NOAA Satellite performance in estimating rainfall over the Island of Lombok

Humairo Saidah¹, Rini S Saptaningtyas¹, Lilik Hanifah¹, IDG Jaya Negara¹, Desi Widyanıty¹

¹Civil Engineering, Faculty of Engineering, Mataram University, Jln. Majapahit 62 Mataram, Indonesia

E-mail: h.saidah@unram.ac.id

Abstract. Rainfall data is the main input data for use in hydrological analysis, especially for planning purposes. Rainfall data is usually obtained from observations through rain stations. However, the availability of rainfall recording stations is sometimes very limited and located far from the project location. Water resource projects often face this obstacle. A technique is needed for estimating rainfall data using rain data from remote sensing through weather satellite data information providers. Since rainfall data from satellites is not the real amount of measured rainfall on the land, the rainfall data from the satellite must still be evaluated as to whether it can be directly used to represent the rainfall data of a region. The aim of this study is to evaluate the validity of NOAA CPC daily total precipitation data by comparing it to measured rainfall data and to analyse the lag time using the cross-correlation method. The results showed that the NOAA data has a lag time of +1 day over Lombok Island. The rain estimated by NOAA positively correlated with gauged rainfall on the next day. However, the performance of the NOAA data was very poor and for prediction purposes, the data still need to be calibrated.

Keywords: Rainfall data, satellite, NOAA CPC, cross correlation

1. Introduction

Rainfall data is the main input data for use in hydrological analysis, especially for civil planning purposes. Rainfall data is usually obtained from observations through rain gauge stations. However, the availability of rain gauge stations is sometimes very limited and located far from the project location. Water resource projects often face this obstacle. A technique is then needed for predicting rainfall data by using rain data from the results of remote sensing through satellite weather data information providers, as a solution for this lack of rain data.

There are many providers of satellite weather data information, including NOAA AVHRR, TRMM 3B42, TRMM 3B42RT, GPM, PERSIANN CCS, CMORPH, and others. Considering that the location of a region to the position of a satellite differs from one another, and the rainfall data from satellites is not the real amount of measured rainfall on the land, the rainfall data from a satellite must still be evaluated as to whether it can be directly used to describe the rainfall data of the region.

The aim of this study is to evaluate the validity of NOAA CPC daily total precipitation data and its possibility to fulfil the needed rainfall data in many ungauged regions of the island of Lombok, not
only for civil planning purposes but also for other uses such as creating planting schedules, water resource management, and even early disaster mitigation.

2. Materials and Methods

2.1. Study Location
The study takes the location of the island of Lombok in West Nusa Tenggara, one of the provinces of Indonesia. It is located between 115° E and 117° E and between 8° S and 9° S. The elevation range varies from below 0 m up to 3100 m at the peak of Mount Rinjani. The southern and north-eastern parts of the island are dominated by dry land and rainfed agriculture. The middle and western parts of the island have many rivers and available surface water. The distribution of rain gauge stations over the island is shown in Figure 1.

2.2. Rain Gauge Data
The primary rainy season mostly occurs from May to October. However, the rainfall period is actually more complex, as shown in Figure 2. The recorded data show that the eastern part of the island of Lombok is the driest region, with a rainfall of less than 1000 mm throughout 2016. The Pringgabaya rain gauge station recorded the lowest amount of rainfall (approximately 588 mm) in 2016, followed by Sepit, Rembitan, and Loang Make with 721 mm, 801 mm, and 828 mm respectively. The Sesaot rain gauge station recorded the highest amount of rainfall (approximately 3806 mm), followed by Lingkuk Lime (3549 mm) and Santong (3384 mm) in 2016.

2.3. Satellite Data
NOAA satellites are American-owned satellites that have the mission to save the environment and to record weather on earth. NOAA satellites carry five types of sensors, one of which is the AVHRR (Advanced Very High Resolutions Radiometer) sensor. NOAA satellites produce image data that can be used to provide meteorological parameters to create cloud maps, determine comparisons between
rain and cloud types and cloud coverage, determine annual variations in cloud coverage, and make weather maps and other weather forecasts [1]. NOAA also provides gridded climate datasets, including for precipitation, from the CPC Unified Precipitation Project that is underway at the NOAA Climate Prediction Centre (CPC). The datasets of daily rainfall were produced at a spatial resolution of 0.25º and presented as plots and arrays for the years from 1979 to present.

