Combination of machine learning model (LR-FR) for flash flood susceptibility assessment in Dawuan Sub watershed Mojokerto Regency, East Java

L Y Irawan¹*, Sumarmi¹, D Panoto¹, Nabila¹, I H Pradana¹, and A C Darmansyah¹
¹Department of Geography, State University of Malang, Jl. Semarang 05, Malang, Indonesia
E-mail: listyo.fis@um.ac.id

Abstract. The Dawuan Sub-watershed in Mojokerto Regency is a prone area to floods. There were flash floods in this area in 2002 and 2019, which caused casualties and property losses. As one of the mitigation efforts, this study aims to map a flash flood's susceptibility using the LR-FR combination machine learning technique (logistic regression and frequency ratio). 11 conditioning factors are used to assess landslide susceptibility, namely: slope, aspect, TWI (Topographic Wetness Index), TPI (Topographic Position Index), SPI (Stream Power Index), profile curvature, distance to drainage, rainfall, geological unit, and land use. The results of the flash flood susceptibility mapping show that areas with very high levels of susceptibility have the following characteristics: slope < 8-35°; aspect east and southwest; TWI >16; TPI (<-3.39)-(-0.06); SPI <50-200; profile curvature (-0.001)-0.0; distance to drainage <10-40; rainfall <2000; geological unit Qvwl, Qvlw3, Qvlp3, Qvlp4, Qvwl, Qvf3, Qvf4 and Qvf8; and agricultural land use. The validation results show that the quality of the LR-FR model used has very good quality, as indicated by the AUC value = 0.93.

1. Introduction
Flash flood is high speed water flow in a steep slope from a river's upstream [1, 2], carrying materials such as mud, stones, logs, sands, and gravel [3]. The flash flood's high speed water flow is really destructive that might cause materials loss and casualties [4-6]. The damage of the flash flood depends on the water flow speed, and the materials carried. The high speed of the flash flood water flow cause limited time and effective decision making for emergency response efforts, and the damage power becomes higher [5,7].

Topography, geology, hydrology, and climatology are all factors that trigger flash floods [2, 8]. The high intensity of rainfall results in high river flow rates [9], while the volume of the reservoir is small or narrow [2, 10] and carries material from landslides upstream of the river [1,3]. Flash flood causes the river to overflow, causing puddles in the area around the river [3, 11].

The Dawuan sub-watershed, which includes several sub-districts in Mojokerto Regency, is one of the areas in Indonesia that is threatened by flash flood [12,13]. Flash flood were recorded in 2002 and 2019. Flash flood in 2002 caused 128 fatalities and damaged of various infrastructure such as hot spring tourist attractions and bridges. Meanwhile, the flash flood incident in 2019 did not result in casualties, but resulted in damage to several infrastructures such as bridges and pipelines for tourist hot water pipelines [14].

Flash flood in the Dawuan Sub-watershed are caused by high intensity of rainfall followed by topographical conditions with the very steep slopes. In addition, other causes are due to the upstream
river or the upper slopes of Mount Welirang experiencing deforestation due to forest fires and the formation of natural dams. The high intensity of rainfall on the upper slopes of Mount Welirang results in high river flow rates and causes the breakdown of natural dams and landslides. So that wood, stone, and other sedimentation materials are carried away causing narrowing of the river body. As a result, the flash flood flow cannot be accommodated by the river body and overflows around the Padusan tourist hot spring.

As one of the disaster mitigation efforts, this study aims to mapping the level of flash flood susceptibility. The assessment of the flash flood susceptibility level in this study uses a machine learning technique with a combination of LR-FR (logistic regression and frequency ratio). This combination model produces the best level of accuracy based on previous research, it is shown that an AUC value of 0.91 is much better than the combination of the LR-WoE (logistic and weigh of efficiency) model, SVM-FR (support vector machine and frequency ratio), and SVM-WoE (support vector machine and weigh of efficiency) [15].

2. Methods

2.1 Research Area

This research area is in the Dawuan sub-watershed with an area of 62.09 km². The research location is in Mojokerto Regency, including Pacet, Trawas, Kutorejo, Pungging and Mojosari Districts. In the research area there were flash flood in 2002 and 2019, as shown in the Figure 1, the occurrence of flash flood is shown as a torrentially area while the sample area characteristics without flash flood events are non-torrential areas.

