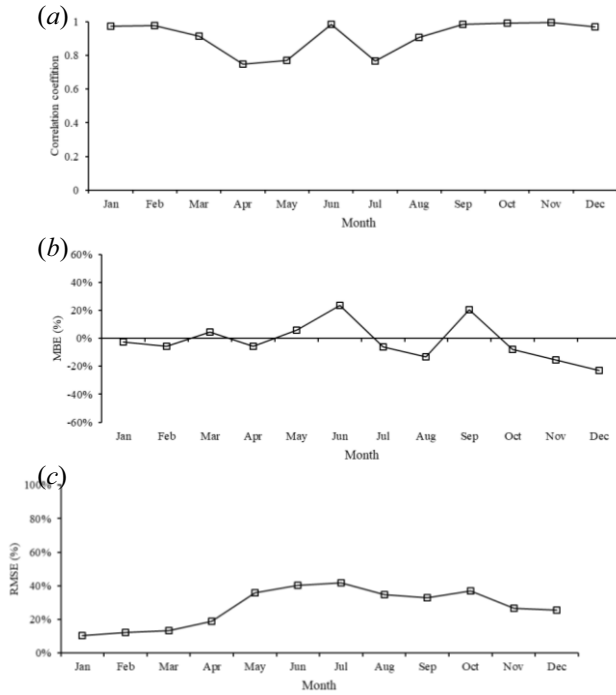


Fig 6. Monthly error statistics values from IMERG 2014 - 2019 : (a) coefficient of correlation, (b) MBE, (c) RMSE

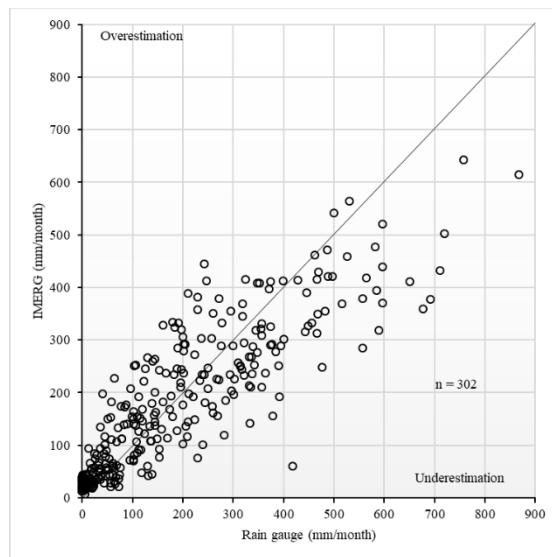


Fig 7. Scatter plots of IMERG product compared to the gauge data for monthly rainfall (2014–2019)

The third result of this study was the IMERG ability test score in describing seasonal rainfall in Bali. Figure 8 presented a distribution of monthly rainfall from IMERG product versus gauge data for two distinct seasons of the year. There were different characteristic of IMERG data compared to rain gauge data in each seasons. The distribution of data was mostly underestimate during wet season whereas overestimate during dry seasons (see figure 8). The agreement showed by coefficient of correlation during the rainy period were higher

than in the droughty period (see figure 9(a)). The correlation values on wet season and dry season were 0.80 and 0.71, respectively. The MBE in wet season and dry season were -7% and 4% (see figure 9 (b)). The value of RMSE from IMERG data in wet season (37%) was lower than dry season (64%). The RMSE values of IMERG data during two different seasons indicated that this data is more reliable on wet season than dry season (see figure 9 (c)).

