Estimation of TRMM rainfall for landslide occurrences based on rainfall threshold analysis

Noraisyah Tajudin1, Norsuzila Ya’acob2, Darmawaty Mohd Ali3, Nor Aizam Adnan4
1,2,3Wireless Communication Technology (WiCoT), Faculty of Electrical Engineering, Universiti Teknologi MARA, Malaysia
4Faculty of Architecture, Planning and Surveying, Universiti Teknologi MARA, Malaysia

ABSTRACT
Landslide can be triggered by intense or prolonged rainfall. Precipitation data obtained from ground-based observation is very accurate and commonly used to do analysis and landslide prediction. However, this approach is costly with its own limitation due to lack of density of ground station, especially in mountain area. As an alternative, satellite derived rainfall techniques have become more favorable to overcome these limitations. Moreover, the satellite derived rainfall estimation needs to be validated on its accuracy and its capability to predict landslide which presumably triggered by rainfall. This paper presents the investigation of using the TRMM-3B42V7 data in comparison to the available rain-gauge data in Ulu Kelang, Selangor. The monthly average rainfall, cumulative rainfall and rainfall threshold analysis from 1998 to 2011 is compared using quantitative statistical criteria (Pearson correlation, bias, root mean square error, mean different and mean). The results from analysis showed that there is a significant and strong positive correlation between the TRMM 3B42V7 and rain gauge data. The threshold derivative from the satellite products is lower than the rain gauge measurement. The findings indicated that the proposed method can be applied using TRMM satellite estimates products to derive rainfall threshold for the possible landslide occurrence.

Keywords: Landslide Rain gauges Rainfall threshold Satellite rainfall estimates TRMM

1. INTRODUCTION
Rainfall is known as one of the main factor which contribute to landslide occurrence, thus causing a harmful event to people, societies and economies worldwide under climate change conditions [1-3]. Relationships between rainfall and landslides are frequently complex due to inconsistent and incomplete landslide records, and the availability of rainfall data [4]. In Malaysia, the comparison of landslide triggering factor shows rainfall has a major contribution to the event, which is about 58% with an average of 2550 mm of rainfall per year [5]. Currently, the prediction of landslide triggered by rainfall is based on examining the empirical relationship between rainfall characteristics and past landslide occurrence [6]. The predefined thresholds were set to define the best separators for triggering and non-triggering rainfall condition to forecast an early warning tools [3, 4] and landslide hazard assessment.

Precipitation data can be obtained in three ways in the urban area which are ground observations, weather radar observations, and satellite monitoring techniques. Ground observations provide the most direct and accurate measurements at the gauge locations, but it is difficult to obtain long-term precipitation data
because of the geographical restriction of the station location and the shortage of meteorological observation equipment. Maintaining these stations are also not cost effective. While precipitation of radar observations have large uncertainties related to error in electronic signal under challenging operating environments. In contrast, satellite remote sensing can provide observations of global precipitation and clouds [7-10].

Satellite derived rainfall products are the most accepted alternative source to overcome the limitations of ground techniques. Recently, satellite information has become available at high spatial and temporal resolution and over large area. However, all satellite precipitation products provide indirect estimates of precipitation which are subjected to bias and stochastic errors, depend highly on the hydro-climatic characteristics of a region. Therefore, satellite-derived precipitation products need to be assessed by comparing the data with ground observations [7, 11, 12]. The performance of rainfall product variation varies by data sources and retrieving algorithms mode. The rainfall product performance for the same type of data also shows differences in different regions and seasons. These factors indicate that the performance of rainfall satellite products are affected by the season, topography, location and hydro-climatic characteristics of the study area. Therefore, the reliability of the satellite rainfall product needs to be validated and compared to the ground-based measurements in a specific area and temporal scales before it can be used in any subsequent application.

Numerous studies on the evaluation of satellite precipitation products with ground measurements have been conducted at the regional scale [4, 13, 14]. The authors in [15] have validated that TRMM satellite rainfall estimates can be used to determine cumulated estimated rainfall – event duration (ED) threshold for the forecast of the possible landslide occurrence in the Umbria region, Central Italy. A similar study has been conducted, while the paper in [16] has validated the threshold and tested the effectiveness of the different threshold definition methods with independent landslide information. The result shows the GPM satellite rainfall products can be applied to derive rainfall threshold for the possible landslide occurrence. In Malaysia, studies on the evaluation of satellite precipitation products are very limited.

Most of the studies conducted are based on the comparison of multiple temporal (daily, monthly and annual) rainfall between satellite products and ground observation [8, 17, 18]. The aim of this study was to evaluate the potential used of TRMM-3B42V7 data in comparison to the rain-gauge data based on rainfall threshold analysis for the prediction of landslide occurrence in Ulu Kelang, Selangor. The evaluation was conducted by using five statistical indicators to compare the average monthly rainfall, cumulative 30-day and 3-day rainfall based on 14 historical landslide events and rainfall threshold for both derivations from TRMM 3B42V7 and Rain gauge data series.

