

Predictive Landslide Susceptibility Modelling Using Statistical Approaches in Mantsa River catchment, Dawro Zone, South-western Ethiopia

Mihretu Fikre

mihretuf@wcu.edu.et

Wachemo University

Meseret Zinabu

Arba Minch University

Leulalem Shano

Arba Minch University

Muralitharan Jothimani

Arba Minch University

Berhanu Erbelo

Wachemo University

Ephrem Getahun

Arba Minch University

Shankar Karuppanna

Saveetha University

Research Article

Keywords: Dawro, Frequency ratio, Logistic regression, Landslide susceptibility, Mantsa river catchment, ROC curve

Posted Date: September 23rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11115873/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Predictive Landslide Susceptibility Modelling Using Statistical Approaches in Mantsa River catchment, Dawro Zone, South-western Ethiopia

Abstract

Landslides are among the most destructive geo-hazards in mountainous regions, threatening human lives, infrastructure, and sustainable land-use planning. Accurate landslide susceptibility mapping is therefore crucial for hazard assessment and mitigation. This study evaluates landslide susceptibility in the Mantsa River Catchment, Dawro Zone, southwestern Ethiopia, an area frequently affected by slope instability, using two statistical approaches: the Frequency Ratio (FR) and Logistic Regression (LR) models. A landslide inventory consisting of 280 historical landslide events was developed through detailed interpretation of Google Earth imagery and extensive field investigations. The inventory was randomly divided into training (70%, $n = 196$) and validation (30%, $n = 84$) datasets. Thirteen landslide conditioning factors were selected based on their geomorphological, geological, and hydrological significance, including land-use/land-cover, lithology, rainfall, distance to lineaments, distance to streams, stream power index, soil type, elevation, aspect, distance to roads, topographic wetness index, slope gradient, and curvature. These factors were processed and standardized as raster layers within a GIS environment. The results reveal strong relationships between landslide occurrence and the conditioning factors, with land-use/land-cover, lithology, rainfall, and distance to lineaments identified as the most influential variables. Besides, areas underlain by mafic volcanic rocks in the study area exhibit higher susceptibility due to intensive weathering, clay mineral development, and the consequent reduction in rock strength, which enhances slope failure during prolonged rainfall events. The model validation using Receiver Operating Characteristic (ROC) analysis produced Area Under the Curve (AUC) values of 0.874 for the FR model and 0.850 for the LR model, demonstrating good predictive performance for both models, with FR showing slightly higher accuracy. The LSI map shows that low-susceptible (37.59%) and very low-susceptible (8.12%) zones are mainly distributed around the north and central directions of the study area. Very high susceptible (8.2%), high susceptible (16.37%) zones are mainly dominated around southwest (Shanda Hawusha, Daligashe, Tima, Okiliand Aocha Bobo), northeast (Gezo Koyisha and Bosa Tedefa), and southeast (Gozobamushi, Tulema Tema Chich Haye, Bosa Shoga) directions of the study area. The moderate-susceptible (29.72%) zone is distributed throughout the rest part of the study area. The findings provide valuable insights for landslide risk reduction, land-use planning, and

sustainable development in Ethiopia and other mountainous regions with similar environmental conditions.

Keywords: Dawro; Frequency ratio; Logistic regression; Landslide susceptibility; Mantsa river catchment; ROC curve.

1. Introduction

Across Ethiopia, landslide impacts have intensified over the past decade, especially in its highland regions where steep terrain, deeply incised valleys, rapid weathering, and high rainfall converge (Shano et al., 2021; Dam et al., 2022; Kebeba et al., 2024). The south-western and southern highlands—including landslides—downslope movement of rock, debris, or earth—rank among the most destructive geological hazards globally. Between 2004 and 2020, over 55,000 fatalities were attributed to rainfall- and earthquake-triggered landslides (Froude and Petley 2021). Risks are escalating: 2024 saw 1,570 global fatalities, driven by intensified precipitation (Yang et al. 2025); August 2025 alone recorded 104 fatal landslides causing 2,365 deaths (Petley 2025a, 2025b). These figures underscore the urgent need for predictive susceptibility science. Across Africa, landslide impacts have intensified, especially in highlands with steep terrain, deep valleys, rapid weathering, and high rainfall. Ethiopia, within the East African Rift, exemplifies this vulnerability. Its highlands—rugged topography, deeply incised valleys, extensive volcanic sequences—are highly prone to slope failures (Shano 2021; Dam et al. 2022). The south-western and southern highlands (Dawro, Gamo, Gofa, Wolaita, Sidama zones) experience recurrent earthflows, debris slides, and rockfalls during prolonged rainy seasons, causing loss of life, home destruction, displacement, and agricultural degradation (Shano et al. 2022; Wubalem et al. 2025). The July 2024 Gofa Zone catastrophe was a watershed event. Unusually heavy rainfall triggered a series of landslides in Kencho Shacha Gozdi village, killing >250 people, affecting >15,000, and displacing thousands (OCHA/WHO 2024). Investigation identified lithology, geological structures, groundwater, slopes, and human activities as conditioning factors, with heavy rainfall as the trigger (Nwankwo et al. 2024). This disaster highlighted a critical gap: many Ethiopian hazard-prone zones—including the Dawro highlands—lack validated landslide susceptibility maps, hampering targeted mitigation.

A growing body of research has assessed landslide susceptibility across Ethiopia using statistical approaches, especially Frequency Ratio (FR) and Logistic Regression (LR). Along the Debretabor-Alember road, FR and Information Value models found 72.16% of landslides in high/very high zones (Tesfaye et al. 2024). In the Muger watershed (upper Blue Nile), FR-LR fusion showed ~185 km² (40.2%) as high-very high susceptible, with FR

average predictive performance of 81.45% (Asmare et al. 2024). In Awabel Woreda (east Gojjam), FR and AHP achieved 89.73% validation accuracy, identifying 6.09% and 16.9% of the area as very high and high susceptibility (Sisay et al. 2024). In the Maze catchment (Omo Valley), FR and AHP were integrated, using 70% inventory for training (Kebeba et al. 2024). Other studies in Bonga, Addi Arkay, and the Blue Nile Gorge similarly applied GIS-based statistical methods. Collectively, these studies confirm the effectiveness of statistical approaches in data-scarce regions, but none specifically addressed the Dawro Zone—a clear knowledge gap. The scientific merit of this work is its focused contribution to a region with no prior systematic susceptibility assessment, not methodological novelty. While advanced machine learning and deep learning (e.g., convolutional neural networks) have shown superior performance (Chang et al. 2023; Gui et al. 2024), statistical models remain widely used for their simplicity, interpretability, and robust predictive capability in data-limited settings (Li et al. 2024; Singh et al. 2024). Pei et al. (2024) showed that XGBoost and LightGBM consistently outperform across sample sizes, but FR is most affected by sample size yet still valuable, especially when integrated. In Ethiopia, where high-quality, temporally consistent landslide inventories are still developing, the interpretability and straightforward implementation of FR and LR are particularly appropriate. Comparative assessments confirm FR, LR, and weight-of-evidence can achieve $AUC > 0.80$ when properly calibrated (Chen et al. 2020; Yu et al. 2023). The combined use of FR and LR provides complementary perspectives—empirical correlations plus probabilistic estimates—and results are readily understood by local disaster management authorities. Selecting appropriate conditioning factors is critical. A systematic review identified slope, lithology, LULC, elevation, aspect, stream proximity, and rainfall as most common (Chicas et al. 2024). In the Ethiopian highlands, dominant factors include weathered volcanic lithology (basalt, rhyolite, trachyte, pyroclastics), slope angle, lineaments, land use, and rainfall (Shano 2021; Dam et al. 2022). The Gofa investigation specifically highlighted geological structures, slopes, and human activities (Nwankwo et al. 2024). Climate change is expected to exacerbate hazard: studies in California (Zhang et al. 2025) and Japan (Tsunetaka et al. 2025) project increased susceptibility through 2100 under moderate to high emission scenarios due to more frequent intense rainfall. The Mantsa River catchment in the Dawro Zone exemplifies the convergence of high-susceptibility factors: rugged, steep topography (slopes 0–70.6°), deeply incised terrain, highly weathered volcanic sequences (basalt, rhyolite, trachyte, tuff, ignimbrite), and intensive rainfall. Field reports from 2017–2025 document numerous destructive landslides causing fatalities, livestock losses, road blockages, crop destruction, and household displacement (Shano et al. 2024; Kebeba et al. 2024). Heavy/prolonged rainfall is the primary trigger, exacerbated by soil saturation, land-use disturbances (especially moderate cultivation on steep slopes), and dry-season cracking. Most residents occupy steep mid-slope and foot-slope positions, where agricultural expansion, road construction, and land-use change further destabilize

slopes. Despite these risks, no systematic landslide susceptibility investigation has been conducted in this catchment—a critical obstacle to disaster risk management.

The primary objective is to develop and validate high-resolution landslide susceptibility models for the Mantsa River catchment using FR and LR, supported by a landslide inventory (n=280) and 13 conditioning factors: LULC, lithology, rainfall, proximity to lineaments, proximity to streams, stream power index, soil type, elevation, aspect, proximity to roads, topographic wetness index, slope, and curvature. The specific contributions are: (1) the first systematically validated landslide inventory for the Mantsa River catchment; (2) identification of the most influential conditioning factors in this geological setting; (3) susceptibility maps to inform local disaster risk reduction in a region with no prior assessment; and (4) a replicable methodological framework for similar data-scarce regions in Ethiopia and beyond. The results will provide robust scientific evidence for landslide risk management, spatial planning, early warning systems, and sustainable development in landslide-prone areas of Ethiopia and comparable mountainous environments worldwide.

2. Methods and Materials

2.1. Study Area

The study area encompasses the Mantsa River catchment in the Dawro Zone, south-western Ethiopia (Latitude: 774423 mN to 798748 mN; Longitude: 281338 mE to 317037 mE, UTM Zone 37). The region features rugged topography with elevations ranging from 951 m to 2,782 m, characterized by mountains, plateaus, deep gorges, and a dendritic drainage pattern, rendering it highly susceptible to slope instability. This region forms part of the Dawuro Highlands, encompassing some of its highest peaks, such as Gazo, Hayo, Esara, and the Atso Saddle Mountains. The lowest elevation is found at the Gojeb River, near its confluence with the Omo River. The Mantsa River catchment within the area exhibits a dendritic drainage pattern (Figure 1). The Mantsa River catchment encompasses three climatic zones: Kolla, Woinadega, and Dega, classified based on elevation, rainfall, and temperature. The region experiences a predominantly hot and humid climate with distinct cold periods. The cold season, typically occurring in April and May, is characterized by heavy rainfall, which increases landslide susceptibility. In contrast, the dry season spans January, February, November, and December, marked by low rainfall and higher temperatures. Climatic data for the study area,

including rainfall, minimum, and maximum temperatures from 2011 to 2021, were obtained from NASA's POWER Project (<https://power.larc.nasa.gov/data-access-viewer/>). The highest annual precipitation recorded during this period was 2,194 mm in 2014, while the peak monthly rainfall of 8.47 mm occurred in May. The Mantsa River catchment exhibits significant elevation-related temperature variations, with annual temperatures ranging from 10.55°C to 35.34°C in 2019. May is the hottest month, with an average minimum temperature of 14.58°C and an average maximum temperature of 34.06 °C.

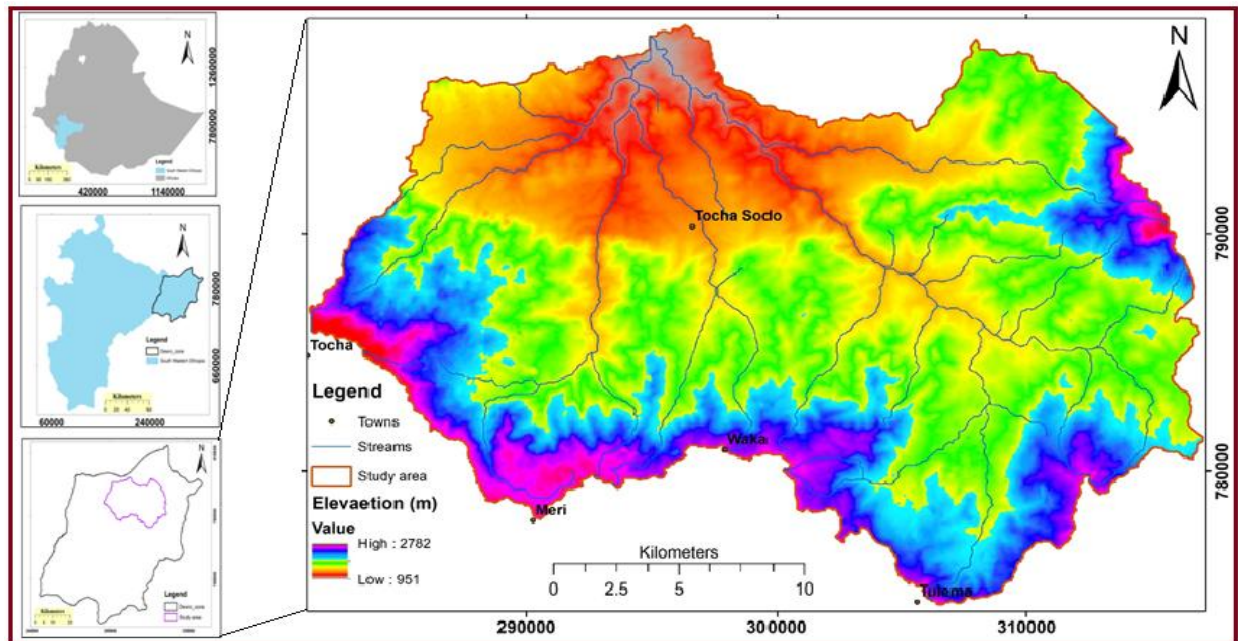


Figure 1: Study area

2.2.1 Geomorphologic and Geological Description of the Study Area

The Mantsa River Catchment exhibits a highly diverse geomorphology that reflects prolonged volcanic activity, tectonic deformation, and fluvial erosion. As shown in the geomorphological map, the landscape is characterized by volcanic mountains, escarpments, dissected uplands, rolling uplands, plateaus, valley bottoms, alluvial fans, and gullied terrains. Volcanic mountains are concentrated in the southwestern and southeastern parts of the catchment, where high-relief cones, ridges, and peaks create steep slopes that are highly susceptible to weathering and slope failure. Escarpments and steep slopes occupy the margins of these volcanic highlands, while structural hills and dissected uplands dominate the eastern and southern sectors. The central catchment is characterized by gently undulating plateaus and rolling uplands dissected by an extensive dendritic drainage network. Valley bottoms and floodplains

occur along the Mantsa River and its tributaries, where alluvial deposition has formed relatively flat surfaces. In contrast, gullied and eroded lands are mainly developed along steep escarpments, reflecting intense surface runoff, active erosion, and geomorphic instability. Geologically, the study area is dominated by volcanic rocks associated with syn-rift magmatism of the Ethiopian Rift System. The stratigraphic succession comprises aphyric basalt, rhyolite, porphyritic trachyte, unwelded tuff, and welded ignimbrite, recording alternating effusive and explosive volcanic episodes. Early fissural eruptions produced highly fractured aphyric basalts intruded by basaltic dikes, followed by rhyolitic lava flows and domes, trachytic flows with well-developed flow banding and cooling joints, and pyroclastic deposits consisting of friable tuffs and strongly welded ignimbrites. Quaternary alluvial, fluvial, and lacustrine sediments overlie these volcanic rocks, occupying valley floors and depositional plains. Regionally, the Mursi Formation is composed mainly of fine-grained mafic volcanic rocks generated by fissural lava outpourings, whereas the overlying Pliocene–Pleistocene Nkalabong Formation contains volcanic ash and tephra deposits important for regional stratigraphic correlation. Structurally, the area is controlled by dominant NE–SW and subordinate NW–SE fault systems related to the Ethiopian Rift, which govern drainage orientation, promote intense fracturing and weathering of volcanic rocks, and significantly increase slope instability and landslide susceptibility throughout the catchment (Martínek et al., 2021; Vidal et al., 2022).