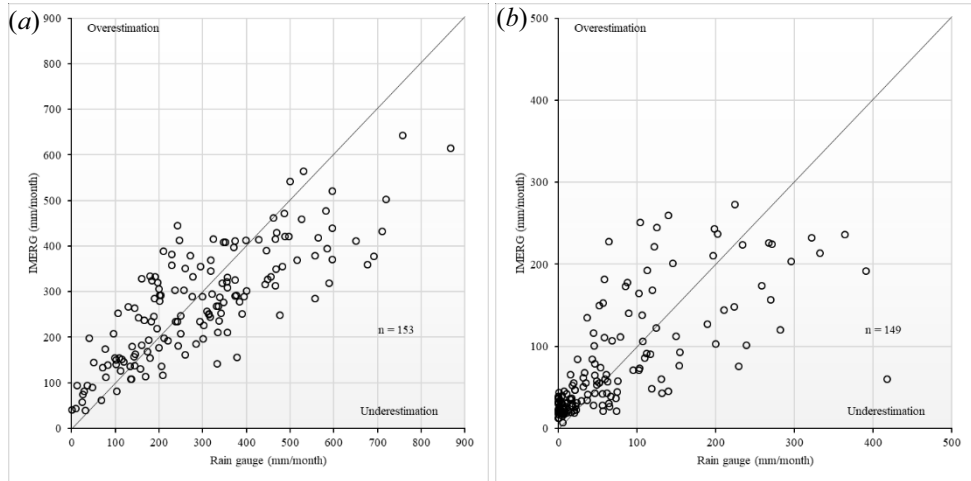


Fig 8. Scatter plots of monthly IMERG estimations versus ground references in wet and dry seasons : (a) wet season ; (b) dry season

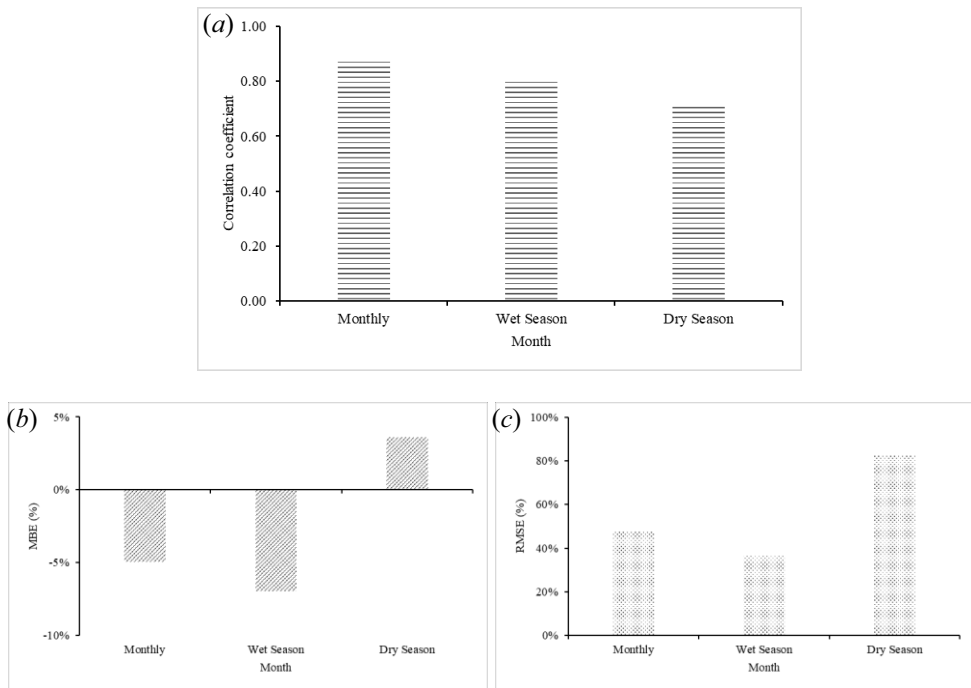


Fig 9. Wet and dry seasons assessment scores between the IMERG and rain references (2014–2019) : (a) Correlation coefficient, (b) MBE and (c) RMSE.

5 Summaries

An overview of an assessment of IMERG rainfall products over Bali at multiple time scale is shown here. A fairly comprehensive comparison of observations data and IMERG estimation in daily, monthly, and seasonal rainfall over the five years at the study location showed that the satellite gives a lower value than ground references data. Statistical analysis showed a moderate relationship between the daily data from IMERG compared to the measurement data. Furthermore, validation analysis showed that the IMERG rainfall product has a excellent conformity to measurement data in Province of Bali on monthly to seasonal time scale.

Daily rainfall data from IMERG have a good performance to detect the occurrence of rainfall but not very accurate when it is not raining. The correlation value of IMERG rainfall product in daily time scale was better by averaging all points rather than in the results of the individual point analysis. Spatial average analysis would be greater for analysing climate product [20]. The monthly long-term and annual pattern from spatially averaged rainfall provided by IMERG indicated near perfect capability of this data to figure out the main patterns of monthly rainfall which significantly similar to ground data.