![Flash flood Inventory Data](image)

**Figure 1. Flash flood Inventory Data**

2.2 Spatial Dataset

2.2.1 Flash flood Data Inventory

The flash flood prediction using a machine learning approach requires inventory data of flash flood in the research area [16]. The inventory of flash flood data includes affected and non-affected areas by flash flood. The flash flood affected areas were obtained from BPBD Mojokerto Regency and field surveys. Where the flash flood incident refers to the flash flood incident in 2002 and 2019. Flash flood data inventory processing using a machine learning approach requires 70% data for training datasets.
and 30% data for validation [15]. The training data format has a two categories value which is 1 as the torrential area and a value of 0 as a non-torrential area [17]. The flash flood data inventory was obtained through BPBD data from Mojokerto Regency and 2020 field Survey. Based on the results of the flash flood data inventory, the area of torrentiality area is 6,300217 km$^2$ and non-torrential area is 9,431966 km$^2$.

![Figure 2. Research Flow Chart](image)

### 2.2.2 Flash Flood Conditioning Factor

There were twelve triggering factors to flash flood used in this research, which includes topographic, geologic, hydrologic and topographic conditions. Topographic parameters were TWI (Topographic Wetness Index), TPI (Topographic Position Index), SPI (Stream Power Index), slope, elevation, profile curvature and distance to drainage. The geologic parameter was geological unit, while hydrologic parameter was rainfall intensity. For the anthropogenic parameter, this research accounted land use and distance from the nearest road. Distance from the nearest road was used on the account of some landslide events which are triggered by a cut slope.

The rainfall data parameter used is the average rainfall within 10 years (2010-2019) obtained from Public Works Office of Mojokerto Regency. The Topography Parameter is processed from DEM Alos Palsar data which is analyzed using software of Q.GIS 3.12 and SAGA GIS 2.3.2. Parameters analyzed using Q.GIS 3.12 software are elevation, slope and distance to drainage and those using SAGA GIS 2.3.2 software are TWI, TPI, SPI, aspect, and profile curvature. Geological unit data is obtained based on the Geological Map classification of Pandaan Sheet scale 1: 50,000 published by Center of Volcanology and Geological Disaster Mitigation (PVMBG). Land Use parameters are obtained through a digitizing process from World View-2 image data.

Rainfall becomes one of the main factors of flash flood. High intensity of rainfall in the short period is the cause of flash flood [7, 16]. The decision of the flash flood threat can be observed and determined from the analysis of rain data [7]. The rain data used were processed using the IDW (Inverse Distance Weighting) method. The IDW method was chosen because it is suitable to be applied to areas with mountainous morphology with rain types with an uneven distribution of rain stations. Rainfall in the Dawuan Sub-watershed has an intensity of <2000, 2000 - 2300, 2300 - 2500, and 2500 - 2900 mm/year. Most of the Dawuan sub-watershed area has a high rainfall intensity, namely 2300-2900 mm/year.

Slope is a significant morphometric factor that most influences the surface runoff and infiltration rate [7, 15]. Areas with a high slope where surface runoff occurs have the possibility of flash flood,
while areas that are relatively flat have the potential for flooding due to inundation processes [6, 7, 17]. Slope of the Dawuan Sub-watershed is classified into 7 (seven) classes, namely <= 2˚, 2˚ - 4˚, 4˚ - 8˚, 8˚ - 16˚, 16˚ - 35˚, 35˚ - 55˚, and > 55˚ [18]. The classification is adjusted based on the morphological conditions of the study area. The condition of the Dawuan Sub-watershed is dominated by mountainous morphology.

Elevation is an analysis of the flood-affected area from a height. The water flow that moves from a higher place to a lower place is affected by the force of gravity [7, 16]. The lower elevation area the quicker the flood compared to the higher elevations [7]. In general, the research area is located in the highlands. The elevation of the Dawuan Sub-watershed is located at an altitude between 50 and 3150 meters above sea level.

