EEET ECOLOGICAL ENGINEERING & ENVIRONMENTAL TECHNOLOGY

Ecological Engineering & Environmental Technology, 2026, 27(12), 229–250
https://doi.org/10.12912/27197050/240382
ISSN 2719–7050, License CC-BY 4.0

Received: 2026.09.19
Accepted: 2026.10.20
Published: 2026.11.02

Comparative performance of maximum entropy, frequency ratio, and logistic regression models for landslide susceptibility mapping in Semarang Regency, Indonesia

Muhammad Fauzan Ramadhan[1*], Sriyono[1], Ayu Wulansari Pramita[1], Halim Sukma Aji[1], Andi Irwan Benardi[1], Verin Kalina[1], Ayu Nur Rofi[1], Katerine Karen Kahyono[1]

[1] Department of Geography, Faculty of Social and Political Sciences, Universitas Negeri Semarang, Kampus Sekaran, 50229 Semarang, Indonesia
* Corresponding author's e-mail: muhammad.fauzan.r@mail.unnes.ac.id

ABSTRACT

Landslide susceptibility mapping in many regions must rely on inventories compiled from community reporting, which provide presence data only and no field-verified non-occurrence points. Which modelling approach best suits such data remains unclear. This study compares Maximum Entropy (MaxEnt), Frequency Ratio (FR), and logistic regression (LR) for landslide susceptibility mapping in Semarang Regency, Indonesia, using an identical set of conditioning factors and an identical validation scheme so that differences in performance reflect the methods themselves. An inventory of 445 landslide points, compiled from the records of the Regional Disaster Management Agency for 2021-2025 and from field survey, was modelled against 13 continuous conditioning factors at 30 m resolution and evaluated through spatial-block cross-validation (5 km blocks, five folds, ten random block assignments), with random five-fold cross-validation as a comparison. Discrimination was assessed with AUC, TSS at a threshold fixed on the training data, and McClish-standardised partial AUC restricted to the high-specificity range (0.8-1.0), with differences tested using the DeLong test and a paired bootstrap. MaxEnt and FR performed equivalently on the full AUC (0.688 and 0.691, mean of the ten block assignments; DeLong $p < 0.05$ in one of ten) and on TSS (0.278 and 0.269), both within the range of moderate discrimination. LR scored lower (AUC 0.653; TSS 0.222) and differed significantly from both, a gap that persisted across 1,000 pseudo-absence datasets and when pseudo-absences were kept at least 500 m or 1000 m from landslide points. MaxEnt outperformed FR only over the high-specificity range (pAUC 0.610 against 0.578; bootstrap $p < 0.005$ in all ten block assignments) and in the most susceptible class, where it captured 57.5% of events against 52.4% for FR and 45.6% for LR. The ranking of the models was unchanged when the inventory was thinned to a minimum of 100 m or 250 m between points. Slope was the strongest predictor in both models that produce importance scores. Presence-based modelling is therefore adequate for report-based inventories, with FR suited to rapid and transparent mapping and MaxEnt to sharpening the most susceptible zone.

Keywords: presence-only data; pseudo-absence; partial AUC; detection bias; landslide inventory; Central Java.

INTRODUCTION

Landslides are among the natural hazards that occur almost every year in Indonesia. Their impacts include fatalities, infrastructure damage, and considerable property loss. Indonesia Disaster Agency or Badan Nasional Penanggulangan Bencana, (2025) recorded 958 landslide events across Indonesia in 2025. Landslides in Indonesia are generally triggered by a combination of steep topography, high rainfall, and susceptible geological and land cover conditions (Hadmoko et al., 2017; Ngadisih et al., 2017; Samodra et al., 2020). According to the Disaster Risk Assessment

(KRB) document of Central Java Province, Semarang Regency is among the areas with a high level of landslide risk. The Regional Disaster Management Agency (BPBD) of Semarang Regency recorded 463 landslide events in the regency between 2021 and 2025 (BPBD Semarang Regency, unpublished records). Susceptibility assessment is therefore required as a basis for spatial planning and mitigation.

Geomorphologically, Semarang Regency is shaped by a range of landform-forming processes. Volcanism is the dominant process shaping this area, marked by the presence of the Ungaran, Telomoyo, and Merbabu mountains (Koesoemadinata, 2020; Van Bemmelen, 1949). Structural influence is also evident, among others through the Rawa Pening Fault and the Merapi-Merbabu structural lineament (Soehaimi, 2008). Denudational processes develop on steeply sloping areas, while fluvial processes are widespread along the Tuntang River and the upper reaches of the Garang River (Hanafi & Pamungkas, 2021). The combination of these geomorphological processes produces a complex landscape with varied slope characteristics. These complex geomorphological conditions contribute to the high landslide susceptibility in Semarang Regency. Understanding the morphological characteristics of the area is therefore important for determining landslide susceptibility spatially.

The development of Geographic Information Systems and remote sensing has encouraged the application of various approaches to landslide susceptibility modelling, which can broadly be divided into heuristic, deterministic, and data-driven approaches. Heuristic approaches rely on expert judgement for weighting (Ado et al., 2022; Liu et al., 2023), but their application depends heavily on the subjective judgement of experts. Statistical approaches such as Logistic Regression (Cemiloglu et al., 2023; Sujatha & Sridhar, 2021; Sun et al., 2021), Weights of Evidence (Gentilucci et al., 2023; Turahmansyah et al., 2026), and Frequency Ratio (Çağan et al., 2026) have also been widely used and provide reasonably good results. With the advance of machine learning, methods such as Deep Learning (Habumugisha et al., 2022)*, Artificial Neural Network (ANN)* (Azarafza et al., 2021; Nikoobakht et al., 2022; Soma et al., 2026), Support Vector Machine (SVM) (Kamran et al., 2021; Ullah et al., 2022; Xu et al., 2024), and Random Forest (RF) (He et al., 2021; Li et al., 2024; Md.Sharafat et al., 2024; Zul et al., 2026) show higher accuracy. Machine learning methods generally report higher accuracy but require larger training datasets and produce models that are more difficult to interpret physically.

