Original Research 🔓 Open access

Mapping Landslide Vulnerability using Machine Learning Approach along the Taba Penanjung-Kepahiang Road, Bengkulu Province

Camelia B. Abrar[1], Ashar M. Lubis[2*], Darmawan I. Fadli[2], Arya J. Akbar[1], Rida Samdara[1]

1. *Physics Study Program, University of Bengkulu, Indonesia*
2. *Geophysics Study Program, University of Bengkulu, Indonesia*

Abstract

Landslides occur when masses of rock, debris or soil move due to various factors and processes that cause land movement. The Taba Penanjung-Kepahiang route is one of the areas in Bengkulu Province that is highly prone to landslides. This causeway is the only fastest land route connecting the Bengkulu-Kepahiang area. In recent years, the road area has often been cut off due to landslides and fallen trees, which have caused road access to be cut off and obstructed and claimed lives. This study uses a Machine Learning (ML) and GIS approach with Variable Frequency Ratio using 16 independent factors obtained from the spatial database and DEM, which correlate with landslide events. This research aims to gain an in-depth understanding of the factors that cause landslides. In addition, the research focus is the development of a Disaster Mitigation Model to design and implement effective strategies to reduce the risk and impact of landslide disasters through in-depth analysis The dependent factor is the location of the landslide from the historical landslide area for the last five years, with a distribution of 70/30%. Furthermore, frequency ratio is used to analyze the correlation between conditioning factors and historical landslides. Then, the independent and dependent factors were normalized to create a landslide susceptibility map. Frequency Ratio (FR) indicates the likelihood of an event occurring, with drainage density (FR= 0.69), shear wave velocity (Vs30) (FR= 0.66), slope (FR= 0.60), and rainfall (FR= 0.55). The output of the processed data is in the table below.

1. Introduction

Landslides are one of the geological disasters that often occur in Indonesia. Landslides occur when a mass of rock or soil moves due to controlling factors and triggering processes. A landslide is a movement of the ground with a slope direction and moves it on an avalanche (Fadilah et al., 2019). Landslide is a process of mass wastage that occurs on slopes formed naturally or engineered by the movement of rock masses, debris, or soil down the slope, which is influenced by gravity (Cruden & VanDine, 2013). Landslides occur continuously from year to year and make landslide disasters the center of attention and become a severe problem in almost all parts of the world because they cause economic or social losses to private and public property (Rotaru et al., 2007).

One of the areas in Bengkulu with a high level of vulnerability to landslides is the Bengkulu-Kepahiang Route; this is because the Bengkulu-Kepahiang Route is an area with diverse geomorphological conditions. This area is the only fastest connecting land road that connects the Bengkulu-Kepahiang area, but the area is a forest area. In recent years, fallen trees and landslides have been obstructed and have caused road access to be cut off and obstructed, claiming lives. Previous cases show that landslides occurred at two or more points on the same day, causing several vehicles to get stuck between the 2 points. Landslides can occur with coverage of more than one slope, triggered by the same phenomenon (Froude & Petley, 2018).

The area has no alternative roads that allow motorists to turn around until the landslide, trees, or other materials from the road can be cleared. Therefore, research related to analyzing and determining landslide hazards on the Bengkulu-Kepahiang Route must be carried out as a non-structural mitigation effort. This effort helps identify relatively safer areas from landslides so that material and non-material losses can be minimized. This research aims to identify landslide-prone areas on the Bengkulu-Kepahiang Route, with the primary objective of taking mitigation steps to enhance the safety of people around the area. According to Alcántara-Ayala & Sassa, (2023) by mapping landslide-prone areas, we can focus on risk management and prevention. Additionally, this research offers insight into the vulnerability of landslides on the Bengkulu-Kepahiang Route, the region's primary land route. With a better understanding of the potential for landslides, we can take action to optimize transportation access and minimize disruption due to landslides, which can harm local communities and the economy.

Several previous studies regarding landslides on the Bengkulu-Kepahiang Cross Route, such as Suhendra & Sugianto, 2018; Hadi et al., 2021; Hadi & Siswanto, 2016; and Sugianto, 2021. In general, studies on landslides that previous researchers have carried out show that the area on the Bengkulu-Kepahiang Cross Route has the potential to experience ground movement. However, the previous study was still carried out deterministically, and the parameters used were still very few, so it is necessary to map landslide-prone areas on the Bengkulu-Kepahiang Cross Road. To overcome this issue, achine Learning (ML) method is one solution that can solve problemsThe ML model is considered essential in disaster mitigation and an ideal landslide management plan in landslide modeling for disaster mitigation and disaster management as a mitigation effort (Pourghasemi et al., 2018). Making a vulnerability map requires data with high accuracy. The more input parameters, the more accuracy and sensitivity analysis in mapping landslide vulnerability will be more accurate (Ghorbanzadeh et al., 2019). The ML landslide detection study uses different classifications (Roodposhti et al., 2019) to increase the efficiency of the outputs.

Mapping of landslide-prone areas will be calculated using Frequency Ratio (FR) and parameters used such as elevation (topography), geological conditions, slope aspect (Slope Aspect), slope (Slope), rainfall, plan curvature, distance from faults/faults (Kavzoglu et al., 2019) as well as other supporting factors. The interconnection of these factors makes a holistic approach important in dealing with and preventing landslides. The factors mitigate the adverse impacts of landslides, and it is essential to implement a comprehensive strategy that includes prevention and preparedness. Sustainable land use planning, reforestation, early warning systems, and advanced technology integrating vulnerable areas are essential to an effective landslide mitigation plan. Developing a Disaster Mitigation Model in the Jalan Taba Penanjung-Kepahiang area aims to design and implement an effective strategy to reduce the risk and impact of landslides through an ML approach. Additionally, we hope that information from this research can be used to increase awareness of community preparedness, help design risk mitigation programs, and facilitate faster responses in emergencies. This is expected to reduce the impacts materially and non-materially so that the road can be traversed safely.

## 2.   Data and Methods

2.1.   Study Area and Geological Setting.

The Taba Penanjung-Kepahiang route (Fig. 1) is an area that connects Central Bengkulu Regency and Kepahiang Regency, which are in Bengkulu Province, with an area of 1124.44 ha of the research area. This area is where landslides frequently occur in Bengkulu province, and narrow access roads do not allow motorists to turn around and cause long traffic jams. Based on DEM data, the topography on the Bengkulu-Kepahiang Cross Route is at an elevation of 96 to 880 masl (Suhendra & Sugianto, 2018) with high rainfall with an average of 235-280 mm/year, so that it will increase the potential for ground movement or landslides (Natasya et al., 2022). Landslides occur due to factors arising from the internal geology of the slope as well as the external environment. The internal factors, which include geomorphology, stratum lithology, and topography, will control the occurrence of landslides. The main external factors in landslides are environmental factors, hydrogeology, and

human engineering activities (Xiao et al., 2019). These factors will be input and extracted as the final input on ML and soil susceptibility index (LSI) in the form of FR values (Zhu et al., 2021).