2.2 Methods

2.2.1 Data Preparation and Landslide Inventory

A landslide inventory was developed by interpreting Google Earth imagery and conducting field surveys (August 2022–February 2023), resulting in 280 landslide locations (active and inactive). Landslide types were classified according to Cruden and Varnes (1996): earth flows (65%), debris slides (25%), and rock falls (10%). The inventory was randomly split into 70% (196 landslides) for training and 30% (84 landslides) for validation. Thirteen causative factors were selected based on a literature review, data availability, and field observations: slope, elevation, aspect, and curvature; proximity to streams, roads, and lineaments; LULC; soil type; lithology; rainfall; SPI; and TWI. A 30m resolution DEM was used to derive topographic parameters. Lithological and soil maps were obtained from the Ethiopian Geological Survey,

LULC data from satellite imagery, and rainfall data from NASA's POWER project (Table 1). All data were unified into a 30m raster grid within a GIS platform. Natural Breaks are used for all continuous landslide factors.

Table 1: Data sets used to prepare landslide factors, along with the respective sources

Data type and Parameters	Source	Scale or resolution	Data format	Purpose
Landslide inventory	Google Earth image and Field work		Tabular	Landslide inventory map
DEM(elevation,slope,aspect, curvature,SPI, TWI)		30m x 30m	Raster	Elevation,slope,aspect, curvature, SPI and TWI map
Proximity to streams	https://earthexplorer.usgs.gov/			Stream
Geological-map (proximity to roads and-lineaments, and, lithology)	Ethiopian Geological Survey	1:50,000		Road,and lineament proximity maps and Lithological map
Climate Data(rainfall, temperature (min. and max.))	https://power.larc.nasa.gov/data-access-viewer/		Vector	Rainfall map
Soil data (Soil layer)	Ethiopian Geological Survey		Soil layer	Soil map
Landuse landcover data	https://search.asf.alaska.edu/	30m x 30m	LULC layer	LULC map

2.2.2 Landslide Susceptibility Modelling

a. Frequency Ratio (FR) Model

The FR method evaluates the spatial relationship between landslide occurrence and each factor class. The FR for a class is calculated as the ratio of the landslide density within the class to the landslide density in the entire study area. An $FR > 1$ indicates a higher correlation with landslides. The Landslide Susceptibility Index (LSI) was computed by summing the FR values of all factors for each pixel. The frequency ratio is defined as the ratio of the probability of an event occurring to the probability of it not occurring (Erener & Düzgün, 2010). It calculates the proportion of landslide-affected cells to unaffected cells within predetermined classes of independent variables, such as slope, elevation, and land cover (Biswajeet Pradhan, 2013). Additionally, the FR model is valuable for ranking causative factors by their influence on

landslide incidents (Kannan et al., 2013). By clearly delineating the differences in scores among causal factors, the FR method facilitates a better understanding of landslide occurrences. To compute the frequency ratio, the number of landslide occurrence pixels is combined with the causal factors. The ratio for each factor is calculated by dividing the landslide occurrence ratio by the ratio for each class within the causal factors (Lee & Talib, 2005). The association between landslide locations and the relevant factors in the study area is determined using the frequency ratio technique, based on observed relationships (Lee & Pradhan, 2007). The findings from this study are then utilised to evaluate the correlation between the causal factors and the spatial distribution of landslides, providing a robust framework for understanding landslide susceptibility. Equation 1 and 2 shows the class ratio and slide ratio calculations.

$$\text{Class Ratio} = \frac{\text{Number of pixel in individual class}}{\text{Total number of pixel in whole class}} \quad (1)$$

$$\text{Slide Ratio} = \frac{\text{Number of landslide pixel in class}}{\text{Total number of landslide pixel}} \quad (2)$$

Where $N_{\text{pix}(si)}$ is the number of landslide pixels in class i , $N_{\text{pix}(Ni)}$ is the total number of landslide pixels having class i in the watershed, $\sum_i N_{\text{pix}(Si)}$ is the total number of pixels containing landslide, and $\sum_i N_{\text{pix}(Ni)}$ is the total number of pixels in the study area. The Frequency Ratio (FR) is a quantitative measure of the correlation between landslides and their associated conditioning variables. A score of one indicates an average level of correlation, while values exceeding one suggest a strong correlation with landslide occurrences. Conversely, a value of one or less indicates a weaker relationship (Çevik & Topal, 2003). In this context, the ratio value for each class illustrates the extent to which specific factors contribute to landslide susceptibility. Ratios greater than 1 indicate a stronger correlation with landslide events, whereas ratios below 1 imply a weaker correlation (Lee & Pradhan, 2006). To compute the frequency ratio, the following formula (Equation 3) is employed, providing a systematic approach to quantifying the relation between landslide occurrences and the identified conditioning factors. This method enhances our understanding of the dynamics that influence landslide susceptibility, facilitating more informed decision-making in susceptibility assessment and management.

$$\text{FR} = \frac{N_{\text{pix}(si)} / \sum_i N_{\text{pix}(Si)}}{N_{\text{pix}(Ni)} / \sum_i N_{\text{pix}(Ni)}} = \frac{\text{SR}}{\text{CR}} \quad (3)$$

The landslide susceptibility index is computed using spatial analysis tools in a raster calculator, which aggregates the frequency-ratio ratings of all causative components (Mohammady et al.,

2012). The calculation method for the Landslide Susceptibility Index (LSI) is detailed in Equation 4). To derive the LSI, the frequency ratios from all factor maps are integrated using the raster calculator found in the spatial analysis toolset (Equation 4). This process allows for a comprehensive assessment of the factors contributing to landslide susceptibility across the study area. Once the LSI is calculated, it is categorised into five distinct classes based on landslide susceptibility. This classification system enhances the interpretability of the index, enabling stakeholders to identify areas at varying degrees of susceptibility. By summing the raster map parameters for frequency ratios, the LSI provides a strong framework for understanding and effectively managing landslide susceptibility.

$$LSI = FR_1 + FR_2 + \dots + FR_n \quad (4)$$

b. Logistic Regression (LR) Model

LR is a multivariate statistical method used to predict the probability (P) of a binary outcome (landslide presence/absence) based on multiple independent variables (causative factors). The relationship is expressed via the logistic function (Equation 5), where Z is a linear combination of the factors and their coefficients (Equation 6). Coefficients were estimated using SPSS software, indicating the direction and magnitude of each factor's influence on landslide probability. This characteristic makes logistic regression particularly suitable for modeling binary outcomes. The independent variables in logistic regression can be continuous, discrete, or a combination of both, as noted by Xu et al. (2013). The probabilities generated by the logistic regression model represent the likelihood of one state of the dependent variable, ranging from 0 to 1. According to Dai (2004), a probability of 0 indicates a 0% chance of the event occurring (a landslide), while a probability of 1 indicates a 100% chance. Thus, the probabilities predicted by the logistic regression model can be interpreted as the likelihood or chance of the binary outcome assuming a particular value, providing valuable insights into landslide susceptibility evaluation. The logistic regression model is based on the logistic function, which is expressed as follows using equation 5:

$$p = \frac{1}{1 + e^{-Z}} \quad (5)$$

Where P is the probability of landslide occurrence, estimated values range from 0 to 1. Z is a landslide causal factor and is assumed to be a linear combination of the causal factors (Equation 6). Probability varies from zero to one. The closer it is to 1, the higher the susceptibility; the closer it is to 0, the lower the vulnerability.

$$Z = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n, \quad (6)$$

2.3 Modelling of Landslide Susceptibility

The predictive novel model of the landslide susceptibility in the study area was formulated in equation (7). A frequency ratio for a parameter class greater than 1 indicates a strong correlation for landslide occurrence, whereas a ratio less than 1 indicates a low correlation of the factor class with landslides. To create landslide susceptibility index levels, all the ratios of raster map landslide causal factors were summed as using equation 7:

$$\begin{aligned} \text{LSI} = & \text{PR}_{\text{Slope}} + \text{PR}_{\text{Elevation}} + \text{PR}_{\text{Curvature}} + \text{PR}_{\text{Aspect}} + \text{PR}_{\text{Stream_prox.}} + \text{PR}_{\text{Road_prox.}} + \\ & \text{PR}_{\text{Lineament_prox.}} + \text{PR}_{\text{LULC}} + \text{PR}_{\text{Soil}} + \text{PR}_{\text{Lithology}} + \text{PR}_{\text{RF}} + \text{PR}_{\text{SPI}} \\ & + \text{PR}_{\text{TWI}} \end{aligned} \quad (7)$$

Where LSI denotes the landslide susceptibility index, PR denotes the prediction rate of each causal factor.

2.4 Validation and Comparison of Models

The predictive performance of both models was validated using the reserved 84 (30%) of the landslide inventory data (Figure 2). Receiver Operating Characteristic (ROC) curves were constructed, and the Area Under the Curve (AUC) was calculated. An AUC value closer to 1 indicates superior predictive accuracy (equation 8). The models were compared based on their AUC values and the spatial distribution of predicted susceptibility classes (Tables 4-6).

$$\text{Tolerance} = 1 - C^2 \quad (8)$$

In this context, C^2 denotes the coefficient of determination for the regression of a specific explanatory variable against all other independent variables. The variance inflation factor (VIF) (equation 9) was used to calculate as the inverse of tolerance.

$$\text{VIF} = \frac{1}{\text{tolerance}} \quad (9)$$

Frequency ratio (FR) greater 1 indicates a higher susceptibility to landslides, while a ratio less than 1 suggests a lower susceptibility. This method influences the relationship between previously recorded landslides and the factors contributing to their occurrence, enabling researchers to identify correlations between landslide prone areas and those that remain unaffected (Lee et al., 2004). The frequency ratio method is favoured for its simplicity and ease of application, making it a widely utilized probability model in landslide susceptibility studies.

The frequency ratio is defined as the ratio of the probability of an event occurring to the probability of it not occurring (Erener & Düzgün, 2010). It calculates the proportion of landslide-affected cells to unaffected cells within predetermined classes of independent variables, such as slope, elevation, and land cover (Biswajeet Pradhan, 2013). Additionally, the FR model proves valuable in ranking the causative factors based on their influence on landslide incidents (Kannan et al., 2013). By clearly delineating the differences in scores among causal factors, the FR method facilitates a better understanding of landslide occurrences. To compute the frequency ratio, the number of landslide occurrence pixel is combined with the causal factors. The ratio for each factor is calculated by dividing the landslide occurrence calculated by dividing the landslide occurrence ratio by the ratio of each class within the causal factors (Lee & Talib, 2005). The association between the locations of landslide and the relevant factors in the study area is determined using the frequency ratio technique, based on observed relationships (Lee & Pradhan, 2007). The findings from this study are then utilized to evaluate the correlation between the causal factors and the spatial distribution of landslides, providing a robust framework for understanding landslide susceptibility.

The Frequency Ratio (FR) value serves as a quantitative measure of the correlation between landslides and their associated conditioning variables. A score of one indicates an average level of correlation, while values exceeding one suggest a strong correlation with landslide occurrences. Conversely, a value of one or less indicates a weaker relationship (Çevik & Topal, 2003). In this context, the ratio value for each class illustrates the extent to which specific factors contribute to landslide susceptibility. Ratios greater than one indicate a stronger correlation with landslide events, whereas ratios below one imply a lower correlation, (Lee & Pradhan, 2006). To compute the frequency ratio, the following formula (Equation below) is employed, providing a systematic approach to quantifying the relation between landslide occurrences and the identified conditioning factors. This method enhances our understanding of the dynamics influencing landslide susceptibility, facilitating more informed decision making in susceptibility assessment and management.

Where $N_{pix(Si)}$ is the number of landslide pixels containing in class i , $N_{pix(Ni)}$ is the total number of landslide pixel having class i in the watershed, $\Sigma_i N_{pix(Si)}$ is the total number of pixels containing landslide and $\Sigma_i N_{pix(Ni)}$ is the total number of pixels in the study area. The landslide susceptibility index is computed by employing spatial analysis tools within a raster calculator, which aggregates the frequency ratio ratings of all causative components (Mohammady et al.,

2012). The calculation method for the Landslide Susceptibility Index (LSI) is detailed in Equation.

To derive the LSI, the frequency ratios from all factor maps are integrated using the raster calculator found in the spatial analysis toolset (Equation below). This process allows for a comprehensive assessment of the various factors contributing to landslide susceptibility across the study area. Once the LSI is calculated, it is subsequently categorized into five distinct classes based on levels of landslide susceptibility. This classification system enhances the interpretability of the index, enabling stakeholders to identify areas at varying degrees of susceptibility. By summing all raster map parameters of the frequency ratios, the LSI provides a strong framework for understanding and managing landslide susceptibility effectively.

$$LSI = FR_1 + FR_2 + \dots + FR_n$$

Where LSI is landslide susceptibility index value and FR_1 , FR_2 , FR_n is the frequency ratio of the raster data layers and n is the total number of factors for the study ((J. Chen et al., 2018). Higher value indicates high landslide susceptibility, and lower value indicates low susceptibility and vice versa.

Landslide susceptibility using logistic regression analysis method

Logistic regression is a powerful statistical modelling technique used to predict the probability of a binary outcome, such as the occurrence or non-occurrence of a landslide, based on the set of independent variables. This method operates on the premise of the logistic function, which estimates the likelihood of a specific outcome specifically, whether a landslide is present or absent. The logistic regression model yields regression coefficients that quantify the influence of each independent variable on the probability of landslide occurrence. In contrast to the Frequency Ratio method, which evaluates each independent variable in isolation by calculating the ratio of landslide areas to non-landslide areas for each factor class, logistic regression simultaneously assesses the relationships among multiple variables. It provides a more nuanced understanding by controlling for the effects of other variables in the model allowing for a comprehensive analysis of their collective impact on landslide susceptibility. As defined by (Chau & Chan, 2005), logistic regression models the probability of landslide occurrence relative to the probability of non-landslide occurrence. This statistical technique is particularly effective when the dependent variable is binary, meaning it can only assume two possible values: occurrence and non-occurrence of an event. According to (Shirzadi et al., 2018), the

spatial prediction of a binary outcome is achieved through a combination of a dependent variable and various independent variables.

According to (Lee & Talib, 2005) one of the notable advantages of logistic regression is its ability to incorporate an appropriate link function into the standard linear regression model. Typically, this link function is the logit function, which transforms the linear combination of independent variables into a probability value constrained between 0 and 1. This characteristic makes logistic regression particularly suitable for modeling binary outcomes. The independent variables employed in logistic regression can be continuous, discrete, or a combination of both, as noted by (Xu et al., 2013). The probabilities generated by the logistic regression model represent the likelihood of one state of the dependent variable, confined within the range of 0 to 1. According to (F. C. Dai, 2004), a probability of 0 indicates a 0% chance of the event occurring (e.g., landslide occurrence), while a probability of 1 signifies a 100% chance. Thus, the probabilities predicted by the logistic regression model can be interpreted as the likelihood or chance of the binary outcome assuming a particular value, providing valuable insights into landslide susceptibility evaluation. The logistic regression model is based on the logistic function, which is expressed as follows:

Where P is the probability of landslide occurrence that estimated values varies from 0 to 1. Z is landslide causal factor and assumed as a linear combination of the causal factors. Probability varies from zero to one. Whenever it is nearer to one, it indicates high susceptible, whenever it is nearer to zero, it indicates very low vulnerable.

$$Z=b_0+b_1X_1+b_2X_2+\dots+b_nX_n,$$

b_0 is the model intercept, n is the number of independent variables, $b_1, b_2, \dots, b_n$ are the coefficients and $X_1, X_2, \dots, X_n$ are the landslide causing factors.