The choice of method is influenced by the availability of landslide inventory data. Landslide inventory data from the Regional Disaster Management Agency (BPBD) originate from community reporting, so almost all recorded landslides are located near settlements and close to roads. In data terms, such information constitutes presence data that can be used to assess landslide susceptibility. Modelling that makes use of event data generally produces more accurate estimates than heuristic approaches that rely on expert judgement for weighting (Taloor et al., 2026; Vujović et al., 2025). In data-driven approaches, the relationship between conditioning factors such as slope, rainfall, and vegetation cover and landslide occurrence is derived directly from observed locations rather than specified beforehand (Barman et al., 2023). Several methods commonly used for this purpose are Maximum Entropy, Frequency Ratio, and Logistic Regression. All three can be applied to event data, but they treat the absence of non-occurrence data in different ways (Chen et al., 2017). This study therefore aims to identify these differences.

The three methods differ fundamentally in how they derive landslide susceptibility. Frequency Ratio is a bivariate method that compares the proportion of events to the proportion of area in each class of a conditioning factor (Silalahi et al., 2019; Yalcin et al., 2011). Its computational simplicity makes the method easy to replicate and widely applied in landslide susceptibility studies (Çağan et al., 2026). Logistic regression is a multivariate method that models the probability of occurrence as a function of all factors simultaneously, and remains one of the most widely used statistical approaches in landslide susceptibility mapping (Cemiloglu et al., 2023; Sujatha & Sridhar, 2021; Sun et al., 2021). Maximum Entropy was developed for species distribution modelling based on presence data (Phillips et al., 2006, 2017) and has more recently been adapted for disaster studies with promising results (Boussouf et al., 2023). Unlike the other two, this method is able to capture non-linear relationships between conditioning factors and landslide occurrence (Huang et al., 2025; Ilinca et al., 2025; Orefice & Innocenti, 2025). Although all three are commonly used, direct comparison between bivariate, multivariate,

and maximum-entropy-based methods on identical data and identical validation schemes remains limited, so the basis for selecting a method for report-based inventories is not yet clear.

This study compares the performance of three landslide susceptibility modelling methods (Maximum Entropy, Frequency Ratio, and logistic regression) using the landslide inventory of Semarang Regency compiled by BPBD. The three models are built using an identical set of conditioning factors and an identical cross-validation scheme, so that differences in their performance reflect differences between methods rather than differences in data treatment. The evaluation covers discrimination ability overall as well as over the high-specificity range, the range that determines the assignment of the highest susceptibility class on the final map. The results are expected to provide a sounder basis for method selection in producing susceptibility maps from report-based landslide inventories, while also producing a landslide susceptibility map of Semarang Regency.

METHODS

Study area

Semarang Regency lies in the central part of Central Java Province, Indonesia, extending between 110°15′ and 110°40′ E and between 7°05′ and 7°30′ S, and is divided into 19 districts. The area displays diverse topography, from lowlands in the north to volcanic slopes in the south, with elevations ranging from 41 to 3,072 m above sea level. This geographic position and topographic diversity make Semarang Regency a representative study area for landslide susceptibility assessment (Figure 1).

Lithologically, Semarang Regency is dominated by volcanic rocks of the Merbabu and Ungaran volcanic complexes. Volcanic rock is the dominant material in Semarang Regency (Koesoemadinata, 2020), and it largely forms terrain that tends to be steep. The abundance of volcanic rock can produce thick soils through intensive weathering, particularly in tropical areas (Noviyanto et al., 2020). These conditions make landslides more likely to occur. Alluvial sedimentary material is mostly found in areas with flat and concave topography. Slope in Semarang Regency is dominated by gentle to flat terrain, but shows contrasting variation because of several high-relief areas. Mount Ungaran and Mount Merbabu contribute steep slopes (>15°) as well as the highest elevations (>1,500 m asl). In addition, the rivers in this area are generally headwater streams, so steep slopes are also common around river channels. Landslide events are mostly found on relatively steep terrain.

Annual rainfall in Semarang Regency ranges from about 2,749 mm/year to 4,077 mm/year,

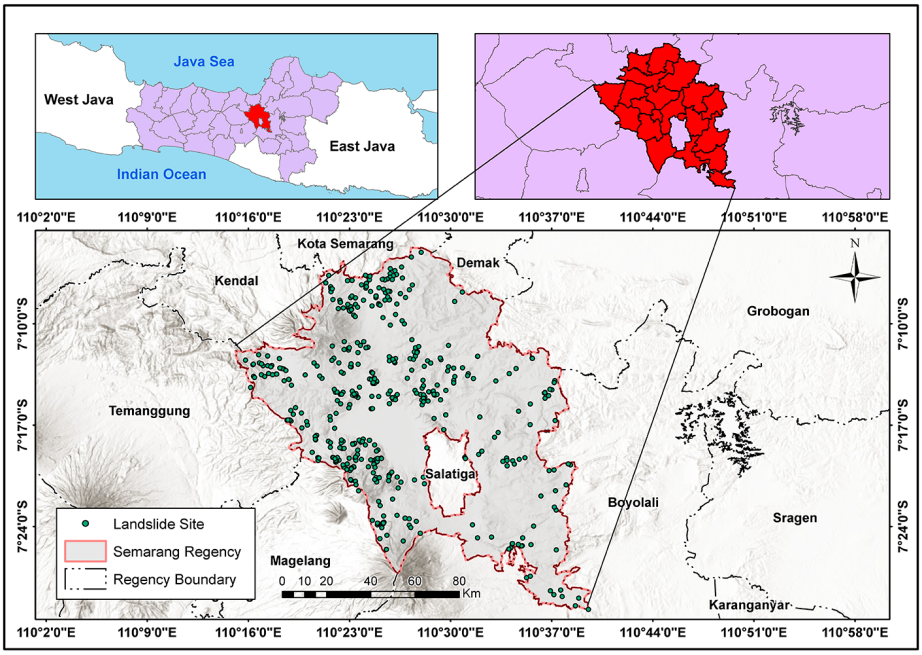

Figure 1. Study area and landslide inventory points

with a mean of 3,031 mm/year. The highest rainfall occurs on the upper volcanic slopes, following the orographic pattern common in volcanic areas. The rainy season in this area lasts from about November to April. Based on the landslide inventory, most landslide events occurred within that period. In terms of vegetation cover, Semarang Regency remains largely dominated by vegetation, as also indicated by a mean NDVI of 0.69. Land use is dominated by dryland agriculture, plantations, and forest on the upper slopes, while settlements are concentrated in the valleys and along the main road corridors connecting Semarang, Ambarawa, and Salatiga.