Figure 1. (i). Jalan Lintas Taba Penanjung – Kepahiang is using Google Satellite. (ii), (iii), (iv). Landslides on the Taba Penanjung Route - Kepahiang in 2021-2022

The Hulusimpang Formation (Tomh) (Fig.2), with an orange symbol flanked by andesitic basalt volcanic rocks, is green. Tomh is around the late mid-Miocene Oligocene, with the most extensive zone mainly along the Ketahun-Musikeruh fault zone and several places in the western part of Lambar.
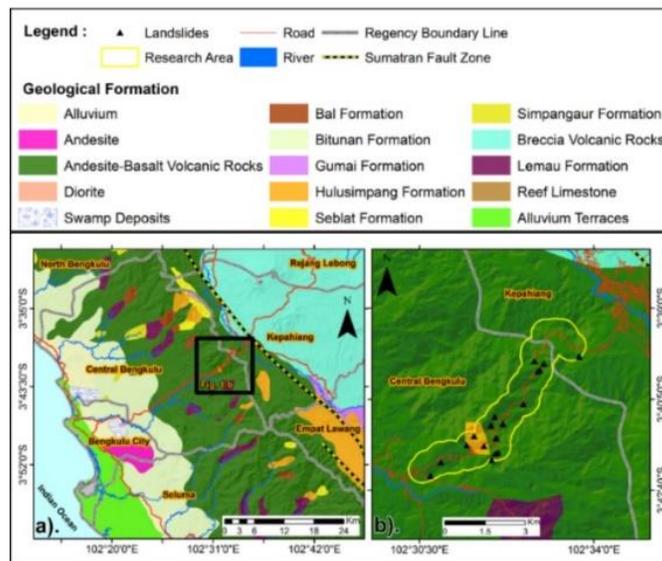
Source: Analysis, 2022

Figure 2. Geological Conditions and Study Area of the Taba Penanjung-Kepahiang Route

2.2. Frequency Ratio (FR) Calculation

FR is a well-known method for mapping landslides (Ozdemir & Altural, 2013). FR measures the degree of correlation between landslide locations according to their independent factors (Solaimani et al., 2013). FR is a method for calculating the effect of subclasses of conditioning factors on landslides (He et al., 2012). From FR for each class from all data layers, it will be combined with the landslide inventory map independent factor map using the equation (Eq.1):

$$FR = \frac{N_{pix(Si)}/N_{pix(Ni)}}{\sum_i N_{pix(Si)}/\sum_i N_{Pix(Ni)}} \ldots\ldots\ldots\ldots [Eq.1]$$

Npix (Si) is the number of pixels of landslides, and Npix (Ni) is the number of pixels of a class. The Landslide Susceptibility Index (LSI) is calculated by the sum of each factor ratio value using the equation (Eq.2):

$$LSI = Fr1 + Fr2 + Fr3 + Frn \ldots\ldots\ldots\ldots Eq.2]$$

LSI will be obtained by adding up each FR value for each conditioning factor correlated with landslide history, where FR is the level of each type of factor. Landslide vulnerability maps are obtained from the correlation between the factors overlaid using spatial weight overlay analysis (Okoli et al., 2023).

2.3. Landslide Inventory and Casual Factor

Data processing uses the QGIS application and Software R. Processing data on Software R uses a grid base with factor attribute values from ML that differ from specific grid sizes (Wei et al., 2022). Input data is in the form of local characterization of the avalanche geometry and internal structure, which is used to further describe slope stability in modeling (Dou et al., 2020). For vulnerability analysis, a correlation is depicted between predisposing causes and triggering factors using a numerical model (van Asch et al., 2007). Data processing uses 16 independent factors (Table 1.), continuous and categorical scale factors, and dependent factors obtained from landslide inventory maps: location data, places, dates, and other information regarding landslides in an area (Guzzetti et al., 2012). In this study, 16 conditioning factors were used, which were classified using different methods, specifically manual, equal interval, and natural breaks (Arabameri et al., 2017).

The slope aspect parameter (Fig. 3.a) correlates closely with weather conditions (Bednarik et al., 2010). It determines its exposure to wind and sunlight, with vulnerability affecting soil moisture and vegetation. While in general, the curvature (Fig. 3.b) is the number of surface defects in an area. The greater the surface defects, the greater the degree of curvature. Curvature can map stratigraphic features using structural deformation models to predict natural fractures and paleo stress (Lisle, 1994; Roberts, 2001; Chopra & Marfurt, 2007).
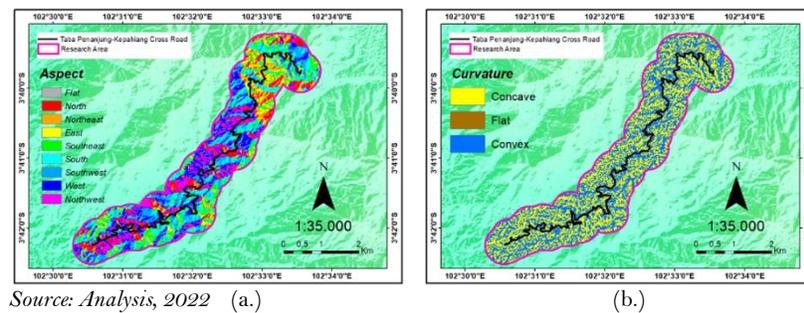

Source: Analysis, 2022   (a.)   (b.)

Figure 3. (a.) Slope Aspect and (b.) Curvature of Research Area

One of the critical environmental factors in landslide mapping is the elevation (Fig.4.a) (Marjanović et al., 2011). Figure 4. a shows elevations ranging from 112-918 m. In the research area, Musi and Manna Segments (Fig. 4.b) exist in the Taba Penanjung and Kepahiang Regency areas.
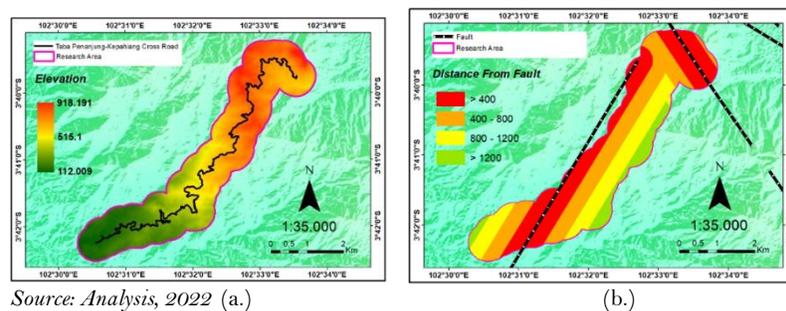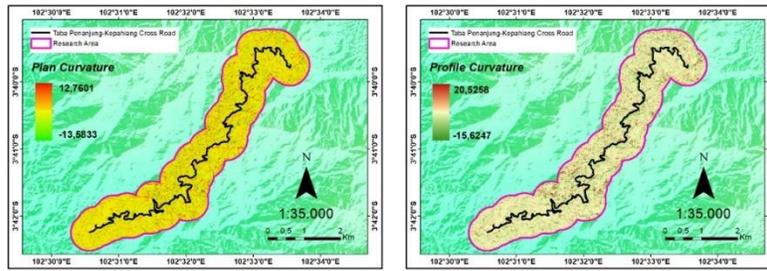

Source: Analysis, 2022 (a.)   (b.)