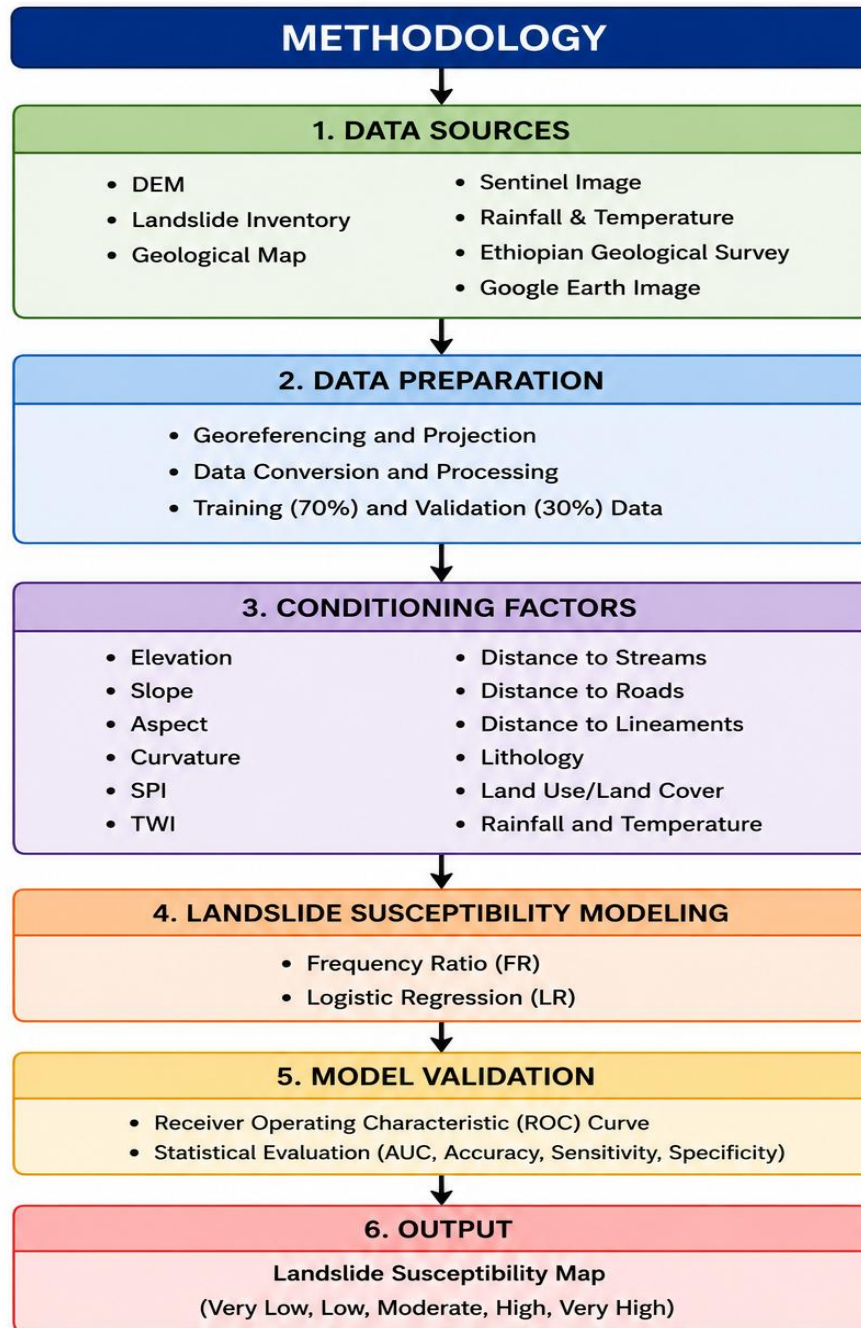


Figure 2. Methodological flow chart for preparation of landslide susceptibility map of the study area.

3. Results

3.1 Geology

The geology of the study area is dominated by extensive volcanic sequences composed of aphyric basalt, rhyolite, porphyritic trachyte, unwelded tuff, and welded ignimbrite, collectively recording alternating effusive and explosive volcanism associated with syn-rift tectonics in the Ethiopian Rift (Vidal, C.M. et al., 2022). The aphanitic basalts, highly fractured and intruded by basaltic dikes, represent early fissural eruptions, overlain by rapidly cooled aphanitic rhyolite flows and domes, followed by trachyte with flow banding and cooling joints, and capped by friable pyroclastic tuff and strongly welded ignimbrite. Overlying these volcanic sequences are undivided alluvial, fluvial, and lacustrine sediments: well-sorted gravel–sand–silt alluvium deposited on low-lying plains; poorly sorted angular fluvial talus fans along escarpments; and welded lacustrine pyroclastics exhibiting high strength and a bell-like sound upon impact. The Mursi Formation consists of fine-grained mafic volcanic rocks formed by fissural lava outpourings, whereas the overlying Pliocene–Pleistocene Nkalabong Formation contains volcanic ash and tephra crucial for stratigraphic correlation and paleoenvironmental reconstruction; NE dominates the structural framework–SW and secondary NW–SE fault systems aligned with the Ethiopian Rift, controlling drainage and promoting rock-mass weakening along fault planes through fracturing and weathering (Martínek, K., et al., 2021).

3.2 Landslide inventory

In this study, data were collected through a combination of Google Earth Pro analysis and field surveys. A total of 280 active and inactive landslides were recorded, forming the basis of the inventory map. From these, 196 landslides (70%) were randomly selected to develop the landslide susceptibility index, while the remaining 84 (30%) were reserved for model validation.

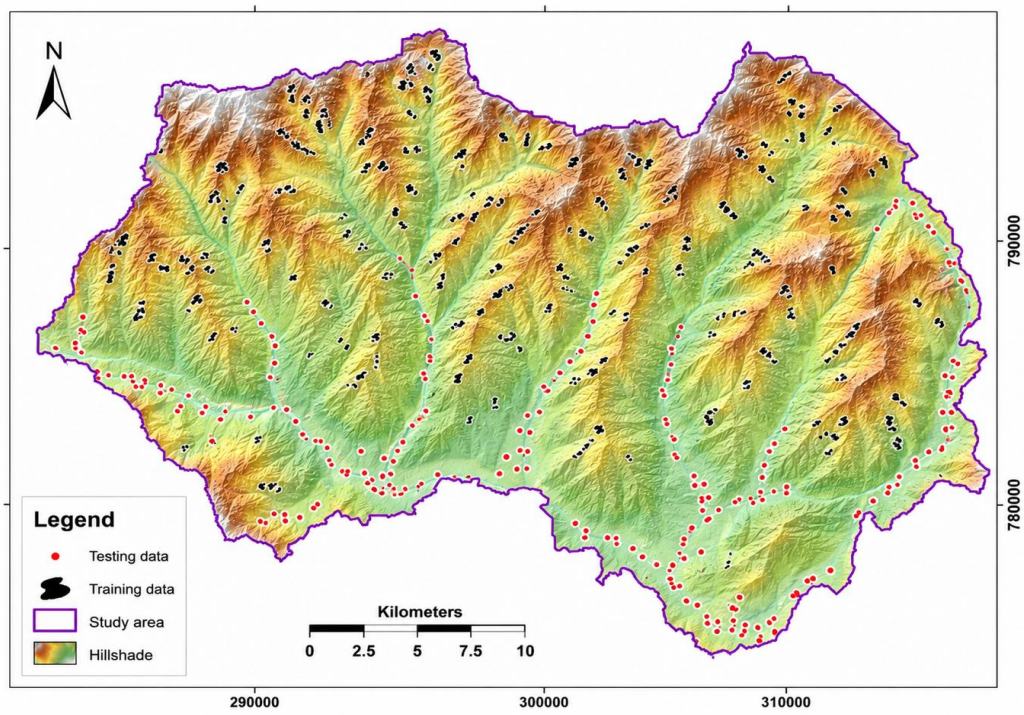


Figure 3: Landslide inventory map of the study area

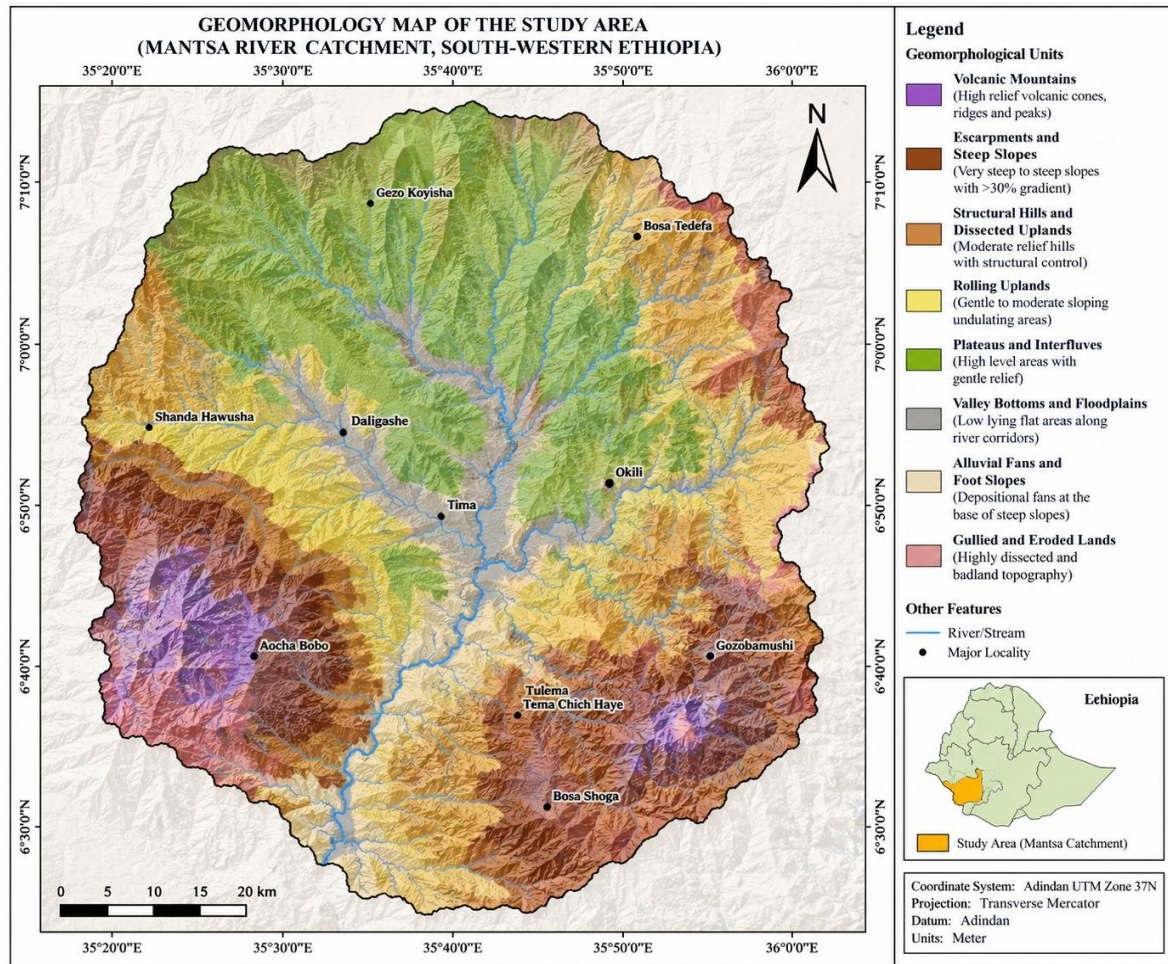


Figure 4. Geomorphology map of the study area

3.2.1 Type of landslide in the study area

Landslides can be triggered by a variety of landslide-causative factors, including geomorphometric factors, geologic units, distance factors, groundwater, land use/land cover, and rainfall, as further elaborated by Shano L. (2021). According to Cruden and Varnes (1996), landslide classification systems employ multiple criteria, including movement pattern characteristics the type of material involved, and the nature of triggering activities. In the present study area, three main types of landslide failure mechanisms were identified: rock falls, debris slides, and earth flows. Rock falls are predominantly composed of individual rock fragments and are mainly caused by the presence of vertical and horizontal joints and by highly weathered materials with fractures. Debris slides involve alluvial, fluvial, and lacustrine sediments during intense rainfall. Earth flows occur in decomposed basalt, rhyolite, trachyte, tuff, and ignimbrite, as well as in alluvial sediments, and

are often associated with intensive farming, high rainfall, and extensively weathered slope materials. These failure mechanisms reflect the complex relationship between geological materials and environmental and human-induced factors in the study area (Figure 5).



Figure 5. Examples of past and recent landslides in the study area include debris slides (A), earthflows (B), and rockfalls (C). There were several similar landslides in the study area (Dawro Zone) caused by factors with predictive rates explained in Figure 19.

3.3 Characterisation of Landslide Factors and their Effects

Landslide factors were arranged and characterized according to predictive rate (Figure 8) of the model (Most to Least: LULC, Lithology, Rainfall, Proximity to Lineament, Proximity to stream, Stream power index (SPI), Soil, Elevation, Aspect, Proximity to road, TWI, Slope, and Curvature), respectively.

3.3.1 Land use land cover (LULC)

In the present analysis, LULC classification distinguishes Cultivation land, Moderate cultivation land, Forest, and Woodland (Figure 6). The frequency-ratio results indicate that areas under moderate cultivation have the highest landslide likelihood. In contrast, the Cultivation, Forest, and Woodland classes show FR values below 1, indicating relatively low susceptibility. These patterns suggest that moderate cultivation, likely representing intensive or poorly managed agricultural use on slopes, may exacerbate slope instability by removing vegetation cover, disrupting soil structure, and increasing surface runoff. In contrast, forested and woodland areas provide stabilising effects through root reinforcement and increased soil cohesion. Therefore, the LULC distribution in the study area plays a critical role in influencing slope stability, with moderate cultivation zones constituting the most vulnerable LULC class for landslide occurrence (Appendix A). Besides, this factor would be found to be the most (4.9) predictive variable out of 13 factors (Figure 6) (Figure 19). This is interpreted as loss of forest cover. Expansion of anthropogenic land uses (e.g., agriculture, built-up areas) is strongly linked to rising landslide susceptibility because the reduction of vegetation and root anchoring weakens slopes and increases erosion (Depicker et al., 2021; Pacheco Quevedo et al., 2023; Zhou et al., 2024).