Data collection

Landslide inventory

Landslide events in this study were compiled from the records of the Regional Disaster Management Agency (BPBD), combined with Google Earth observation and direct field recording. The BPBD landslide inventory is the product of community reports of landslide events in Semarang Regency. Spatial information from the BPBD reports was obtained through two approaches, namely reading the EXIF metadata of the field photographs accompanying each report and refining the position through interpretation of landslide scarps on historical satellite imagery. The inventory draws on the BPBD records for 2021 to 2025, which list 463 landslide events with their date, village, and district. Records were excluded when their location description was incomplete or their photographs were insufficient to identify and locate the failure. Because the relocated points were not stored with the identifier of their source record, the number of excluded records cannot be counted exactly, 58 of the 463 records lie in villages that contain no relocated point. Relocation yielded 357 points from these records; where one record showed several separate failures, each failure was relocated as a separate point. Duplicate reports were screened during relocation records from the same hamlet (dusun) with close event dates were compared through their photographs, and when the photographs showed the same failure the records were treated as one event and relocated as a single point, whereas records showing different failures were relocated separately. Field recording between 21 June and 2 July 2026 yielded 120 GPS-located points, 49 at the sites of events listed in the BPBD records and 71 at additional landslide sites without a BPBD record. The survey team visited the sites of BPBD-recorded events and recorded landslides when visible from the roads travelled that were not in the BPBD records. Each point was recorded with the GPS Map Camera smartphone application, which stores the coordinates in a geotagged photograph.

The combined 477 points were thinned to one point per 30 m cell of the analysis grid. Seven points fell outside the regency boundary and one field record had an invalid coordinate; these eight were removed. A further 24 points shared a 30 m cell with another point (23 cells), and in each such cell one point was kept, with a field-surveyed point retained in preference to a relocated one. The points within a shared cell were at most 21 m apart and in 16 of the 23 cells less than 3 m apart. Eight of these cells held a BPBD record and a field point of the same event, so thinning also removed double counting between the two sources. There are eleven point held two or more relocated BPBD points, which may be repeated reports of one event, separate failures of one event, or repeated events at the same site. The final inventory contains 445 points, 333 from the BPBD records and 112 from field recording. Thinning by cell leaves close points in adjacent cells: after thinning, 21 pairs of points were less than 50 m apart and 46 pairs less than 100 m apart. Because some of these may still be reports of one event, the analysis was repeated with the inventory thinned to a minimum distance of 100 m (401 points) and 250 m (347 points) between points, retaining field points in preference to relocated points. Landslide coordinates were recorded at the crown, as this is the part where the landslide movement originates. All landslide types were treated as a single group, so that the inventory forms a single landslide class without subtypes. Merging the various landslide types was intended to simplify the dataset for the sake of modelling effectiveness and efficiency. Forming the data into binary groups allows the model to recognise spatial patterns more clearly, so that the analysis of conditioning factors becomes more interpretable (Orefice & Innocenti, 2025).

Environmental covariates

Thirteen conditioning factors were used as predictors, all of them continuous data. Continuous data were used so that the comparison could be made on equal terms. The predictor data were

harmonised to a spatial resolution of 30 m, following the native resolution of the Landsat 8/9 imagery underlying NDVI and NDBI. The finer DEMNAS elevation model was aggregated to 30 m, while CHIRPS rainfall, which has a coarser resolution (~5.5 km), was resampled to the same grid. Its spatial detail remains coarse because no finer-resolution rainfall data representative of this area were available. Terrain derivatives were computed in SAGA GIS 9.7.2 from DEMNAS (Badan Informasi Geospasial, 2026) after aggregation to 30 m, using the SAGA tools Slope, Aspect, Curvature (slope, aspect, plan and profile curvature), Topographic Wetness Index (One Step) (TWI), Topographic Position Index (TPI), and Vertical Distance to Channel Network (HAND). NDVI and NDBI were computed in Google Earth Engine from a median composite of Landsat 8 and Landsat 9 Collection 2 Level-2 (Tier 1) surface reflectance scenes acquired between 1 January 2021 and 31 December 2025, the same period as the BPBD records (84 Landsat 8 and 74 Landsat 9 scenes). Scenes with more than 30% cloud cover were excluded, and cloud and cloud-shadow pixels were masked with the QA_PIXEL band before compositing. NDVI was computed from the red and near-infrared bands (B4, B5) and NDBI from the near-infrared and shortwave-infrared bands (B5, B6). Rainfall is the mean annual CHIRPS precipitation for 2013-2025, obtained by summing the daily CHIRPS data over the 13 years and dividing by 13. The year 2026 was excluded because it was still incomplete at the time of processing, and a partial year would bias the annual mean downward. As a multi-year mean, the rainfall layer represents the long-term climatic setting of each location rather than the rainfall that triggered individual events. All layers were aligned to a common 30 m grid (WGS 84 / UTM zone 49S, EPSG:32749), using bilinear resampling for layers not already aligned, and masked to the regency boundary. Cells lacking a value for any covariate were excluded, leaving 1,121,289 valid cells, and a background sample of 10,000 points was drawn at random from these cells. A more detailed explanation of each factor is presented below.

Elevation

Elevation describes the vertical position of a landslide event. This condition affects surface temperature as well as the geomorphological processes that operate. Elevation data can be used as one of the controlling factors of landsliding when the study area is mountainous (Cellek, 2020). Because Semarang Regency also has mountainous terrain, elevation was used as a parameter in determining landslide susceptibility. Elevation data were obtained from DEMNAS and resampled to a cell size of 30 x 30 m.

Slope

Slope is a variable frequently used to identify landslide susceptibility (Saleem et al., 2019). This is because slope is a controlling factor of the geomorphological processes that take place (McColl, 2022). Landsliding is one form of such geomorphological process, so this variable was used to assess landslide susceptibility in Semarang Regency. Slope data were derived from the elevation data obtained previously.

Plan curvature

Plan curvature describes the convergence or divergence of surface flow in the direction perpendicular to the slope. Curvature may be concave, convex, or flat. Concave slopes can be one cause of landsliding because they tend to collect water, which makes the soil heavier (Ohlmacher, 2007). These data were also obtained by deriving them from the elevation data.

Profile curvature

Profile curvature describes the change in gradient along the direction of slope flow. This value governs the acceleration and deceleration of surface flow as well as the pattern of erosion and deposition of material along the slope. A concave slope form can be an indicator of landsliding, particularly of the rotational type, because it produces a change in slope form into a concave shape (Rajabian, 2023). These data were derived together with plan curvature.