Figure 4. (a.) Elevation. of Research Area in m and
(b.) Distance from Fault and Fault Location on the Research Area in km

In the horizontal direction, Plan Curvature (Fig. 5.a) reflects ridges and valleys on a surface, affecting flow dispersion and convergence. In the standing order, the profile curvature (Fig. 5.b) can reflect the degree of slope transformation involving the flow's acceleration and deceleration (Lee et al., 2018).
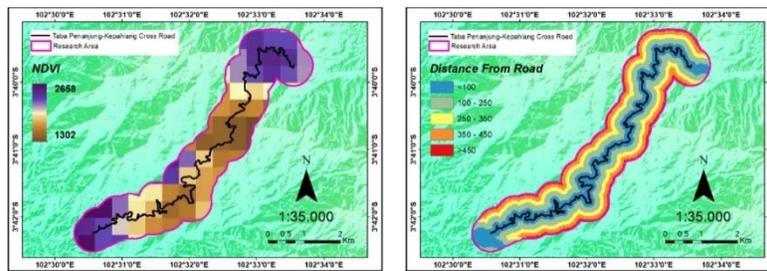
Source: Analysis, 2022 (a.)   (b.)

Figure 5. (a.) Plan Curvature and (b.) Profile Curvature of Research Area

The slope of the slope (Fig. 6.a) is a factor that significantly determines the occurrence of landslides in an area and affects the level of soil slides (Fan et al., 2022). The greater the degree of slope, the higher the level of vulnerability in the area. However, steep slopes naturally formed due to bedrock outcrops are not prone to landslides (Mohammady et al., 2012). Shear wave velocity Vs30 (Fig. 6.b) is the average shear wave velocity with a depth of up to 30 m from the ground surface (Hadi et al., 2018).
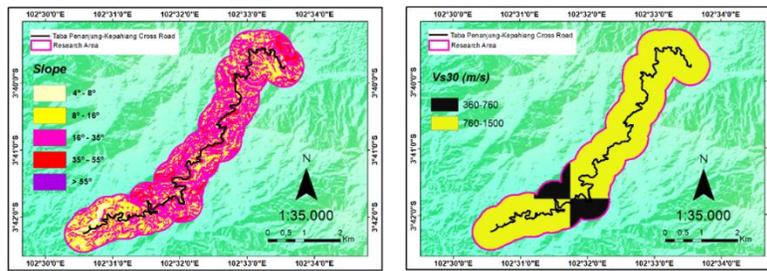
Source: Analysis, 2022 (a.)   (b.)

Figure 6. (a.) Slope and (b.) Shear Wave Velocity (Vs30) of Research Area

Normalized Difference Vegetation Index (NDVI) refers to the active vegetation biomass or forest cover (Fig. 7.a). Landslides usually occur on bare land and grasslands (Wang et al., 2020). Road construction is one of the factors controlling slope stability, with the hypothesis that landslides occur more frequently along the road. This is due to cutting drainage and cutting slopes from making roads that are not suitable (Dahal et al., 2008). This research focuses only on the Taba Penanjung-Kepahiang Route. The relationship between distance from roads (Fig. 7.b) and landslide risk can be influenced by several factors, including an area's geological and topographic characteristics.

Source: Analysis, 2022 (a.)   (b.)

Figure 7. (a.) Normalized Difference Vegetation Index (NDVI) and
(b.) Distance from road at research area

Other topographical factors such as Topographic Wetness Index (TWI) (Fig. 8.a) and Sediment Phosphorous Index (SPI) (Fig. 8.b) use processed dem data in fill dem, flow direction, slope (o), and flow accumulation using the Jenks natural breaks method (Ciurleo et al., 2016) (Equation 3 and 4):

$$TWI = \log e(\frac{A}{tan\beta}) \dots\dots\dots\dots [Eq.3]$$

$$SPI = A \cdot \tan\beta \ldots\ldots\ldots\ldots \ [Eq.4]$$

A is the flow Accumulation value in meters squared, and β is the slope(o). TWI accurately describes how topographic changes impact land runoff, while SPI reflects the ability of a water system to erode the soil surface (Moore & Grayson, 1991; Xiao et al., 2019).


*Source: Analysis, 2022* (a.)   (b.)
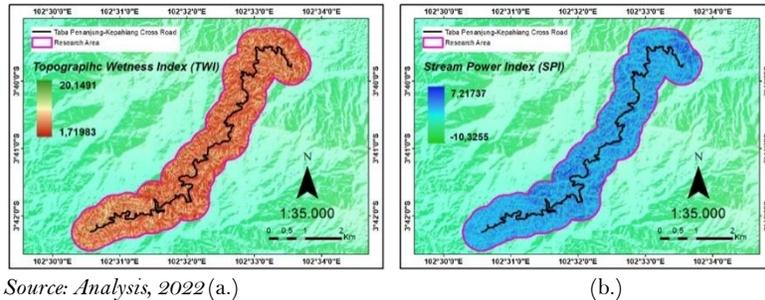
Figure 8. (a.) Topographic Wetness Index (TWI) and (b.) Sediment Phosphorous Index (SPI) of research area

Geological factors in the form of lithology, where there are only two classifications of geological conditions, Hulusimpang formations, and Andesit-Basalt Volcanic Rocksstructures, are described in (Fig. 9.a). The supporting factor used is the average rainfall data in the last ten years obtained from the Data Center for River Region VII (BWSVII) Bengkulu City (Fig. 9.b). The rainfall data is allocated and applied in analyzing recurring periods (Koutsoyiannis, 2004; Wallis et al., 2007; Shou & Lin, 2020).

Source: Analysis, 2022 (a.)   (b.)
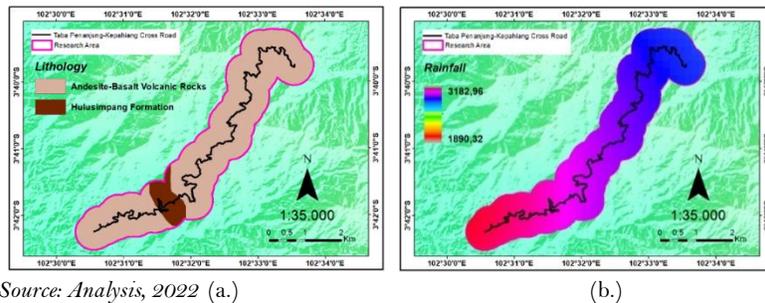
Figure 9. (a.) Lithology at Research Area; (b.) Rainfall at Research Area in mm/year

The slope control factor to the ratio of the total length of the river basin is called the drainage density (Fig. 10.b). In general, the higher the density of infiltration drainage, the lower it will be, and the movement of the soil surface will be faster. Drainage density indicates the degree of saturation with the flow, which can adversely affect slope saturation (Pachauri et al., 1998; Nagarajan et al., 2000; Çevik & Topal, 2003). The distance between drainage systems (Fig. 10.a) and landslide risk is also essential in planning infrastructure and minimizing landslide risk. Good drainage can help reduce groundwater levels around slope areas.

