JOURNAL OF APPLIED GEOSPATIAL INFORMATION

Vol 10 No 1 2026

http://jurnal.polibatam.ac.id/index.php/JAGI
ISSN Online: 2579-3608

Spatial Assessment of Flash Flood Susceptibility in a Steep Tropical Watershed: The Banyuputih Case Study, Indonesia

Gandhi Teguh Lesmana[1], Entin Hidayah[1*], Retno Utami Agung Wiyono[1], Fidyasari Kusuma Putri [2], Hilma Wasilah Robbani [1], Jagat Adi Samudra [1]

[1] Department of Civil Engineering, Faculty of Engineering, Universitas Jember, 68121 Jember, Indonesia
[2]Program of Professional Engineering, Faculty of Engineering, Universitas Jember, 68121 Jember, Indonesia
*Corresponding author e-mail: entin.teknik@unej.ac.id

Received: January, 19 2026
Accepted: May 28, 2026
Published: June 11, 2026

Abstract
Flash flood susceptibility in upstream watersheds is influenced not only by rainfall intensity but also by the spatial configuration of physiographic characteristics and land use. The Banyuputih Watershed in East Java has experienced recurrent flash floods, emphasizing the need for spatial assessment to support effective mitigation planning. This study develops a flash flood susceptibility map using a GIS-based multi-criteria approach integrated with logistic regression. The analyzed factors include DEM-derived topographic parameters (elevation, slope, Topographic Position Index (TPI), Topographic Wetness Index (TWI), and plan curvature), Hydrologic Soil Group (HSG), rainfall, river density, land cover, and NDVI. The relative influence of each factor was determined from logistic regression coefficients. The results classify the watershed into five susceptibility levels: very low, low, moderate, high, and very high. High to very high susceptibility zones are spatially limited and mainly concentrated in upstream and parts of midstream areas characterized by flow-convergent topography and less protective land cover. Most of the watershed is dominated by very low to moderate vulnerability, indicating that flash flood potential is spatially localized. The resulting map provides a scientific basis for watershed management, land-use planning, and targeted nature-based mitigation strategies.

Keywords: Flash flood; Vulnerability; Banyuputih watershed; Information geographic system; Logistic Regression.

1. Introduction

Flash flood susceptibility in upstream watersheds is influenced not only by rainfall intensity but also by the spatial configuration of physiographic factors and land use that control hydrological response. In watersheds characterized by steep slopes and short flow paths, even minor land cover changes can significantly increase surface runoff and shorten the time of concentration, thereby intensifying the potential for flash flood occurrence (Chow & Maidment, 1998; Maidment, 1993; Montgomery & Dietrich, 1994). These characteristics make upstream watersheds highly sensitive systems to landscape alteration, particularly in tropical regions where high-intensity rainfall frequently occurs (Bruijnzeel, 2004; Sidle, 2006; Tarolli, 2016).

Flash floods are among the most destructive hydrometeorological hazards due to their sudden onset and extremely limited warning time. In upstream areas, flash floods are typically characterized by rapid rises in water level and discharge, high flow velocity, and severe, destructive impacts on settlements, infrastructure, and human safety (Brázdil et al., 2024; Sadkou et al., 2024). This risk is further exacerbated by land use and land cover changes, such as the conversion of forests into open land, intensive agriculture, or built-up areas, which reduce soil infiltration capacity and increase direct surface runoff.

The Banyuputih Watershed in East Java represents an upstream basin with such physiographic characteristics and has experienced recurrent flash flood events. Steep slopes, dense drainage networks, and increasing land use pressure in the upstream area make this watershed highly susceptible to rapid hydrological responses. Therefore, a spatial understanding of flash flood susceptibility and its controlling factors is essential to support effective and sustainable watershed management and mitigation planning.

Geographic Information System (GIS)-based flash flood susceptibility mapping has been widely applied to integrate multiple controlling factors, including slope, land use/land cover, soil characteristics, drainage density, rainfall intensity, and proximity to river networks (Abeysiriwardana & Wijesekera, 2022). However, many previous studies still rely on linear or semi-subjective weighting approaches, such as simple overlay techniques or the Analytic Hierarchy Process (AHP) (Abeysiriwardana & Wijesekera, 2022; Saaty, 1980). These approaches generally assume linear relationships among parameters and may not fully capture the nonlinear interactions and hydrological complexities occurring in steep upstream watersheds.

Moreover, numerous flash flood susceptibility studies produce zonation maps without quantitatively evaluating the relative contribution of each controlling factor using data-driven methods (Merghadi et al., 2020). This limitation makes it difficult to link susceptibility results to dominant physical mechanisms governing flash flood occurrence (Islam et al., 2025; Kaya & Derin, 2023), thereby constraining their application in more precise and adaptive mitigation planning, particularly in small to medium-sized tropical watersheds.

To address these limitations, data-driven statistical approaches, particularly Logistic Regression, provide advantages in modeling the relationship between environmental predictors and flood occurrence probability quantitatively. This method estimates the relative contribution of each factor through regression coefficients, enabling explicit interpretation of both the direction and magnitude of influence (Hosmer et al., 2013). Logistic Regression is a binary classification technique used to estimate event probability based on environmental controlling factors and has been demonstrated to be effective in landslide and flood susceptibility assessments (Gigović, L., Pourghasemi, H. R., Drobnjak, S., & Bai, S. 2019), This method has proven effective in landslide and flood susceptibility analysis and remains a relevant conventional statistical approach in GIS-based hazard mapping studies. The integration of Logistic Regression with GIS enables spatially accurate and process-transparent flash flood susceptibility mapping, as the relationships among variables can be traced through measurable model parameters.

Based on this background, this study aims to develop a flash flood susceptibility map of the Banyuputih Watershed by integrating a GIS-based multi-parameter approach with Logistic Regression modeling. The study is expected to generate a robust susceptibility map while quantitatively identifying the dominant controlling factors of flash flood occurrence. The results are intended to provide a scientific basis for sustainable watershed management, land-use regulation, and the development of nature-based flash flood mitigation strategies in upstream basins with similar physiographic characteristics.

2. Materials and Methods

2.1 Study Area

This study was conducted in the upper Banyuputih Watershed, located in Situbondo and Bondowoso Regencies, East Java Province, Indonesia. The watershed represents a mountainous upstream catchment within a volcanic landscape influenced by the Ijen volcanic complex and is characterized by a highly responsive hydrological regime typical of steep tropical watersheds (Wahyuni & Sachro, 2024).

The upper Banyuputih Watershed is dominated by hilly to mountainous terrain, with slope gradients ranging from moderate to very steep. Such steep topography significantly accelerates surface runoff, shortens the time of concentration, and enhances susceptibility to flash flood generation, particularly under conditions of vegetation degradation (Suprayogi et al., 2022))

Watershed boundaries were delineated using the Indonesian National Digital Elevation Model (DEMNAS) provided by the Badan Informasi Geospasial (BIG), which also served as the basis for drainage network extraction and slope analysis (BIG, 2018). Land use/land cover conditions were derived from ESRI Land Cover data based on Sentinel-2 imagery, which has been widely applied in hydrological and environmental studies for monitoring land cover dynamics (Soliman et al., 2022; Zachrani et al., 2024).

The selection of the upper Banyuputih Watershed was motivated by the recurrent occurrence of flash floods, substantial damage to settlements and infrastructure, and its representativeness as a volcanic mountainous watershed in Indonesia, making it highly suitable for evaluating land-cover-based flash flood mitigation strategies (Akuatiklestari et al., 2022; Qatrinnada et al., 2024).