![Figure 2. Monthly rainfall data over the island of Lombok throughout 2016](image)

The data used in this study consist of daily total precipitation data directly downloaded from the NOAA web site at [https://www.esrl.noaa.gov/psd/](https://www.esrl.noaa.gov/psd/) as the results of the NOAA CPC Unified Precipitation Project. The data series was then processed by interpolating the values of the gridded daily total precipitation data to the rain gauge station location coordinates to obtain the basic data of NOAA daily rainfall at the same location. The generated data was then accumulated and presented by months, as in Figure 3.

![Figure 3. NOAA CPC precipitation data over the island of Lombok throughout 2016](image)
2.4. Methods
The steps of the study are described by the following flowchart (Figure 4).

![Flowchart of the study](image)

**Figure 4. Flowchart of the study**

3. Results and Discussion
3.1. Lag time analysis
Cross-correlation analysis was used to determine the delay time of satellite rainfall data in linear relation to the observational data. The assumption taken is that the satellite rain data will be recorded earlier than the results of measurements, because the travel of water from the atmosphere to the land surface will require time.

Cross-correlation is a method for predicting the relationship of an input data series (data x) to output data (data y) on a system. The cross-correlation method is widely applied to see the relationship of two groups of data in time-series analysis. This method requires the user to have a sample of two data series at the same time and is assumed to be stationary in means and variances. The cross-correlation method is often used in analysing the linear relationship between input and output in the study of hydrology and hydrogeology, such as the analysis of the delay time between rainfall and changes in spring discharge, or the linear relationship between satellite rainfall and measured rainfall [3, 4].

Cross-correlation is a standard method for estimating the degree of two correlated data series. For two series x (i) and y (i) with i = 0, 1, 2 ... N - 1, the cross-correlation (r) after delay d is stated as [5]:

\[
r(d) = \frac{\sum_{i}(x(i) - mx)(y(i-d) - my)}{\sqrt{\sum_{i}(x(i) - mx)^2 \sum_{i}(y(i-d) - my)^2}}
\]  

(1)
where \( m_x \) and \( m_y \) are the average values of each data row. Series \( x(i) \) is the series of satellite rain data and series \( y(i) \) is the series of observed rainfall data. This correlation will show that the satellite rain data and observation rainfall data have a correlation of \( r(d) \). If \( r(d) \) shows a positive value, then the satellite data can be used to predict rainfall observation data. The results of lag time analysis between NOAA CPC rainfall data and rain gauge data are shown in Table 1 and Figures 5 and 6.

### Table 1. Correlation coefficients between NOAA CPC daily precipitation to observations for many lag times

| Rain gauge Station | Lag time (day) |
|--------------------|----------------|
|                    | -3  | -2  | -1  | 0   | 1   | 2   | 3   |
| Gunungsnari        | 0.12| 0.09| 0.12| 0.13| 0.13| 0.09| 0.09|
| Kuripan            | 0.13| 0.09| 0.13| 0.14| 0.12| 0.09| 0.09|
| Sesaot             | 0.13| 0.09| 0.12| 0.14| 0.12| 0.09| 0.08|
| Jurangsate         | 0.14| 0.09| 0.12| 0.14| 0.13| 0.10| 0.08|
| Lingkuk Lime       | 0.13| 0.09| 0.11| 0.13| 0.13| 0.10| 0.09|
| Loang Make         | 0.14| 0.10| 0.12| 0.14| 0.14| 0.11| 0.10|
| Rembitan           | 0.14| 0.10| 0.12| 0.13| 0.14| 0.12| 0.09|
| Ijo Balit          | 0.12| 0.12| 0.12| 0.13| 0.17| 0.12| 0.10|
| Perian             | 0.10| 0.11| 0.11| 0.13| 0.18| 0.13| 0.11|
| Pringgabaya        | 0.09| 0.14| 0.11| 0.14| 0.20| 0.18| 0.13|
| Sepit              | 0.20| 0.11| 0.16| 0.12| 0.23| 0.13| 0.12|
| Santong            | 0.11| 0.15| 0.11| 0.16| 0.25| 0.21| 0.14|
| Average            | 0.13| 0.11| 0.12| 0.13| 0.16| 0.12| 0.10|