Topographic wetness index

TWI measures the potential accumulation of water at the soil surface. The TWI variable is often used as an index of the interaction between topographic influence and hydrological processes (Rózycka et al., 2017). High TWI values indicate potential water accumulation,

marking areas that are easily saturated, and such conditions make landsliding more likely to occur (Conoscenti et al., 2008).

Topographic position index

TPI measures the relative position of a point with respect to its surroundings. TPI can therefore be used to distinguish relief features such as ridges, summits, slopes, and plains. TPI values can be used to identify landslides because they are able to detect areas with steep slopes (Nseka et al., 2019).

Aspect

Aspect is the compass direction that a slope faces. This study divided aspect into two components, northness and eastness, so that the processed data remain continuous. Slope aspect influences landslide occurrence because it is related to geomorphological processes and to climatic conditions (Gorokhovich & Vustianiuk, 2021).

Height above nearest drainage (HAND)

HAND expresses the vertical height of a point above the nearest channel. Lower HAND values indicate positions close to surface flow, while larger values indicate elevations far above surface flow. High values close to a river channel indicate the presence of steep slopes. HAND was therefore used because it can characterise the location of steep terrain .(Nobre et al., 2011)

NDVI

NDVI describes the density of vegetation in an area. Low NDVI values represent open conditions and are often associated with a higher probability of landsliding (Zhang et al., 2023). The opposite condition is shown by high NDVI values, which indicate areas covered by vegetation. Vegetated areas may reflect good land conservation practice and therefore tend not to experience landsliding (Qiu et al., 2024).

NDBI

NDBI describes built-up surfaces and buildings. This index serves as a proxy for the anthropogenic influence on landslide occurrence in Semarang Regency, because the landslide inventory database originates from community reporting. High NDBI values can be one cause of high landslide susceptibility (Thomas et al., 2021). Together with NDVI, this parameter was used to approximate land cover in Semarang Regency.

Rainfall

Rainfall is one of the main triggers of landsliding, particularly in tropical regions (Amarasinghe et al., 2024). Rainfall is also regarded as one of the factors with the strongest influence on landslide occurrence because it adds load to the soil, which then triggers slope failure (Liu et al., 2024). The varied topography of Semarang Regency also produces variation in rainfall that can contribute to landslide events. Mean annual rainfall was obtained from the CHIRPS climatology, with a range of 2,749 to 4,077 mm/year.

Distance to river

Distance to river reflects the influence of fluvial processes on slope stability. Slope stability is strongly affected by rivers, because a river is an active geomorphological agent in causing bank erosion and can trigger slope failure (Zhou et al., 2025). Distance was computed as the Euclidean distance to the river network from the Indonesian Topographic Map (Rupabumi Indonesia) at 1:25,000 (sheets 1408-521 Tegalrejo, 1408-522 Ngablak, 1408-523 Grabag, 1408-524 Ambarawa, 1408-541 Sumowono, 1408-542 Ungaran, 1408-611 Ampel, 1408-613 Salatiga, and 1408-631 Wiru) using SAGA GIS.

Multicollinearity

Multicollinearity testing was carried out to ensure that the predictors did not overlap excessively with one another. Highly correlated predictors can inflate the variance of the coefficients and reduce model stability. This testing refers to the interdependence among the parameters used. Interdependence among the 13 conditioning factors was examined in two ways: Pearson correlation for pairwise relationships, and the variance inflation factor (VIF) for joint collinearity. All testing was carried out in the R software environment. Both were computed on the covariate values at the 10,000 background points, and VIF was screened with the vifstep function of the usdm package at a threshold of 5. All 13 predictors had VIF values below 5 (Table 1; maximum 2.58, for elevation),

Table 1. Variance inflation factor (VIF) and tolerance (1/VIF) of the 13 conditioning factors, computed with usdm::vifstep on 10,000 random background points

Conditioning factor	VIF	Tolerance
Elevation	2.58	0.39
Slope	1.75	0.57
Plan curvature	1.20	0.83
Profile curvature	1.70	0.59
TWI	1.91	0.52
TPI	2.18	0.46
Northness	1.02	0.98
Eastness	1.01	0.99
HAND	1.56	0.64
NDVI	2.36	0.42
NDBI	2.15	0.47
Rainfall	2.35	0.43
Distance to river	1.24	0.80

so all of them were retained for the MaxEnt, FR, and LR modelling. One pair nevertheless had |r| above 0.7, rainfall and elevation (r = 0.741), followed by NDVI and NDBI (-0.680) and TPI and profile curvature (0.630). The full correlation matrix and the VIF of every predictor are given in Supplementary Table S1. Because these correlated predictors were retained, the importance of each member of a pair is interpreted with caution.

Data processing

Landslide susceptibility mapping was tested using three different forms of modelling, namely MaxEnt, FR, and LR. All testing was carried out in the R software environment using the landslide point data and the environmental covariates. The models were evaluated by spatial-block cross-validation, which is the primary validation of this study. The study area was overlaid with a grid of 5 km by 5 km blocks, the 67 blocks that contained landslide or background points were assigned at random to five folds, and every point took the fold of its block, so that the training and test points of a fold never share a block. Because a single random assignment can produce unbalanced folds, the assignment was repeated ten times (assignments r = 1-10, random seed 20260 + r); across assignments, a test fold held between 27 and 151 landslide points. The same folds were used for all three models: MaxEnt and FR were trained on the training presences and the training background, LR on the training presences and pseudo-absences drawn from the training background, and all three were evaluated on the same held-out landslide and background points. Random five-fold cross-validation (folds assigned at random with seed 42, separately for presence and background points, so that each fold held 89 presence points and 2,000 background points) is reported for comparison, because random folds can place training and test points close together and inflate performance when observations are spatially autocorrelated.

MaxEnt

Maximum Entropy (MaxEnt) modelling is a form of modelling that makes use of the available landslide event data. MaxEnt modelling was originally used to predict the habitat suitability of a species based on presence data for that species (Phillips et al., 2006). Such presence data are equivalent in nature to landslide event data, because both are presence-only data, so the approach can also be used to identify the probability of landslide susceptibility in an area (Huang et al., 2025; Ilinca et al., 2025). In principle, MaxEnt modelling computes the probability of landslide susceptibility by maximising entropy. By maximising entropy, MaxEnt selects the most uniform probability distribution that remains consistent with the observed data, so that partial information from the event points and the environmental factors is assembled into a coherent estimate of the distribution. The aim is to estimate the probability distribution (π) of landslide occurrence across all locations (X) defined within the study area, with each conditioning factor acting as a moment constraint that equates the expected value of the factor under the distribution π with the value observed at the event points. Relative susceptibility is expressed as an exponential function of the combination of covariates,

$$\eta(x) = \exp\left(\sum_j \beta_j \, f_j(x)\right) \quad (1)$$

where: $f_j(x)$ – function of the covariate and β_j – weight, which is then rescaled into a relative susceptibility index ranging from 0 to 1.