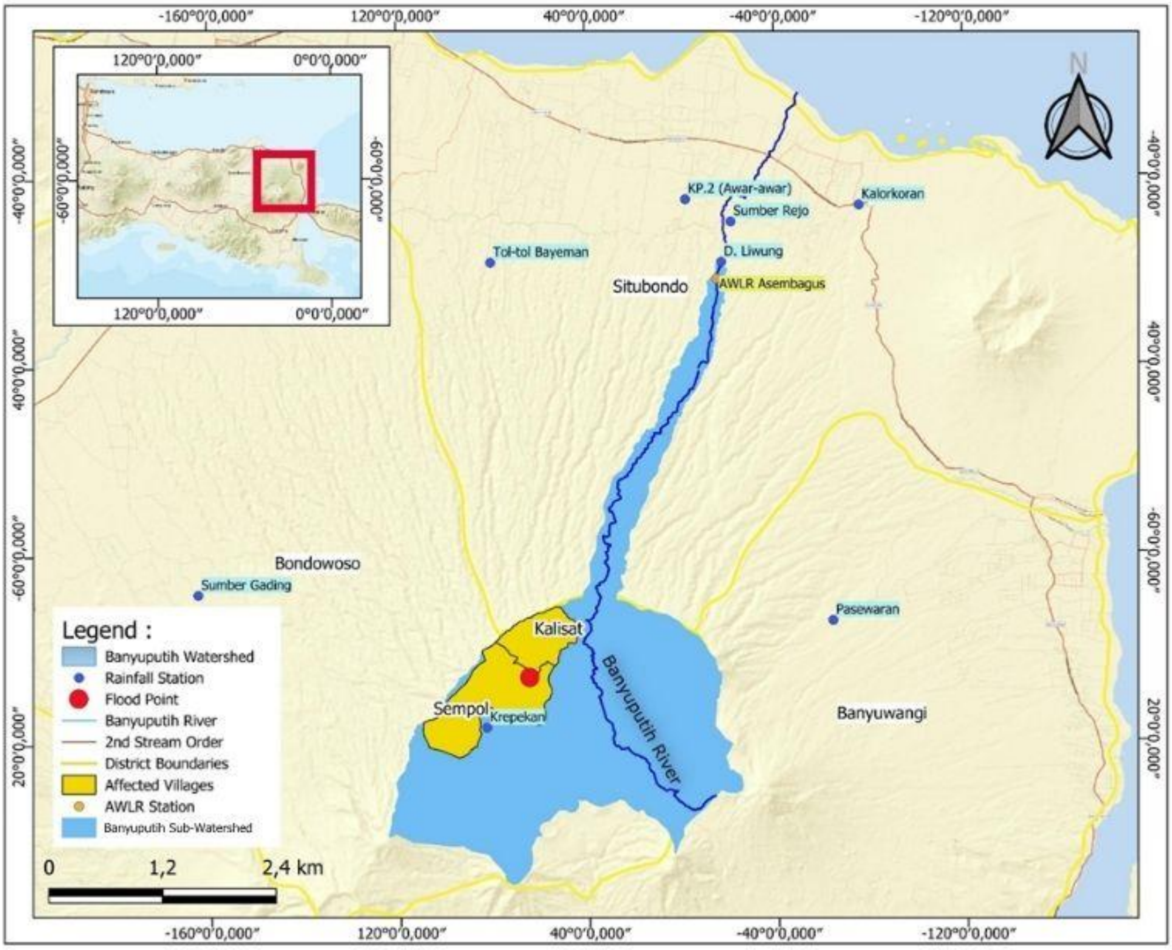

Figure 1 Study area

2.2 Research Steps

This study was conducted through several main stages to map flash flood susceptibility using a Logistic Regression approach. The initial stage involved the preparation of flash flood conditioning factors, consisting of topographic parameters derived from the Digital Elevation Model (DEM), namely elevation, slope, Topographic Position Index (TPI), Topographic Wetness Index (TWI), and plan curvature. Additional factors included Hydrologic Soil Group (HSG) representing soil characteristics, rainfall as the hydrometeorological triggering factor, river density representing drainage network structure, and land surface conditions represented by land cover and NDVI.

All spatial parameters were pre-processed and standardized to ensure consistent spatial resolution and coordinate systems. Flood and non-flood occurrence data were compiled and divided into training and testing datasets. The training dataset was used to construct the Logistic Regression model, which estimates the probability of flash flood occurrence based on the relationship between the conditioning factors and observed flood events. Model performance was evaluated using the testing dataset through the Area Under the Curve (AUC) derived from the Receiver Operating Characteristic (ROC) curve.

The calibrated model was subsequently applied to generate a flash flood susceptibility map, classifying the study area into several susceptibility levels.

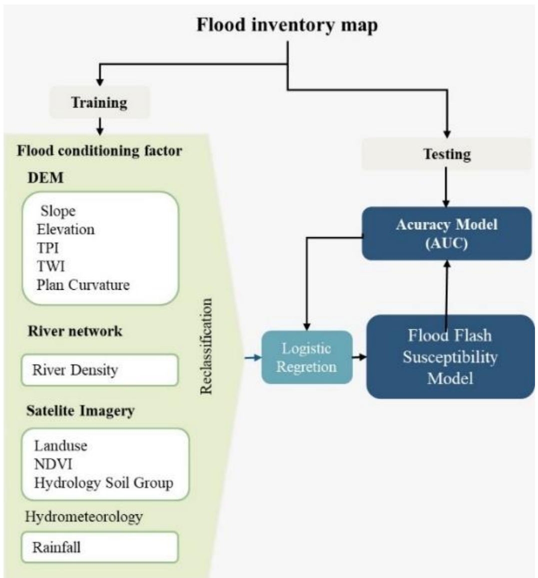

Figure 2 Flowchart of the study

2.3 Data Collection and Processing

The data used in this study consist of spatial and hydrometeorological datasets. Spatial data include the Indonesian National Digital Elevation Model (DEMNAS), which was used to delineate the Banyuputih Watershed and to derive topographic parameters such as elevation, slope, Topographic Position Index (TPI), Topographic Wetness Index (TWI), plan curvature, and river density. These topographic parameters play an important role in controlling surface flow direction, water accumulation, and flash flood potential.

Hydrologic Soil Group (HSG) data were obtained from the NASA HYSOGs250m dataset to represent soil infiltration characteristics and runoff potential. Land cover data were derived from the ESRI Global Land Cover dataset based on Sentinel-2 imagery, providing information on surface conditions influencing infiltration and overland flow. The Normalized Difference Vegetation Index (NDVI) was calculated from Sentinel-2 imagery processed using the Google Earth Engine (GEE) platform to represent vegetation density within the study area.

Hydrometeorological data include rainfall data obtained from the Public Works and Water Resources Agency (PU-SDA), which were used to represent the spatial distribution of rainfall as the primary triggering factor of flash floods. All spatial and hydrometeorological datasets were integrated and analyzed to generate the flash flood susceptibility map of the Banyuputih Watershed.

Table 1 summarizes the data sources and their respective applications in this study.

No	Data Type	Data Source	Function
1	DEM	DEMNAS (Indonesian National DEM)	Used to derive topographic parameters, including slope, elevation, Topographic Position Index (TPI), Topographic Wetness Index (TWI), plan curvature, and river density
2	HSG (Hydrologic Soil Group)	Global Hydrologic Soil Group dataset (250 m resolution)	Used to generate Hydrologic Soil Group (HSG) map representing soil infiltration and runoff potential
3	Land Cover	ESRI Land Cover Sentinel https://livingatlas.arcgis.com/landcoverexplorer/	Used to generate a land cover map indicating surface characteristics influencing runoff and infiltration
4	NDVI	Multispectral satellite imagery (10 m resolution) Sentinel-2	Used to calculate the Normalized Difference Vegetation Index (NDVI) representing vegetation density
5	Rainfall Data (PU-SDA)	Observed rainfall data from the water resources agency	Used to generate a rainfall distribution map as the main triggering a factor of flooding

2.4 Selection of Flash Flood Conditioning Factors

The selection of flash flood conditioning factors represents a crucial step in developing a flood susceptibility model. These factors were chosen based on their strong theoretical and empirical relationships with surface runoff dynamics, flow accumulation, and flood intensity within a watershed. Based on the workflow illustrated in Figure 2 and a review of recent literature, the following factors were incorporated into this study.