Distance to Drainage is the slope stability factor caused by the drainage structure distance. This condition affects runoff, infiltration and soil moisture. The distance between the slope and the drainage system results in lower slope stability. As a result, if the intensity of rainfall is high, it can cause landslides and erosion [19].

Material deposition due to erosion and landslides that settles in the river results in the formation of natural river weirs containing solid material. This is the initial mechanism for flash flood. Distance to drainage in the Dawuan Sub-watershed area is divided into 5 classes, including <= 10, 10 - 20, 20 - 30, 30 - 40, and > 40 meters from the existing drainage area or river flow.

Aspect is a factor that has an indirect effect on flash flood. The intensity of sunlight and rainfall are factors that affect its value. Slope conditions with less sunlight intensity result in high soil moisture and reduced water infiltration into the soil. This condition causes high surface runoff [15].

The aspect conditions of the Dawuan Sub-watershed are classified into nine classes, namely: flat, north, northeast, east, southeast, south, southwest, west, and northwest. Aspects of the study area affect the intensity of erosion, run-off, and material deposition on a slope when a flash flood occurs. The nine parameters are influenced by variations in the microclimate in the study area.

Topographic Wetness Index is a hydrological parameter that is influenced by flow intensity and water flow accumulation at a hydrographic catchment point [7, 15, 16]. The TWI value describes the amount of water content stored in slope constituent materials which can affect slope instability. Slope stability factor has a big influence on the mechanism of flash flood. This is influenced by the drainage system and also the rainfall that occurs. Classified into 5 (five) TWI classes, namely: <= 6, 6 - 10, 10 - 12, 12 - 16, and > 16.

Topographic Position Index (TPI) is a morphometric index value that reflects the difference in height between a place and a place around it. The values are grouped based on the same value and classified into between classes. Interval class describes the values from lowest to highest [15].

TPI classification of the Dawuan Sub-watershed is classified into 5 classes, <= -3.39, -3.39 - (-0.06), (-0.06) - 0.80, 0.80 - 1, and > 1. Positive TPI values indicate that the elevation of an area is higher than that of the region. others around it (upper slope). Meanwhile, a negative TPI value indicates that the area has a lower elevation (lower slope), and if it is close to zero it indicates that the area tends to be flat (valley).

The Stream Power Index (SPI) is a representation of the strength of water flow in influencing erosion [7]. The SPI value indicates erosion strength, the higher the SPI value the higher the slope erosion strength [18,20]. The SPI value is associated with water mass conservation, gravity, basin hydrology, hydraulic geometry, shear stress, climate, basin, flood interval, fracture distance in bedrock, and bedrock erodibility. SPI for the Dawuan River Basin is classified into 5 (classes), namely: <= 50, 50 - 100, 100 - 200, 200 - 400, and > 400. The classification is based on the results of DEM (Digital Elevation Model) processing.

Profile Curvature is a morphometric factor that affects surface runoff. The curvature value greatly affects slope stability [7,15]. The curvature profile of the Dawuan Sub-watershed is classified into three classes, namely: -(0.07) - (-0.001), (-0.001) - 0.0, and 0.0 - 0.13. Profile curvature is one of the parameters for assessing the susceptibility of flash flood which shows the slope gradient values such as convex, flat, and concave.

Volcano and geological condition of Sub Watershed of Dawuan was created because of the influence of Welirang Volcano during geology quarter period. The volcanic activity affects the
condition of the existing main constituent stones. The rock found on the slopes of Welirang Volcanoes are igneous, pyroclastic and clastic volcanic rocks that are Quaternary in age. The geological units found in the Dawuan Sub-watershed include, Qvwl, Qvlw3, Qvok, Qvw, Qvlp1, Qvlp2, Qvlp3, Qvlp4, Qvw1, Qvf3, Qvf4, Qvf8, and Qvf.

Land use is one of the factors that cause flash flood that affect the surface flow process and the infiltration process [15]. Land use describes the condition of land cover that affect the infiltration process [7,17]. Land use in the Dawuan sub-watershed area is grouped into tigs, namely building area, agricultural, and forest. Most of the land use is dominated by agricultural areas.