Because it models non-linear relationships between each factor and landslide occurrence, MaxEnt is able to capture susceptibility patterns that are not simple (Boussouf et al., 2023). MaxEnt

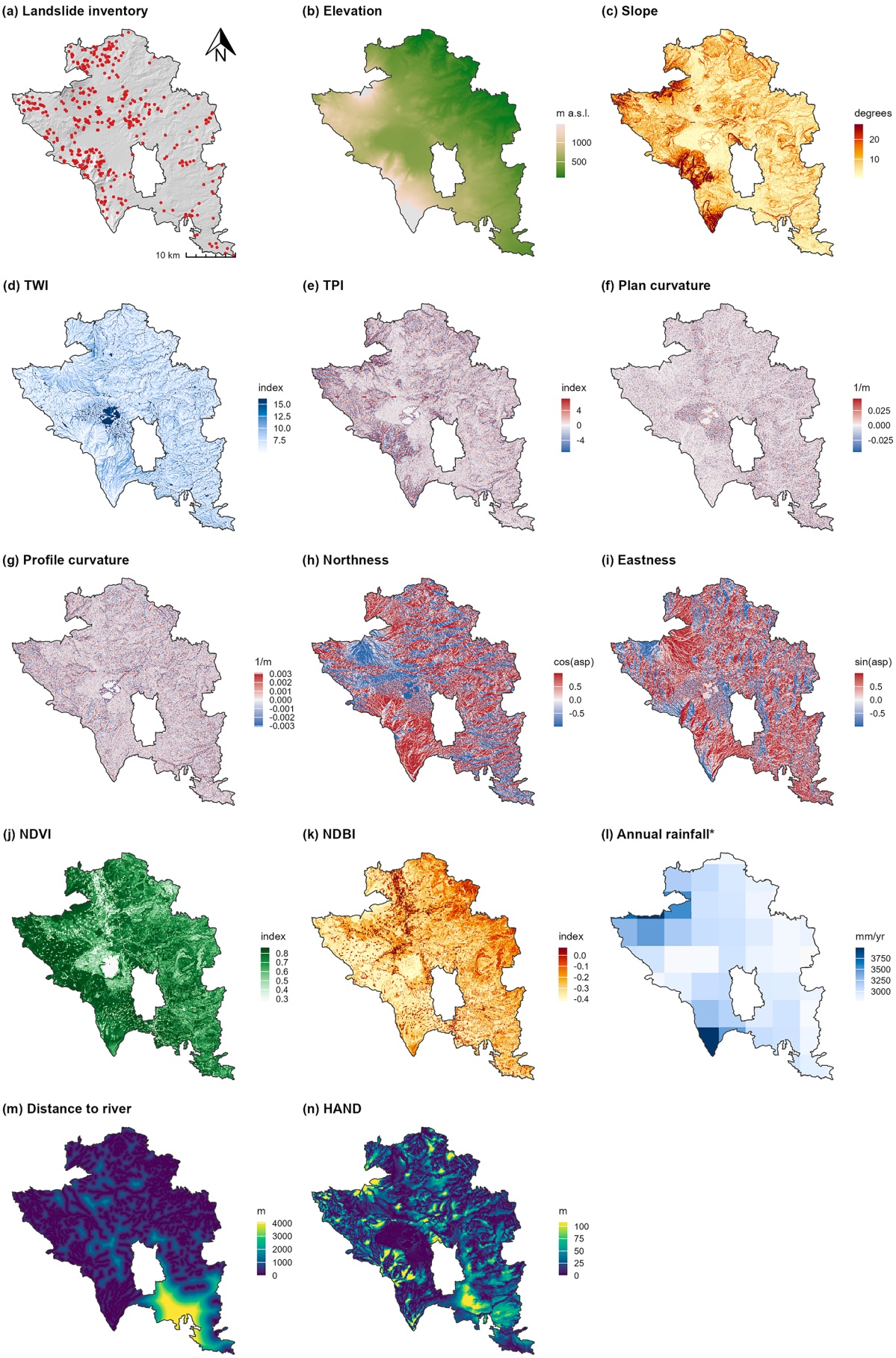

Figure 2. Environmental covariates

was fitted with the maxnet package (R) using linear, quadratic, and hinge features and the default regularisation, and predictions were expressed on the cloglog scale. maxnet implements the MaxEnt formulation as a penalised generalised linear model and is not the Java MaxEnt software, so results may differ slightly from that implementation.

Frequency ratio

Frequency ratio (FR) is a simple and transparent bivariate statistical method, and is therefore widely used in landslide susceptibility mapping (Çağan et al., 2026). The method computes a frequency ratio for each class of a factor and produces a susceptibility map from those ratios. Within the frequency ratio model, the spatial relationship between landslide locations and each factor contributing to landslide occurrence can be obtained. The method assesses the spatial relationship between the landslide distribution and each conditioning factor by comparing the density of events in each factor class with the proportion of area that class occupies. Each factor was divided into ten classes, and for the i-th class the frequency ratio was computed as

$$FR_i = \frac{N_i/N}{A_i/A} \tag{2}$$

where: N_i and N – are the number of landslides in the class and in total, and A_i and A – are the area of the class and the total area.

An FR value of 1 indicates an average association (events proportional to area), FR > 1 indicates a more susceptible class (landslides exceeding the area proportion), while FR < 1 indicates a safer class. Class breaks and frequency ratio values were computed from the training data of each fold only, because deriving class breaks from the full dataset would leak test information and inflate the AUC of the FR model. Class breaks were the deciles of the training background values of each factor, with the lowest and highest classes left open-ended. The susceptibility index of a cell is the sum of the FR values of the classes into which it falls across the 13 factors. The class breaks, counts, and FR values of every fold are given in Supplementary Table S2.