Source: Analysis, 2022 (a.)   (b.)
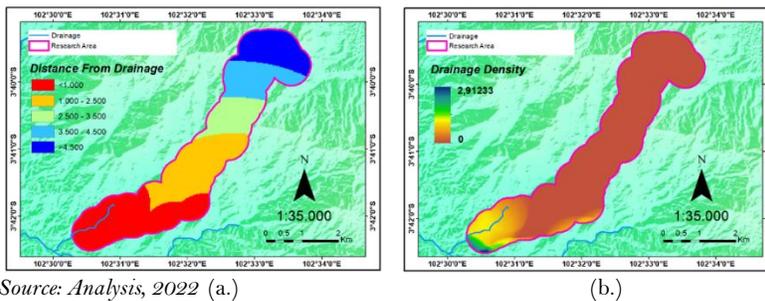
Figure 10. (a.) Distance from Drainage and (b.) Drainage Density at Research Area

This research began with a literature study, carried out by studying previous studies related to this proposed research. The research data collection carried out in this research was a secondary collection in the form of DEMNAS data, Geological maps, and administrative maps. Next, data processing uses the FR algorithm. The next stage is an analysis of the results of existing ML algorithms. Analysis was carried out using the ROC curve to evaluate and validate the data obtained to obtain the best landslide susceptibility model. The higher the accuracy value of the ROC curve, the better the model produced, and vice versa. The analysis results will be depicted as a landslide susceptibility map on the Bengkulu-Kepahiang Route. In general, the research stages are as shown in the flow diagram below (Fig. 11)
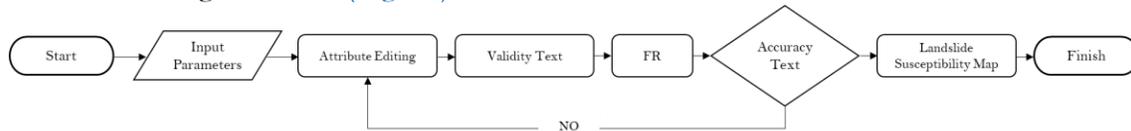

Figure 11. Research Diagram

Table 1. Conditioning Factors and Their Classification

Factors	Classes	Data Scale	Techniques
Aspect (A)	F (−1); N (0–22.5; 337.5–360); NE (22.5–67.5); E (67.5–112.5); SE (112.5–157.5); S (157.5–202.5); SW (202.5–247.5); W (247.5–292.5); NW (292.5–337.5)	https://earthexplorer.usgs.gov/ (22 December 2022) Geospatial Data Cloud	8 x 8 m DEM
Curvature (C)	-319-0; 0; 0-268;	https://earthexplorer.usgs.gov/ (22 December 2022) Geospatial Data Cloud	8 x 8 m DEM
Elevation (E)	112-150; 250-450; 450-650; 650-850; 850-918	https://earthexplorer.usgs.gov/ (22 December 2022) Geospatial Data Cloud	8 x 8 m DEM
Distance From Fault (DF)	0-100; 100–250; 250-350; 350- 450; >450	https://geologi.esdm.go.id/geomap Indonesia Catalogue Service For Geographic Information	Buffering
Plan Curvature (PLC)	(-13)-(-10); (-10)-0; 0-12	https://earthexplorer.usgs.gov/ (22 December 2022) Geospatial Data Cloud Continue	8 x 8 m DEM
Profile Curvature (PRC)	(-15)-(-10); (-10)-0; 0-5; 5-10; 10-20	https://earthexplorer.usgs.gov/ (22 December 2022) Geospatial Data Cloud	8 x 8 m DEM
Slope (°) (S)	4°-8°; 8°-16°; 16°-35°; 35°-55°; >55°	https://earthexplorer.usgs.gov/ (22 December 2022) Geospatial Data Cloud	8 x 8 m DEM
SPI	(-10)-(-6); (-6)-(-2); (-2)-2; 2-6; 6-7	https://earthexplorer.usgs.gov/ (22 December 2022) Geospatial Data Cloud	8 x 8 m DEM
NDVI	1300-1500; 1500-1700; 1700-1900; 1900-2100; 2100-2685	https://earthexplorer.usgs.gov/ (22 December 2022) Geospatial Data Cloud	Extract By Mask
Drainage Density (DRD)	0-0.5; 0.5-1; 1-1.5; 1.5-2; 2-2.91	https://earthexplorer.usgs.gov/ (22 December 2022) Geospatial Data Cloud	8 x 8 m DEM
TWI	1.7-8; 8-16; 16-20	https://earthexplorer.usgs.gov/ (22 December 2022) Geospatial Data Cloud	8 x 8 m DEM
Vs30	360-760; 760-1500	https://earthexplorer.usgs.gov/ (28 December 2022) Geospatial Data Cloud	Extract By Mask
Lithology (LIT)	1;2	https://www.indonesia-geospasial.com/ National Catalogue Service For Geographic Information	Digitization process
Rainfall (R)	2500-2800; 2800-3165	Data Curah Hujan 10 tahun Balai Wilayah Sungai VII Kota Bengkulu.	Kriging Interpolation method
Distance From Road (DR)	0-100; 100-250; 250-350; 350-450; >450	https://tanahair.indonesia.go.id/portal-web Peta AOI	Buffering
Distance From Drainage (DD)	0-1000; 1000-2500; 2500-3500; 3500-4500; >4500	https://earthexplorer.usgs.gov/ (22 December 2022) Geospatial Data Cloud	Buffering

Source: Analysis, 2022

3. Result and Discussion

3.1 Correlation Analysis between Landslide and Independent Factor

The ratio between the slide classification, a percentage of the overall failure, and the class area as a percentage of the entire map is called FR (Nourani et al., 2014). FR is generated from each conditioning factor with weights in each sub-class (Umar et al., 2014). The correlation of each element with landslide history determines the FR value (Shahabi et al., 2014). The conditioning factor sub-class is used in the input variables of the data-based model, with the classification of parameters used in determining the value of the distribution of attribute intervals related to the sub-class (Huang et al., 2020). The FR results are a probability comparison value between safe and landslide-prone areas, normalized to FRn with a range of 0-1 to see the correlation between conditioning factors and landslide history (Chen & Chen, 2021). To show the correlation between conditioning factors, pairwise value analysis is used to ensure each index factor's independence level between each conditioning factor.

Table 2. Pairwise Comparison Matrix of Conditioning Factors

Factors	S	E	DF	SPI	A	C	PRC	PLC	R	DR	TWI	DD	NDVI	DRD	Vs30	LIT
S	1.00															
E	0.51	1.00														
DF	0.68	1.34	1.00													
SPI	0.45	0.89	0.66	1.00												
A	0.37	0.72	0.54	0.81	1.00											
C	0.21	0.41	0.30	0.46	0.56	1.00										
PRC	0.35	0.69	0.51	0.78	0.95	1.69	1.00									
PLC	0.04	0.09	0.07	0.10	0.12	0.22	0.13	1.00								
R	0.16	0.31	0.23	0.35	0.43	0.76	0.45	3.48	1.00							
DR	0.61	1.21	0.90	1.36	1.67	2.97	1.76	13.69	3.93	1.00						
TWI	0.64	1.26	0.94	1.42	1.74	3.10	1.83	14.27	4.10	1.04	1.00					
DD	0.45	0.89	0.66	1.00	1.22	2.17	1.28	10.01	2.88	0.73	0.70	1.00				
NDVI	0.48	0.95	0.71	1.07	1.32	2.34	1.38	10.79	3.10	0.79	0.76	1.08	1.00			
DRD	1.11	2.18	1.62	2.46	3.02	5.37	3.17	24.70	7.10	1.80	1.73	2.47	2.29	1.00		
Vs30	0.54	1.07	0.80	1.20	1.48	2.63	1.55	12.10	3.48	0.88	0.85	1.21	1.12	0.49	1.00	
LIT	0.02	0.04	0.03	0.05	0.06	0.10	0.06	0.48	0.14	0.04	0.03	0.05	0.04	0.02	0.04	1.00