2.3 Data Normalization

This study used data normalization using the ratio frequency (FR). FR is a bivariate statistical model that can be understood and used easily. The weight assessment in FR applies the ratio between the number of pixels with the phenomenon of torrentiality in the factor class and the total number of pixels with phenomenon of torrentiality in the research area [7,15].

\[ FR = \frac{\sum_{i=1}^{m} Np(LXi)}{\sum_{j=1}^{n} Np(LXj)} \]

Explanation: FR- Class I frequency ratio of j, Np (LXi) - Number of pixels, Np (Xj) - Number of pixels in factor variable Xj, m- Number of classes in factor variable Xi, and n- Number of factors in regional research.

The weight estimation of the FR is the result of a combination of the flash flood condition factors in the form of a raster with a raster that is suitable for the area affected by torrentiality. The results of these combinations are calculated and categorized in each class according to the factors [15]. Then the data normalization was carried out by applying the min-max method which produced values in the range 0.01-0.99 [6,7,21].

\[ Value_{new} = \frac{value_{real} - min(v)}{max(v) - min(v)} \times (upper - lower) + lower \]

Explanation: Value_{new} – value of normalization result, value- real value of the class, upper-constant0.99, and lower-constant 0.01.

2.4 Flash Flood Modeling

Logistic Regretion Analysis (LR) is a multivariate statistical mode used to analyze the dependent relationship between binary variables (Y) and several independent variables (Xi) describing the spatial distribution of the independent variables. In logistic regression the dependent variable is binary so that the class consists of 2, namely the value of 1 or the existence of the phenomenon being analyzed and the value of 0 where the phenomenon is not analyzed. Logistic regression models predict flash flood vulnerability from the relationship between phenomena and factors in each class [15,22]. In order to obtain the LR value for the probability of flash flood using the following equation:

\[ P = \frac{1}{1+e^{-z}} = \frac{e}{1+e^{z}} \]

Explanation: P- Probability of flash flood, e- exponential, z- a combination of the independent variables obtained from the following equation

\[ Z = b_0 + b_1X_1 + b_2X_2 + \cdots + b_nX_n \]

Explanation: b_0 – intercept of the model, b_n (n= 0, 1, 2, …, n)- steep coefficient of LR model, and X_n(n= 0, 1,2,…, n)- independent variable [15].

2.5 Model Validation
The model validation used in this research is Receiver Operating Characteristics (ROC) and Area Under Curve (AUC). This validation is used to assess the level of validation of the modeling results depicted in the ROC curve with the AUC value. The purpose of this validation is to look at the capacity to predict the occurrence of an event. The AUC value indicates the effectiveness of the model, if the value is close to 1 then the model will be better [15]. The equation for getting the AUC value is as follows:

$$AUC = \frac{\sum TP + \sum TN}{(P+N)}$$  \hspace{1cm} (5)

Explanation: $TP$- True Positive and $TN$-True Negative are the number of pixels that are classified correctly, $P$- the number of pixels for which there are torrential phenomena, and $N$- the number of pixels for which there are no terrestrial phenomena.

The statistical index of this study uses sensitivity and accuracy. In order to get sensitivity and accuracy values, you can use the following equation:

$$Sensitivity = \frac{TP}{TP+FN}$$  \hspace{1cm} (6)

$$Specificity = \frac{TN}{FP+TN}$$  \hspace{1cm} (7)

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN}$$  \hspace{1cm} (8)

Explanation: $FP$-False Positive and $FN$-False Negative is the number of misclassified pixels.

3. Results and Discussion

3.1 Analysis of the Flash Flood Frequency Studies in Each Factor Class

Analysis of the flash flood frequency events from the summary of the calculation of the FR value to the Normalized Value (Nv) which is shown in each class factor. The number of pixels for flash flood events in the study area ($L_S$) is 40300, while the number of pixels for the entire research area ($A_S$) is 397338.