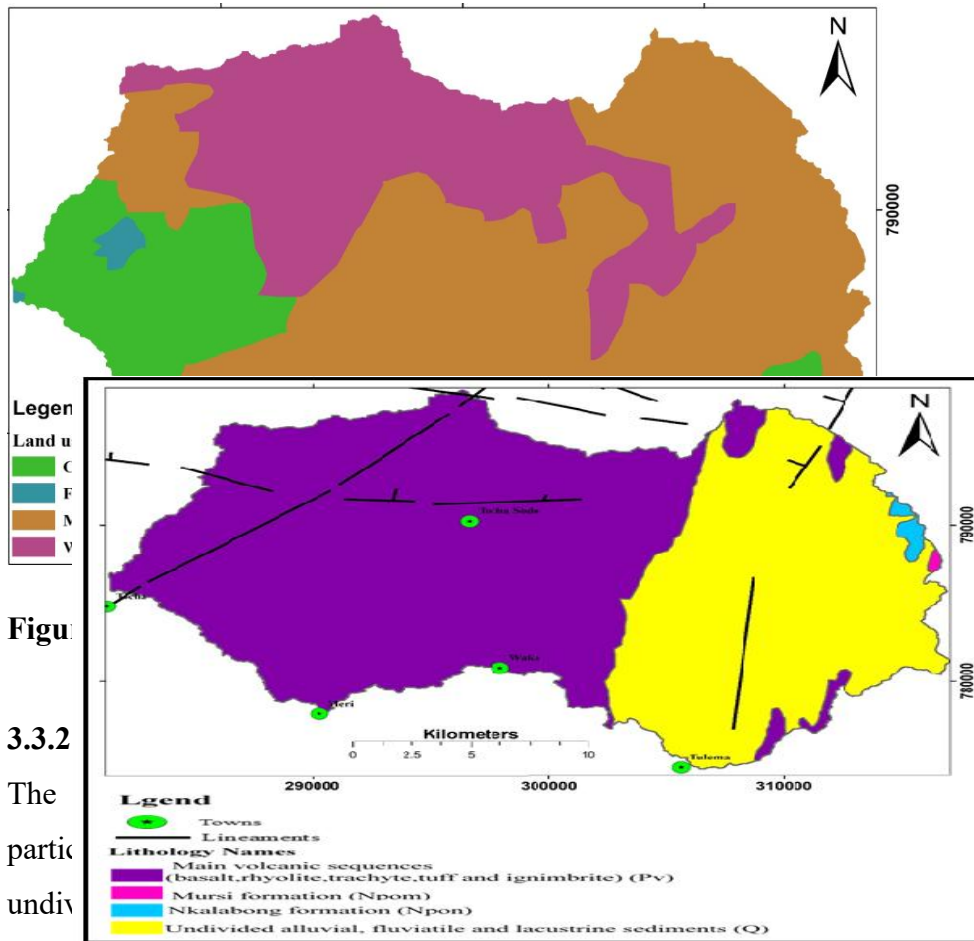
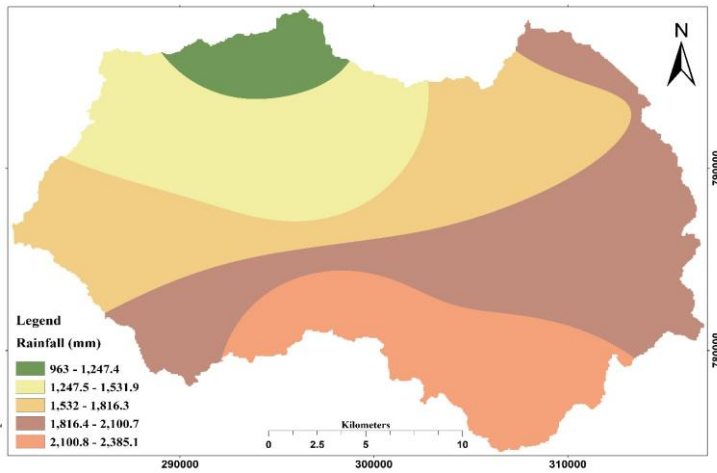


Figure 3.3.2

The partic undiv

thin the study area, and ignimbrite) and

landslide susceptibility. In contrast, the Mursi and Nkalabong formations show no recorded landslide events (Figure 7)(Appendix A). Elevated values for these units reflect their increased vulnerability, primarily due to extensive weathering, high fracturing, and their occurrence on moderately high to steep slopes, which collectively reduce slope stability, and it is the second most predictive variable of the 13 factors within the landslide model(Equation 11) for the present study. The combination of fractured volcanic rocks, weathered igneous deposits, and lineaments further increases slope failure potential by facilitating groundwater infiltration and reducing material cohesion . These findings align with recent studies demonstrating that lithology, particularly fractured and weathered rocks, is a dominant conditioning factor in landslide susceptibility mapping, influencing both slope stability and failure mechanisms (Gao et al., 2021; Sharma et al., 2022; Mengistu et al., 2023). Understanding lithological controls (rock genesis and geo-mechanical behaviour) is therefore essential for delineating high-risk zones and guiding targeted landslide mitigation strategies in complex geological terrains.

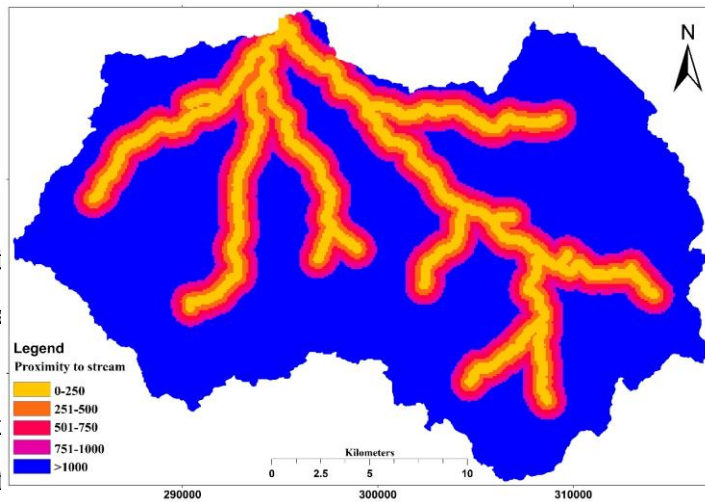


3.3.3 Rainfall

Rainfall patterns derived from 2011–2021 gauging data and interpolated using the Kriging method show that the southern, south-western, and southeastern regions receive higher precipitation. In contrast, the northern and northwestern regions experience relatively low rainfall. Frequency-ratio analysis indicates that higher-elevation zones, particularly between 2100.8 m and 2385.1 m, exhibit the highest landslide susceptibility, with a FR value of 2.96, reflecting the strong influence of rainfall as a triggering factor (Figure 8 & Appendix A). This relationship demonstrates a clear positive correlation between rainfall, elevation, and landslide occurrence, underscoring the importance of rainfall in landslide susceptibility modeling and risk mitigation (Meng et al., 2021; Li et al., 2022; Wubalem et al., 2023). The estimated predictive rate for this landslide-causing factor is 3.6, ranking it third among the remaining 13 variables in the recent study area (Figure 8). Elevated rainfall increases soil moisture and pore-water pressure, reducing shear strength and increasing the likelihood of slope failures.

3.3.5 Proximity to stream

Proximity-to-stream map showing stream channels and indicating a marked increase in landslide activity while distances of

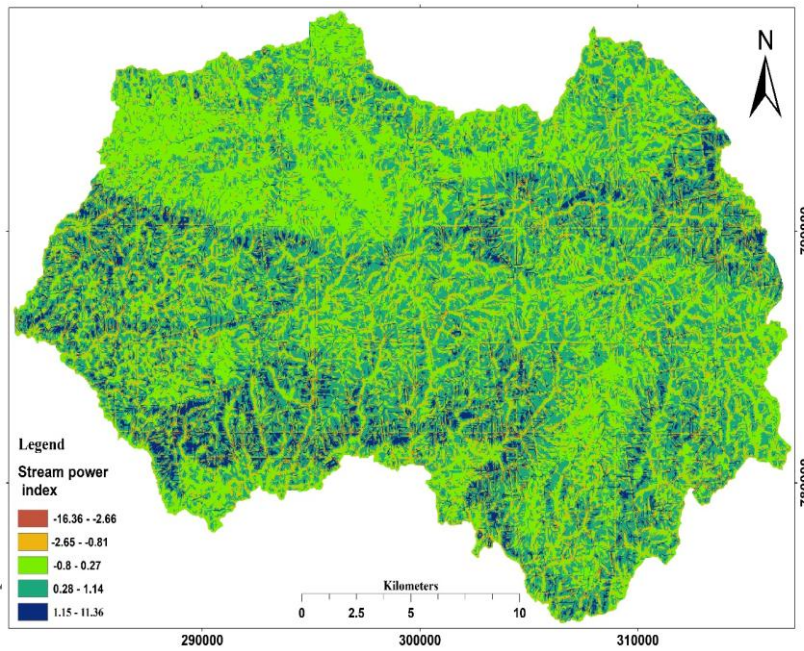


and 501–750 m from stream channels, respectively, indicating a marked increase in landslide activity while distances of

instability with predictive potential of 2.16 (Figure 19). This pattern demonstrates that landslide occurrence increases sharply in near-stream environments where continuous fluvial undercutting, gully incision, and riverbank erosion progressively weaken slope foundations and enhance moisture infiltration, ultimately reducing shear strength. Recent research confirms that stream-adjacent zones exhibit intensified geomorphic disturbance, elevated pore-water pressures, and accelerated mass wasting due to concentrated hydrological loading and toe erosion (Yalcin et al., 2021; Fekadu & Raghuvanshi, 2022; Li et al., 2023). Therefore, proximity to streams emerges as a dominant conditioning factor in the study area, with short-distance classes representing critical hotspots of landslide activity due to the combined effects of mechanical erosion and hydrological destabilization.

3.3. 6

The S



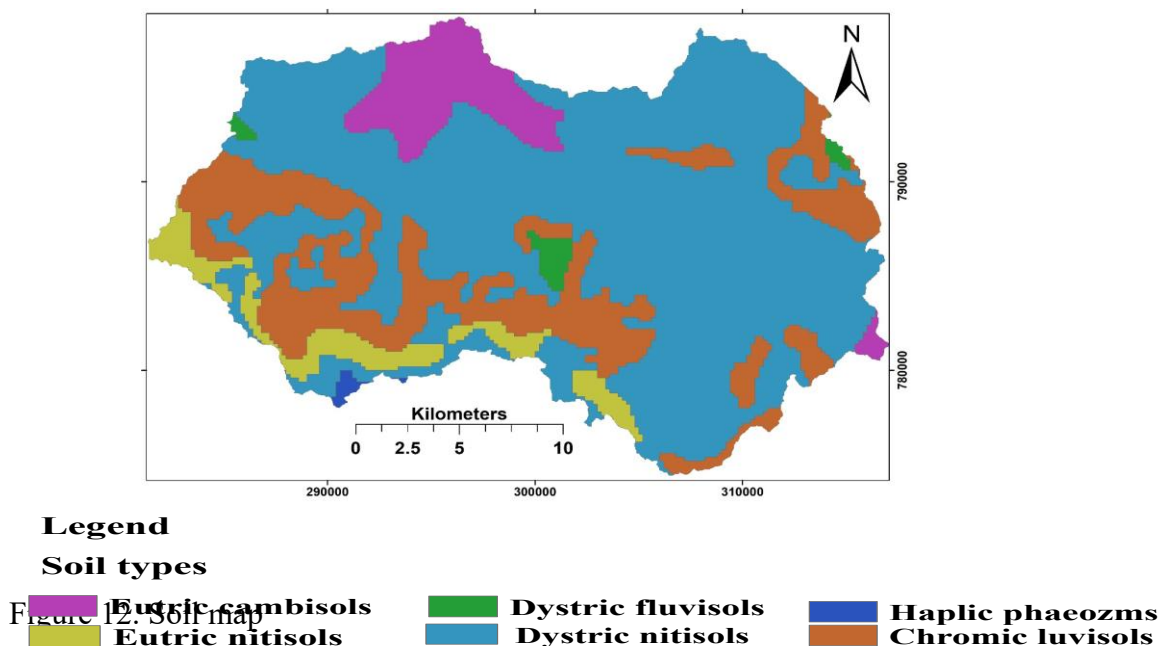
erosive potential, strongly

influences landslide occurrence. Frequency-ratio analysis indicates that the fifth SPI class, representing the highest erosive potential, has an FR of 1.52. In contrast, the third and fourth classes have FR values of 1.00 and 1.04, respectively (Appendix A), with a predictive rate of 2.09 in the novel model (Figure 19). Higher SPI values correspond to increased water flow velocity and volume, intensifying erosion and slope destabilisation. The positive correlation between SPI and landslide occurrence highlights that areas with greater erosive energy are more susceptible to slope failures, making SPI a critical metric for identifying high-risk zones and guiding mitigation strategies in highly rugged landscapes (Li et al., 2021; Chen et al., 2022; Tesfaye et al., 2023).

Figure 11. Stream power index map of the study area

3.3.7 Soil

The soil types within the study area are dominated by Haplic Phaeozems, Chromic Luvisols, Dystric Nitisols, Eutric Cambisols, and Fluvisols (Figure 12 and Appendix A), with frequency-ratio analysis indicating that Haplic Phaeozems soils are most susceptible to landslides, exhibiting FR values of 1.52, having predictive rate (2.08) (Figure 19). These FR amounts suggest that slopes under these soil types are particularly prone to failure due to their physical properties, such as lower cohesion, higher erodibility, and greater sensitivity to moisture changes, which reduce slope stability under environmental and anthropogenic stresses. Soil characteristics, including texture, permeability, and organic matter content, in combination with topography, land use, and rainfall, strongly influence landslide occurrence (Kumar et al., 2021; Zhang et al., 2022; Tesfaye et al., 2023). Consequently, understanding the spatial distribution of these vulnerable soil types is critical for effective landslide hazard assessment and management, as these soils serve as key conditioning factors controlling slope instability in the region.

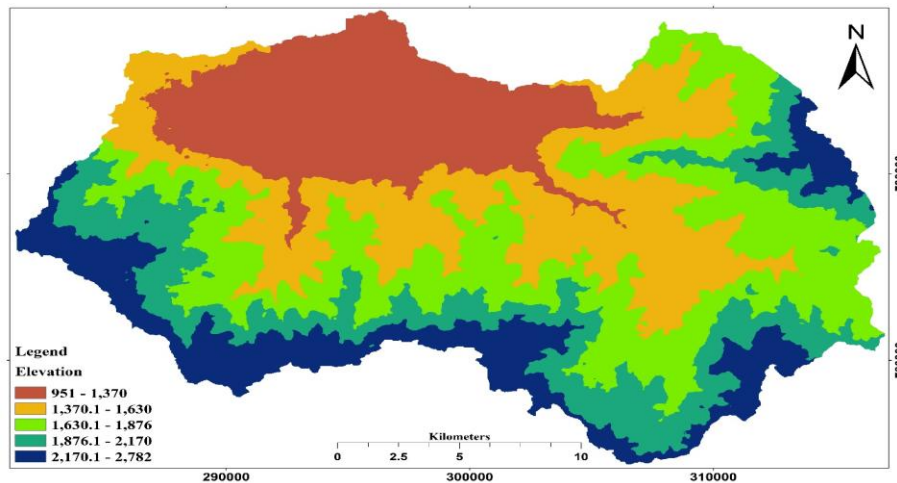


3.3.8 Elevation

Elevation classes in the study area range from 951 to 2170 m. The frequency-ratio analysis shows that elevations of 951–1370 m, 1630.1–1876 m, and 1876.1–2170 m exhibit FR values <1 , indicating low landslide probability, whereas the 1370.1–1630 m and 1630–1876 m classes record FR values >1 , with the latter reaching a maximum FR of 1.5 and Predictive power of 1.75. This pattern suggests that mid-elevation zones experience disproportionately higher landslide occurrence than lower and upper terrain levels. Recent investigations confirm that intermediate altitudes often concentrate landslide activity because these zones typically coincide with steep valley flanks, deeply weathered regolith, and intense human or hydrological disturbances that enhance slope instability (Rahman et al., 2021; Fekadu et al., 2022; Sun et al., 2023). Thus, the results demonstrate that elevation exerts a strong control on landslide distribution, with mid-altitude belts representing critical hotspots of instability within mountainous landscapes.

