Flash Flood Susceptibility Mapping at Andungbiru Watershed, East Java Using AHP-Information Weighted Method

Devi Ratna Handini, Entin Hidayah* (D), Gusfan Halik

Master Program in Civil Engineering, Department of Civil Engineering, Faculty of Engineering, University of Jember, Jl. Kalimantan No. 37 Jember, 68121, Indonesia

Abstract

Flash floods are among the most frequent natural disasters caused by heavy rain associated with a severe thunderstorm, which leads to social and economic losses in infrastructure and agriculture. Therefore, this research aims to map flash flood potential susceptibility (FFPS) in the Pekalen watershed, using Geographic Information System (GIS) technology and statistical analysis to reduce the risk of flooding. The opinion and experience of an expert on the weight assessment method were carried out using the Analytical Hierarchy Process (AHP). Furthermore, the probability statistical methods and GIS were used in flash flood areas in the Pekalen watershed in Andungbiru, Probolinggo village. This study was carried out using geomorphological factors, namely elevation, slope, stream power index, and topographic wetness index, with a resolution of 30 m. Thematic map scale of the land use, river density, distance to the river, rainfall, and geology is in the ratio of is in a ratio of 1:25,000. Imagery processing was carried out using Landsat 8 30 m x 30 m resolution imagery, such as the Normalized Difference Vegetation Index. The result showed that the model map of FFPS obtained low 8%, low 23%, moderate 27%, moderate to high 26%, high 13%, and very high 2% index values. The next stage of modeling analysis led to validation using statistic receiver operating Characteristic Curve (ROC) of area Under Curve (AUC) with a value of 90.15. In conclusion, the factors that significantly trigger flash floods are distance to the river, land use, and slope.

Keywords: AHP-weighted; Information content; FFSP; GIS; Geomorphology

1. Introduction

Flash floods are one of the natural disasters affecting human life, leading to death and natural and artificial environmental damage (Guzzetti et al., 2005; Penning-Rowsell et al., 2005; Salvati et al., 2010). According to Du et al. (2013), the process of measuring runoff areas and damage caused by flash floods is challenging. This is because flash flood disasters generally have varying intervals and durations (Tehrany et al., 2015).

*Corresponding author. Email address : entin.teknik@unej.ac.id (Entin Hidayah) Several research methods use AHP as a disaster modeling map extension. AHP is also used to calculate weight assessment based on the opinion and experience of experts (Bui et al., 2019; Elkhrachy, 2015; He et al., 2019). The research on potential mapping disasters using opinions and combination statistics produced good validation values, namely AHP-Weighted Information content AUC 0.893, AHP method 0.821, and information Content 0.842 (He et al., 2019). This study used the AHP-weighted information content method, which combines the expert experience method and frequency distribution statistics. The model results are validated using statistics, while the area below the curve (AUC) is used to calculate the accuracy of the flash flood model. Mapping potential flash flood-prone areas using the AHP-Weighted Information Content method is essential for sustainable planning (Costache et al., 2020), protecting life, and maintaining a river condition environment (Zhang et al., 2015).

AHP is a method used for various kinds of disasters, such as landslides, flash floods, etc. This is an interesting research due to its complex nature, making it possible to carry out a spatial approach during the flash flood modeling process (He et al., 2019). There has never been a mapping of the potential for flash floods in the Pekalen watershed, therefore, it is imperative to use the AHP method combined with information content weights to determine the factors that influence the accuracy of the modeling results.

Vulnerability is needed to determine the appropriate methods and analysis needed to map a potential flash flood (Cloke & Pappenberger, 2009). Flash flood modeling utilizes GIS mapping and remote sensing methods (Pradhan & Shafie, 2009). The purpose of this study is to apply the AHP-weighted Information Content method at in the Pekalen watershed using a geographic information system (GIS). This method was used to determine the distribution of potential flash flood events based on factors that have been calculated from their various weights using expert knowledge and mathematical analysis.