Overall, the selected factors encompass topographic, hydrological, hydrometeorological, and land surface characteristics, collectively providing a comprehensive spatial representation of the physical and environmental conditions influencing flash flood occurrence.

2.4.1 Digital Elevation Model (DEM)

The Digital Elevation Model (DEM) is a fundamental dataset widely used in hydrological analysis because it represents the spatial characteristics of terrain (O'Neil et al., 2019). In flash flood susceptibility studies, DEM is commonly processed to derive several topographic indicators that control surface runoff behavior and flow accumulation patterns. Common DEM-derived parameters include slope, Topographic Wetness Index (TWI), Topographic Position Index (TPI), aspect, surface curvature, and convergence index. Topography is considered a primary controlling factor as it governs flow direction, runoff velocity, and water accumulation within a watershed (Tsumita et al., 2025).

Slope is one of the most influential factors in flash flood modeling. Gentle slopes tend to slow down surface flow and promote water accumulation, whereas steep slopes accelerate runoff and may trigger rapid hydrological responses characteristic of flash floods (Manopkawee et al., 2025).

TPI measures the relative topographic position of a location compared to its surrounding terrain; extreme TPI values indicate ridges or depressions that influence localized surface water concentration (Li et al., 2025).

The Topographic Wetness Index (TWI) combines slope and flow accumulation to represent soil moisture potential and water accumulation zones, with higher values typically associated with valleys or areas prone to ponding (Ma'rufah et al., 2024).

Surface curvature is divided into plan curvature and profile curvature: plan curvature controls lateral flow convergence or divergence, while profile curvature affects vertical acceleration or deceleration of flow. Both parameters are widely applied in hydrological analysis to estimate flow direction and surface water dynamics (Pham et al., 2020).

Overall, these DEM-derived parameters are essential because topography fundamentally determines surface runoff patterns and water accumulation zones, forming the core basis of flash flood susceptibility modeling algorithms.

2.4.2 River Density

River density is an important parameter as it represents the structural characteristics of the drainage network and the local hydrological response to rainfall. Drainage density influences flow pathways and runoff conveyance capacity; areas with denser river networks generally have more developed flow channels that affect the spatial distribution and concentration of floodwaters. Due to its strong relationship with runoff dynamics and watershed response, river density is widely incorporated into flood susceptibility modeling studies.(Shrestha et al., 2025).

2.4.3 Satellite Imagery and Remote Sensing

Land use and NDVI represent surface cover conditions that influence infiltration capacity and surface runoff processes, thereby contributing to either the increase or reduction of flood vulnerability. Land use reflects surface permeability to water; areas with dense vegetation generally exhibit higher infiltration rates and lower surface runoff compared to built-up or impermeable areas. The Normalized Difference Vegetation Index (NDVI) is a vegetation index that reflects vegetation density and health conditions. In this study, NDVI is used as a proxy for vegetation cover that affects infiltration capacity, evapotranspiration, and the potential accumulation of surface water. (Pham et al., 2020).

2.4.4 Hydrometeorology

Rainfall is the most critical hydrometeorological input in triggering floods, as its intensity and spatial distribution directly influence the volume of surface runoff generated. High rainfall intensity within a short duration can rapidly exceed infiltration capacity, leading to increased runoff and a higher likelihood of flash flood occurrence. Therefore, rainfall data are widely incorporated in flood susceptibility studies to capture the spatial and temporal variability of hydrological processes. (Ashfaq et al., 2025).

2.5 Flash Flood Susceptibility Mapping Using Logistic Regression

2.5.1 Concept of Logistic Regression in Flash Flood Susceptibility Mapping

Logistic Regression is a statistical method used to model the relationship between a binary dependent variable and multiple independent variables. In flash flood susceptibility mapping, the dependent variable is classified into two categories: flood occurrence (1) and non-flood occurrence (0), while the independent variables consist of conditioning factors such as elevation, slope, TWI, TPI, NDVI, land cover, Hydrologic Soil Group (HSG), river density, and rainfall.

The Logistic Regression model transforms the linear combination of predictor variables into a probability value using the logit function, as expressed in Equation (1).

$$P(Y=1) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}} \quad (1)$$

In this equation, P(Y = 1) represents the probability of flood occurrence, β_0 is the intercept, β_i denotes the regression coefficients, and X_i represents the predictor variables. The coefficient values indicate the direction and magnitude of the influence of each variable on the likelihood of flood occurrence. A positive coefficient increases the probability of flooding, whereas a negative coefficient decreases it. The model output is a probability map that can subsequently be classified into different susceptibility levels (Hosmer et al., 2013).

Due to its probabilistic nature and straightforward interpretability, Logistic Regression is considered an effective method for flash flood susceptibility analysis based on spatial data.

2.5.2 Training and Testing Data Preparation

Flood occurrence data (class 1) and non-flood data (class 0) were divided into training and testing datasets to develop and evaluate the Logistic Regression model. The split was conducted proportionally using a specific ratio (e.g., 70% for training and 30% for testing), allowing the model to learn the relationship between conditioning factors and flood occurrence during the training phase, and subsequently be validated using independent data not involved in model calibration. This procedure aims to reduce overfitting and ensure the model's predictive capability when applied to unseen data. (James et al., 2021).

2.5.3 Selection and Transformation of Predictor Variables

Predictor variables were selected based on a literature review and their relevance to flash flood hydrological processes, including topographic, hydrological, land use, soil, and rainfall factors. All variables were standardized to the same spatial resolution and coordinate system, and their values were extracted at flood and non-flood locations. If necessary, normalization or reclassification was performed to improve model stability and minimize the influence of differing variable scales. (Waiyasusri et al., 2023).

2.5.4 Development of the Logistic Regression Model

The Logistic Regression model was constructed by incorporating the predictor variables into a logit function to estimate the probability of flood occurrence. Parameter estimation was performed using the Maximum Likelihood Estimation (MLE) method to obtain the optimal regression coefficients. The model produces probability values ranging from 0 to 1, representing the likelihood of flash flood occurrence for each pixel. (Hosmer et al., 2013).

2.5.5 Analysis of Coefficients and Direction of Variable Effects

The regression coefficients were analyzed to determine the direction and magnitude of each variable's influence on the probability of flood occurrence. A positive coefficient indicates an increase in flood probability, whereas a negative coefficient indicates a decrease in probability. This analysis was also used to identify the dominant factors contributing to flash flood susceptibility in the study area. (Hosmer et al., 2013).

2.5.6 Mapping the Flash Flood Susceptibility Index

The probability values generated by the Logistic Regression model were converted into a flash flood susceptibility index map. These probability values were subsequently classified into several susceptibility classes, ranging from very low to very high, using an appropriate classification method. The resulting map illustrates the spatial distribution of flash flood susceptibility levels across the study area. (Hosmer et al., 2013).

2.5. Model Validation Using ROC and AUC

The performance of the Logistic Regression model was evaluated using the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) value. The ROC curve illustrates the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR) across various threshold values. TPR and FPR were calculated using Equations (2) and (3), respectively.

$$TPR = \frac{TP}{TP+FN} \quad (2)$$

$$FPR = \frac{FP}{FP+TN} \quad (3)$$

where TP (True Positive) represents the number of correctly predicted flood occurrences, FN (False Negative) denotes flood events that were not detected by the model, FP (False Positive) refers to non-flood locations incorrectly classified as floods, and TN (True Negative) indicates non-flood locations correctly predicted as non-flood.

The AUC value is calculated as the area under the ROC curve, ranging from 0 to 1. Mathematically, the AUC can be expressed as shown in Equation (4).

$$AUC = \int_0^1 TPR(FPR)\ d(FPR) \quad (4)$$

The closer the AUC value is to 1, the better the model's ability to distinguish between flood-prone and non-flood areas(Fawcett, 2006; Hosmer et al., 2013).

3. Result and Discussion

3.1 Characteristics of Factors Conditioning Flash Floods

3.1.1 Slope

Slope is a primary topographic factor controlling surface runoff velocity and hydrological response in the upstream watershed. The slope map of the Banyuputih watershed was derived from DEMNAS and classified into several slope classes ranging from gentle to very steep Figure 3.