Source: Analysis 2022

The relationship between conditioning factors is shown in Table 2. The vertical table shows the plan curvature as the indicator with the highest correlation among other factors. While horizontally, the drainage density indicator has the highest value among other factors. The highest correlation is owned by drainage density and plan curvature, with a correlation of 24.70. Previous case studies, by Hadi et al. (2018) and Sugianto (2021) regarding landslide-prone mapping using different factors. Sugainto's 2020 study revealed the structure of the shear wave velocity (Vs) or subsurface structures along the Bengkulu Kepahiang Causeway based on measurements and inversion of microtremor data. This study showed a correlation between the rate of the Vs30 shear waves on the Taba Penanjung - Kepahiang Cross Road, the results of which were associated with the potential for landslides. Whereas Hadi et al.'s research, applied the HVSR and SAW methods related to the potential for landslides in the Kepahiang district. Both of these studies still use deterministic methods using only a few parameters.

This research uses 16 parameters from processed DEM data extracts and processed catalog map data. The area of this study is only 1124.44 (ha), with the minimal classification of conditioning factors due to the small space. These conditioning factors greatly influence the FR results, where a few combinations will result in a low FR. Classifying many conditioning factors with a large area is necessary to increase the FR output. FR analysis shows (Table 3.) that in Aspect indications, the highest probability is in the North West class, while the lowest possibility is in the East class. This is related to the vegetation index (NDVI) in the northwest direction, which is dominated by a high vegetation index. Following Lee & Min (2001), there is a correlation between the vegetation index and the slope. As for the elevation indicator, the highest probability is at an altitude of 450-650 with FR=0.39 and a massive difference with the lowest FR in the 112-250 altitude class with FR=0.08. The

relationship between the fault and the landslide shows that class 800-1200 has the highest probability value with FR=0.49, while the fault distance with the lowest FR is in class >1200 with FR=0.02.

In the Curvature indication, the highest probability is in the concave and convex class with FR=0.38. The Plan Curvature indication has the highest probability value in 2 categories, specifically (-10)-0 and 0-12 with FR=0.34, and the lowest FR is (-13) -(-10) with FR=0.32. As for the Profile Curvature indicator, the highest probability is in class 10-20 with FR = 0.32, while the lowest is in class 0-5 with FR = 0.11. These three indicators are extracted from the same input data. With spatial analysis, the extraction of the three indicators is produced simultaneously. In the Slope indicator, the highest probability is in the class with the highest degree of slope, specifically >55 with FR=0.60, while the lowest FR is in the slope with the lowest degree, specifically 4-8 degrees with FR=0.003. This is consistent with a general aspect, the shear stress slope material will increase according to the increase in slope degrees, and landslides are likely to occur on the steepest slopes (Yilmaz, 2009).

The highest probability of SPI is in class 6-7, where FR=0.35, while the lowest chance is in class (-2)-2, with FR=0.08. As for NDVI, the highest probability is in the vegetation class 1300-1500 with FR = 0.29 and in class 2100-2658 showing no likelihood of landslides with that class, where the results of the data allocation in that class show FR = 0. At Drainage Density, the highest probability is at a density of 0.5-1 with FR=0.69. At a density of 1.5-2.2-2.91, it shows that there is no probability in that class with FR=0. For the TWI indicator, the probability of landslides is in class 8-16 with FR=0.46, followed by class 1.7-8 with FR=0.45, and the lowest probability is in class 16-20 with FR=0.08. In the Vs30 indicator, there are only two classes. The highest probability is found at the shear wave speed with a value of 360-760 FR=0.66, and the lowest is at 760-1500 with FR=0.34. Lithological indicators in this area only have two types of rocks: Quaternary Volcanoes and Andesite. The highest probability is in Andesitic rocks with FR=0.51, while in Quaternary volcanoes, the probability differs significantly from Andesitic rocks with FR=0.49.

The Rainfall Indicator is allocated from rainfall data for the previous ten years, so 2-factor classifications are obtained. The highest probability is in class 2500-2800 with FR=0.55, not different from the probability in class 2800-3165 with FR=0.45. According to Regmi et al. (2014), landslides usually occur along cut roads and road construction processes that damage the natural conditions of the slopes. In this indicator, the distance of the research area from the station causes the minimum classification it can obtain. On the distance from the road indicator, the highest probability is the shortest distance, specifically 0-100 with FR=0.43, while the lowest probability is at distances of 340-450 and >450 with FR=0.06 and 0.08. Distance From Drainage indicator, the highest probability is at a distance of 1000-2500 with FR = 0.35, while the lowest probability is at a distance > 4500 with FR = 0.08, which indicates that the frequency of landslides decreases with increasing distance to the drainage canal and can be attributed to the fact that during rainstorms the groundwater level will rise and the initiation of landslides is affected by the modified terrain conditions by ditch erosion (Dai & Lee, 2001).

The landslide vulnerability map on the Taba Penanjung – Kepahiang Cross Road is divided into five areas landslide areas with low, medium, high, and very high vulnerability. The results show that a very low classification has an area of 8% of the total area, low and high with an area of 25%, medium with 28%, and very high with 14%. As shown in Figure 13 the final map of the landslide vulnerability mapping shows that the area symbol in red is for an area very prone to landslides. In contrast, the area with a blue sign indicates that the area has a very low landslide vulnerability. This research aims to identify landslide vulnerability in road areas through correlation analysis between images and diagrams. Image diagram in Figure 12, displays a comprehensive presentation of the area in Figure 13 where the red area indicates a high level of vulnerability to landslides, supported by the significant frequency of landslides in that area. Meanwhile, the blue area is considered safe against landslides, even though it has experienced such events because the factors that have been identified indicate a low level of vulnerability.