2. Methods

2.1 Study Area

This research was carried out in Pekalen Watershed, Probolinggo, East Java, with coordinates 1130.23 E-1130.35 E and 70.55'0 - 80.2'3. Figure 1 shows the flash flood event area is in the village, with the area of the Upper Pekalen watershed covering 8,530 hectares. The flash flood event occurred at an elevation of 608 meters above sea level.



Figure 1. The study area

2.2 Research Design

The predetermined parameters used to determine flash floods are Elevation, Slope, SPI, TWI, Land use, River Density, Distance to River, NDVI, short-term heavy rain, and geology. These parameters are essential because each region has different regional morphological characteristics (Elkhrachy, 2015). Data were collected from the inventory of flash flood events obtained from the survey at the site. The results consist of 70% training data (63 ha) and 30% validation (27 ha) (Kia et al., 2012).

The selection of factors mapping the potential of flash floods is essential in creating an insecurity map to determine the Elevation, Slope, TWI (Costache et al., 2020), SPI, River Density, rainfall, NDVI (Bui et al., 2019), Land use (Costache et al., 2020), Geology (Cao et al., 2016), and distance from the river (Popa et al., 2019). Table 1 is a general description of data source and resolution information.

Flash Flood Factor	Source	Resolution
Elevation	www.usgs.gov (USGS, 2020)	30 X 30 m
Slope	Derivated from SRTM (NASA, 2020)	30 X 30 m
SPI	Derivated from SRTM (NASA, 2020)	30 X 30 m
TWI	Derivated from SRTM (NASA, 2020)	30 X 30 m
Land use	tanahair.indonesia.go.id (BIG, 1999)	1:25,000
River Density	Derivated from layer of river	30 X 30 m
Distance To River	Derivated from layer of river	30 X 30 m
NDVI	Landsat 8 (USGS, 2020)	30 X 30 m
Rainfall	Derivated from layer of rainfall	30 X 30 m
Geology	tanahair.indonesia.go.id (BIG, 1999)	1:25,000

Table 1. Factors used in flash flood mapping modeling

Furthermore, this study comprises geomorphological and hydrological factors, with Figures 2-11 showing the classification of each factor. The spatial data types are grid and vector layers.



Figure 6. River density map

Figure 7. Distance to river map



Figure 10. Geology map

Figure 11. Land use map

- The slope (degree) (Figure 2) factor is derivated from the elevation data and marbles, a. which significantly influence the process of flow (Khosravi et al., 2016).
- Elevation (meter) (Figure 3) data is obtained from NASA imagery with SRTM radar b. resolution of 30 x 30 (m). In general, flash floods usually occur in areas with plateaus.
- c. Topographic Wetness Index (TWI) and Stream Power Index (SPI) imagery (Figures 4 and 5) are derived from DEM factors considered as variables that describe the hydrological arrangement of an area. In general, a high TWI value indicates the presence of prone areas with wet topography and potential to surface runoff. Similarly, SPI is a topographic prediction used to model the strength of a water flow. The topographic wetness index (TWI) is a spatial concept that models the accumulation of flow from each basin and the ability of water to move due to its slope and gravity. SPI and TWI modeling is derived from DEM and processed using tools in SAGA GIS, along with similarities from TWI (Eq.1) and SPI (Eq. 2) modeling (Moore et al., 1991).

Devi Ratna Handini et al. / Geosfera Indonesia 6 (2), 2021, 127-142

$$TWI = Ln\left(\frac{As}{\beta}\right) = Ln\left(\frac{As}{\beta}\right)$$
(1)

$$SPI = Asx \tan \beta = Asx \tan \beta \tag{2}$$

where the As and β denote specific water catchment areas (m²/m) and slope angles in degrees. TWI shows the amount of flow accumulated at any point in the catchment area and the tendency of water to descend to the slope using gravitational force (Moore et al., 1991).