![Figure 5](image_url)

**Figure 5.** Correlation coefficients of many lag times

The lag time between NOAA CPC daily precipitation and measurement rainfall is determined by the best correlation coefficient at the time interval. Table 1 and Figure 5 illustrate that, spatially, the island of Lombok has lag times that vary from D-3 to D+1. Figure 6 shows the average correlation coefficient over the island of Lombok. Based on the highest correlation coefficient, it can be concluded that the average lag time of satellite rainfall product to rain gauge measurement over the island of Lombok is D+1. This means that the rain data produced by NOAA CPC on one day positively corresponds with the rainfall data by rain gauge measurement on the next day.
Figure 6. Average correlation coefficients of lag times over the island of Lombok

3.2. Comparison and Validation
The validation test was used to find out the validity of NOAA data compared to the observation data. Furthermore, this was used to consider whether the satellite rainfall data can be used directly to predict the amount of rainfall of the study location, or if adjustments are still needed.

The accuracy test of satellite rainfall data can be seen from the values of normalized root-mean-square error (NRMSE), correlation coefficient (r), mean error (ME), and efficiency (E) [2,6,7].

\[
\begin{align*}
    r &= \frac{\sum_{i=1}^{n} (S_i - \bar{S})(G_i - \bar{G})}{\sqrt{\sum_{i=1}^{n} (S_i - \bar{S})^2 \sum_{i=1}^{n} (G_i - \bar{G})^2}} \\
    ME &= \frac{1}{n} \sum_{i=1}^{n} (S_i - G_i) \\
    NRMSE &= \sqrt{\frac{\sum_{i=1}^{n} (S_i - G_i)^2}{\sum_{i=1}^{n} G_i^2}} \\
    E &= 1 - \frac{\sum (S_i - G_i)^2}{\sum (G_i - \bar{G})^2}
\end{align*}
\]

Here, \( G \) = gauged data, \( \bar{G} \) = average of gauged data, \( S \) = satellite data, and \( n \) = number of data pairs. ME and NRMSE are in millimetres, and \( E \) and \( r \) have no units. ME and NRMSE measure the average magnitude of satellite prediction error referring to the observed rainfall. Both values can range from 0 to infinity. These “statistical goodness of fit” tests are widely used to measure errors and are simple in their operation.

The value of efficiency (E), also known as the coefficient of efficiency or Nash-Sutcliffe efficiency (NSE), shows the quality of an estimate relative to a reference. The value varies from minus infinity to one. The negative values mean that the gauged mean is a better estimate. An efficiency of 0 indicates that the satellite data is as accurate as the mean of the gauged data. A value of 1 corresponds to a perfect match of satellite data to gauged data. The results of the statistical values of comparison are shown in Table 2 and Figure 7.
Table 2. Statistical values of comparison between NOAA CPC daily total precipitation data and gauged rainfall data over the island of Lombok in 2016

| Rain Gauge Station | r   | ME  | NRMSE | E    |
|--------------------|-----|-----|-------|------|
| Gunungsari         | 0.06| -1.16| 2.79  | -0.57|
| Kuripan            | 0.15| 2.18 | 4.22  | -1.11|
| Sesaot             | 0.14| -4.68| 2.08  | -0.32|
| Jurangsate         | 0.15| -2.47| 2.58  | -0.26|
| Lingkuk Lime       | 0.21| -4.06| 2.17  | -0.21|
| Loang Make         | 0.02| 3.35 | 6.85  | -2.36|
| Rembitan           | 0.18| 0.26 | 3.18  | -0.47|
| Ijo Balit          | 0.16| 3.32 | 6.65  | -2.75|
| Perian             | 0.10| -1.71| 2.66  | -0.50|
| Pringgabaya        | 0.10| 3.85 | 9.21  | -5.92|
| Sepit              | 0.12| 3.61 | 7.07  | -4.50|
| Santong            | 0.21| -3.56| 2.39  | -0.15|