3.3.9 Aspect

In this study

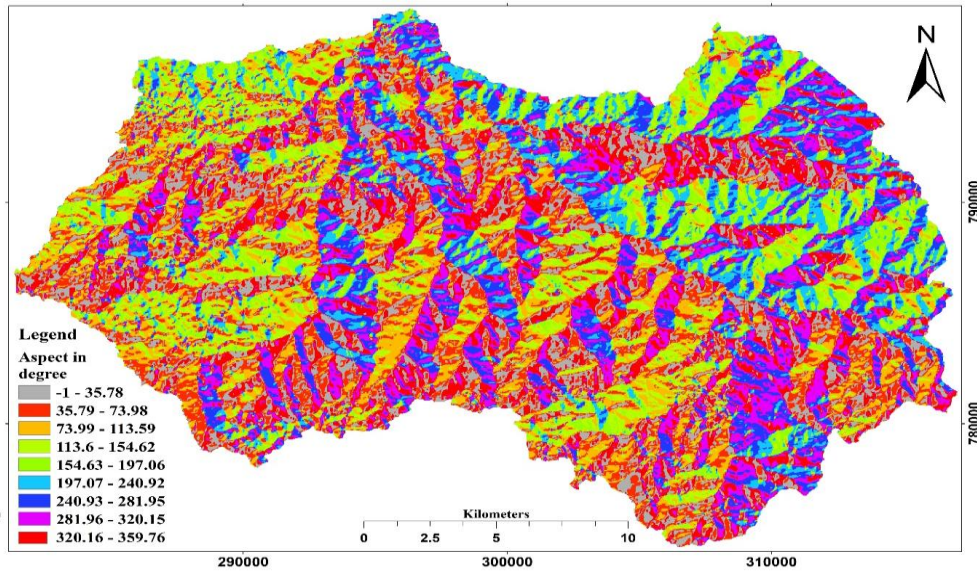


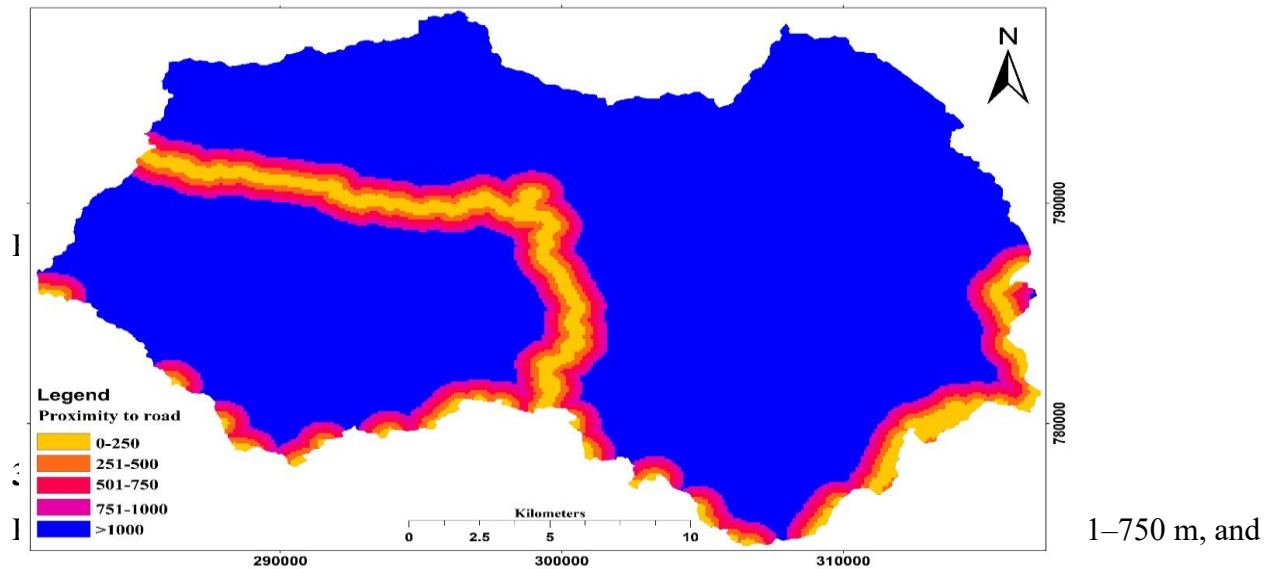
Global Digital Elevation Model (GDEM) of the area, featuring a spatial resolution of 30m by 30m. The generated aspect map was classified into different categorical ranges: Flat (-1-35.78), North (35.78-73.98), Northeast (73.99-113.59), East (113.6-154.62), Southeast (154.63-197.06), South (197.07-240.92), Southwest (240.93-281.95), West (281.96-320.15), and Northwest (320.16-359.76) with Predictive rate: 1.18 (Figure 19). Areas shaded by curvature can intensify the effects of accumulated rainwater; when additional rainfall occurs, these conditions significantly increase the

likelihood of landslides (Wubalem, 2021). The analysis of the aspect reveals varying impacts on landslide occurrences, with distinct frequency ratios across different aspect classes. Specifically, the south slopes ori South (1 frequency the impo significant particular enhanced material s 2020; Wu

ating that s, such as ve lower highlight ation can slopes are ig winds, ly reduce & Meten, ions such

as insolation, moisture retention, and drying rates, which, in turn, modify the hydro-mechanical behavior of soils and contribute to slope instability (Kumar et al., 2021; Shano et al., 2022; Tesfaye & Raghuvanshi, 2023). Thus, aspect constitutes an essential topographic variable in landslide susceptibility assessment, with specific orientations serving as critical hotspots of instability in mountainous terrain.





751–1000 m exhibit elevated frequency-ratio values of 1.92, 1.771, 1.409, and 1.336, respectively (Appendix A) possessing the predictive potential of 1.31 (Figure 19) whereas distances greater than 1000 m record FR values <1 , indicating a reduced likelihood of landslides at farther distances from road networks. These results clearly demonstrate that landslide susceptibility increases as proximity to roads decreases, reflecting the strong destabilising effects of road excavation, cut-slope steepening, removal of toe support, and concentrated surface runoff generated along road corridors. Contemporary research confirms that road construction and associated earthworks amplify geomorphic disturbance, modify natural drainage pathways, and significantly lower the shear strength of adjacent slope materials, thereby enhancing landslide occurrence in near-road zones (Kavzoglu et al., 2021; Raghuvanshi & Ayalew, 2022; Zhang et al., 2023). Thus, road proximity emerges as a critical anthropogenic factor in the study area, with short-distance classes acting as pronounced hotspots of slope instability.

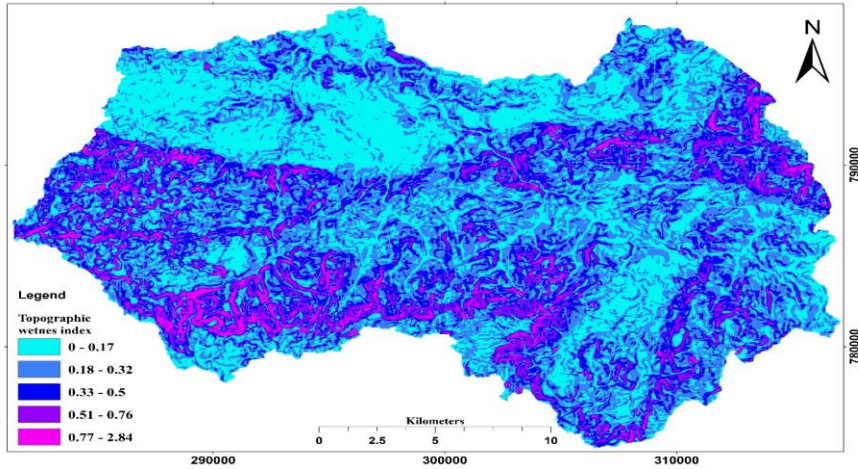
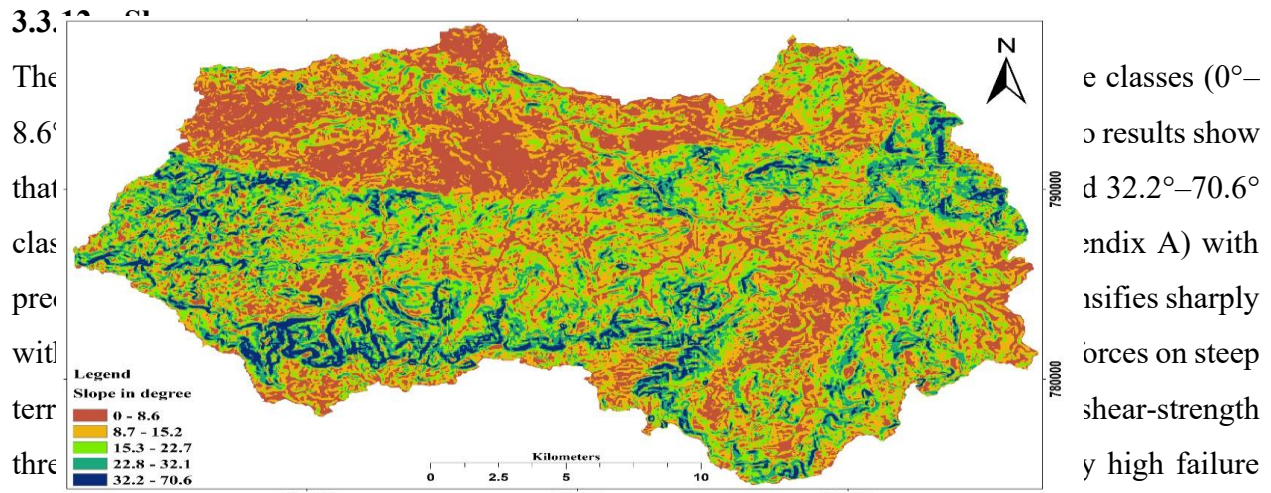


Figure 15 Proximity to road map

3.3.11 Topographic Wetness Index (TWI)

The TWI, which quantifies the influence of topography on soil moisture accumulation, is strongly associated with landslide susceptibility. The class (0.51–0.76) exhibits the highest FR of 1.18, indicating that areas with higher potential wetness are more prone to slope failures (Appendix A) with a predictive value of 1.05. Increased TWI values represent zones where water and weathered materials accumulate, reducing shear strength and destabilizing slopes. This positive relationship between TWI and landslide occurrence demonstrates that high-moisture areas, particularly on steeper slopes, are inherently more vulnerable to landslides. Incorporating TWI into susceptibility assessments improves the spatial prediction of landslides and informs targeted mitigation planning (Zhang et al., 2021; Mengistu et al., 2022; Li et al., 2023).

Figure 16 Topographic wetness index map

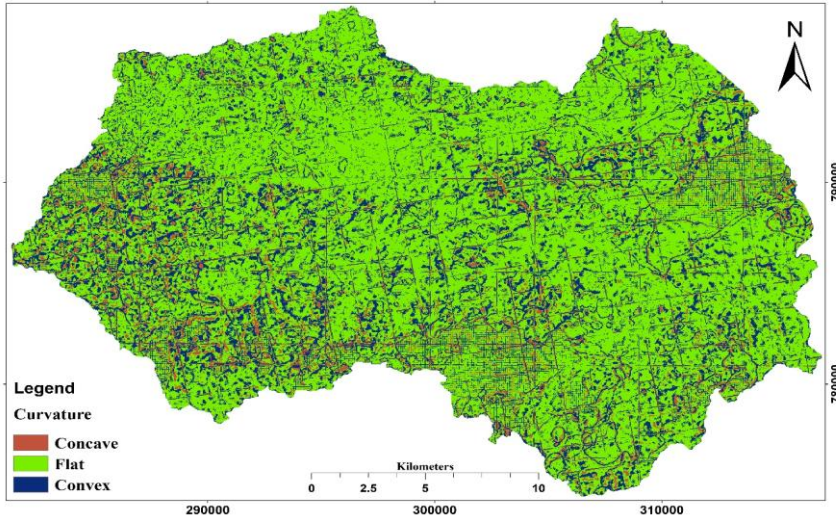


rates (Zhang et al., 2021; Shano et al., 2022; Wang et al., 2023). Within the study area, the minimum slope angle is 0° , while the maximum is 70.6° . Research by Kebeba et al. (2024) underscores the correlation between slope angle and landslide probability, noting that as the slope steepens, the gravitational force acting on the materials increases, thereby increasing susceptibility to landslides. Steeper slopes are inherently more prone to sliding, as highlighted by Gómez & Kavzoglu (2005) and further elaborated by Firomsa & Abay (2019). Landslide occurrence is directly influenced by the slope angle of a given area, as demonstrated by Lee et al. (2004). Consequently, the observed pattern confirms that slope angle is a primary predictor of landslide susceptibility in mountainous landscapes, particularly where highly weathered geological materials amplify the reduction in resisting forces.

Figure 17 S

3.3.13 Cu

Curvature i
frequency-r
prediction c
curvature c
pattern sug



ies (Figure 1). The
pendix A) with a
t, concave and flat
e occurrence. This
water, the convex

slopes in this terrain, characterized by weathered soils and fractured rock, are more prone to failure due to intensified surface runoff and the concentration of erosional forces along divergent flow paths. (Duman et al., 2005; Riaz et al., 2018). Recent studies support that convex topography enhances shear stress and promotes material detachment, while concave zones primarily influence subsurface water convergence, collectively shaping distinct landslide responses across curvature types (Chen et al., 2021; Shano & Raghuvanshi, 2022; Guo et al., 2023). Accordingly, curvature exerts a crucial geomorphological control on slope instability, with convex landforms representing critical hotspots of landslide initiation in the study area. (Chen et al., 2020; Shano et al., 2021; Wubalem, 2021)

Figure 18 Curvature map of the study area

3.4 Landslide Susceptibility Mapping by Frequency Ratio Method

In the frequency ratio model, the landslide susceptibility index (LSI) in the study area ranges from 0 to 2.962 (Table 2). To prepare the landslide susceptibility map (Fig. 37), the landslide susceptibility index maps for each conditioning factor were reclassified using the natural break classification method. The landslide susceptibility index values were classified into five susceptibility classes (Table 4): very high (8.20%), high (16.37%), moderate (29.72%), low (37.59%), and very low (8.12%). The LSI map shows that low-susceptible (37.59%) and very low-susceptible (8.12%) zones are mainly distributed around the north and central directions of the study area. Very high susceptible (8.2%), high susceptible (16.37%) zones are mainly dominated around southwest (Shanda Hawusha, Daligashe, Tima, Okiliand Aocha Bobo) northeast (Gezo Koyisha and Bosa Tedefa), and southeast (Gozobamushi, Tulema Tema Chich Haye, Bosa Shoga) directions of the study area. The moderate-susceptible (29.72%) zone is distributed throughout the study area.

Table 2 Landslide susceptibility classes and corresponding area coverage based on the frequency ratio (FR)

No.	Landslide-Susceptible in frequency ratio	Classes	Frequency Ratio Model		Area in Percentage
			Area (Pixel)	Area in km ²	
1	Very Low Susceptible		49696.00	44.73	8.12
2	Low Susceptible		230092.00	207.08	37.59
3	Moderate Susceptible		181912.00	163.72	29.72
4	High Susceptible		100200.00	90.18	16.37
5	Very High Susceptible		50201.00	45.18	8.20

3.5 Landslide Susceptibility Modeling

Logistic Regression (LR) was applied to evaluate the relationship between landslide presence (1) and absence (0) using 13 conditioning parameters (Table 4), following the classical formulation (Equation 10 & Equation 11). Model fitting was performed in SPSS 24, and significance testing revealed that slope, proximity to lineament, LULC, lithology, and rainfall exert the strongest influence on landslide occurrence ($p < 0.05$) (Table 4 and Table 5). The final LR model (Equation 11) integrates hydrological, topographic, geological, soil, and climatic variables, demonstrating that higher SPI, TWI, rainfall, slope, and convex curvature increase susceptibility, consistent with established findings (Chen & Wang, 2007; Shano et al., 2021).

Positive coefficients indicate increased landslide probability, whereas negative values (e.g., aspect, proximity to road) reduce likelihood when other variables are constant (equation 11). The computed Z-values were converted into a Landslide Susceptibility Index (LSI) raster and classified into five susceptibility classes using natural breaks (Jenks) (Figure 23). The spatial outputs indicate that very high (8.29%) and high (18.68%) susceptibility zones cluster along steep volcanic highlands of the southwest, northeast, and southeast. In contrast, low and very low susceptibility zones dominate central and northern low-relief areas (Figure 23). The distribution aligns with lithologic exposures of basalt–rhyolite complexes, fault-lineament density, and cultivated slopes, supporting conclusions from recent susceptibility studies in similar terrains (Arabameri et al., 2021; Pham et al., 2023; Meten et al., 2025). Overall, the LR model effectively captures the multivariate interactions driving slope instability. Its susceptibility patterns correspond closely with mapped historical landslides and geomorphological controls.

$$Z = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n \quad (10)$$

Where Z represents the linear combination, b_0 is the intercept of the equation and $b_1, b_2, \dots, b_n$ are the coefficients of variables $X_1, X_2, \dots, X_n$. It follows that logistic regression (LR) involves fitting the above equation to the data, where b_0 is the intercept of the model, b_i ($i = 0, 1, 2, \dots, n$) represents the coefficients of the logistic regression model, and X_i ($i = 0, 1, 2, \dots, n$) denotes the independent (Lee & Sambath, 2006). The Z value, derived from the coefficients of these independent variables, aids in predicting the probability of landslide occurrences across all pixels in the study area, as expressed in Equation 11. This process culminated in the creation of a landslide susceptibility index (LSI) map by converting the results to raster format (Figure 23) and (Equation 11).

$$Z = 1.223 * \text{Slope} + 0.127 * \text{Elevation} + 0.816 * \text{Curvature} + (-0.225) * \text{Aspect} + 0.635 * \text{Stream} + (-0.654) * \text{Prox. Road} + 1.616 * \text{Prox. Lineament} + 0.998 * \text{LULC} + 0.173 * \text{Soil} + 0.685 * \text{Lithology} + 0.831 * \text{RF} + 0.037 * \text{SPI} + 0.150 * \text{TWI} - 22.581 \quad (11)$$

3.5.1 Validation and Comparison of Models

Model evaluation used 70% of the landslides for training and 30% for testing (Figure 2). Predictive accuracy was assessed using Receiver Operating Characteristic (ROC) curves, Area Under the Curve (AUC), multicollinearity diagnostics, and overall goodness-of-fit measures (Figures 20 and 21) and (Tables 4, 5, and 6).