- An NDVI (ratio) is a measure that describes the vegetation characteristics of an area, factors that affect the surface flow, and the infiltration capabilities of a site (Figure 8). According to Tehrany et al. (2013), areas with less dense vegetation are considered more prone to flooding.
- e. Distance to the river (meter) is a factor that is lower from the thematic map layer of the river (Figure 7). This is a factor that shows the distribution of flash flood events to the position of the river (Popa et al., 2019).
- f. The River density (km/m) that flows in watersheds influences the level and intensity of flash floods (Figure 6). The river network and the surrounding areas are highly vulnerable to the manifestation of flash flood events (Fernández & Lutz, 2010).
- g. The geology map (Figure 10) obtained from Rupa Bumi Indonesia (RBI) is one of the factors contributing to the response to water flow when there is high rain (Cao et al., 2016).
- h. Rainfall (mm) is also considered an essential variable in the assessment of flash flood vulnerability (Figure 9). Therefore, precipitation is used to produce a map of rainfall distribution using the Inverse Distance Weighted (IDW) method (Bartier & Keller, 1996).



Figure 12. Flow chart of the methodology

2.3 AHP-Weighted Information Content Method

AHP-weighted information content method is a technique that combines expert monitoring and statistical calculation. The calculated value is entered into the class of each factor, while the weight is obtained from used distribution frequency formulas implemented into GIS. Eq. 3 is a formula for determining each class.

$$I(X_{i}, H) = ln \frac{N_{i}/N}{S_{i}/S} I(X_{i}, H) = ln \frac{N_{i}/N}{S_{i}/S}$$
(3)

where S is the area of study, N is the total area of flash floods, Si is the number of disaster-causing factors, and Ni is the number of flash floods distributed in the causative factors of the disaster.

AHP weight assessment is calculated based on the opinion of experts, while the probability statistic model is used to determine the weight of each factor class. FFSP is calculated using the Eq. 4 (He et al., 2019).

$$FF = \beta 1G1 + \beta 2G2 + \beta 3G3 + \beta 4G4 + \beta 5G5 + \beta 6G6 + \beta 7G7 + \beta 8G8 + \beta 9G9 + \beta 10G10FF = \beta 1G1 + \beta 2G2 + \beta 3G3 + \beta 4G4 + \beta 5G5 + \beta 6G6 + \beta 7G7 + \beta 8G8 + \beta 9G9 + \beta 10G10.$$
(4)

Eq. 2 is the sum of the weight of each class for each factor. This summation produces a flash flood potential layer next in the process of re-class in GIS to determine the index of potential flash floods (He et al., 2019).

AHP-weighted information content method is used to develop the AHP method by combining it with the statistic probability of frequency distribution. The assessment process provides constant value on each factor and frequency distribution assessment to give weight to each class. Furthermore, the calculated value is converted into a GIS layer, then to the numeric form and spatial data. After each factor is converted, their sum is weighed to create a potential flash flood map before proceeding to the reclassification stage to create an index class. According to Khosravi et al. (2016), the selection of index class for the map of FFSP used the trusted natural breaks method because it is most suitable for the classification process of index map insecurity.

Validation is a critical stage in creating and identifying maps of areas vulnerable to flash floods and determining the model's quality. Therefore, the modeling results are unusable, assuming they do not have validation values (Biswajeet Pradhan & Lee, 2010). ROC is a method widely used to evaluate the accuracy level of model results. Tehrany et al. (2013) stated that using this method is better because the results are generally complete with good visuals. Many studies use this method as an option for their validation process (5).

Furthermore, in the potential mapping of flash flood disasters, the prediction model results need to be evaluated in the modeling process.

This study uses the statistical analysis method to quantitatively calculate the relationship between flash floods and causal factors based on a spatial modeling approach. It also focuses on the probability of frequency distribution with the approach of geodynamic environmental data. The method combines probability data with the opinions of experts that need to be verified.

3. Results and Discussion

Calculating AHP results using weighted Information content determined the weight value of each class factor. AHP is a method that uses expert opinion, such as individuals that are professionals in the field of disaster and GIS, to determine the input from the weight of each factor. The opinions of these experts include the National Disaster Relief Agency, National Disaster Management Authority, hydrologists, GIS experts, and the Development Planning Agency at Sub-National Level. Figure 13 is the result of an AHP weighted information content calculation.