Table 1. Correlation Between Landslide and Its Triggering Factors Using FR

| Conditioning Factors | Classes | $L_{ci}$ | $A_{ci}$ | FR     | Nv     |
|----------------------|---------|----------|----------|--------|--------|
| **Elevation**        | 50 – 670| 34780    | 231876   | 1.478866 | 1.00   |
|                      | 670 – 1290 | 5212     | 115227   | 0.44597 | 0.30   |
|                      | 1290 – 1910 | 308      | 35954    | 0.084461 | 0.06   |
|                      | 1910 – 2520 | 0       | 10843    | 0       | 0.00   |
|                      | 2520 – 3150 | 0       | 3438     | 0       | 0.00   |
| **Slope**            | <2      | 924      | 10434    | 0.873125 | 0.71   |
|                      | 2-4     | 4921     | 57602    | 0.842308 | 0.69   |
|                      | 4-8     | 10370    | 108129   | 0.945566 | 0.77   |
|                      | 8-16    | 14886    | 119711   | 1.226024 | 1.00   |
|                      | 16-35   | 9106     | 94576    | 0.949296 | 0.77   |
|                      | 35-55   | 93       | 6809     | 0.134665 | 0.11   |
|                      | >55     | 0        | 77       | 0       | 0.00   |
| **Aspect**           | Flat    | 2982     | 63078    | 0.466106 | 0.00   |
|                      | North   | 1880     | 36963    | 0.501471 | 0.02   |
|                      | Northeast | 7005    | 67954    | 1.016361 | 0.26   |
|                      | East    | 6699     | 30227    | 2.185093 | 0.82   |
|                      | Southeast | 2063   | 7902     | 2.574052 | 1.00   |
|                      | South   | 816      | 4243     | 1.896148 | 0.68   |
|                      | Southwest | 4356   | 22179    | 1.936426 | 0.70   |
|                      | West    | 7619     | 65133    | 1.153326 | 0.33   |
|                      | Northwest | 6880  | 99659    | 0.680655 | 0.10   |
| Topographic Wetness Index | < 6 | 14999 | 189684 | 0.779627 | 0.00 |
|---------------------------|-----|-------|--------|-----------|-----|
| 6 – 10                    | 13280 | 164981 | 0.793632 | 0.00 |
| 10 – 12                   | 1823 | 20609 | 0.872137 | 0.01 |
| 12 – 16                   | 5171 | 16026 | 3.181299 | 0.32 |
| > 16                      | 5027 | 6038 | 8.208633 | 1.00 |
| Topographic Position Index | < -3.39 | 136 | 583 | 2.299987 | 1.00 |
| (-3.39) - (-0.06)         | 31182 | 174359 | 1.763253 | 0.74 |
| (-0.06) - 0.80            | 7578 | 168030 | 0.444655 | 0.10 |
| 0.80 – 1                  | 517 | 16447 | 0.309927 | 0.04 |
| > 1                       | 887 | 37919 | 0.230633 | 0.00 |
| Stream Power Index        | < 50 | 19593 | 137704 | 1.402844 | 1.00 |
| 50 – 200                  | 7734 | 72932 | 1.045541 | 0.72 |
| 100 – 200                 | 8547 | 69673 | 1.209495 | 0.85 |
| 200 – 400                 | 3566 | 54755 | 0.642115 | 0.40 |
| > 400                     | 860 | 62274 | 0.136159 | 0.00 |
| Profile Curvature         | -0.07 - (-0.001) | 28498 | 191229 | 1.469318 | 0.07 |
| (-0.001) - 0.0            | 11433 | 5351 | 21.06591 | 1.00 |
| 0.0 - 0.13                | 369 | 200758 | 0.018122 | 0.00 |
| Distance to Drainage      | < 10 | 3327 | 4089 | 8.022149 | 1.00 |
| 10 – 20                   | 9084 | 11411 | 7.848894 | 0.98 |
| 20 – 30                   | 6448 | 8088 | 7.860297 | 0.98 |
| 30 – 40                   | 5895 | 8025 | 7.242589 | 0.90 |
| > 400                     | 15546 | 365725 | 0.419101 | 0.00 |
| Rainfall                  | < 2000 | 8423 | 28281 | 2.93648 | 1.00 |
| 2000 – 2300               | 14827 | 107511 | 1.359739 | 0.32 |
| 2300 – 2500               | 14048 | 220373 | 0.628509 | 0.00 |
| 2700 - 2900               | 3002 | 41173 | 0.718875 | 0.04 |
| Geological Unit           | Qvwl | 1568 | 8016 | 1.928606 | 0.77 |
| Qvlw3                     | 3290 | 31014 | 1.045907 | 0.42 |
| Qvok                      | 0 | 28159 | 0 | 0.00 |
| Qvw                       | 3097 | 92055 | 0.331703 | 0.13 |
| Qvlp1                     | 1008 | 14827 | 0.670289 | 0.27 |
| Qvlp2                     | 298 | 17515 | 0.167749 | 0.07 |
| Qvlp3                     | 1772 | 15149 | 1.15328 | 0.46 |
| Qvlp4                     | 2273 | 11405 | 1.964985 | 0.78 |
| Qvw1                      | 5881 | 61530 | 0.942365 | 0.38 |
| Qvf3                      | 7944 | 43925 | 1.783128 | 0.71 |
| Qvf4                      | 5257 | 21408 | 2.421123 | 0.97 |
| Qvf5                      | 5104 | 20099 | 2.503752 | 1.00 |
| Qvf                        | 2808 | 32236 | 0.858838 | 0.34 |
| Land Use                   | Building Area | 936 | 32330 | 0.285447 | 0.00 |
| Agricultural              | 28006 | 216443 | 1.275741 | 1.00 |
| Forest                    | 11358 | 148565 | 0.753773 | 0.47 |