Spatially, the Banyuputih watershed is dominated by moderate to steep slopes, particularly in the upstream and middle parts of the watershed. Steep slopes accelerate surface flow, shorten concentration time, and increase the potential for rapid peak discharge formation, which is a key characteristic of flash floods. Meanwhile, gentle slope areas are more common in the downstream regions and along river valleys.

These results indicate that slope plays a crucial role as a conditioning factor for flash floods in the Banyuputih watershed and serves as one of the key parameters in modelling flash flood vulnerability.

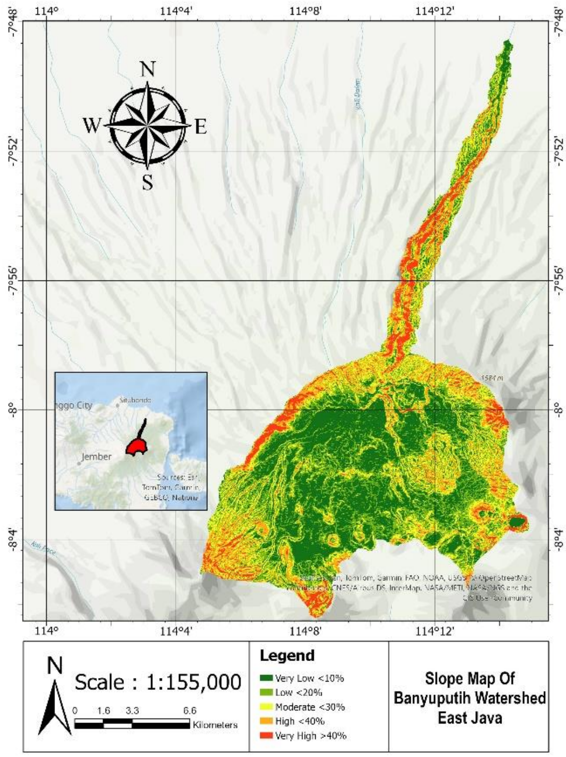

Figure 3 Slope Map Of Banyuputih Watershed

3.1.2 Elevation

Elevation plays a role in controlling flow patterns, runoff accumulation, and the hydrological response of a watershed. The elevation map of the Banyuputih watershed was obtained from DEMNAS, illustrating the variation in terrain height from the downstream to upstream areas of the watershed Figure 4.

In general, high elevations dominate the upstream part of the watershed, characterized by volcanic mountain topography. Areas with high elevation typically have steep slopes and dense drainage networks, contributing to rapid flow response and an increased potential for flash flood occurrence. Conversely, low elevations are located in the downstream regions and river valleys, serving as receiving areas for flows from upstream.

This elevation distribution indicates a strong relationship between high-relief upstream areas and the potential for flash flood formation, making elevation a key factor in modelling flash flood susceptibility in the Banyuputih watershed.

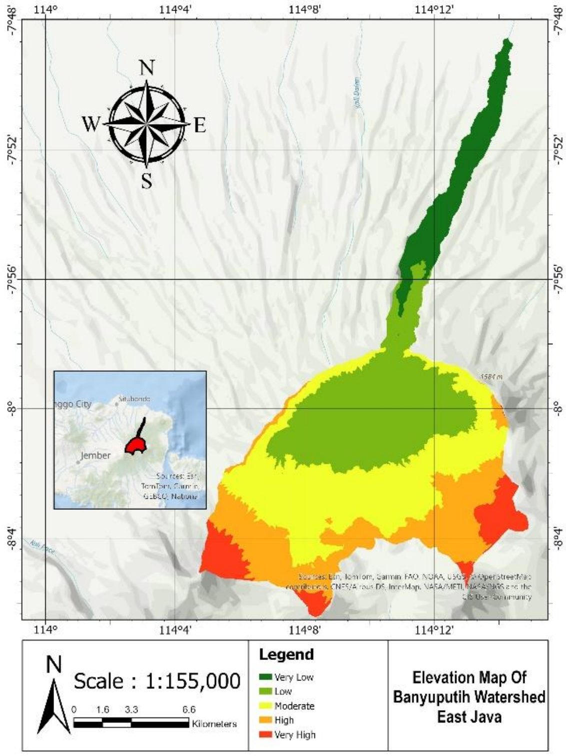

Figure 4 Elevation Map Of Banyuputih Watershed

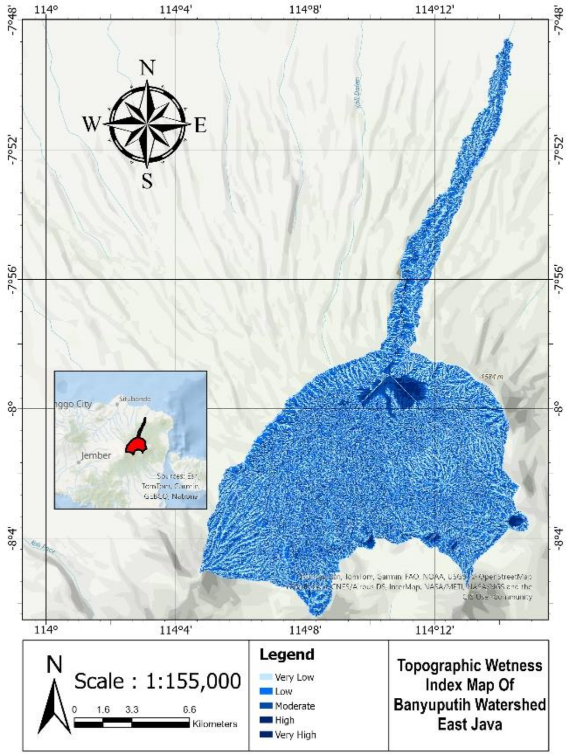

Figure 5 Topographic Wetness Index Map Of Banyuputih Watershed

3.1.3 TWI

The Topographic Wetness Index (TWI) is used to represent the potential for surface water accumulation based on the relationship between slope and flow accumulation. High TWI values indicate areas with a greater tendency for water pooling.

The TWI map of the Banyuputih watershed was derived from DEMNAS and shows that high TWI values are generally distributed along river valleys and depression zones, particularly in the middle to downstream parts of the watershed **Figure 5**. These areas have the potential for flow concentration and increased soil moisture, making them more susceptible to flash floods.

Conversely, low TWI values are mostly found on ridges and steep slopes, which tend to have rapid surface flow without significant water accumulation. This spatial distribution of TWI highlights its role as an important indicator of flow accumulation locations and flash flood susceptibility in the Banyuputih watershed.

3.1.4 TPI

The Topographic Position Index (TPI) is used to indicate the relative position of a location within the surrounding topography, whether it is situated on a ridge, slope, or depression. This parameter plays a role in identifying zones of accumulation and surface flow paths.

The TPI map of the Banyuputih watershed shows that negative TPI values, representing depressions and valleys, are predominantly found along the main river network and its tributaries **Figure 6**. Areas with negative TPI values tend to serve as flow accumulation locations and have a higher potential for flash flood occurrence.

Conversely, positive TPI values, which indicate ridges and relatively higher topographic areas, generally act as runoff source zones. The spatial distribution of TPI confirms that topographic position is an important factor in controlling flow paths and flash flood susceptibility locations in the Banyuputih watershed.