Logistic regression

Logistic regression (LR) is a multivariate statistical method that models the relationship between a binary variable (landslide present or absent) and several independent variables. Its advantage is that the predictors may be continuous or discrete and need not be normally distributed, which suits the varied nature of landslide conditioning factors (Budimir et al., 2015). Unlike MaxEnt and FR, LR requires non-landslide data as a comparison, so non-landslide (pseudo-absence) points were drawn at random from the study area. The probability of landslide occurrence is modelled as a logistic function of a linear combination of the predictors

$$P = \frac{1}{1 + e^{-(\beta_0 + \sum_i \beta_i x_i)}} \tag{3}$$

where: β_0 – is the intercept and β_i – the coefficient of each predictor estimated from the data.

The value of P ranges from 0 to 1 and expresses the probability of landslide susceptibility. Pseudo-absences were generated with the following algorithm. In fold k, the number of pseudo-absences equalled the number of training presences (ratio 1:1; 294 to 418 points, depending on the fold), and they were drawn at random without replacement from the training background points of the fold, that is, from the background points lying in the blocks of the other four folds. The background points are a uniform random sample of 10,000 of the 1,121,289 valid 30 m cells of the regency (spatSample in terra, seed 42), so the pseudo-absences are spread uniformly over the study area outside the test blocks. In the primary model no exclusion zone was applied: pseudo-absences were not restricted by distance from landslide points, by slope, or by any other criterion. The random seed of the draw was 1000 + k in every block assignment. A balanced ratio was chosen because it has been shown to yield a more equitable comparison (Dou et al., 2023). The 1:1 ratio was fixed as the primary model before the results were examined, and ratios of 1:2 and 1:5 were fitted under random folds only as a sensitivity check. Because random pseudo-absences can fall on unrecorded landslide-prone ground, which the detection bias of the inventory makes likely, two further checks were run in every block assignment. The draw was repeated 100 times (seed 1,000,000 + 10,000r + 10d + k for dataset d), giving 1,000 pseudo-absence datasets over the ten assignments, and the same 100 draws were repeated with pseudo-absences restricted to training background points lying at least 500 m, and in a second variant at least 1,000 m, from the nearest

training presence (Euclidean distance). These restrictions left on average 6,346 and 4,211 of about 8,000 training background points eligible.

Validation and accuracy

Accuracy was measured with two metrics, AUC and TSS. AUC (area under the ROC curve) assesses threshold-independent discrimination ability, while TSS (true skill statistic) is a threshold-based measure (Allouche et al., 2006). For a classification threshold t, the TSS is

$$\mathrm{TSS}(t) = \mathrm{Se}(t) + \mathrm{Sp}(t) - 1 \quad (4)$$

where: $Se(t)$ – is the sensitivity, the proportion of landslide points with a predicted value of at least t, and $Sp(t)$ – the specificity, the proportion of background points (pseudo-absences for LR) with a predicted value below t.

The threshold was determined solely from the training fold. In each fold, the Youden-optimal threshold t^* was found on the predictions of the fitted model for its own training data:

$$t^* = \arg\max_t [\mathrm{Se}_{\mathrm{train}}(t) + \mathrm{Sp}_{\mathrm{train}}(t) - 1] \quad (5)$$

where: $Se_{train}(t)$ and $Sp_{train}(t)$ – are computed from the training presences and training background for MaxEnt and FR, and from the training presences and pseudo-absences for LR.

The threshold t^* was then applied unchanged to the held-out fold, on which TSS was computed as $Se_{test}(t^*) + Sp_{test}(t^*) - 1$, so that no test-fold information entered the threshold. For each block assignment, AUC and TSS were averaged over the five folds; the mean and the range of these averages over the ten assignments are reported, together with the mean between-fold standard deviation.

The significance of the differences in AUC between models was evaluated using the DeLong test (DeLong et al., 1988) on the out-of-fold predictions pooled across the five folds of each block assignment, as a paired test that uses the covariance of the two AUCs estimated on the same observations. Model performance was also evaluated using the partial AUC (pAUC) (McClish, 1989), restricted to the high-specificity range (0.8-1.0). With $Se(f)$ the sensitivity at false positive rate $f = 1 - Sp$, the partial area over the range $0 \le f \le 0.20$ is

$$\mathrm{pAUC} = \int_0^{0.2} \mathrm{Se}(f)\,\mathrm{d}f \quad (6)$$

computed with the trapezoidal rule on the empirical ROC curve. It lies between $\mathrm{pAUC}_{min} = 0.02$, the area under the chance diagonal over the same range, and $\mathrm{pAUC}_{max} = 0.2$, the area under the ROC curve of a perfect classifier. The partial area was standardised with the McClish correction:

$$\mathrm{pAUC_s} = \frac{1}{2}\left[1 + \frac{\mathrm{pAUC} - \mathrm{pAUC_{min}}}{\mathrm{pAUC_{max}} - \mathrm{pAUC_{min}}}\right] \quad (7)$$

so that $\mathrm{pAUC}_s = 0.5$ corresponds to chance and $\mathrm{pAUC}_s = 1$ to perfect discrimination. Because the specificity is computed on background points spread over the study area, the range 0.8-1.0 corresponds approximately to the 20% of the area with the highest scores. The pAUC approach is needed because the full AUC evaluates the ROC curve as a whole, which risks masking differences in performance within particular regions of the curve. Through pAUC, the assessment is focused exclusively on the portion of the curve with a very low false alarm rate. Over this range the model behaves very strictly and predicts only the most convincing locations. The significance of the differences in pAUC between models was then tested using a bootstrap procedure. The bootstrap was chosen for the pAUC test because the variance formula of the DeLong test is defined only for the full AUC, not for a truncated (partial) area under the curve. For models A and B, evaluated on the n_1 landslide and n_0 background observations pooled across the five folds of one block assignment, the paired, stratified bootstrap proceeds as follows (roc.test in the pROC package): (i) draw n_1 landslide and n_0 background observations with replacement, each observation keeping its predictions from both models, so that the correlation between the two models is preserved; (ii) compute the standardised pAUC of A and B on resample b and their difference $\Delta_b = \mathrm{pAUC}_{s,A}^{(b)} - \mathrm{pAUC}_{s,B}^{(b)}$; (iii) repeat for $b = 1, \ldots, B$, with $B = 2000$; (iv) compute the test statistic and its two-sided p-value:

$$D = \frac{\mathrm{pAUC_{s,A}} - \mathrm{pAUC_{s,B}}}{s_\Delta}, \quad p = 2[1 - \Phi(|D|)] \quad (8)$$

where: $\mathrm{pAUC}_{s,A}$ and $\mathrm{pAUC}_{s,B}$ – are computed on the original (not resampled) observations, s_Δ – is the standard deviation of the

B bootstrap differences Δ_b, and Φ – is the standard normal distribution function.