Based on the normalization value (Nv) in Table 1 shows various number between 0 – 1. This number indicates the magnitude of the influence of each factor on the level of flash flood
susceptibility. The closer to the number 0, the lower the effect and the closer to the number 1, the bigger the effect.

3.2 Model Construction of LR-FR

The combination of the LR-FR model using equation 3 (probability) to produce a probability value for flash flood vulnerability. The use of a soft computing approach using the python programming language in the construction process of the LR-FR model. The landslide probability value is obtained by calculating the $Z$ value or weight of the independent variable combination.

The results of LR modeling are obtained from the weight value of each factor ($b_n$), this value indicates the magnitude of the influence of the level of flash flood susceptibility. The higher the $b_n$, value the higher the effect on the level of flash flood susceptibility and the lower the $b_n$ value, the lower the effect on the level of flash flood susceptibility. Distance to Drainage is the most influencing factor on the level of flash flood susceptibility with a value ($b_n$) of 9.3338 and SPI is the least influencing factor for flash flood susceptibility with a value ($b_n$) -2.125.

$$Z = -8.6445 + 0.522[TWI] + 1.0306[TPI] - 4.3773[SPI] + 4.5014[Slope] + 1.9826[Rainfall] - 0.5798[Profile Curvature] - 0.7608[Land Use] + 4.9497[Geology] + 4.8713[Elevation] + 9.3338[Distance to Drainage] + 3.0003[Aspect]$$

3.3 Model Validation

Model validation was held after obtaining the value of flash flood susceptibility using a combination of the machine learning technique model LR-FR (Logistic Regression - Frequency Ratio). Model validation is done to test the quality of the prediction results. The accuracy of the prediction model is determined by the AUC value which ranges from 0.5 to 1.0 where the AUC value approaches 1.0, the accuracy of the prediction model for the probability of flash flood events will be better. If the AUC value is close to 0.5, the accuracy of the prediction model for the probability of flash flood events will get worse.

The accuracy value of the validation of the flash flood susceptibility model is 0.93, which means that the AUC value shows the accuracy of the prediction model for the probability of flash flood events is very good. The ROC curve is obtained from the process of plotting the True Positive Rate value and the False Positive Rate value.

3.4 Analysis of Flash Flood Susceptibility Map

The classification results of flash flood susceptibility consist of 5 classes, namely very low, low, medium, high and very high Figure 9. Based on the distribution of very low and low susceptibility classes, it is in an area with high elevation covering the upper slopes of Mount Welirang.