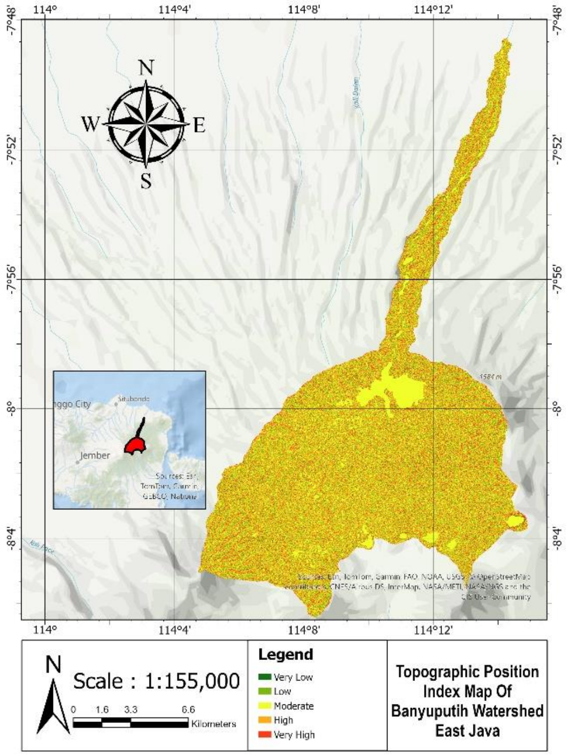

Figure 6 Topographic Position Index Map Of Banyuputih Watershed

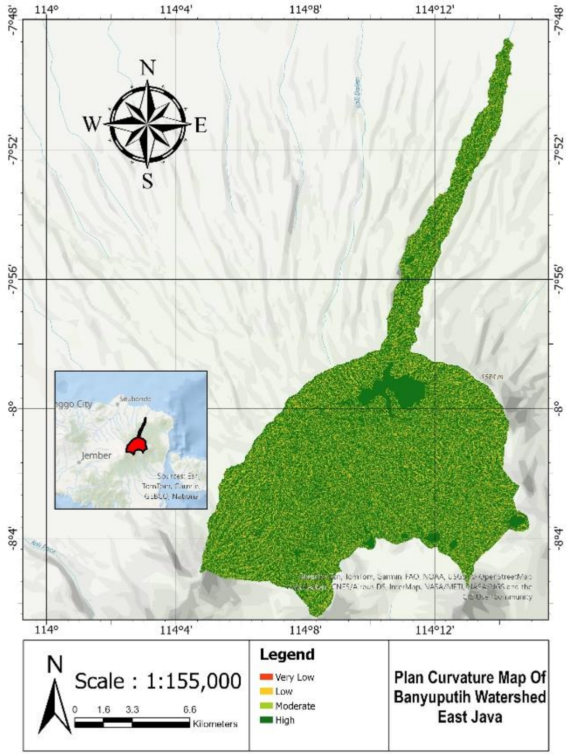

Figure 7 Plan Curvature Map Of Banyuputih Watershed

3.1.5 Plan Culvature

Plan curvature is used to describe the lateral spreading or concentration of flow across the land surface. Negative plan curvature values indicate convergent surface conditions, while positive values indicate divergent surfaces.

The plan curvature map of the Banyuputih watershed shows that areas with negative plan curvature are widely distributed along valleys and river networks Figure 7. These zones reflect surface flow convergence, which can increase local flow volume and enhance the potential for flash flood occurrence.

Conversely, areas with positive plan curvature are generally located on ridges and open slopes, which tend to disperse flow and reduce water accumulation. The spatial distribution of plan curvature emphasizes its important role in controlling flow concentration and flash flood susceptibility in the Banyuputih watershed.

3.1.6 Landcover

Land use and land cover directly influence infiltration capacity and the magnitude of surface runoff. The land use/land cover map of the Banyuputih watershed was obtained from the ESRI Global Land Cover dataset based on Sentinel-2 imagery, illustrating the variation in land surface conditions across the study area. Figure 8.

In general, the upstream areas of the watershed are dominated by vegetative cover, such as forests and agricultural land, whereas built-up areas and open land are more prevalent in the middle to downstream parts of the watershed. The conversion of land cover from natural vegetation to open or built-up areas can potentially reduce infiltration capacity and increase surface runoff.

The spatial distribution of land use indicates that areas with degraded land cover and more impermeable surfaces contribute more significantly to heightened flash flood vulnerability. Therefore, land use/land cover represents an important factor in modeling flash flood susceptibility in the Banyuputih watershed.

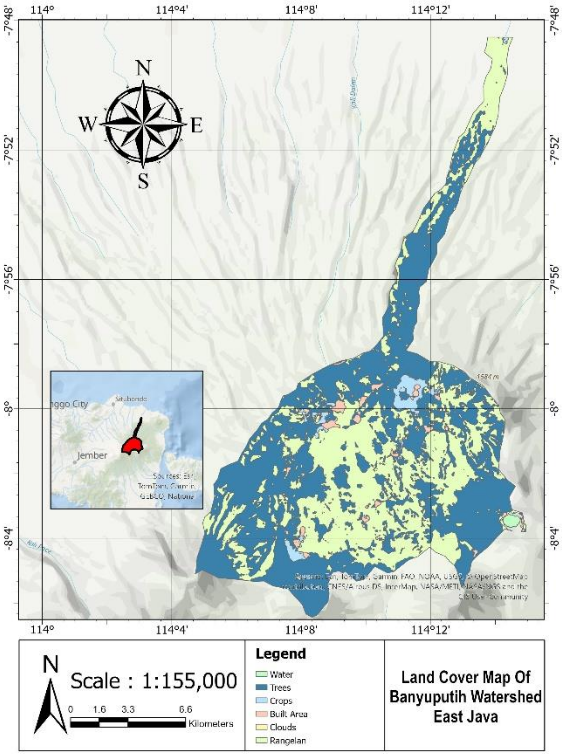

Figure 8 Land Cover Map Of Banyuputih Watershed

3.1.7 NDVI

The Normalized Difference Vegetation Index (NDVI) is used to represent vegetation density and condition, which influence infiltration, evapotranspiration, and surface runoff. The NDVI map of the Banyuputih watershed was calculated from Sentinel-2 imagery, showing the variation in vegetation density across the study area Figure 9

High NDVI values are generally distributed in the upstream parts of the watershed, which are still dominated by dense vegetation cover, enhancing infiltration and reducing surface runoff. Conversely, low NDVI values are mostly found in open land areas, intensive agricultural fields, and built-up zones, which have lower infiltration capacity.

This NDVI distribution indicates that areas with low vegetation density are more vulnerable to rapid runoff formation and contribute to flash flood occurrence. Therefore, NDVI represents an important parameter in identifying the influence of vegetation conditions on flash flood susceptibility in the Banyuputih watershed.

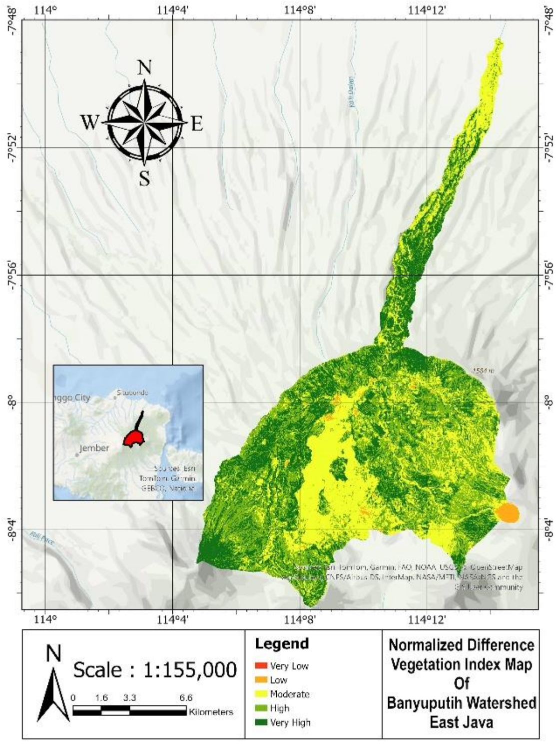

Figure 9 Normalized Difference Vegetation Index Map Of Banyuputih Watershed

3.1.8 HSG

The Hydrologic Soil Group (HSG) is used to represent soil infiltration capacity and the potential for surface runoff formation. The HSG map of the Banyuputih watershed was obtained from the global HYSOGs dataset, classifying soils based on their permeability characteristics Figure 10.