**Table 3.** FR Value Calculation on Factor Classification

Factor	Classification	Classified Area/Km2	Proportion of Classified Area/%	Number of Landslide Points/pts	Proportion of the Number of Landslide Points/%	Density of Landslide Points/(pts/km2)	Frequency Ratio (FR)
Aspect	(-1)	7070	4.07	960	4.38	1.08	0.10
	(-1)-22.5	11162	6.44	2496	11.40	1.77	0.17
	22.5-67.5	13282	7.66	1088	4.97	0.65	0.06
	67.5-112.5	22373	12.91	960	4.38	0.34	0.03
	112.5-157.5	30960	17.86	1856	8.47	0.47	0.05
	157.5-202.5	28036	16.17	3776	17.25	1.07	0.10
	202.5-247.5	25426	14.67	3648	16.66	1.14	0.11
	247.5-292.5	25063	14.46	3840	17.54	1.21	0.12
	292.5-337.5	9923	5.72	3264	14.91	2.60	0.25
Curvature	-319-0	31293	17.81	4928	22.06	1.24	0.38
	0	97435	55.45	10368	46.41	0.84	0.26
	0-268	46971	26.73	7040	31.51	1.18	0.36
Elevation	112-150	40348	22.96	2432	10.88	0.47	0.08
	250-450	28087	15.98	4864	21.77	1.36	0.24
	450-650	23581	13.42	6528	29.22	2.18	0.39
	650-850	41270	23.48	5440	24.35	1.04	0.18
	850-918	42413	24.13	3072	13.75	0.57	0.10
Distance From Fault	<400	63850	36.34	5760	26.01	0.72	0.26
	400-800	57157	32.53	6464	29.19	0.80	0.29
	800-1200	39505	22.48	9472	42.77	1.18	0.43
	>1200	15172	8.63	448	2.02	0.06	0.02
Plan Curvature	(-13) -(-10) (-10)	33273	18.94	3968	17.76	0.94	0.32
	(-10)-0	96843	55.14	12480	55.87	1.01	0.34
	0-12	45511	25.91	5888	26.36	1.02	0.34
Profile Curvature	(-15) -(-10)	8724	4.96	2112	9.45	1.90	0.27
	(-10)-0	44124	25.11	5568	24.92	0.99	0.14
	0-5	81519	46.39	8128	36.38	0.78	0.11
	5-10	36651	20.86	5184	23.20	1.11	0.16
	10-20	4681	2.66	1344	6.01	2.26	0.32
Slope (°)	4°-8°	15990	9.22	64	0.29	0.03	0.003
	8°-16°	45271	26.12	3328	15.20	0.58	0.06
	16°-35°	99830	57.60	15104	69.00	1.20	0.12
	35°-55° >55°	12120	6.99	3328	15.20	2.17	0.22
		84	0.04	64	0.29	6.03	0.60
SPI	(-10) -(-6)	6261	3.56	960	4.29	1.21	0.20
	(-6) -(-2)	18724	10.65	2432	10.88	1.02	0.17
	(-2)-2	60076	34.19	3712	16.61	0.49	0.08
	2-6	71244	40.54	10112	45.27	1.12	0.19
	6-7	19394	11.03	5120	22.92	2.08	0.35
NDVI	1300-1500	27087	15.41	4864	21.83	1.42	0.29
	1500-1700	37659	21.43	6080	27.29	1.27	0.26
	1700-1900	37766	21.49	6400	28.73	1.34	0.27
	1900-2100	44657	25.41	4928	22.12	0.87	0.18
	2100-2685	28515	16.23	0	0	0	0
Drainage Density	0-0.5	142290	80.98	18432	82.75	1.02	0.28
	0.5-1	11577	6.58	3712	16.66	2.53	0.69
	1-1.5	9443	5.37	128	0.57	0.11	0.03
	1.5-2	9817	5.58	0	0	0	0
	2-2.91	2567	1.46	0	0	0	0
TWI	1.7-8	105265	59.91	13952	62.46	1.04	0.45
	8-16	59795	34.03	8128	36.38	1.07	0.46
	16-20	10639	6.05	256	1.14	0.19	0.08
Vs30	360-760	17861	10.16	4032	18.15	1.79	0.66
	760-1500	157865	89.83	18176	81.84	0.91	0.34
Lithology	1	162323	92.37	20480	92.21	1.00	0.49
	2	13348	7.59	1728	7.78	1.02	0.51
Rainfall	2500-2800	119300	67.90	16000	71.83	1.06	0.55
	2800-3165	56391	32.09	6272	28.16	0.88	0.45
Distance From Road	0 - 100	45297	25.81	11136	50.43	1.95	0.43
	100 - 250	37072	21.12	3328	15.07	0.71	0.16
	250 - 350	32586	18.56	5184	23.47	1.26	0.28
	350 - 450	31378	17.88	1088	4.92	0.28	0.06
	>450	29156	16.61	1344	6.08	0.37	0.08
Distance From Drainage	0 - 1000	57274	32.63	6400	28.98	0.89	0.17
	1000 - 2500	32863	18.72	7360	33.33	1.78	0.35
	2500 - 3500	27695	15.78	2176	9.85	0.62	0.12
	3500 - 4500	25365	14.45	4480	20.28	1.40	0.27
	>4500	32292	18.40	1664	7.53	0.41	0.08
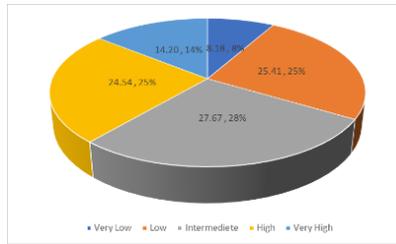
Source: Analysis, 2022

Figure 12. Diagram of Comparison of Landslide Susceptibility Area in (%)

In this research, focus is given to a relatively small research area that has a high probability of landslides. This aims to provide a more specific analysis compared to previous research conducted at the sub-district or district level. By utilizing ML method, this research is one of the first to apply this approach to a highway area that frequently experiences landslides, including the 12 most frequently landslide areas, according to the Regional Disaster Management Agency (BPBD).
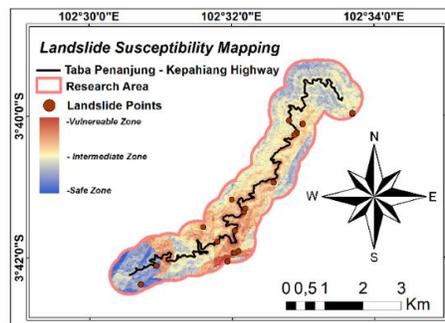
Source: Analysis, 2022

Figure 13. Landslide Susceptibility Mapping

In this research we have been successfully implementing the ML approach to access and to map landslide vulnerability along Taba Penanjung Kepahiang road. Previously The ML has been widely used in landslide mapping in several regions in Indonesia, such as Aldiansyah & Wardani (2024), Darminto et al. (2021) and Irawan et al., (2021). Previous researches used ML with different algorithms, with almost the same input parameters. The broad field of study makes it easier to process data with good output, which is different from this research. This research only focuses on the road area, so it only maximizes the FR algorithm, but the results of this research can be used in the landslide mitigation process. Finally, our results can also become a reference for the government and stakeholders in regional development, planning, and disaster management. It is also hoped that the resulting mapping can become an effective pre-disaster tool for carrying out specific and optimal mitigation in the area so that losses due to landslides can be minimized. It is hoped that this research can significantly contribute to efforts to prevent and manage disasters in highway areas that are vulnerable to landslides.

4. Conclusion

Landslides are described as rock or soil movements influenced by various factors, causing economic and social losses. The Bengkulu-Kepahiang route has been identified as vulnerable due to diverse geomorphological conditions and landslides that disrupt road access and cause casualties The lack of alternative ways to advertise the impact of landslides on transportation has prompted the need for research to analyze and determine the dangers of landslides. This research aims to identify areas prone to landslides using ML as an effective disaster mitigation tool. ML models are essential for landslide modeling (Purwanto, 2021), because they provide accurate vulnerability maps by combining several input parameters. The study area of the Taba Penanjung-Kepahiang Road has varying levels of vulnerability to landslides caused by several factors such as altitude, geological conditions, slope aspect, rainfall, curvature of the land, and distance from the fault. This research uses the FR

method to map landslide-prone areas, emphasizing the need for a holistic approach to handling and preventing landslides. Vulnerability analysis integrates various factors, including topography, lithology, vegetation, distance from roads, and drainage characteristics. The Landslide Susceptibility Index (LSI) is calculated by adding up the FR values for each factor thereby contributing to developing a comprehensive Disaster Mitigation Model using ML.