3.5.2 Landslide Inventory and FAD Completeness

A total of 280 active and inactive landslides were recorded. Earth flows (65%) dominate, occurring in decomposed volcanic rocks and often associated with intensive farming and high rainfall. The observed frequency-area distribution conformed well to the theoretical inverse-gamma model, with $R^2 = 0.87$ and a Kolmogorov–Smirnov D-statistic of 0.09 ($p = 0.24$), indicating no statistically significant difference between observed and theoretical distributions. The rollover point (the most frequent landslide area) was $\sim 300 \text{ m}^2$, below which under-sampling of very small landslides is possible. The power-law exponent for landslides larger than the rollover was $-2.25 (\pm 0.18)$, consistent with global inventories (range -2.2 to -2.5). These metrics confirm that the inventory is sufficiently complete for regional-scale susceptibility modeling (Korup et al., 2024; Huang et al., 2024).

Integration for FR and LR : In this study, the Frequency Ratio (FR) model was used to generate empirical landslide susceptibility indices by quantifying the spatial association between historical landslide pixels and individual factor classes, while the Logistic Regression (LR) model was applied to calculate the net probability of landslide occurrence by estimating the multivariate relationship between all conditioning factors simultaneously. The integration of FR and LR provides a complementary framework: FR offers a transparent, class-based ranking of factor importance for spatial planning, whereas LR yields a probabilistic susceptibility map that accounts for interactive effects and allows for formal statistical validation, thereby strengthening the

scientific basis for risk-informed decision-making. FR results are presented separately while LR focuses exclusively on Logistic Regression modeling, equations, coefficients, and susceptibility outputs.

3.5.3 ROC of AUC Evaluation

The Frequency Ratio (FR) model achieved an AUC of 0.874 (Figure 20), indicating very good predictive performance. The LR model achieved an AUC of 0.85 (Figure 10), indicating it is a good model. Both exceed the 0.80 threshold commonly recognised as very good predictive strength (Yesilnacar & Topal, 2005). Although FR slightly outperformed LR, as also reported in recent comparative studies (Fanos et al., 2022; Hong et al., 2023), both models consistently distinguish unstable from stable zones (Figures 20 and 21).

The final LR validation results have now been incorporated. The predictive performance of the LR model achieved an AUC value of 0.850 (95% CI: 0.804–0.896), while the FR model achieved an AUC of 0.874. Both models therefore demonstrate very good predictive capability, with FR showing only marginally higher performance. Landslide inventory points have now been overlaid on the susceptibility maps to facilitate visual comparison between observed landslides and predicted susceptibility zones while maintaining overall map readability.

3.5.4 Multicollinearity and Goodness of Fit

Tolerance and VIF values, computed from equations 8 & 9, confirm the absence of multicollinearity, with all tolerance values > 0.10 and $VIF < 10$, validating the independence of the variables (Table 7). (The Omnibus test (Table 4) yielded $\chi^2 = 403.781$, $p < 0.001$, indicating that the LR model significantly improves over the null model. The Hosmer–Lemeshow test ($p = 0.643$; Table 5) confirms an excellent fit between predicted and observed landslides. The Nagelkerke R^2 of 0.885 (Table 6) indicates that 88.5% of landslide variability is explained by the model, consistent with the high explanatory power observed in recent LR-based susceptibility assessments (Zhao et al., 2024) (Shano et al., 2022). Therefore, the models' performance has been evaluated and interpreted as follows: the frequency ratio (FR) model identifies 16.37% (high and very high susceptibility) zones, whereas Logistic regression (LR) identifies 18.68% and 8.29% for the high and very high classes, respectively (Tables 3–7). Spatially, both models consistently map

high susceptibility along structural lineaments, steep volcanic escarpments, and intensively cultivated slope areas characterized by geomorphological convergence of runoff, weakened lithology, and human disturbance. Moderate zones (28.87%) are more widely dispersed across transitional landscapes. Together, the results demonstrate that both FR and LR models exhibit vigorous predictive capability, with FR performing slightly better but LR offering clearer interpretability of factor contributions. These assets support their application for regional planning, hazard mitigation, and prioritisation of early warning.

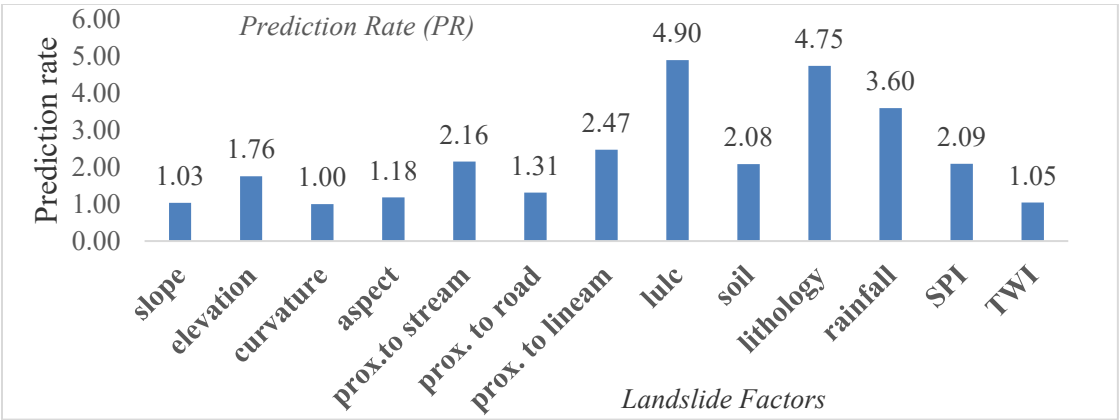


Figure 19: Landslide causative factor influence on landslide in the study area using the frequency ratio model

Table 3 Landslide susceptibility classification and area coverage based on FR

No	Landslide Susceptible Classes	Frequency Ratio Model		
		Area (Pixel)	Area in km ²	Area in Percentage
1	1	40198	36.18	6.56
2	2	230227	207.22	37.59

3	3	176847	159.20	28.87
4	4	114422	102.60	18.68
5	5	50407	45.69	8.29
Total		612101	550.89	100.00

Table 4 Data fit and model validation in Omnibus tests of model coefficients

Omnibus Tests of Model Coefficients				
Chi-				
square				
df				
Sig.(P)				
Step 1	Step	403.781	13	.000
	Block	403.781	13	.000
	Model	403.781	13	.000

Table 5: Data fit and model validation in the Hosmer & Lemeshow test

Hosmer and Lemeshow Test			
Step	Chi-square	Df	Sig.(P)
1	6.034	8	.643

Table 6: Data fit and model validation in Nagelkerke R Square

Model Summary				
-2 Log Likelihood				
Step	-2 Log Likelihood	Cox & Snell R Square	Nagelkerke R Square	

1	49.851 ^a	.664	.885
---	---------------------	------	------

Table 7 Data fit and model validation in collinearity statistics

Factors	Coefficients	
	Collinearity Statistics	
	Tolerance	VIF
Slope	.376	2.660
Elevation	.500	2.002
Curvature	.665	1.504
Aspect	.942	1.062
Proximity to stream	.697	1.436
Proximity to road	.616	1.623
Proximity to lineament	.351	2.852
Landuse landcover	.676	1.480
Soil	.842	1.187
Lithology	.871	1.149
Rainfall	.740	1.352
Stream power index	.624	1.604
Topographic wetness index	.734	1.363

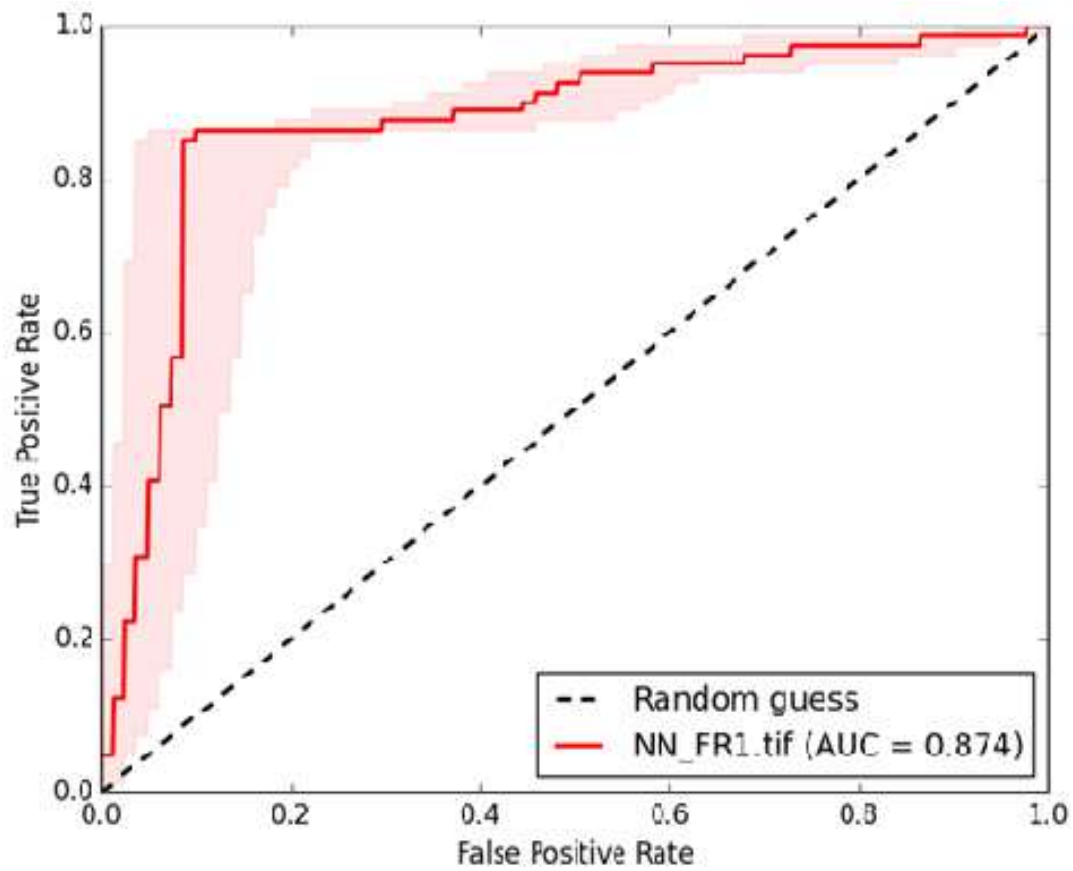


Figure 20 Model performance of the ROC curve of the landslide susceptibility index map for frequency ratio (FR)

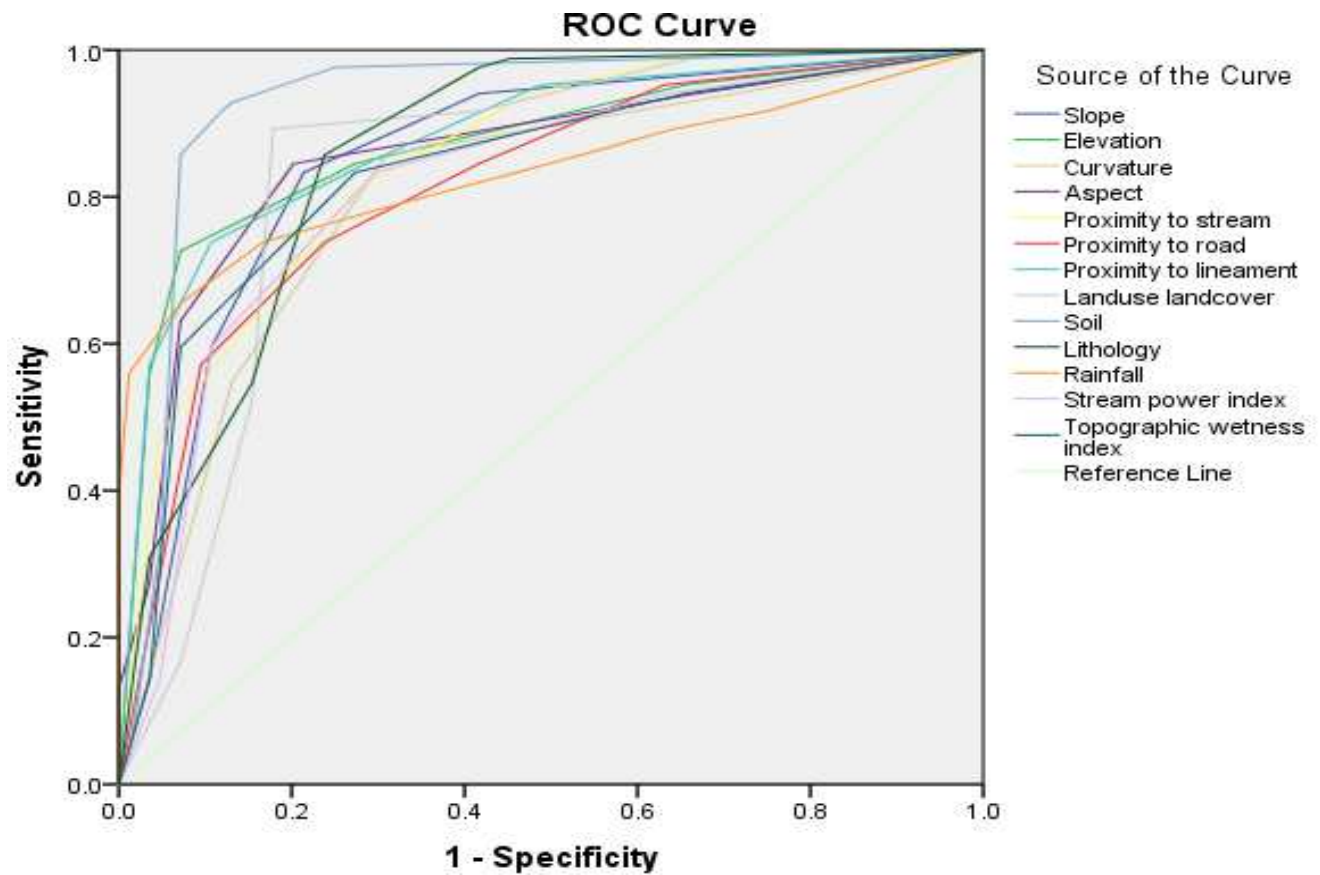


Figure 21 Model performance of the ROC curve of the landslide susceptibility index map for logistic regression (LR)

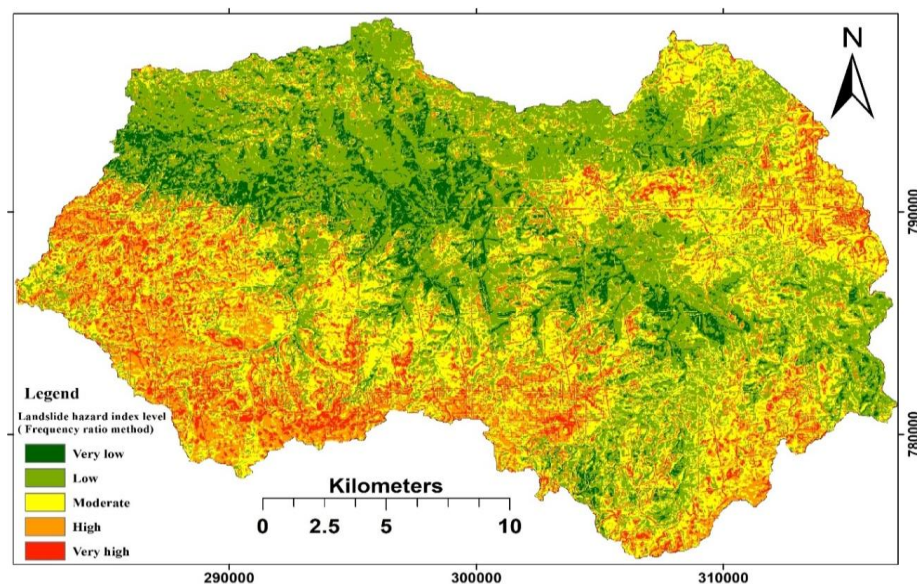


Figure 22 Landslide susceptibility map generated by FR

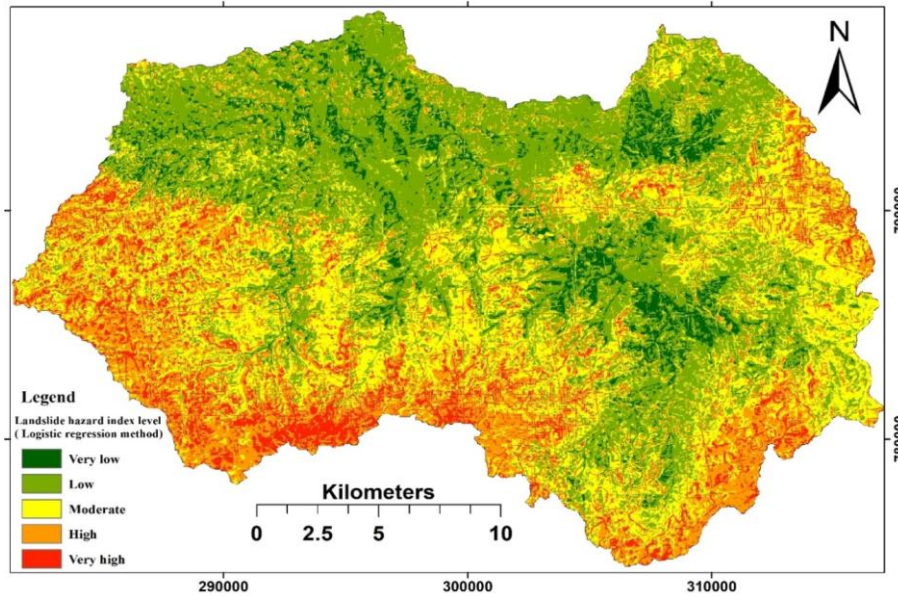


Figure 23 Landslide susceptibility map produced by LR.