In seasonal time scale, satellite performed better estimation of rainfall for the wet season than the dry season period. The distribution of data was mostly underestimate during wet season whereas overestimate during dry seasons. Higher error percentage also found during dry season then during the wet season. This pattern has been also found in Singapore [4], and similar result in Brazil [12]. IMERG encountered difficulties in estimating rainfall in the dry period, when rainfall events characteristics were generally less intense, lower volume, and more sparsely distributed across the territory. In general, IMERG data has good prospects as a source of rainfall data for remote areas that are difficult to reach or locations that do not yet have rain observation network infrastructure. But, error correction is still needed to improve the accuracy of this data especially on a daily scale.

Acknowledgments

This research was supported by Indonesian Agency for Meteorology Climatology and Geophysics (BMKG), Centre for Remote Sensing and Ocean Sciences (CREOS) Udayana University, and Yamaguchi University. The authors are thankful to GPM-IMERG data providers for making it available to carry on this study (<https://pmm.nasa.gov/gpm>).

References

1. S. Prawirowardoyo, *Meteorologi* (Institut Teknologi Bandung, Bandung, 1996)
2. A. Y. Hou, G. Skofronick-Jackson, C. D. Kummerow, J. M. Shepherd, *Global Precipitation Measurement*, In : S. Michaelides s. (eds) *Advances in Measurement, Estimation, and Prediction* (Springer, Berlin, 2008)
3. H. Feidas, *Theoretical and Applied Climatology*, **99**, 193–216 (2010)
4. M.L Tan, Z. Duan, *Remote Sens*, **9**, 720 (2017)
5. G.J Huffman, D.T. Bolvin, E.J. Nelkin, (2017). URL: https://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf
6. O. Sungmin, P. Kirstetter. *Q. J. R. Meteorol. Soc*, **144**, 270-281 (2018)
7. M.F.R. Gaona, A. Overeem, H. Leijnse, R. Uijlenhoet, *J. Hydrometeorol*, **17**, 2799-2814 (2016)

8. R. Xu, F. Tian, L. Yang, H. Hu, H. Lu, A. Hou, J. Geophys. Res. Atmos, **122**, 910-924 (2017)
9. D. Sahlu, E.I. Nikolopoulos, S.A. Moges, E.N. Anagnostou, D. Hailu, J. Hydrometeorol, **17**, 2875-2882 (2016)
10. S. Prakash, M.R. Ramesh Kumar, S. Mathew, R. Venkatesan, Theor. Appl. Climatol, **134**, 467-475 (2017)
11. W.R. Huang, Y.H. Chang, P.Y. Liu, Atmos. Res, **214**, 239-249 (2018)
12. L. Salles, F. Satge, H. Roig, T. Almada, D. Olivetti, W. Ferreira, Water, **11**, 668 (2019)
13. A.R. As-Syakur, Jurnal bumi Lestari, **7(2)**, 123-129 (2007)
14. E. Aldrian, Y.S. Djamil, Int J Climatol, **28**, 435-448 (2008)
15. E. Aldrian, R.D. Susanto, Int J Climatol, **23**, 1435-1452 (2003)
16. I.W.A. Yuda, R. Prasetya, *Seminar Nasional dan Rapat Tahunan Bidang MIPA*, 495-503 (2014)
17. NASA, *Global Precipitation Measurement (GPM) Mission : Algorithm Theoretical Basis Document*, (2017)
18. H.H. Hendon, Journal of Climate, **16**, 1775-1790 (2003)
19. D.S. Wilks, *Statistical Methods in the Atmospheric Science*, (San Diego : Academic Press (1995)
20. A. R. As-Syakur, T. Tanaka, R. Prasetya, I. K. Swardika & I. W. Kasa, Int J Remote Sens, **32(24)**, 8969-8982 (2011)
21. C.J. Willmott, Bull. Amer. Meteor, **63(11)**, 1309 – 1313 (1982)
22. M.D. Setiawati and F. Miura, Journal of Geoscience and Environment Protection, **4**, 101-117 (2016)