Mapping landslide-prone areas can use a ML method approach with Variable FR to help find the spatial relationship between landslide events and conditioning factors extracted using the weights of each class of each conditioning factor with an area. Data processing uses 16 independent and dependent factors as historical points in the occurrence of landslides in the last five years. Data processing uses pixel or grid methods with a specific grid size for each factor attribute value of the ML model. Modeling-based representation of slope stability requires several indicators involving local characterization of landslide geometry and internal structure. The data processing results are a landslide hazard map with five classifications: very low, low, medium, high, and very high (very vulnerable). This research is still dominated by medium areas, precisely 28%. For future research, adding more parameters with more detailed classification and higher correlation can make the output results more accurate. This study emphasizes the importance of multidimensional landslide mitigation strategies, including sustainable land use planning, reforestation, early warning systems, and advanced technologies. The research aims to reduce the risk and impact of landslides, optimize transportation access, and increase community awareness and preparedness This research contributes valuable information to designing risk mitigation programs, facilitating rapid response to emergencies, ensuring safer road traffic, and minimizing material and non-material impacts.

5.   Acknowledgments

6.   References

Alcántara-Ayala, I., & Sassa, K. (2023). Landslide risk management: from hazard to disaster risk reduction. *Landslides, 20*(10), 2031-2037. [Crossref]

Aldiansyah, S., & Wardani, F. (2024). Assessment of resampling methods on performance of landslide susceptibility predictions using machine learning in Kendari City, Indonesia. *Water Practice & Technology, 19*(1), 52-81. [Crossref]

Arabameri, A., Pourghasemi, H. R., & Yamani, M. (2017). Applying different scenarios for landslide spatial modeling using computational intelligence methods. *Environmental earth sciences, 76*, 1-20. [Crossref]

Bednarik, M., Magulová, B., Matys, M., & Marschalko, M. (2010). Landslide susceptibility assessment of the Kraľovany–Liptovský Mikuláš railway case study. *Physics and Chemistry of the Earth, Parts A/B/C, 35*(3-5), 162-171. [Crossref]

Çevik, E., & Topal, T. (2003). GIS-based landslide susceptibility mapping for a problematic segment of the natural gas pipeline, Hendek (Turkey). *Environmental geology, 44*, 949-962. [Crossref]

Chen, X., & Chen, W. (2021). GIS-based landslide susceptibility assessment using optimized hybrid machine learning methods. *Catena, 196*, 104833. [Crossref]

Chopra, S., & Marfurt, K. (2007). Seismic curvature attributes for mapping faults/fractures, and other stratigraphic. CSEG Recorder, (November), 38–42.

Ciurleo, M., Calvello, M., & Cascini, L. (2016). Susceptibility zoning of shallow landslides in fine grained soils by statistical methods. *Catena, 139*, 250-264. [Crossref]

Cruden, D., & VanDine, D. F. (2013). Classification, description, causes and indirect effects-Canadian technical guidelines and best practices related to landslides: a national initiative for loss reduction. *Geological survey of Canada, Open file, 7359*, 22. [Crossref]

Dahal, R. K., Hasegawa, S., Nonomura, A., Yamanaka, M., Masuda, T., & Nishino, K. (2008). GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. *Environmental Geology, 54*, 311-324. [Crossref]

Dai, F. C., & Lee, C. F. (2001). Terrain-based mapping of landslide susceptibility using a geographical information system: a case study. *Canadian Geotechnical Journal, 38*(5), 911-923. [Crossref]

Darminto, M. R., Widodo, A., Alfatinah, A., & Chu, H. J. (2021). High-resolution landslide susceptibility map generation using machine learning (Case Study in Pacitan, Indonesia). *International Journal on Advanced Science, Engineering and Information Technology, 11*(1), 369-379. [Crossref]

Dou, J., Yunus, A. P., Bui, D. T., Merghadi, A., Sahana, M., Zhu, Z., ... & Pham, B. T. (2020). Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. *Landslides, 17*, 641-658. [Crossref]

Fadilah, N., Arsyad, U., & Soma, A. S. (2019). Analisis tingkat kerawanan tanah longsor menggunakan metode frekuensi rasio di Daerah Aliran Sungai Bialo. *Perennial, 15*(1), 42-50. [Crossref]

Fan, H., Lu, Y., Hu, Y., Fang, J., Lv, C., Xu, C., ... & Liu, Y. (2022). A landslide susceptibility evaluation of highway disasters based on the frequency ratio coupling model. *Sustainability, 14*(13), 7740. [Crossref]

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences, 18*(8), 2161-2181. [Crossref]

Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S. R., Tiede, D., & Aryal, J. (2019). Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sensing, 11*(2), 196. [Crossref]

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K. T. (2012). Landslide inventory maps: New tools for an old problem. *Earth-Science Reviews, 112*(1-2), 42-66. [Crossref]

Hadi, A. I., Farid, M., Harlianto, B., & Sari, J. I. (2021). Landslide Potential Investigation for Disaster Risk Reduction in Central Bengkulu Regency, Bengkulu Province, Indonesia. *Indonesian Journal on Geoscience, 8*(3). [Crossref]

Hadi, A. I., & Siswanto, B. K. (2016). Landslide potential analysis using microtremor and slope data on Bengkulu-Kepahiang Main Road at Km 31–60. *IOSR Journal of Applied Geology and Geophysic, 4*, 9-14.

Hadi, A. I., Brotopuspito, K. S., Pramumijoyo, S., & Hardiyatmo, H. C. (2018). Regional landslide potential mapping in earthquake-prone areas of Kepahiang Regency, Bengkulu Province, Indonesia. *Geosciences, 8*(6), 219. [Crossref]

He, S., Pan, P., Dai, L., Wang, H., & Liu, J. (2012). Application of kernel-based Fisher discriminant analysis to map landslide susceptibility in the Qinggan River delta, Three Gorges, China. *Geomorphology, 171*, 30-41. [Crossref]

Huang, F., Cao, Z., Guo, J., Jiang, S. H., Li, S., & Guo, Z. (2020). Comparisons of heuristic, general statistical and machine learning models for landslide susceptibility prediction and mapping. *Catena, 191*, 104580. [Crossref]

Irawan, L. Y., Bachri, S., Panoto, D., Pradana, I. H., Faizal, R., Devy, M. M. R., & Prasetyo, W. E. (2021). The Use of Machine Learning for Accessing Landslide Susceptibility Class: Study Case of Kecamatan Pacet, Kabupaten Mojokerto. In *IOP Conference Series: Earth and Environmental Science* (Vol. 884, No. 1, p. 012006). IOP Publishing. [Crossref]

Kavzoglu, T., Colkesen, I., & Sahin, E. K. (2019). Machine learning techniques in landslide susceptibility mapping: a survey and a case study. *Landslides: Theory, practice and modelling*, 283-301. [Crossref]

Koutsoyiannis, D. (2004). Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: II. Recherche empirique sur de longues séries de précipitations. *Hydrological Sciences Journal, 49*(4). [Crossref]

Lee, J. H., Sameen, M. I., Pradhan, B., & Park, H. J. (2018). Modeling landslide susceptibility in data-scarce environments using optimized data mining and statistical methods. *Geomorphology, 303*, 284-298. [Crossref]

Lee, S., & Min, K. (2001). Statistical analysis of landslide susceptibility at Yongin, Korea. *Environmental geology, 40*, 1095-1113. [Crossref]

Lisle, R. J. (1994). 高斯曲率与高应变被引400, 12(12), 1811–1819.