In general, the Banyuputih watershed is dominated by soil groups with moderate to low infiltration capacity. Areas with low-infiltration HSG tend to generate higher surface runoff during rainfall events, thereby increasing the potential for flash floods. Conversely, soils with higher infiltration capacity help reduce runoff volume.

The spatial distribution of HSG indicates that soil characteristics play an important role in controlling watershed hydrological response and serve as a significant conditioning factor in modelling flash flood susceptibility in the Banyuputih watershed.

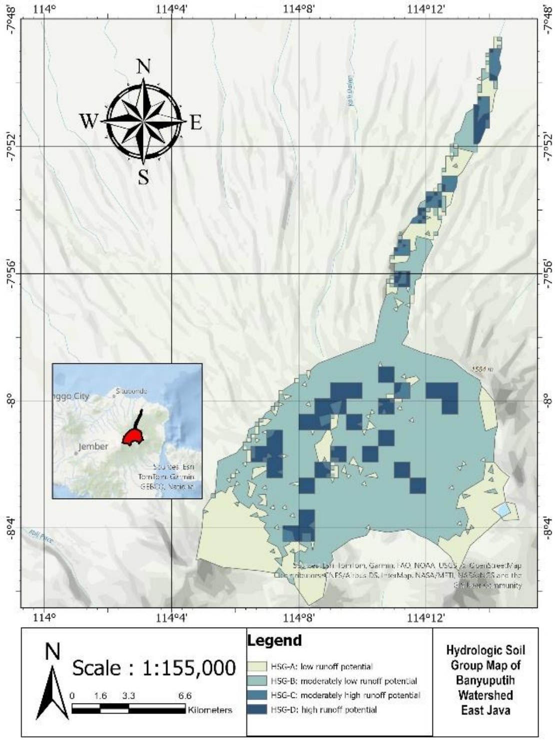

Figure 10 Hydrologic Soil Group Map Of Banyuputih Watershed

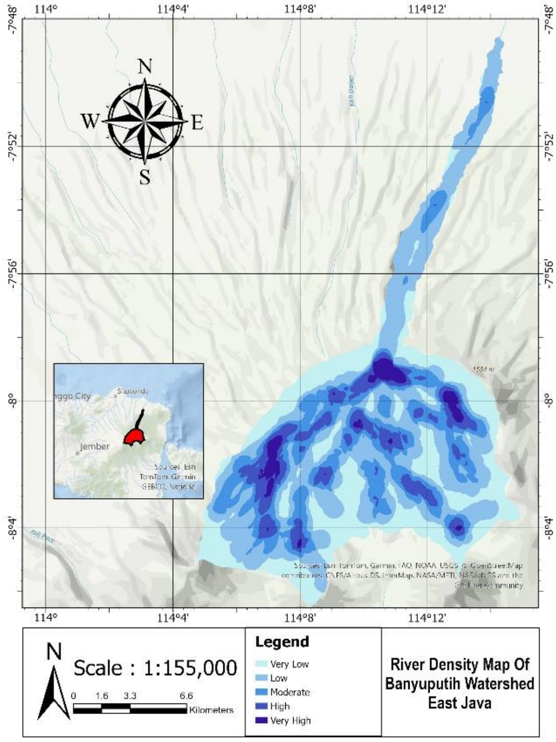

Figure 11 River Density Map Of Banyuputih Watershed

3.1.9 River Density

River density is used to represent the intensity of the drainage network within a watershed and the hydrological response to rainfall. The river density map of the Banyuputih watershed was derived from DEMNAS-based river network extraction Figure 11.

Spatially, high river density values are mostly found in the upstream and middle parts of the watershed, characterized by steep topography and dense drainage networks. This condition causes surface runoff to converge more rapidly into the river channels, increasing the potential for rapid peak discharge formation.

Conversely, areas with low river density are generally located in the downstream regions and relatively flat plains, exhibiting slower flow response. The distribution of river density indicates that regions with dense drainage networks have higher flash flood susceptibility and play an important role in modeling flash flood susceptibility in the Banyuputih watershed.

3.1.10 Rainfall

Rainfall is a primary triggering factor for flash floods, as it determines the volume of water entering the watershed system. The rainfall distribution map of the Banyuputih watershed was compiled based on observed rainfall data from the PU-SDA agency, illustrating the spatial variation of precipitation across the study area Figure 12.

The mapping results indicate that areas with relatively high rainfall are mostly distributed in the upstream parts of the watershed. When combined with steep slopes and dense river networks, these conditions can generate substantial surface runoff and rapid flow response.

Conversely, the middle to downstream parts of the watershed generally receives relatively lower rainfall. This spatial distribution of rainfall highlights its role as a primary triggering factor that, together with topographic and surface characteristics, controls the level of flash flood susceptibility in the Banyuputih watershed.

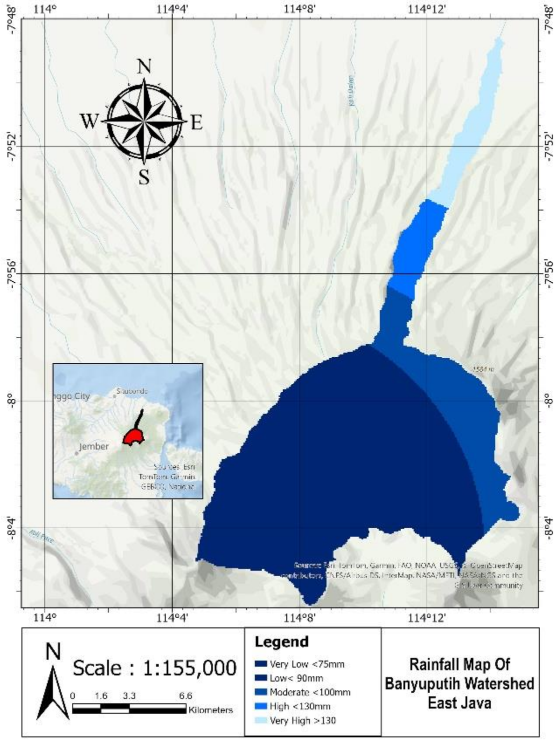

Figure 12 Rainfall Map of Banyuputih Watershed

3.2 Logistic Regression

Based on the logistic regression results Table 2, each variable exhibits a different contribution to the probability of flash flood occurrence. The Topographic Position Index (TPI) has the largest positive coefficient (8.658513), indicating that topographic positions favouring flow convergence significantly increase flood vulnerability. The Topographic Wetness Index (TWI, 0.536619) also shows a positive influence, suggesting that areas with high potential for water accumulation have a higher likelihood of flooding. River density and HSG contribute positively as well, although their effects are relatively small.

In contrast, NDVI (-9.017295) shows the strongest negative effect, indicating that vegetation plays a crucial role in reducing flood risk through enhanced infiltration and decreased surface runoff. Other variables, such as slope, elevation, land cover, and IDW, also have negative effects on flood probability. Plan curvature does not exhibit a significant influence, while the intercept value (8.773172) represents the baseline log-odds of the model.

Overall, topographic factors (particularly TPI and TWI) and vegetation conditions (NDVI) are the dominant variables in determining the flash flood susceptibility of the study area.

Table 2 Logistic Regression Coefficients of Flood Susceptibility Conditioning Factors

Variables	Coefficient
Land Cover	-0.137115
TWI	0.536619
NDVI	-9.017295
Plan Curvature	0.000000
River Density	0.005356
TPI	8.658513
Slope	-0.033407
HSG	0.000181
Elevation	-0.005317
IDW	-0.053370
Intercept	8.773172

3.3 Susceptibility Map

The flash flood susceptibility map of the Banyuputih watershed **Figure 13** , shows varying risk levels, ranging from Very Low to Very High. Areas coloured in red (Very High) are mainly located in the upstream and middle parts of the watershed, influenced by topographic factors and river flow patterns, consistent with the logistic regression results. Zones with Moderate–High risk are distributed along river valleys, while green areas (Low–Very Low) are found on elevated plains with well-established drainage. This map can serve as a basis for flood mitigation planning and prioritization in vulnerable areas.