2. RESEARCH METHOD

2.1. Study area

This study was conducted in Ulu Kelang, Selangor, located at the latitude of 3°12’30”N and longitude of 101° 45’ 28” E with a distance of 5 km from Kuala Lumpur city center. Ulu Kelang, Selangor was known as the most susceptible region of landslides in Malaysia. Commonly, the climate of this area is hot and humid as it is in the tropical monsoon region and pronounced wet or rainy seasons started from February until May and September until December each year. According to the Malaysian Meteorological Department (MMD), the accumulated monthly rainfall for Ulu Kelang is between 58 and 420 mm per month, with an average of 200 rainy days per year and 2440 mm average annual rainfall. The Ulu Kelang area has suffered several fatal landslides cases due to rainfall events [19].

2.2. Validation process

There are two independent data sources of rainfall information from 1998 to 2011 were used in this study. The first is a daily rainfall measurement obtained by 2 rain gauges, namely the Empangan Genting Kelang Station and JPS Ampang station located in Ulu Kelang Area, Selangor. The second source of rainfall information is a satellite rainfall estimate product, which is TRMM-3B42V7 with a resolution of 0.25O x 0.25O, provided by NASA TRMM. The used of TRMM-3B42V7 data in this study is based on the adequate reliability of this product to estimate the monthly and annual precipitation over the tropical region [20, 21]. According to the J. Paska at el, a moderate correlation between TRMM-3B42V7 and ground observation was found for daily precipitation in Peninsular Malaysia [22]. The validation was conducted by comparing three analyses, which are the average monthly rainfall, cumulative 30-day and 3-day rainfall based on 14 historical landslide events and rainfall threshold analysis in Ulu Kelang, Selangor. The average monthly rainfall was analyzed over 14 years from 1998 to 2011. The period was selected based on the 14 historical landslide events which occurred from 1998 to 2011. The next comparison was based on the cumulative 30-day and 3-day rainfall parameter which involved in the construction of rainfall threshold for Ulu Kelang, Selangor. Finally, the rainfall threshold (E3-E30 diagram) derived from both
TRMM-3B42V7 and Rain gauge was compared to identify the reliability of rainfall threshold based on TRMM-3B42V7 product for the prediction of landslide occurrence. Figure 1 shows the flowchart of validation process.

2.3. Validation statistics

The Mean (M), Mean Difference (MD), Root Mean Square Error (RMSE), Bias and Pearson correlation coefficient (R) were used to compare the differences and correlation between TRMM-3B42V7 and rain gauges. The point-to-pixel assessment for comparing the pixels and the corresponding rain gauges was used. The Mean is used to define the central tendency of rainfall data set for both rain gauges and TRMM-3B42V7, while MD is the mean difference between the two data sets. RMSE measures the mean error between rain gauges and TRMM-3B42V7 rainfall, in millimeter (mm). The Bias is a measurement of how the average satellite rainfall estimates the magnitude as compared to the ground-based rainfall observation, where is the perfect score is 1. A bias value above or below 1 indicates an aggregate satellite rainfall overestimation or underestimation of the ground-based precipitation amounts respectively. The Pearson correlation is used to show the correlation between spatial distribution of the TRMM-3B42V7 and rain gauges, with the value ranges between 0 – 1. 0 indicates no correlation while the value of 1 shows a perfect correlation. The equations for the statistical metrics that were used for validation are shown in Table 1.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{flowchart.png}
\caption{Flowchart of validation process}
\end{figure}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
Statistical Metrics & Equation & Perfect Score \\
\hline
Mean (M) & $M = \frac{1}{n} \sum_{i=1}^{n} x_i$ & - \\
Mean Difference (MD) & $MD = \frac{1}{n} \sum_{i=1}^{n} (X_i - Y_i)$ & 0 \\
BIAS & $BIAS = \frac{\sum Y_i}{\sum X_i}$ & 1 \\
Root Mean Square Error (RMSE) & $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - Y_i)^2}$ & 0 \\
Pearson Correlation (R) & $R = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2 \sum_{i=1}^{n} (Y_i - \bar{Y})^2}}$ & 1 \\
\hline
\end{tabular}
\caption{List of the statistical metrics employed to validate the performance of TRMM-3B42V7 [23]}
\end{table}

Where $X_i$ is the representation of rain gauged rainfall, at a time $i$, $Y_i$ is the representation of TRMM-3B42V7 at a time $i$.
3. RESULTS AND ANALYSIS