Figure 13. AHP-Weighted information content calculation results

In Figure 13, the weight calculation shows that the factor with the most significant negative value is the distance to river factor, which is in the second class with a weight value of -5.42. This is followed by the land-use and marble factors with negative values -3.22 and -2.70.

The AHP-Weighted Information Content Calculation results consist of positive and negative values. The positive values indicate that the probability of events is more significant in each class than in the factor class. Algorithm natural (LN) formula produces both negative and positive values. The negative value indicates that the probability of events is less than the value of a factor class. The factor with the highest positive value, namely the land-use in the rice field feature, has a value of 3.69, followed by the rainfall factor in the sixth grade worth 2.73 and the fifth rate at 1.49. In the class of marbles, the positive score is first class with a value of 1.30, which ranges from zero to seven degrees. The case at the research site based on the probability of events and factors used in this study stated that flash floods occurred at a reasonably high elevation of 475-621 m, at a slope of 0-7 and 17-21 degrees.

The incident occurred in rice fields using settlements and fields with river runoff up to 177 m during the 15 mm continuous rainfall. The GIS software is used to determine the overlay of each factor. The natural interval statistical method is used to determine the classification results, consisting of six classes, namely low, low, medium, medium to high, high, and very high.



The following figures (Figure 14 - Figure 23) illustrate the weighting in each class.

In the elevation factor (Figure 14), the probability value of events in the first class with a value of one ranges from 475 - 621 m. Meanwhile, in the slope factor (Figure 15), the probability of the highest occurrence in the first, second, third, and fourth classes are 1.30, 0.36,-0.1 and -2.70.



In the SPI factor (Figure 16), the probability value of events in class one has a value of 0.01 with a range of SPI 0-324 and -0.37 in class two. In the TWI factor (Figure 17), the probability value of its appearance in classes one, two, three, four, five, and six are -1.31, -0.87, -0.01, 0.56, 0.80, and 1.38.



On the land use factor (Figure 18), the value of the probability of forest occurrence is 0.00 for plantations at 1.44. The highest negative value for land use is -3.22 and -199 for settlements, while the highest for rice fields is 3.69 and -1.90 for shrubs. For the river density factor (Figure 19), the value of the probability of occurrence in classes one, two, three, four, five, and six are 0, -1.38, 1.10, -0.87, 0.07, and -0.75.



The distance factor to the river (Figure 20) and the probability value in class one and are 0.95 and -5.42. Meanwhile, in classes three, four, five, and six, it is 0.00. In the NDVI factor (Figure 21), the probability event value in classes one, two, three, four, five and six are 1.27, 0.74, 0.74, -0.23, -0.58, and -0.78.



In the rainfall factor (Figure 22), the probability event value in grades one, two, three, and four have a score of 0.00. Meanwhile, in the fifth and sixth grades, the scores are 1.49 and 2.73, respectively. In the geology factor (Figure 23), the probability event value in Argopuro Volcanic and Lamong volcanic rock are 0.23 and 0.00.

The class weight value of each layer factor needs to be converted into data with a grid resolution of 30 m x 30 m using the GIS. The results of the map of potential flash floods are shown in Figure 24.



Pekalen Sub-Watershed, Probolinggo East Java

Figure 24. Map of potential flash floods

Table 3. Table of percentage and broad proportional values on models in each class

Class	Count	Area (m ²)	Percentage (%)
Very Low	7748	697.32	8
Low	21577	1941.93	23
Moderate	25545	2299.05	27
Moderate to high	24574	2211.66	26
High	11729	1055.61	13
Very High	2268	204.12	2

The calculation obtained a map of the spread of flash flood insecurity with each index class, as shown in Figure 24. The study area in Table 3 shows that the highest percentage value is moderate 27%, followed by moderate to high 26%, Low 23%, High 13%, Very Low 8%, and very high with only 2%. The spatial distribution shows that the research area is fairly vulnerable because moderate to high values dominate 27% and 26%, respectively.

The results of the flash flood encoding using the AHP method can carry out the validation process using the graph below the AUC curve, as shown in Figure 25.