Based on the classification of the flash flood susceptibility map Table 2, the area distribution of each class indicates varying levels of susceptibility across the study area. The Very Low class occupies the largest area, 71.72 km², indicating that most of the watershed has relatively low susceptibility to flash floods.

The Low class covers 53.49 km², followed by the Moderate class at 38.98 km². This suggests that although the majority of the area falls under low-risk categories, a significant proportion of the watershed has moderate susceptibility that warrants attention in spatial planning and disaster mitigation.

The High susceptibility class occupies 23.31 km², whereas the Very High class represents the smallest area at 6.73 km². Despite its relatively small extent, zones with High to Very High susceptibility are priority areas for mitigation and risk reduction, as they are likely to experience more severe flash flood impacts.

Overall, this distribution indicates that the study area is dominated by Low to Very Low susceptibility classes, yet the presence of high-risk areas still requires special attention in flash flood risk management strategies.

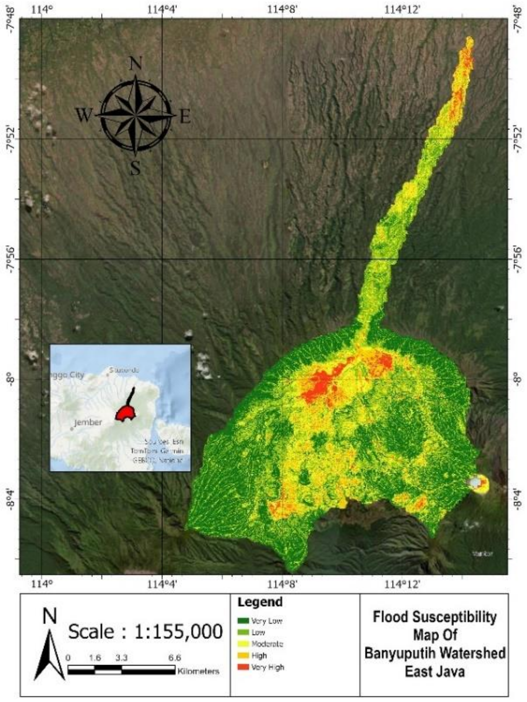

Figure 13 Flood Susceptibility Map Of Banyuputih Watershed

Table 3 Area Distribution of Flash Flood Susceptibility Classes

Class	Area (km^2)
Very Low	71.72
Low	53.49
Moderate	38.98
High	23.31
Very High	6.73

3.4 Accuracy of the Flash Flood Map

The model evaluation results Figure 14 show a ROC curve positioned above the diagonal line, indicating that the model performs better than random guessing in distinguishing vulnerable and non-vulnerable areas. The AUC value of 0.845 suggests that the model has moderate predictive capability, with approximately an 85% probability of correctly identifying vulnerable locations compared to non-vulnerable ones. These results demonstrate that the model is reasonably reliable, although there remains potential to improve its accuracy through the inclusion of additional data or more representative variables.

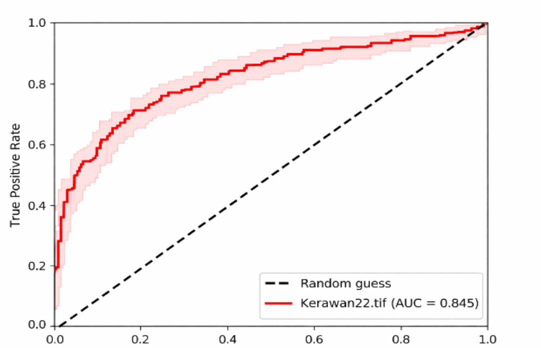

Figure 14 AUC Graph

3.5 Discussion

In line with the research objectives and conceptual framework outlined in the introduction, the logistic regression results indicate that topographic and vegetation parameters are the dominant factors controlling flash flood susceptibility in the Banyuputih watershed. The highest positive coefficient for TPI and the positive contribution of TWI confirm that valley zones and areas with high flow accumulation potential have a greater probability of flood occurrence. These findings are consistent with previous flood susceptibility mapping studies, which identify DEM-derived topographic indicators such as TPI, TWI, and slope as key factors in explaining flow convergence mechanisms and the formation of concentrated runoff (Waiyasusri et al., 2023). Geomorphologically, areas with low topographic positions and high accumulation capacity naturally tend to be primary flash flood locations.

In contrast, the largest negative coefficient for NDVI highlights the protective role of vegetation in reducing flood probability through enhanced infiltration and decreased surface runoff. This result aligns with previous studies reporting a negative relationship between vegetation density and flood vulnerability, where well-vegetated areas can buffer the watershed’s hydrological response to intense rainfall (Rahman et al., 2025).

Therefore, vegetation-based land cover management is an important component of flash flood risk mitigation strategies. The positive contributions of river density and HSG support literature findings that dense drainage networks and soils with low infiltration capacity accelerate runoff concentration processes and increase peak discharge. Meanwhile, the non-significance of plan curvature suggests that lateral flow convergence in this watershed is more effectively represented by a combination of other topographic parameters, particularly TPI and TWI.

An AUC value of 0.845 indicates good predictive capability, falling within the performance range reported in Logistic Regression-based flood susceptibility studies (Waiyasusri et al., 2023). This reinforces that the integration of Logistic Regression and GIS in this study not only produces spatially representative susceptibility maps but also allows for transparent identification and quantification of the relative contributions of each factor, as emphasized in data-driven statistical approaches (Hosmer et al., 2013).

4 Conclusion

This study successfully developed a flash flood susceptibility map for the Banyuputih Watershed by integrating a multi-parameter GIS-based approach with Logistic Regression modelling. The results indicate that topographic factors, particularly the Topographic Position Index (TPI) and Topographic Wetness Index (TWI), are the most influential variables in increasing flash flood probability. These findings confirm that topographic convergence zones and areas with high flow accumulation potential play a critical role in controlling flash flood formation.

Conversely, NDVI exhibits the strongest negative coefficient, highlighting the protective role of vegetation in reducing flood probability through enhanced infiltration and reduced surface runoff. Additional variables such as river density and Hydrologic Soil Group (HSG) contribute positively to flood vulnerability, although their influence is relatively smaller.

The spatial distribution of susceptibility classes reveals that most of the watershed falls within the Very Low to Low categories; however, High to Very High susceptibility zones are concentrated in upstream areas and along valley corridors, making them priority areas for mitigation efforts. The model performance, with an AUC value of 0.845, indicates good predictive capability in distinguishing flood-prone from non-prone areas.

Overall, this study demonstrates that the integration of Logistic Regression and GIS provides a robust and transparent framework for quantifying the relative contribution of conditioning factors in flash flood susceptibility mapping. The findings offer a scientific basis for sustainable watershed management, land-use planning, and the implementation of nature-based solutions in physiographically similar upstream catchments.

Future studies should integrate high-resolution dynamic rainfall data to better capture extreme rainfall triggers and compare Logistic Regression with advanced machine learning models to evaluate potential improvements in predictive performance. In addition, coupling susceptibility mapping with hydrodynamic simulations and future land-use or climate change scenarios would enhance the reliability and long-term applicability of flash flood risk assessments.

5 Acknowledgement

The authors gratefully acknowledge the support provided by the local Public Works and Water Resources Agency (PU-SDA) for supplying rainfall data used in this study. We also thank the Geospatial Information Agency of Indonesia for providing access to DEMNAS data, as well as the developers of the ESRI Global Land Cover dataset and Sentinel-2 imagery that supported the spatial analysis.