3.1. Comparison of average monthly rainfall

Figure 2 shows the average monthly rainfall for the case-study of over 14-years period, by comparing the TRMM-3B42V7 and gauge data. It also shows the distribution of rainfall data and landslide events occurrences in 1998 to 2011. The comparison between two rainfall data reveals that the TRMM rainfall product has overestimated rainfall distribution in January, February, and December. In the meantime, between March to November, the TRMM product shows an underestimate rainfall distribution. However, the highest amount of TRMM rainfall product is recorded during this period. Even though there is a difference in the rainfall value of TRMM product, but it the average monthly rainfall derived by rain gauge rainfall data shows similar trend. From the analysis, TRMM-3B42V7 has demonstrated a significant and strong positive correlation to the rain gauge rainfall data with R equivalent to 0.88. The RMSE between rain-gauge and TRMM data has lower value, which is 12.97 mm.month⁻¹ that is close to the rain-gauge data. This finding is in agreement with the study in [8, 24] which reported the good performance of TRMM products. However, the TRMM 3B42V7 has underestimated the rain-gauge precipitation with a BIAS of -0.0985. The maximum monthly average rainfall for rain-gauge data is 382.875mm in November and 321.7475 mm for TRMM product was recorded in April. Meanwhile, the minimum values were 133.5 mm (February) and 141.8675 mm (July) respectively. This analysis also shows that the highest number of landslide events occurred when there are highest amount of rainfall during the rainy season obviously in May and November. Therefore, the possibility of rainfall as a triggering factor for landslide occurrences in wet or rainy season is higher in Ulu Kelang, Selangor.

![Figure 2. Monthly average rainfall (1998-2011) for rain gauge, TRMM-3B42V7 and the number of landslide events](image)

3.2. Comparison of cumulative 30-day rainfall and 3-day rainfall

The next analysis was referring to the parameter involved in the derivation of the rainfall threshold model for landslide occurrences in Ulu Kelang Selangor. The cumulative 30-day and 3-day rainfall were used to compare the rainfall intensity between rain gauge and TRMM-4B42V7 during landslide occurrences. The cumulative rainfall data is based on 14 historical landslides that occurred between 1998 to 2011. Figure 3 shows the rainfall intensity for cumulative 30-day and cumulative 3-day rain for gauge and TRMM-3B42V7. Both histograms in Figure 3 (a) and (b) show that TRMM-3B42V7 data has underestimated as compared to the rain gauge rainfall intensity. The differences in percentage of rain gauge to TRMM-3B42V7 is 63% (Cumulative 30-day) and 45% (Cumulative 3-day) respectively. The maximum cumulative 30-day rainfall intensity for TRMM-3B42V7 is 387.36 mm, while rain gauge is 759.5 mm; and for cumulative 3-day are 94.89 mm for TRMM and 196 mm for the rain gauge. Overall, these analyses indicate that the rain gauge rainfall has exceeded the TRMM-3B42V7 rainfall.

Each TMPA-3B42V7 product was plotted against gauge data in scatter plots based on cumulative 30-day and 3-day rainfall during landslides occurrences from 1998 to 2011 at Ulu Kelang Selangor, as shown in Figure 4. This analysis is carried out to analyze the correlation between TRMM-3B42V7 and rain gauge. From the analysis, both TRMM-3B42V7 cumulative 30-day and 3-day rainfall presented a significant and
strong positive correlation to the rain gauge rainfall data. The correlation coefficient of TRMM-3B42V7 versus rain gauge is 0.66 for cumulative 30-day rainfall and 0.63 for cumulative 3-day rainfall. By referring to RSME value for both cumulative 30-day and 3-day rainfall, it shows that TRMM-3B42V7 data is closer to rain gauge data set. The RMSE value is 5.81 mm for cumulative 30-day rainfall while for cumulative 3-day rainfall is 5.31 mm. The result of the statistical metrics for TRMM-3B42V7 versus rain gauge event rainfall is shown in Table 2.

3.3. Comparison of rainfall threshold analysis for TRMM-3B42V7 and rain gauge rainfall data

To define the rainfall that triggered the landslide, it requires the identification of rainfall time or condition and other related information of the landslide occurrences. Based on the rainfall analysis, the rainfall conditions that had triggered the landslides in Ulu Kelang are either short and long duration or prolonged continuous period of rainfall. Therefore, the rainfall threshold that incorporates cumulative 3-day versus 30-day rainfall for fourteen selected landslides event was derived for the best rainfall threshold condition of Ulu Kelang, Selangor. This method is appropriate for the rainfall condition that had triggered landslides in the area [25]. The plotted rainfall threshold (E₃ – E₃₀) diagram for historical rainfall that has resulted in landslide using TRMM-3B42V7 product and rain gauge rainfall data is shown in Figure 5(a) and (b) respectively. Both rainfall threshold lines were generated from the correlation between the 3-day (E₃) and 30-day (E₃₀) antecedent precipitation index based on the occurrence of minor and major landslide events.