The random seed was set before each test (5000 + 10r + j for comparison j), so the p-values are reproducible (Robin et al., 2011). The DeLong test and the pAUC are computed once per block assignment on the pooled out-of-fold predictions, so they carry no between-fold standard deviation; for each comparison the median and range of the differences and p-values over the ten assignments, and the number of assignments with $p < 0.05$, are reported. Both tests treat points as independent; spatial autocorrelation among the points therefore makes their p-values optimistic, and they are interpreted as indicative rather than exact.

The relative importance of the conditioning factors was assessed with the same measure for MaxEnt and LR, permutation importance: the decrease in AUC when the values of one factor are randomly permuted, averaged over ten permutations and expressed as a percentage of the summed decreases. Both models were evaluated on the same data, the 445 presences and 10,000 background points. FR was not included in this assessment because of its bivariate nature. FR values are computed per class within each factor and summed without a per-factor weight, so no standardised per-factor importance measure comparable to the other two models is available. The agreement in factor ranking between MaxEnt and LR was measured using the Spearman rank correlation. Both measures were computed from models fitted on the full dataset rather than per validation fold. The stability of the factor ranking in LR was tested separately by repeating the pseudo-absence draw 50 times at the same ratio. In each draw the model was refitted and the factors ranked by permutation importance (three permutations per factor). The rank range reported in Table 4 represents the minimum and maximum ranks across those 50 draws, while the reported importance values come from the LR model that produced the final map.

For interpretation purposes, the final susceptibility map of each model was classified into five susceptibility levels (from lowest to highest). This classification was applied using a percentile threshold approach, so that each class covers an equal proportion of the area, namely 20% of the total area each. The class thresholds of each map are listed in Supplementary Table S3. Each class was validated by computing the frequency ratio (FR) against the distribution of landslide points. A susceptibility map is considered valid and coherent when the FR values show a monotonic increasing trend from the lowest to the highest class (Sahin, 2020).

All processing was carried out in the R software environment. Spatial data and covariates were handled with the terra and sf packages, while each model used a different implementation: MaxEnt was run with the maxnet package, logistic regression with the built-in glm function of R, and the frequency ratio was computed through purpose-written code. Cross-validation, AUC, TSS, the DeLong test, and the partial AUC were computed with the pROC package, while the multicollinearity test (VIF) used the usdm package. Spatial-block folds were built with purpose-written code. The complete R code, the seeds, and the package versions are provided as supplementary material.

RESULTS

Model performance and significance

All three models discriminated landslide locations better than chance under spatial-block validation, with AUC values between 0.65 and 0.69 that indicate moderate discrimination (Table 2). FR and MaxEnt performed almost identically on the full AUC (0.691 and 0.688, mean of the ten block assignments) and on TSS (0.269 and 0.278). MaxEnt exceeded FR in four of the ten assignments on AUC and in six on TSS, and in 24 of the 50 test folds on AUC. LR recorded the lowest values (AUC 0.653; TSS 0.222) and was below FR in 41 and below MaxEnt in 43 of the 50 test folds. Under random folds all three models scored higher (MaxEnt 0.726, FR 0.713, LR 0.679), and MaxEnt ranked first on the full AUC, an order that spatial blocks did not preserve.

Spatial-block CV: mean of the ten block assignments of the fold-averaged AUC and TSS, with the range across assignments in parentheses; the mean between-fold SD was 0.050 (FR), 0.052 (MaxEnt), and 0.052 (LR) for AUC and 0.087, 0.088, and 0.091 for TSS. Random CV: mean (SD) of five folds. TSS is computed at the threshold fixed on the training data of each fold. [a] pAUC (specificity 0.8-1.0) is McClish-standardised and computed once per assignment on the out-of-fold predictions pooled across the five folds, so it has no

Table 2. Summary of the performance comparison of the three models under spatial-block cross-validation (primary validation; ten block assignments) and random five-fold cross-validation (comparison)

Model	Spatial-block CV (primary)			Random CV (comparison)		
	AUC	TSS	pAUC [a]	AUC	TSS	pAUC [a]
MaxEnt	0.688 (0.671-0.703)	0.278 (0.244-0.310)	0.610 (0.594-0.621)	0.726 (0.028)	0.329 (0.058)	0.633
FR	0.691 (0.682-0.701)	0.269 (0.256-0.295)	0.578 (0.570-0.584)	0.713 (0.007)	0.312 (0.023)	0.594
LR	0.653 (0.640-0.662)	0.222 (0.196-0.241)	0.555 (0.548-0.561)	0.679 (0.023)	0.248 (0.043)	0.566

between-fold SD. The AUC of the pooled predictions, used in the DeLong tests, is 0.697 (FR), 0.692 (MaxEnt), and 0.654 (LR) under spatial blocks and 0.713, 0.725, and 0.678 under random folds.

Variation across folds was larger under spatial blocks than under random folds for all three models (mean between-fold SD of AUC 0.050 to 0.052, against 0.007 to 0.028), because each spatial test fold covers a different part of the regency (Figure 3). Under spatial blocks the three models varied to a similar degree, whereas under random folds FR was clearly the most stable (SD 0.007). Across block assignments, the fold-averaged AUC of MaxEnt ranged more widely (0.671 to 0.703) than that of FR (0.682 to 0.701).

The differences in AUC were tested with the DeLong test on the out-of-fold predictions pooled within each block assignment (Table 3; ROC curves in Figure 4). The full AUC of MaxEnt and FR did not differ in nine of the ten assignments (median p = 0.52); the one significant difference (p = 0.011) favoured FR. LR was significantly below FR in all ten assignments and below MaxEnt in nine. The two presence-only models separated only over the high-specificity range: MaxEnt had a higher pAUC than FR in all ten assignments (median difference 0.030, range 0.023 to 0.052; bootstrap p = 6.6×10^{-9} to 4.9×10^{-3}), and LR had the lowest pAUC, significantly below MaxEnt in all ten assignments and below FR in nine. The high-specificity range is the portion of the ROC curve that determines the assignment of the highest susceptibility class on the final map, so the advantage of MaxEnt lies in the strictest portion of the curve.