Figure 25. Receiver Operating Characteristic (ROC)-based analysis of the Area Under the Curve (AUC) data graph

The advantage of this research is to test the performance model that uses a combination of statistics and expert statements, with the results of a map and the distribution of potential flood events. This research utilized factors that have been used in previous studies related to flash flood disasters in the Pekalen watershed, such as elevation (Bui et al., 2019; Costache et al., 2020), Slope (Khosravi et al., 2016), TWI (Bui et al., 2019), SPI (Bui et al., 2019), Land use (Costache et al., 2020), distance to the river (Popa et al., 2019), River Density (Bui et al., 2019), NDVI (Bui et al., 2019), Rainfall (Bui et al., 2019), Geology (Cao et al., 2016). Previous studies used various approaches to determine geomorphological factors; therefore, the most frequently used factors for modeling were implemented in this study.

The statistic method for the validation process used the analysis method under the curve (AUC) to obtain an event model value of 90.15%, as shown in Figure 25. This value indicates that the modeling made based on the probability of frequency distribution has proximity between flash flood events models.

In addition, the increase in performance (AUC value) is supported by a selection factor that is adjusted to the type of flash flood events from similar studies, which have an AUC value above 80%. The resulting flood hazard map is 30 m x 30 m resolution at a scale of 1:25,000. This map can be used as a reference in making flood risk models and regional spatial alignment, therefore, they are able to offer risk.

4. Conclusion

In conclusion, the flash flood potential vulnerability map using the AHP-Weighted Information Content method gives a significant result. Furthermore, adjusting the resulting map of each layer produced a resolution of 30 m x 30 m on a scale of 1:25,000. Of the ten factors analyzed using AHP-Weighted Information Content, those that significantly trigger flash floods are distance to the river, land use, and slope with a weight value of four. This shows the area with a high probability of occurring in the rice field features. Meanwhile, factors that have weak influence are TWI, SPI, and land-use factors. In this study, land use has a high value because it is a factor that affects runoff on the surface. The flash flood susceptibility potential map can be used as a reference in flood mitigation, with a resolution scale of 30 m x 30 m. Furthermore, research on modeling, it required information on the previous occurrence to determine the temporal phenomena that occur. The group number of each factor needs to be varied to test the dominance of the triggering factors for a flash flood.

Conflict of Interest

The authors declare no conflict of interest.

References

- Bartier, P. M., & Keller, C. P. (1996). Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). *Computers and Geosciences*, 22(7), 795–799. https://doi.org/10.1016/0098-3004(96)00021-0.
- BIG (1999). *Peta RBI Format shp*. Retrieved from https://tanahair.indonesia.go.id/portal-web.
- Bui, D. T., Tsangaratos, P., Ngo, P. T. T., Pham, T. D., & Pham, B. T. (2019). Flash flood susceptibility modeling using an optimized fuzzy rule based feature selection technique and tree based ensemble methods. *Science of the Total Environment*, 668, 1038–1054. https://doi.org/10.1016/j.scitotenv.2019.02.422.
- 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(9). https://doi.org/10.3390/su8090948.
- Cloke, H. L., & Pappenberger, F. (2009). Ensemble flood forecasting: A review. *Journal of Hydrology*, 375(3–4), 613–626. https://doi.org/10.1016/j.jhydrol.2009.06.005.
- Costache, R., Pham, Q. B., Sharifi, E., Linh, N. T. T., Abba, S. I., Vojtek, M., Vojteková, J., Nhi, P. T. T., & Khoi, D. N. (2020). Flash-flood susceptibility assessment using multicriteria decision making and machine learning supported by remote sensing and GIS techniques. *Remote Sensing*, 12(1). https://doi.org/10.3390/RS12010106.