Aspect classes were derived following standard geomorphological and terrain analysis conventions, dividing slope orientation into Flat, N, NE, E, SE, S, SW, W, and NW categories.

4. Discussions

The primary goal of this study was to develop and validate predictive landslide susceptibility models for the Mantsa River catchment using Frequency Ratio (FR) and Logistic Regression (LR). The results demonstrate that both statistical models possess robust predictive capability, with the FR model (AUC = 0.874) marginally outperforming the LR model (AUC = 0.850). This slight superiority of FR in data-scarce, high-relief terrains has been observed in other comparative studies within the Ethiopian highlands (Tesfaye et al., 2024; Sisay et al., 2024), likely due to FR's non-parametric nature which does not assume a linear relationship between predictors and landslide occurrence (Pradhan, 2013). However, the LR model's strength lies in its ability to quantify the net effect of each factor while controlling for interactions, offering clearer interpretability for risk management (Shano et al., 2022).

The study area's geomorphology is dominated by rugged topography with elevations ranging from 951 m to 2,782 m, deeply incised valleys, and a dendritic drainage pattern. The slope map (0° – 70.6°) revealed that 32.2° – 70.6° class has the highest FR (1.44). This is consistent with global findings that steep slopes ($>30^{\circ}$) exceed the shear-strength threshold of weathered soils and fractured bedrock (Wang et al., 2023). The geomorphological map (Figure 2) identifies escarpments, deep gorges, and colluvial footslopes as the primary loci of instability, where gravitational driving forces are maximized.

Geologically, the area is underlain by extensive syn-rift volcanic sequences of the Ethiopian Rift. The presence of mafic rocks (aphyric basalt, porphyritic trachyte) and felsic pyroclastics (unwelded tuff, welded ignimbrite) creates a dual problem for slope stability.

Mafic rocks, particularly the highly fractured basalts, are susceptible to rapid chemical weathering under the region's hot and humid climate (annual rainfall up to 2,194 mm) due to their mafic minerals composition. This weathering transforms competent rock into thick, clay-rich saprolites (smectite and kaolinite) that have low internal friction angles and high shrink-swell capacity (Vidal et al., 2022). When saturated during the heavy April-May rainy season, these clayey soils lose shear strength, leading to earthflows—the dominant failure type (65% of inventory). In contrast, the overlying unwelded tuff and ignimbrite are friable and weakly consolidated, acting as aquifers that percolate water down to the less permeable basalt contacts, creating zones of high pore-water pressure. The undivided alluvial, fluvial, and lacustrine sediments occupy valley bottoms and footslopes, where they are undercut by streams, resulting in debris slides (Mengistu et al., 2023). The NE-SW and NW-SE fault systems, aligned with the Main Ethiopian Rift, further enhance instability by fragmenting rock masses, creating preferential pathways for water infiltration, and serving as planes of weakness for rockfalls (Martínek et al., 2021). The FR analysis confirmed this, with proximity to lineaments (0–250 m) yielding the third-highest FR (2.96), indicating that tectonic weakening is a first-order control.

LULC emerged as the most influential factor (predictive rate 4.9), with moderate cultivation land showing $FR = 1.52$ and accounting for 75.95% of landslide pixels. This is a critical finding: agricultural expansion onto steep slopes ($>15^{\circ}$) removes natural forest root reinforcement, disrupts soil structure through ploughing, and alters drainage pathways, directly increasing susceptibility (Depicker et al., 2021; Pacheco Quevedo et al., 2023). Forest and woodland classes had $FR < 0.25$, underscoring their protective role. This suggests that reforestation and agro-ecological zoning should be high-priority mitigation strategies. Rainfall was the third most predictive factor ($FR = 2.96$ for 2100–2385 mm zone). The May peak rainfall (8.47 mm monthly average) saturates the weathered regolith. This aligns with the Gofa 2024 disaster, where a multi-day extreme rainfall event triggered catastrophic slope failures on similar volcanic terrain (Shano et al., 2024; Nwankwo et al., 2024). The positive coefficients for SPI (1.52) and TWI (1.18) in the

LR model confirm that zones of concentrated overland flow and soil moisture accumulation are hotspots for failure initiation (Li et al., 2021; Zhang et al., 2021).

The ROC-AUC values (FR=0.874; LR=0.850) are comparable to or better than recent studies in Ethiopia: Asmare et al. (2024) reported 81.45% for a fused FR-LR model in the Muger watershed, while Tesfaye et al. (2024) achieved 72.16% using FR along the Debretabor-Alember road. Our higher performance is likely due to the larger, field-verified inventory (n=280) and the inclusion of rainfall as a dynamic factor. The Hosmer-Lemeshow test ($p=0.643$) and Nagelkerke R^2 (0.885) for the LR model indicate an excellent fit, meaning that the 13 selected factors explain nearly 90% of the variance in landslide distribution. The absence of multicollinearity ($VIF < 3$ for all factors) confirms that each factor contributes unique information. The high- and very high-susceptibility zones (24.57–26.97% of the area) are concentrated in the southwest (Shanda Hawusha), northeast (Gezo Koyisha), and southeast (Gozobamushi). These zones coincide precisely with the convergence of: (i) steep slopes ($>22.8^\circ$) on weathered basaltic terrain; (ii) high drainage/lineament density; (iii) moderate cultivation land use; and (iv) high annual rainfall (>1816 mm). The low-susceptibility zones (45.71%) dominate the northern and central low-relief plateaus, where slopes are gentler and land use is less intensive. The current models provide a scientifically robust, first-generation susceptibility map that can guide land-use planning and early warning in the Dawro Zone.

4.1 Limitations and future research recommendations

Although this study presents the first comprehensive landslide susceptibility assessment for the Mantsa River catchment, the following limitations should be addressed in future research. First, the completeness and reliability of the landslide inventory could be enhanced by developing larger multi-temporal datasets through the integration of unmanned aerial vehicles (UAVs), LiDAR, and high-resolution satellite imagery. Second, the present analysis employed a 30 m digital elevation model (DEM) to maintain consistency with the spatial resolution of the other thematic layers used in the susceptibility modelling. Future studies should investigate the use of higher-resolution topographic datasets, such as ALOS PALSAR (12.5 m) or LiDAR-derived DEMs, to better characterize local terrain variability and slope processes. Third, incorporating dynamic triggering factors, including rainfall intensity–duration relationships, antecedent rainfall, groundwater fluctuations, and land-use/land-cover changes, would improve the temporal representation of landslide occurrence and enhance predictive performance. Furthermore, integrating geotechnical parameters such as soil and rock cohesion, internal friction angle, permeability, weathering depth, and pore-water pressure would provide a more realistic representation of slope failure mechanisms. Model evaluation could also be strengthened by

complementing ROC–AUC analysis with additional validation measures, including confusion matrices, precision, recall, F1-score, and cross-validation, thereby providing a more comprehensive assessment of model robustness. Moreover, future studies should explore and compare advanced machine-learning and ensemble techniques, such as Random Forest, Extreme Gradient Boosting (XGBoost), LightGBM, and deep learning models, to further improve susceptibility prediction accuracy. Finally, assessing the transferability of the developed models to neighboring catchments and investigating the influence of projected climate change on future landslide susceptibility patterns would enhance the broader applicability and long-term relevance of landslide susceptibility assessments.

5. Conclusions

- First systematic susceptibility mapping for the Mantsa River catchment (Dawro Zone) was achieved, filling a critical knowledge gap for a region with recurrent deadly landslides but no prior scientific assessment.
- Both Frequency Ratio (FR) and Logistic Regression (LR) models performed very well (AUC = 0.874 and 0.850, respectively), confirming their suitability for data-scarce, high-relief terrains where interpretability is as important as predictive power.
- Approximately 24.6–27.0% of the catchment area is classified as high to very high susceptibility, concentrated in the south-western, north-eastern, and south-eastern sectors, where steep volcanic slopes, dense lineaments, and high rainfall converge.
- Land use/land cover (LULC), specifically moderate cultivation on slopes, is the single most influential factor (predictive rate 4.9), accounting for 75.95% of landslide pixels and highlighting the critical role of anthropogenic pressure in slope destabilization.
- The weathering of mafic volcanic rocks (basalt, trachyte) into expansive clays, combined with steep slopes ($>32^\circ$) and high rainfall (>2100 mm), creates a high-hazard lithological-hydrological syndrome that explains the dominance of earthflows (65% of inventory).
- Proximity to lineaments (0–250 m) and streams (0–250 m) are the strongest structural and hydrological predictors (FR = 2.96 and 2.68, respectively), demonstrating that tectonic fracturing and fluvial undercutting are primary destabilizing processes.
- The FR model is recommended for spatial planning due to its higher AUC and straightforward empirical class-based weighting, while the LR model is better for understanding the net contribution of each factor and for probabilistic risk forecasting.

- The validated susceptibility maps provide an essential evidence base for proactive risk reduction, enabling targeted mitigation (e.g., reforestation of steep cultivated slopes), hazard-informed land-use zoning, and the prioritization of early warning systems in Ethiopia's landslide-prone highlands.

Conflicts of interest

The author declares there is no conflict of interest.

Funding: No funding received.

Ethical approval: Not applicable

Consent to participate: Not applicable

Consent to Publish: not applicable.

Clinical trial number: not applicable.

Data availability: on reasonable request, the corresponding authors will provide the datasets created and/or analyzed during this study.

Authorship contribution statement

Meseret Zinabu: Writing – original draft, Visualization, Methodology, Investigation, Data-curation, Conceptualization, Writing – reviewing and editing. **Leulalem Shano:** Writing – original draft, Visualization, Methodology, Investigation, Data-curation, Conceptualization, Writing – reviewing and editing. **Muralitharan Jothimani:** Writing – original draft, Formal analysis, Writing – reviewing and editing. **Mihretu Fikre:** Writing – original draft, Visualization, Methodology, Investigation, Data-curation, Conceptualization, Writing – reviewing and editing. **Berhanu Erbelo:** Writing – original draft, Visualization, Methodology, Investigation, Data-curation, Conceptualization, Writing – reviewing and editing. **Ephrem Getahun:** Writing-original draft, Visualization, Methodology, Investigation, Data curation, Conceptualization, Writing – reviewing and editing. **Shankar Karuppannan:** Writing – original draft, Visualization, Methodology, Investigation, Data-curation, Conceptualization, Writing – reviewing and editing. Alemayehu Ayele: Writing – original draft, Visualization, Methodology, Investigation, Data curation, Conceptualization, Writing – reviewing and editing.

References

Abdulwahid, W. M., & Pradhan, B. (2017). Landslide vulnerability and risk assessment for multi-hazard scenarios using airborne laser scanning data (LiDAR). *Landslides*, 14(3), 1057–1076. <https://doi.org/10.1007/s10346-017-0804-3>.

- Arabameri, A. (2021). GIS-based landslide susceptibility mapping using hybrid ensemble machine learning. *Landslides*, 18, 2581–2601. <https://doi.org/10.1007/s10346-021-01682-3>.
- Asmare, D.(2024). Spatial assessment employing fusion logistic regression and frequency ratio models to monitor landslide susceptibility in the upper Blue Nile basin of Ethiopia: Muger watershed. *Environmental Systems Research*, 13, 54. <https://doi.org/10.1186/s40068-024-00345-5>.
- Pradhan, B. (2013). A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Catena*, 51, 350–365.
- Çevik, E., & Topal, T. (2003). GIS-based landslide susceptibility mapping for a problematic segment of the natural gas pipeline, Hendek (Turkey). *Environmental Geology*, 44(8), 949–962. <https://doi.org/10.1007/s00254-003-0838-6>.
- Chang, Z., Catani, F., Huang, F., Liu, G., Meena, S. R., Huang, J., & Zhou, C. (2023). Landslide susceptibility prediction using slope unit-based machine learning models considering the heterogeneity of conditioning factors. *Journal of Rock Mechanics and Geotechnical Engineering*, 15(5), 1127–1143. <https://doi.org/10.1016/j.jrmge.2022.12.001>.
- Chen, W., Peng, J., Wang, J., et al. (2020). Landslide susceptibility modeling using GIS-based machine learning. *Remote Sensing*, 12, 356. <https://doi.org/10.3390/rs12030356>.
- Chen, W., et al. (2021). Influence of topographic curvature on landslide initiation in complex mountainous terrains. *Geomorphology*, 387, 107785. <https://doi.org/10.1016/j.geomorph.2021.107785>.
- Chen, X., et al. (2022). Integration of hydrological and topographic indices for landslide hazard assessment. *Natural Hazards*, 112, 1211–1230. <https://doi.org/10.1007/s11069-021-05180-9>.
- Chen, X., Zhou, Y., & Du, P. (2024). Comparative evaluation of statistical models for landslide susceptibility mapping. *Geomorphology*, 439, 108845. <https://doi.org/10.1016/j.geomorph.2024.108845>.
- Chen, Z., & Wang, J. (2007). Landslide hazard mapping using logistic regression model in Mackenzie Valley, Canada. *Natural Hazards*, 42(1), 5–89. <https://doi.org/10.1007/s11069-006-9063-1>.
- Chicas, S. D., Li, H., Mizoue, N., Ota, T., Du, Y., & Somogyvári, M. (2024). Landslide susceptibility mapping core-base factors and models' performance variability: A systematic review. *Natural Hazards*, 120, 12573–12593. <https://doi.org/10.1007/s11069-024-07015-6>.