Figure 7. Values of Mean Error (ME), Normalized Root-Mean-Square Error (NRMSE), Efficiency (E), and Correlation Coefficient (r) between NOAA CPC daily total precipitation results and gauged rainfall data over the island of Lombok in 2016
Table 2 and Figure 7 show high values of NRMSE and ME, which indicate high deviations of the satellite data with respect to gauged data. Sometimes the NOAA data values are underestimates and at other times are overestimates. The most severe underestimated values are for wet regions such as Sesaot, Jurang Sate, and Lingkuk Lime, and the most severe overestimated values are for dry regions such as Pringgabaya, Ijo Balit, Sepit, and Loang Make.

The NOAA satellite rainfall data corresponds poorly to gauged data, as indicated from the values of correlation coefficients and efficiency. The positive weak relationship is shown by a low correlation coefficient of less than 0.21. The negative efficiency values also indicate that NOAA satellite data is of a poor quality, which means that NOAA data is not an accurate predictor compared to observation data.

3.3. Discussion
Research has been carried out to evaluate the performance of NOAA CPC rainfall data by comparing it to rainfall measurement data. The data tested for performance was the total daily rainfall data from the results of the CPC Global Unified Gauge-Based Analysis of Precipitation, with a grid of 0.25° x 0.25°.

The results showed that on average the NOAA CPC rainfall data had a positive correlation with rainfall measured on the next day by rain gauge stations on the island of Lombok. The relationship is weak but positive for all rain gauged data, approximately 0.16. The low correlation coefficient was caused by the use of relatively high temporal (daily) and spatial (0.25°) data resolution. The performance of the satellite product decreases as spatial and temporal resolutions increase [3, 7]. However, with a positive correlation coefficient number, a study [3] stated that TRMM and GPM satellite data are suitable to be used for predicting rainfall data, and another study [7] stated that NOAA-CMAP data performed well for a low temporal (monthly) and spatial (2.5°) resolution, with a bias rate of less than 10%.

Based on the positive correlation coefficient figures throughout rain gauge stations on the island of Lombok, NOAA data has potential to be used to predict rainfall in areas of the island of Lombok that do not have rain gauges. The accuracy test between NOAA rain data and observation data on the island resulted in low efficiency and high error values. This condition presents an opportunity to calibrate the NOAA data or to develop an algorithm to predict and obtain rainfall data based on NOAA rainfall data.

4. Conclusion
The conclusion of this study is that rain data from the NOAA CPC global daily total has a lag time of +1 day from the observation data. The data recorded on one day by the NOAA satellite is rainfall measured on the next day on land. Based on the positive correlation figures throughout rain gauge stations on the island of Lombok, NOAA data has potential to be used to predict rainfall in areas of Lombok that do not have rain gauges, but with prior calibration.

Acknowledgements
This work was funded by a grant from Mataram University.

References
[1] Massinai Muh Altin 2005 Indonesia Prosiding PIT MAPIN XIV Surabaya
[2] Prasetya, R. As-syakur, A.R., Osawa, T 2013 J Theor Appl Climatol p.575
[3] Pratiwi, Destiana Wahyu, Joko Sujono, Adam Pamudji Rahardjo 2017 Prosiding Seminar Nasional Sains dan Teknologi Jakarta
[4] Romadlon M Muzayyin, Teguh Hariyanto 2014 J GEOID Vol.10, No.01
[5] Koriyanti Ery 2009 Jurnal Penelitian Sains Volume 12 Nomor 1(B)
[6] Su, Fengge, Hong Y. Lettenmaier D.P 2008 American Meteorological Society Journal
[7] Dinku T, Ceccato P, Grover-Kopec E, Lemma M, Connor S. J, and Ropelewski, C.F 2007 International Journal of Remote Sensing, 28:7, 1503-1526