Based on the normalization values graph in Figure 7, it can be seen that the characteristics of areas with very low to low levels of flash flood susceptibility have characteristics: slope 35-55 °; aspect flat, north, northwest and northeast; TWI <6-12; TPI -0.06-> 1; SPI> 400; profile curvature -0.07-0.001 and 0.0-0.13; distance to drainage> 40; Rainfall 2300-2900; geological units Qvok, Qvw, Qvlp1, Qvlp2; and land use in the building area. Meanwhile, areas with very high hazard class are in areas with slope characteristics <8-35 °; aspect east and southwest; TWI> 16; TPI <(- 3.39) - (- 0.06); SPI <50-200; profile curvature (-0.001) -0.0; distance to drainage <10-40; rainfall <2000; geological units Qvw1, Qvw3, Qvlp3, Qvlp4, Qvw1, Qvf3, Qvf4 and Qvf8; and land use on agricultural. Flash flood susceptibility map can seen in figure 3.
Figure 3. Flash Flood Susceptibility Map

4. Conclusions
The assessment results of the level of flash flood susceptibility in the research area using machine learning techniques show that the quality is very good, as indicated by the value of AUC = 0.93. The flash flood susceptibility class is divided into five: very low, low, medium, high, and very high. The very high hazard class is in the area with the following characteristics: slope <8-35 °; aspect east and southwest; TWI > 16; TPI <(-3.39) - (-0.06); SPI <50-200; profile curvature (-0.001) -0.0; distance to drainage <10-40; rainfall <2000; geological units Qvwl, Qvlw3, Qvlp3, Qvlp4, Qvwl, Qvf3, Qvf4 and Qvf8; and land use on agricultural.

Acknowledgments
Special thanks to Lembaga Penelian dan Pengabdian (LP2M) Universitas Negeri Malang for funding this research through PNBP UM funding scheme, thus the research, writing and publishing process of this paper can be well-completed.

References
[1] Kim, B. S., & Kim, H. S. 2014 Evaluation of flash flood severity in Korea using the modified flash flood index (MFFI) Journal of Flood Risk Management, 7 344–356
[2] Erwin, R., Azmeri, Ismail, N. 2018 Kajian Kerentanan Bencana Banjir Bandang di Gampong Beureuut Kecamatan Seulimun Kabupaten Aceh Besar Jurnal Teknik Sipil, 1 961-970
[3] Utama, L., & Naumar, A. 2015 Kajian Kerentanan Kawasan Berpotensi Banjir Bandang dan Mitigasi Bencana pada Daerah Aliran Sungai (DAS) Batang Kuranji Kota Padang Jurnal Rekayasa Sipil, 9 21–28
[4] Barasa, B. N., & Perera, E. D. P. 2018 Analysis of land use change impacts on flash flood occurrences in the Sosiani River basin Kenya International Journal of River Basin Management, 16 179–188
[5] Alipour, A., Ahmadalipour, A., & Moradkhani, H. 2020 Assessing flash flood hazard and damages in the southeast United States Journal of Flood Risk Management, 13
[6] Bui, D. T., Tsangaratos, P., Ngo, P. T. T., Pham, T. D., & Pham, B. T. (2019) Flash flood susceptibility modeling using an optimized fuzzy rulebased feature selection technique and treebased ensemble methods Science of the Total Environment, 668 1038–1054
[7] Cao, C., Xu, P., Wang, Y., Chen, J., Zheng, L., & Niu, C. 2016 Flash flood hazard susceptibility mapping using frequency ratio and statistical index methods in coalmine subsidence areas Sustainability (Switzerland), 8
[8] Adi, S. (2013). Karakteristik Bencana Banjir Bandang di Indonesia Jurnal Sains dan Teknologi Indonesia, 15 42–51
[9] Costache, R. 2019 Flash-flood Potential Index mapping using weights of evidence, decision Trees models and their novel hybrid integration Stochastic Environmental Research and Risk Assessment, 33 1375–1402
[10] Pradhan, A. M. S., & Kim, Y. T. 2016 Evaluation of a combined spatial multi-criteria evaluation model and deterministic model for landslide susceptibility mapping Catena, 140 125–139
[11] Van Zuidam R A 1973 Terrain Analysis and Classification Using Aerial Photographs: A Geomorphological Approach (Enschede: International Institute for Aerial Survey and Earth Sciences (ITC)).