- Dai, F. C., Lee, C. F., Tham, L. G., Ng, K. C., & Shum, W. L. (2004). Logistic regression modelling of storm-induced shallow landsliding in time and space on natural terrain of Lantau Island, Hong Kong. *Bulletin of Engineering Geology and the Environment*, 63(4), 315–327. <https://doi.org/10.1007/s10064-004-0259-4>
- Dam, N. T., Shano, L., & Alemu, K. (2022). Landslide susceptibility assessment using geospatial techniques. *Natural Hazards*, 113, 1565–1584. <https://doi.org/10.1007/s11069-022-05347-0>.
- Depicker, A., et al. (2021). Vegetation cover and root reinforcement as stabilizing factors in slope stability assessment: A global meta-analysis. *Landslides*, 18, 2583–2601. <https://doi.org/10.1007/s10346-021-01746-3>.
- Duman, T. Y., Çan, T., Emre, Ö., Keçer, M., Doğan, A., Ateş, Ş., & Durmaz, S. (2005). Landslide inventory of northwestern Anatolia, Turkey. *Engineering Geology*, 77(1–2), 99–114. <https://doi.org/10.1016/j.enggeo.2004.09.005>.
- Erener, A., & Düzgün, H. S. B. (2010). Improvement of statistical landslide susceptibility mapping by using spatial and global regression methods in the case of More and Romsdal (Norway). *Landslides*, 7(1), 55–68. <https://doi.org/10.1007/s10346-009-0188-x>.
- Fanos, A. M. (2022). Performance comparison of statistical models for landslide susceptibility. *Catena*, 213, 106248. <https://doi.org/10.1016/j.catena.2022.106248>.
- Fekadu, D., & Raghuvanshi, T. K. (2022). Hydrological controls of channel proximity on landslide susceptibility in tropical highlands. *Landslides*, 19, 3011–3026. <https://doi.org/10.1007/s10346-022-01941-4>.
- Fekadu, D. (2022). Topographic controls on landslide occurrence in high-relief tropical terrains. *Natural Hazards*, 113, 1557–1578. <https://doi.org/10.1007/s11069-022-05288-y>.
- Ferrari, F. (2021). Rainfall thresholds for landslide initiation. *Natural Hazards and Earth System Sciences*, 21, 1247–1264. <https://doi.org/10.5194/nhess-21-1247-2021>.
- Firomsa, M., & Abay, A. (2019). Landslide assessment and susceptibility zonation in Ebantu district of Oromia Region, western Ethiopia. *Bulletin of Engineering Geology and the Environment*, 78(6), 4229–4239. <https://doi.org/10.1007/s10064-018-1407-0>.
- Froude, M. J., & Petley, D. N. (2021). Global fatal landslide occurrence from 2004 to 2020. *Landslides*, 18, 3101–3115. <https://doi.org/10.1007/s10346-021-01664-6>.
- Gebreyohannes, T., & van Westen, C. J. (2022). Influence of lineament proximity on slope instability in fractured bedrock systems. *Natural Hazards*, 114, 1783–1804. <https://doi.org/10.1007/s11069-022-05472-0>.

- Gómez, H., & Kavzoglu, T. (2005). Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela. *Engineering Geology*, 78(1–2), 11–27. <https://doi.org/10.1016/j.enggeo.2004.11.004>.
- Guo, Z. (2023). Topographic convergence and divergence effects on shallow landslide processes. *Engineering Geology*, 323, 107176. <https://doi.org/10.1016/j.enggeo.2023.107176>.
- Gui, G.(2024). Application of artificial intelligence in three aspects of landslide risk assessment: A comprehensive review. *Natural Hazards Research*. <https://doi.org/10.1016/j.nhres.2024.05.006>.
- Haque, U.(2020). Global landslide hazard distribution and impacts. *Scientific Reports*, 10, 1190. <https://doi.org/10.1038/s41598-020-61146-6>.
- Hong, H. (2023). Improved statistical models for regional landslide susceptibility. *Geomorphology*, 435, 108846. <https://doi.org/10.1016/j.geomorph.2023.108846>.
- Huang, R. (2021). Structural discontinuities and their role in controlling landslide activity in tectonically active terrains. *Landslides*, 18, 3025–3040. <https://doi.org/10.1007/s10346-021-01781-y>.
- Kavzoglu, T.(2021). Evaluating the impact of road-induced disturbances on landslide susceptibility mapping. *Natural Hazards*, 108, 2981–3002. <https://doi.org/10.1007/s11069-021-04878-6>.
- Kebeba, O., Shano, L., Chemdesa, Y., & Jothimani, M. (2024). Integration of geospatial analysis, frequency ratio, and analytical hierarchy process for landslide susceptibility assessment in the Maze catchment, Omo Valley, southern Ethiopia. *Quaternary Science Advances*, 16, 100248. <https://doi.org/10.1016/j.qsa.2024.100248>.
- Kebeba, T., Shano, L., & Mulugeta, A. (2024). Landslide susceptibility mapping in southwestern Ethiopia. *Environmental Earth Sciences*, 83, 124. <https://doi.org/10.1007/s12665-024-11324-7>.
- Li, J. (2021). Stream Power Index as a predictor of landslide susceptibility in mountainous watersheds. *Geomorphology*, 385, 107702. <https://doi.org/10.1016/j.geomorph.2021.107702>.
- Li, Y. (2023). Incorporating topographic wetness index in landslide susceptibility modeling: A case study in complex terrain. *Landslides*, 20, 203–218. <https://doi.org/10.1007/s10346-022-01985-3>.
- Li, Y. (2024). Modelling uncertainties and sensitivity analysis of landslide susceptibility prediction under different environmental factor connection methods and machine learning

models. *KSCE Journal of Civil Engineering*, 28, 45–62. <https://doi.org/10.1007/s12205-023-2430-9>.

Nwankwo, C. N. (2024). Fatal landslides in Kencho, Shacha & Gozdi villages, Gofa Zone, Ethiopia: A detailed investigation. *Quaternary Science Advances*, 16, 100241. <https://doi.org/10.1016/j.qsa.2024.100241>.

OCHA/WHO. (2024). Landslide in Geza Gofa Woreda, Gofa Zone, Southern Ethiopia. WHO Regional Office for Africa.

Pei, X., et al. (2024). Sample size effects on landslide susceptibility models: A comparative study of heuristic, statistical, machine learning, deep learning and ensemble learning models with SHAP analysis. *Computers & Geosciences*, 193, 105723. <https://doi.org/10.1016/j.cageo.2024.105723>.

Singh, K., Bhardwaj, V., Sharma, A., & Thakur, S. (2024). A comprehensive review on landslide susceptibility zonation techniques. *Quaestiones Geographicae*, 43(1), 79–91. <https://doi.org/10.14746/quageo-2024-0006>.

Sisay, T., Tesfaye, G., Jothimani, M., Reda, T. M., & Tadese, A. (2024). Landslide susceptibility mapping using combined geospatial, FR and AHP models: A case study from Ethiopia's highlands. *Discover Sustainability*, 5, 474. <https://doi.org/10.1007/s43621-024-00730-4>.

Tesfaye, B., Jothimani, M., & Dawit, Z. (2024). Mapping landslide susceptibility in the Debretabor–Alember road sector, Northwestern Ethiopia. *Journal of Degraded and Mining Lands Management*, 11(2), 5169–5179. <https://doi.org/10.15243/jdmlm.2024.112.5169>.

Tesfaye, B. (2023). Integrated geophysics for investigation of landslide-affected site located near to Alemketema town, Central Ethiopia. *Journal of Applied Geophysics*, 216, 105160. <https://doi.org/10.1016/j.jappgeo.2023.105160>.

Tsunetaka, H., Murakami, W., Nakao, K., & Mtibaa, S. (2025). Evaluation of the impact of climate change on rainfall for potential landslide triggering in Japan. *Earth Surface Processes and Landforms*, 50(1), e6050. <https://doi.org/10.1002/esp.6050>.

Wang, Z., Li, T., & Zhang, J. (2023). Deep learning ensemble for landslide susceptibility mapping. *Engineering Geology*, 323, 107231. <https://doi.org/10.1016/j.enggeo.2023.107231>.

Yang, J.(2025). Escalating risks and impacts of rainfall-induced geohazards. *Natural Hazards Research*, 5(3), 447–454. <https://doi.org/10.1016/j.nhres.2025.01.002>

Yu, X., Xiong, T., Jiang, W., & Zhou, J. (2023). Comparative assessment of the efficacy of five kinds of models in landslide susceptibility mapping for factor screening: A case study at Zigui–

Badong in the Three Gorges Reservoir Area, China. Sustainability, 15(1), 800. <https://doi.org/10.3390/su15010800>.

Zhang, Y. (2025). Assessing the impact of climate change on rainfall-triggered landslides: A case study in California. Landslides, 22, 1707–1724. <https://doi.org/10.1007/s10346-024-02428-0>.

Appendix A: Annexures

Statistical Evaluation of Landslide Occurrence and Contributing Factors Using the Frequency Ratio (FR) Method

Parameters with class	Number of pixel in individual class (Npix(Ni))	% of Npix(Ni)	Number of landslide pixel in class (Npix(si))	% of Npix(Si)	Slide Ratio (SR)	Class Ratio (CR)	Frequency Ratio (FR)	Relative Frequency ratio	Relative Frequency (RF (non %))
S _{lo} 0°-8.6°	167908	27.43	1291	23.05	0.23	0.27	0.84	0.16	15.73

	8.7°-15.2°	184860	30.20	1667	29.77	0.30	0.30	0.99	0.18	18.45
	15.3°-22.7°	137376	22.44	1416	25.29	0.25	0.22	1.13	0.21	21.08
	22.8°-32.1°	85438	13.96	746	13.32	0.13	0.14	0.95	0.18	17.86
	32.2°-70.6°	36519	5.97	480	8.57	0.09	0.06	1.44	0.27	26.88
	Total	612101		5600				5.34		
Elevation	951-1370	120277	19.65	625	11.16	0.11	0.20	0.57	0.12	11.52
	1370.1-1630	142043	23.21	1954	34.89	0.35	0.23	1.50	0.30	30.50
	1630.1-1876	163306	26.68	1452	25.93	0.26	0.27	0.97	0.20	19.71
	1876.1-2170	104188	17.02	707	12.63	0.13	0.17	0.74	0.15	15.04
	2170.1-2782	82287	13.44	862	15.39	0.15	0.13	1.15	0.23	23.22
Total		612101		5600				4.93		
Curvature	Concave	109778	17.93	893	15.95	0.16	0.18	0.89	0.29	28.59
	Flat	439853	71.86	4007	71.55	0.72	0.72	1.00	0.32	32.02
	Convex	62470	10.21	700	12.50	0.13	0.10	1.22	0.39	39.39
Total		612101		5600				3.11		
Aspect (direction of slope)	Flat (-1-35.78)	79242	12.95	570	10.18	0.10	0.13	0.79	0.09	8.56
	North (35.78-73.98)	71474	11.68	465	8.30	0.08	0.12	0.71	0.08	7.74
	Northeast(73.99-113.59)	71665	11.71	588	10.50	0.11	0.12	0.90	0.10	9.76
	East (113.6-154.62)	63229	10.33	431	7.70	0.08	0.10	0.75	0.08	8.11
	Southeast(154.63-197.06)	56976	9.31	574	10.25	0.10	0.09	1.10	0.12	11.99
	South (197.07-240.92)	55113	9.00	672	12.00	0.12	0.09	1.33	0.15	14.51
	Southwest(240.93-281.95)	64663	10.56	1065	19.02	0.19	0.11	1.80	0.20	19.59
	West (281.96-320.15)	73334	11.98	795	14.20	0.14	0.12	1.18	0.13	12.90

	Northwest(320.16-359.76)	76405	12.48	440	7.86	0.08	0.12	0.63	0.07	6.85
Total		612101		5600				9.19		
Proximity-to stream	0-250	19040	3.11	467	8.34	0.08	0.03	2.68	0.33	32.52
	251-500	34160	5.58	656	11.71	0.12	0.06	2.10	0.25	25.46
	501-750	50792	8.30	844	15.07	0.15	0.08	1.82	0.22	22.03
	751-1000	77871	12.72	631	11.27	0.11	0.13	0.89	0.11	10.74
	>1000	430238	70.29	3002	53.61	0.54	0.70	0.76	0.09	9.25
Total		612101		5600				8.24		
Proximity to road	0-250	2607	4.26	458	8.17	0.08	0.04	1.92	0.26	26.23
	251-500	2502	4.09	405	7.24	0.07	0.04	1.77	0.24	24.21
	501-750	2641	4.32	340	6.08	0.06	0.04	1.41	0.19	19.26
	751-1000	2100	3.43	257	4.59	0.05	0.03	1.34	0.18	18.27
	>1000	51351	83.91	4140	73.92	0.74	0.84	0.88	0.12	12.04
Total		61201		5600				7.32		
Proximity-to lineaments	0-250	16830	2.75	456	8.14	0.08	0.03	2.96	0.34	34.20
	251-500	17801	2.91	455	8.13	0.08	0.03	2.79	0.32	32.27
	501-750	28830	4.71	364	6.50	0.07	0.05	1.38	0.16	15.94
	751-1000	28400	4.64	169	3.02	0.03	0.05	0.65	0.08	7.51
	>1000	520240	84.99	4156	74.21	0.74	0.85	0.87	0.10	10.08
Total		612101		5600				8.66		
LULC	Cultivation land	150068	24.52	1017	18.16	0.18	0.25	0.74	0.28	28.12
	Moderate cultivation land	305880	49.97	4253	75.95	0.76	0.50	1.52	0.58	57.69
	Forest	21718	3.55	25	0.45	0.00	0.04	0.13	0.05	4.78
	Woodland	134435	21.96	305	5.45	0.05	0.22	0.25	0.09	7.5.9
Total		612101		5600				2.63		
Soil	Dystric nitisols	324057	52.94	4013	71.74	0.72	0.53	1.36	0.19	19.33
	Eutric cambisols	113420	18.53	859	15.36	0.15	0.19	0.83	0.12	11.82
	Dystric fluvisols	113718	18.58	550	9.83	0.10	0.19	0.53	0.08	7.55

	Chromic luvisols	10899	1.78	172	3.07	0.03	0.02	1.73	0.25	24.64
	Eutric nitisols	27511	4.49	125	2.23	0.02	0.04	0.50	0.07	7.09
	Haplic Phaeozems	22496	3.68	426	7.62	0.08	0.04	2.07	0.30	29.56
Total		612101		5594				7.01		
Lithology	-Main volcanic sequences (basalt, rhyolite, trachyte, tuff and ignimbrite) (Pv)	375088	61.28	3650	65.18	0.65	0.61	1.06	0.51	51.22
	-Undivided alluvial, fluvial and lacustrine sediments (Q)	210430	34.38	1950	34.82	0.35	0.34	1.01	0.49	48.78
	-Mursi formation (NPom)	17669	2.89	0	0.00	0.00	0.03	0.00	0.00	0.00
	-Nkalabong formation (Npon)	8914	1.46	0	0.00	0.00	0.01	0.00	0.00	0.00
	Total	612101		5600				2.08		
Rainfall	963 - 1,247.4	113548	18.55	700	12.50	0.13	0.19	0.67	0.11	10.81
	1,247.5 - 1,268.3	117180	19.14	575	10.27	0.10	0.19	0.54	0.09	8.60
	1532 - 1816.3	128314	20.96	1043	18.63	0.19	0.21	0.89	0.14	14.25
	1816.4 - 2100.7	219039	35.78	2362	42.18	0.42	0.36	1.18	0.19	18.91
	2100.8 - 2385.1	34020	5.56	920	16.43	0.16	0.06	2.96	0.47	47.42
	Total	612101		5600				6.23		
Stream-power index	16.36 - 2.66	16376	2.68	72	1.29	0.01	0.03	0.48	0.10	10.38
	2.65 - 0.81	67865	11.09	362	6.46	0.06	0.11	0.58	0.13	12.60
	0.8 - 0.27	203767	33.29	1866	33.32	0.33	0.33	1.00	0.22	21.63
	0.28 - 1.14	274909	44.91	2614	46.68	0.47	0.45	1.04	0.22	22.46
	1.15 - 11.36	49184	8.04	686	12.25	0.12	0.08	1.52	0.33	32.94
	Total	612101		5600				4.63		

Topographic wetness index	0 - 0.17	199090	32.53	1535	27.41	0.27	0.33	0.84	0.15	15.11
	0.18 - 0.32	198647	32.45	1821	32.52	0.33	0.32	1.00	0.18	17.97
	0.33 - 0.5	128975	21.07	1271	22.70	0.23	0.21	1.08	0.19	19.32
	0.51 - 0.76	66536	10.87	719	12.84	0.13	0.11	1.18	0.21	21.18
	0.77 - 2.84	18853	3.08	254	4.54	0.05	0.03	1.47	0.26	26.41
Total		612101		5600				5.58		