![Figure 3. Total rainfall (a) cumulative 30-day rainfall and (b) cumulative 3-day rainfall for rain gauge and TRMM based on 14 landslide events in ulu kelang, Selangor](image)
Estimation of TRMM rainfall for landslide occurrences based on rainfall threshold … (Noraisyah Tajudin)

Figure 4. Scatter plot between rain gauge and TRMM 3B42V7 based on 14 landslide events in Ulu Kelang, Selangor, (a) cumulative 30-day rainfall (b) cumulative 3-day rainfall

Table 2. Results of statistical metrics for TRMM-3B42V7 vs rain gauge event rainfall

| Cumulative Rainfall | 30 Day | 3 Day |
|---------------------|--------|-------|
| MD                  | 149.78 | 51.67 |
| BIAS                | 0.3687 | 0.5444|
| RMSE                | 5.81   | 5.31  |
| R                   | 0.6651 | 0.6326|

Figure 5. Plotting of E3-E30 diagram for historical rainfall that has resulted in landslide using (a) TRMM-3B42V7 rainfall data and (b) rain gauge rainfall data

To compare the rainfall threshold diagram between TRMM-3B42V7 and rain gauge rainfall data, both graphs are plotted in the diagram as in Figure 6. The figures show, the rainfall threshold equation for TRMM-3B42V7 is lower than the rain gauge rainfall threshold equation based on the major and minor landslide threshold equation’s analysis. The equation for both rainfall threshold is tabulated in Table 3. By referring to the comparison of rainfall threshold of TRMM-3B42V7 and rain gauge, it shows that the underestimation of the “ground-based” rainfall data is observed [16].
Figure 6. Combination of rainfall threshold derived by rain gauge and TRMM-3B42V7

Table 3. The equation of rainfall threshold for major and minor landslide occurrences derived from rain gauges, TRMM-3B42V7

| Rainfall Product | Major Landslide Threshold | Minor Landslide Threshold |
|------------------|---------------------------|---------------------------|
| Rain Gauges      | $E_3 = 167.39 - 0.21E_{30}$ | $E_3 = 108.73 - 0.21E_{30}$ |
| TRMM-3B42V7      | $E_3 = 140.016 - 0.448E_{30}$ | $E_3 = 104.56 - 0.448E_{30}$ |

4. CONCLUSION

In this study, the statistical method and rainfall threshold analysis is used to compare TRMM 3B42V7 and rain gauge precipitation data sets on daily scale. The analysis is based on cumulative 3-day and 30-day rainfall of historical landslide which is the parameter for rainfall threshold. The result shows, the RMSE of average monthly rainfall for TRMM 3B42V7 is 12.97 mm/month which is close to the rain-gauge data. The TRMM 3B42V7 has overestimated rainfall distribution in January, February, and December which are at the end of Northeast Monsoon season. While the RMSE of cumulative 30-day rainfall and 3-day rainfall show that TRMM 4B42V7 product closer with rain gauge data sets about 5.81 mm (Cumulative 30-day) and 5.31 mm (Cumulative 3-day). On the other hand, the correlation (R) between rain-gauge and TRMM 4B42V7 presented a significant and strong positive correlation, that is 0.66 for cumulative 30-day rainfall and 0.63 for cumulative 3-day rainfall respectively. The comparison of rainfall threshold ($E_3$ – $E_{30}$) diagram for historical rainfall that has resulted in landslide shows that the rainfall threshold equation for TRMM-3B42V7 is lower than the rain gauge rainfall threshold equation based on major and minor landslide threshold equation analysis. Overall, the results show that the comparison of the parameters of rainfall thresholds obtained from the rain gauge and TRMM 3B42V7 using the statistical methods, reveal that the threshold derived from the satellite products has underestimated the rainfall distribution with respect to ground-based observations. These results are in agree accordance the finding in [4, 9, 16] in which the underestimation of the “ground-based” rainfall data observation. Finally, the proposed method which is using the TRMM 3B42V7 products can be used to obtain rainfall threshold for possible landslide.

ACKNOWLEDGEMENTS

The authors would like to thank Faculty of Electrical Engineering, Universiti Teknologi MARA (UiTM) for their valuable support. This research is partly funded by the Malaysian Government through UiTM under 600-IRMI/DANA5/3/BESTARI (122/2018). We are grateful to NASA TMPA for providing the TRMM product version 7 for 3B42(7) data.

REFERENCES

[1] E. Monsieurs, O. Dewitte, and A. Demoulin, “A susceptibility-based rainfall threshold approach for landslide occurrence,” Natural Hazards and Earth System Sciences, vol. 19, no. 4, pp. 775–789, Apr. 2019.
Estimation of TRMM rainfall for landslide occurrences based on rainfall threshold ... (Noraisyah Tajudin)