6 Reference

Abeysiriwardana, H. D., & Wijesekera, N. T. S. (2022). Application of GIS and Logistic Regression for Flood Susceptibility Mapping in Nilwala River Basin, Sri Lanka. *Engineer: Journal of the Institution of Engineers, Sri Lanka*, *55*(2), 1. https://doi.org/10.4038/engineer.v55i2.7503

Akuatiklestari, J., Sabriyati, D., & Hadi, M. P. (2022). Artikel penelitian 2 Kajian Hidrologi Debit Puncak Penyebab Banjir Bandang Menggunakan Pemodelan Hidrograf Satuan Sintesis-SCS (HSS-SCS) Hydrological Assessment of Peak Flood Discharge Causes of Flashflood Using SCS Synthetic Unit Hydrograph Modeling (SCS-SUH). *OPEN ACCESS*, *5*(2), 80–90. https://doi.org/10.31629/akuatiklestari.v5i2.4462

Ashfaq, S., Tufail, M., Niaz, A., Muhammad, S., Alzahrani, H., & Tariq, A. (2025). Flood susceptibility assessment and mapping using GIS-based analytical hierarchy process and frequency ratio models. *Global and Planetary Change*, *251*. https://doi.org/10.1016/j.gloplacha.2025.104831

Brázdil, R., Faturová, D., Šulc Michalková, M., Řehoř, J., Caletka, M., & Zahradníček, P. (2024). Spatiotemporal variability of flash floods and their human impacts in the Czech Republic during the 2001–2023 period. *Natural Hazards and Earth System Sciences*, *24*(10), 3663–3682. https://doi.org/10.5194/nhess-24-3663-2024

Bruijnzeel, L. A. (2004). Hydrological functions of tropical forests: not seeing the soil for the trees? *Agriculture, Ecosystems & Environment*, *104*(1), 185–228. https://doi.org/10.1016/j.agee.2004.01.015

Chow, V. te., & Maidment, D. R. . (1998). *Applied hydrology*. McGraw-Hill.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression*. Wiley. https://doi.org/10.1002/9781118548387

Islam, T., Zeleke, E. B., Afroz, M., & Melesse, A. M. (2025). A Systematic Review of Urban Flood Susceptibility Mapping: Remote Sensing, Machine Learning, and Other Modeling Approaches. *Remote Sensing*, *17*(3), 524. https://doi.org/10.3390/rs17030524

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning*. Springer US. https://doi.org/10.1007/978-1-0716-1418-1

Kaya, C. M., & Derin, L. (2023). Parameters and methods used in flood susceptibility mapping: a review. *Journal of Water and Climate Change*, *14*(6), 1935–1960. https://doi.org/10.2166/wcc.2023.035

Li, K., Guo, L., Wang, G., Gao, J., Ma, J., Li, J., Huang, P., Zhai, B., & Sun, X. (2025). A novel hybrid framework of high-resolution flood susceptibility mapping in ungauged mountainous regions. *Weather and Climate*

Extremes, *50*, 100822. https://doi.org/10.1016/j.wace.2025.100822

Maidment, D. R. (1993). *Handbook of Hydrology*. McGraw-Hill.

Manopkawee, P., Mankhemthong, N., Kanthata, S., & Sooksabai, C. (2025). Quantifying Regional-Scale Flash Flood Susceptibility Using Watershed Geomorphometry-Based Approach along the High Range, West of Chiang Mai Basin, Northern Thailand. *Natural Hazards Research*. https://doi.org/10.1016/j.nhres.2025.07.001

Ma'rufah, W., Ridwan, R., & Amin, M. (2024). Deteksi Kerawanan Banjir Genangan Menggunakan Topographic Wetness Index (TWI) di Sub Das Way Katibung. *Jurnal Agricultural Biosystem Engineering*, *3*(2), 238. https://doi.org/10.23960/jabe.v3i2.9435

Merghadi, A., Yunus, A. P., Dou, J., Whiteley, J., ThaiPham, B., Bui, D. T., Avtar, R., & Abderrahmane, B. (2020). Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Science Reviews*, *207*, 103225. https://doi.org/10.1016/j.earscirev.2020.103225

Montgomery, D. R., & Dietrich, W. E. (1994). A physically based model for the topographic control on shallow landsliding. *Water Resources Research*, *30*(4), 1153–1171. https://doi.org/10.1029/93WR02979

Pham, B. T., Avand, M., Janizadeh, S., Phong, T. Van, Al-Ansari, N., Ho, L. S., Das, S., Le, H. Van, Amini, A., Bozchaloei, S. K., Jafari, F., & Prakash, I. (2020). GIS Based Hybrid Computational Approaches for Flash Flood Susceptibility Assessment. *Water*, *12*(3), 683. https://doi.org/10.3390/w12030683

Qatrinnada, W. F. P., Hidayah, E., Halik, G., & Wiyono, R. U. A. (2024). A literature review: rainfall thresholds as flash flood monitoring for an early warning system. *Water Practice & Technology*, *19*(11), 4486–4498. https://doi.org/10.2166/wpt.2024.271

Saaty, T. L. (1980). *The Analytic Hierarchy Process*. McGraw-Hill.

Sadkou, S., Artigue, G., Fréalle, N., Ayral, P., Pistre, S., Sauvagnargues, S., & Johannet, A. (2024). A review of flash-floods management: From hydrological modeling to crisis management. *Journal of Flood Risk Management*, *17*(3). https://doi.org/10.1111/jfr3.12999

Shrestha, S., Dahal, D., Poudel, B., Banjara, M., & Kalra, A. (2025). Flood Susceptibility Analysis with Integrated Geographic Information System and Analytical Hierarchy Process: A Multi-Criteria Framework for Risk Assessment and Mitigation. *Water*, *17*(7), 937. https://doi.org/10.3390/w17070937

Sidle, R. C. O. H. (2006). *Landslides: Processes, Prediction, and Land Use*. American Geophysical Union.

Soliman, M., Morsy, M., & Radwan, H. (2022). Assessment of Implementing Land Use/Land Cover LULC 2020-ESRI Global Maps in 2D Flood Modeling Application. *Water*, *14*(23), 3963. https://doi.org/10.3390/w14233963

Suprayogi, S., Widyastuti, M., Hadi, M. P., Christanto, N., Tivianton, T. A., Fadhilah, G. O., Rahmawati, L., & Fadlillah, L. N. (2022). Runoff Coefficient Analysis After Regional Development in Tambakbayan Watershed, Yogyakarta, Indonesia. *Jurnal Ilmu Lingkungan*, *20*(2), 396–405. https://doi.org/10.14710/jil.20.2.396-405

Tarolli, P. (2016). Humans and the Earth's surface. *Earth Surface Processes and Landforms*, *41*(15), 2301–2304. https://doi.org/10.1002/esp.4059

Tsumita, N., Piyapong, S., Kaewkluengklom, R., Jaensirisak, S., & Fukuda, A. (2025). Flood susceptibility mapping of urban flood risk: comparing autoencoder multilayer perceptron and logistic regression models in Ubon Ratchathani, Thailand. *Natural Hazards*, *121*(15), 17833–17867. https://doi.org/10.1007/s11069-025-07494-8

Wahyuni, I. T., & Sachro, S. S. (2024). The Impact of Land Cover Changes on Flood Hydrograph Characteristics (Jragung dam Case Study). *MEDIA KOMUNIKASI TEKNIK SIPIL*, *29*(2), 280–288. https://doi.org/10.14710/mkts.v29i2.59877

Waiyasusri, K., Wetchayont, P., Tananonchai, A., & Suwanmajo, D. (2023). Flood Susceptibility Mapping Using Logistic Regression Analysis In Lam Khan Chu Watershed, Chaiyaphum Province, Thailand. *GEOGRAPHY, ENVIRONMENT, SUSTAINABILITY*, *16*(2), 41–56. https://doi.org/10.24057/2071-9388-2022-159

Zachrani, A. O., Rayes, M. L., & Sulaeman, Y. (2024). IDENTIFIKASI LAHAN KRITIS SKALA 1:25.000 BERBASIS CITRA SENTINEL 2 DAN DEMNAS RESOLUSI 8 METER (STUDI KASUS SUB DAS CIKAPUNDUNG JAWA BARAT). *Jurnal Tanah Dan Sumberdaya Lahan*, *11*(2), 327–338. https://doi.org/10.21776/ub.jtsl.2024.011.2.4