Marjanović, M., Kovačević, M., Bajat, B., & Voženílek, V. (2011). Landslide susceptibility assessment using SVM machine learning algorithm. *Engineering Geology, 123*(3), 225-234. [Crossref]

Mohammady, M., Pourghasemi, H. R., & Pradhan, B. (2012). Landslide susceptibility mapping at Golestan Province, Iran: a comparison between frequency ratio, Dempster–Shafer, and weights-of-evidence models. *Journal of Asian Earth Sciences, 61*, 221-236. [Crossref]

Moore, I. D., & Grayson, R. B. (1991). Terrain-based catchment partitioning and runoff prediction using vector elevation data. *Water Resources Research, 27*(6), 1177-1191. [Crossref]

Nagarajan, R., Roy, A., Vinod Kumar, R., Mukherjee, A., & Khire, M. V. (2000). Landslide hazard susceptibility mapping based on terrain and climatic factors for tropical monsoon regions. *Bulletin of Engineering Geology and the Environment, 58*, 275-287. [Crossref]

Natasya, I. D., Larang, M. P., Putri, E. G. G., & Refrizon, R. R. (2022). Upaya Mitigasi Bencana Longsor Jalan Lintas Bengkulu-Kepahiang Berdasarkan Kecepatan Gelombang Geser (Vs). *Newton-Maxwell Journal of Physics, 3*(1), 33-37. [Crossref]

Nourani, V., Pradhan, B., Ghaffari, H., & Sharifi, S. S. (2014). Landslide susceptibility mapping at Zonouz Plain, Iran using genetic programming and comparison with frequency ratio, logistic regression, and artificial neural network models. *Natural hazards*, *71*, 523-547. [Crossref]

Okoli, J., Nahazanan, H., Nahas, F., Kalantar, B., Shafri, H. Z. M., & Khuzaimah, Z. (2023). High-Resolution Lidar-Derived DEM for Landslide Susceptibility Assessment Using AHP and Fuzzy Logic in Serdang, Malaysia. *Geosciences*, *13*(2), 34. [Crossref]

Ozdemir, A., & Altural, T. (2013). A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. *Journal of Asian Earth Sciences*, *64*, 180-197. [Crossref]

Pachauri, A. K., Gupta, P. V., & Chander, R. (1998). Landslide zoning in a part of the Garhwal Himalayas. *Environmental Geology*, *36*, 325-334. [Crossref]

Pourghasemi, H. R., Gayen, A., Park, S., Lee, C. W., & Lee, S. (2018). Assessment of landslide-prone areas and their zonation using logistic regression, logitboost, and naïvebayes machine-learning algorithms. *Sustainability*, *10*(10), 3697. [Crossref]

Regmi, A. D., Devkota, K. C., Yoshida, K., Pradhan, B., Pourghasemi, H. R., Kumamoto, T., & Akgun, A. (2014). Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. *Arabian Journal of Geosciences*, *7*, 725-742. [Crossref]

Roberts, A. (2001). Curvature attributes and their application to 3D interpreted horizons. *First break*, *19*(2), 85-100. [Crossref]

Roodposhti, M. S., Aryal, J., Lucieer, A., & Bryan, B. A. (2019). Uncertainty assessment of hyperspectral image classification: Deep learning vs. random forest. *Entropy*, *21*(1), 1–15. [Crossref]

Rotaru, A., Oajdea, D., & Răileanu, P. (2007). Analysis of the landslide movements. *International journal of geology*, *1*(3), 70-79.

Shahabi, H., Khezri, S., Ahmad, B. B., & Hashim, M. (2014). Landslide susceptibility mapping at central Zab basin, Iran: a comparison between analytical hierarchy process, frequency ratio and logistic regression models. *Catena*, (115), 55-70. [Crossref]

Shou, K. J., & Lin, J. F. (2020). Evaluation of the extreme rainfall predictions and their impact on landslide susceptibility in a sub-catchment scale. *Engineering Geology*, *265*, 105434. [Crossref]

Solaimani, K., Mousavi, S. Z., & Kavian, A. (2013). Landslide susceptibility mapping based on frequency ratio and logistic regression models. *Arabian Journal of Geosciences*, *6*, 2557-2569. [Crossref]

Sugianto, N., & Refrizon, R. (2021). Struktur Kecepatan Gelombang Geser (Vs) di Daerah Rawan Gerakan Tanah (Longsor) Jalan Lintas Kabupaten Bengkulu Tengah-Kepahiang. *Indonesian Journal of Applied Physics*, *11*(2), 134-142. [Crossref]

Suhendra, Z. C. B., & Sugianto, N. (2018). Geological condition at landslides potential area based on microtremor survey. *ARPN Journal of Engineering and Applied Sciences*, *13*(8), 3007-3013.

Umar, Z., Pradhan, B., Ahmad, A., Jebur, M. N., & Tehrany, M. S. (2014). Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in West Sumatera Province, Indonesia. *Catena*, *118*, 124-135. [Crossref]

van Asch, T. W., Malet, J. P., van Beek, L. P., & Amitrano, D. (2007). Techniques, issues and advances in numerical modelling of landslide hazard. *Bulletin de la Société géologique de France*, *178*(2), 65-88. [Crossref]

Wallis, J. R., Schaefer, M. G., Barker, B. L., & Taylor, G. H. (2007). Regional precipitation-frequency analysis and spatial mapping for 24-hour and 2-hour durations for Washington State. *Hydrology and Earth System Sciences*, *11*(1), 415-442. [Crossref]

Wang, Y., Sun, D., Wen, H., Zhang, H., & Zhang, F. (2020). Comparison of random forest model and frequency ratio model for landslide susceptibility mapping (LSM) in Yunyang County (Chongqing, China). *International journal of environmental research and public health*, *17*(12), 4206. [Crossref]

Wei, R., Ye, C., Sui, T., Ge, Y., Li, Y., & Li, J. (2022). Combining spatial response features and machine learning classifiers for landslide susceptibility mapping. *International Journal of Applied Earth Observation and Geoinformation*, *107*, 102681. [Crossref]

Xiao, T., Yin, K., Yao, T., & Liu, S. (2019). Spatial prediction of landslide susceptibility using GIS-based statistical and machine learning models in Wanzhou County, Three Gorges Reservoir, China. *Acta Geochimica*, *38*, 654-669. [Crossref]

Yilmaz, I. (2009). Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat—Turkey). *Computers & Geosciences*, *35*(6), 1125-1138. [Crossref]

Zhu, L., Wang, G., Huang, F., Li, Y., Chen, W., & Hong, H. (2021). Landslide susceptibility prediction using sparse feature extraction and machine learning models based on GIS and remote sensing. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1-5. [Crossref]