- Du, J., Fang, J., Xu, W., & Shi, P. (2013). Analysis of dry/wet conditions using the standardized precipitation index and its potential usefulness for drought/flood monitoring in Hunan Province, China. *Stochastic Environmental Research and Risk Assessment*,27(2), 377–387. https://doi.org/10.1007/s00477-012-0589-6.
- Elkhrachy, I. (2015). Flash Flood Hazard Mapping Using Satellite Images and GIS Tools: A case study of Najran City, Kingdom of Saudi Arabia (KSA). *Egyptian Journal of Remote Sensing and Space Science*, 18(2), 261–278. https://doi.org/10.1016/j.ejrs.2015.06.007.
- Fernández, D. S., & Lutz, M. A. (2010). Urban flood hazard zoning in Tucumán Province, Argentina, using GIS and multicriteria decision analysis. *Engineering Geology*, 111(1– 4), 90–98. https://doi.org/10.1016/j.enggeo.2009.12.006.
- Moore, I. D., Grayson, R. B., & Ladson, A. R. (1991). Digital terrain modeling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, *5*(1), 3–30. https://doi.org/10.1002/hyp.3360050103.
- NASA (2020). *Shuttle Radar Topography Mission (SRTM)*. Retrieved from <u>https://cmr.earthdata.nasa.gov/search/concepts/C1000000240-LPDAAC_ECS.html</u>.
- Guzzetti, F., Stark, C. P., & Salvati, P. (2005). Evaluation of flood and landslide risk to the population of Italy. *Environmental Management*, *36*(1), 15–36. https://doi.org/10.1007/s00267-003-0257-1.
- He, H., Hu, D., Sun, Q., Zhu, L., & Liu, Y. (2019). A landslide susceptibility assessment method based on GIS technology and an AHP-weighted information content method: A case study of southern Anhui, China. *ISPRS International Journal of Geo-Information*, 8(6). https://doi.org/10.3390/ijgi8060266.
- Khosravi, K., Nohani, E., Maroufinia, E., & Pourghasemi, H. R. (2016). A GIS-based flood susceptibility assessment and its mapping in Iran: a comparison between frequency ratio and weights-of-evidence bivariate statistical models with multi-criteria decision-making technique. *Natural Hazards*, *83*(2), 947–987. https://doi.org/10.1007/s11069-016-2357-2.
- Kia, M. B., Pirasteh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A., & Moradi, A. (2012). An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. *Environmental Earth Sciences*, 67(1), 251–264. https://doi.org/10.1007/s12665-011-1504-z.
- Penning-Rowsell, E., Floyd, P., Ramsbottom, D., & Surendran, S. (2005). Estimating injury and loss of life in floods: A deterministic framework. *Natural Hazards*, *36*(1–2), 43–64. https://doi.org/10.1007/s11069-004-4538-7.
- Popa, M. C., Peptenatu, D., Draghici, C. C., & Diaconu, D. C. (2019). Flood hazard mapping using the flood and Flash-Flood Potential Index in the Buzau River catchment, Romania. *Water (Switzerland)*, *11*(10). https://doi.org/10.3390/w11102116.
- Pradhan, B, & Shafie, M. (2009). Flood hazard assessment for cloud prone rainy areas in a typical tropical environment. *Disaster Advances, February*.

- Pradhan, Biswajeet, & Lee, S. (2010). Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. *Environmental Earth Sciences*, 60(5), 1037–1054. https://doi.org/10.1007/s12665-009-0245-8.
- Salvati, P., Bianchi, C., Rossi, M., & Guzzetti, F. (2010). Societal landslide and flood risk in Italy. *Natural Hazards and Earth System Science*, 10(3), 465–483. https://doi.org/10.5194/nhess-10-465-2010.
- Tehrany, M. S., Pradhan, B., & Jebur, M. N. (2013). Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *Journal of Hydrology*, 504, 69–79. https://doi.org/10.1016/j.jhydrol.2013.09.034.
- Tehrany, M. S., Pradhan, B., & Jebur, M. N. (2015). Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. *Stochastic Environmental Research and Risk Assessment*, 29(4), 1149–1165. https://doi.org/10.1007/s00477-015-1021-9.
- USGS (2020). Landsat 8 Imagery. Retrieved from https://www.usgs.gov/.
- Zhang, D. wei, Quan, J., Zhang, H. bin, Wang, F., Wang, H., & He, X. yan. (2015). Flash flood hazard mapping: A pilot case study in Xiapu River Basin, China. *Water Science and Engineering*, 8(3), 195–204. https://doi.org/10.1016/j.wse.2015.05.002.