Differences are the first model minus the second, given as the median and range over the ten block assignments; each test uses the out-of-fold predictions pooled across the five folds of one assignment. Full AUC: DeLong test; pAUC: paired, stratified bootstrap (2,000 replicates). The single

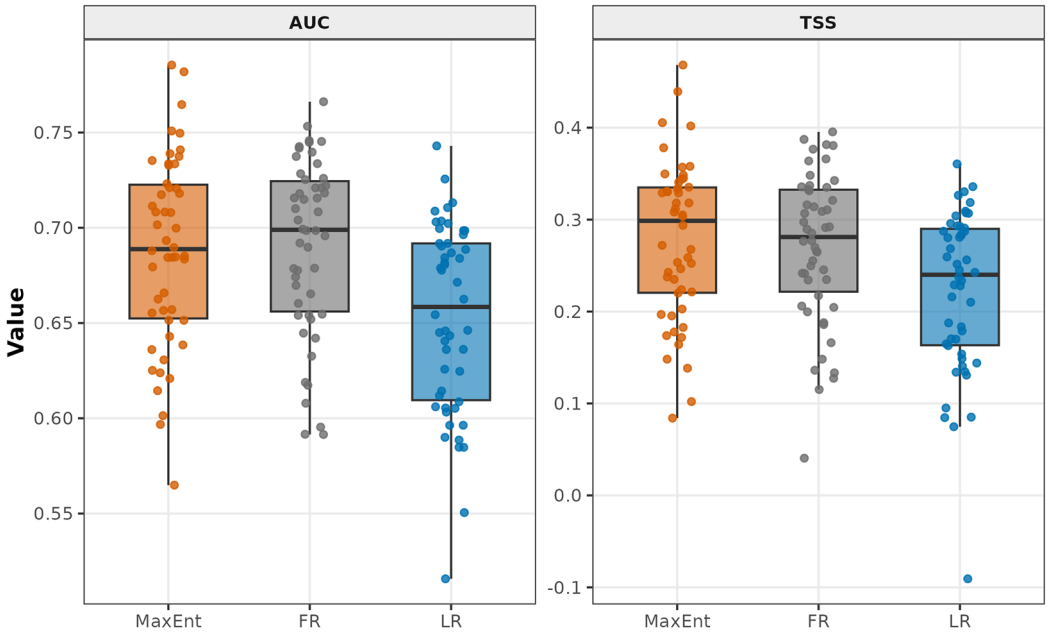

Figure 3. AUC and TSS of the 50 spatial-block test folds (ten block assignments × five folds)

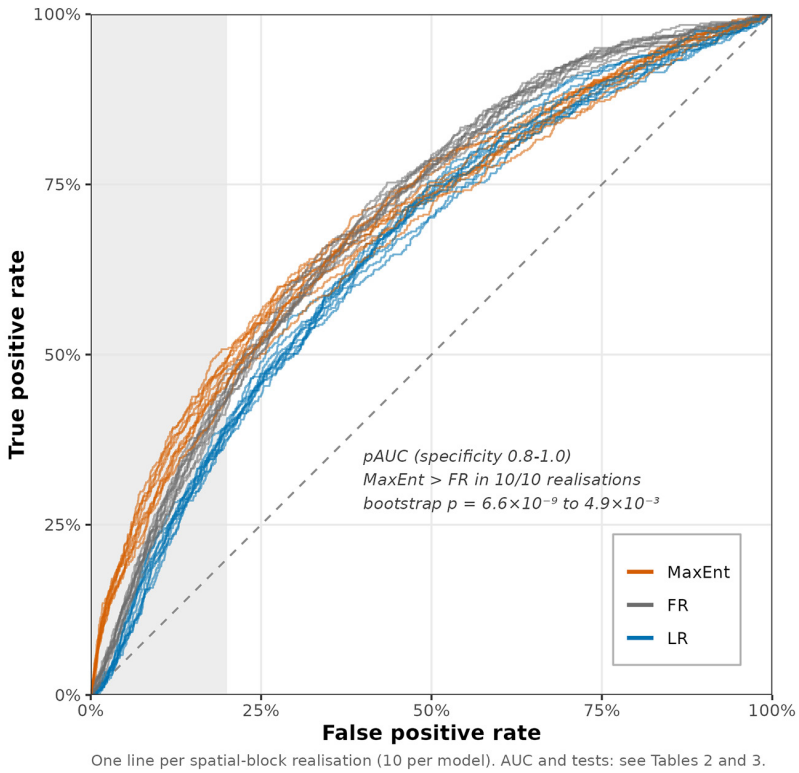

Figure 4. ROC curves of the out-of-fold predictions of the three models under spatial-block cross-validation (one curve per block assignment). The shaded band marks the specificity range of the pAUC

Table 3. Pairwise tests between models under spatial-block cross-validation (ten block assignments)

Comparison	Metric (test)	Difference, median (range)	p, median (range)	p < 0.05 (of 10)
MaxEnt − FR	Full AUC (DeLong)	−0.004 (−0.027 to +0.009)	0.52 (0.01 to 0.95)	1/10
	pAUC (bootstrap)	+0.030 (+0.023 to +0.052)	2.4×10^{-4} (6.6×10^{-9} to 0.005)	10/10
MaxEnt − LR	Full AUC (DeLong)	+0.037 (+0.012 to +0.055)	2.3×10^{-4} (9.1×10^{-8} to 0.30)	9/10
	pAUC (bootstrap)	+0.059 (+0.041 to +0.073)	2.8×10^{-11} (2.9×10^{-16} to 3.9×10^{-6})	10/10
FR − LR	Full AUC (DeLong)	+0.043 (+0.032 to +0.055)	3.9×10^{-5} (3.9×10^{-7} to 4.7×10^{-4})	10/10
	pAUC (bootstrap)	+0.024 (+0.014 to +0.032)	0.003 (3.6×10^{-5} to 0.07)	9/10

significant difference in full AUC between MaxEnt and FR favoured FR. Random-fold tests: Supplementary Table S4.

Random five-fold cross-validation gave higher values for all models and a different order between the two presence-only models (Table 2). MaxEnt ranked first on the full AUC and TSS (0.726 and 0.329, against 0.713 and 0.312 for FR), but the difference in AUC was not significant (DeLong $p = 0.146$), whereas the pAUC advantage of MaxEnt (0.633 against 0.594; $p = 1.8\times10^{-6}$) and the lower performance of LR ($p \leq 1.9\times10^{-4}$ against both) matched the spatial-block results (Supplementary Table S4). The drop in AUC from random to spatial folds (0.038 for MaxEnt, 0.022 for FR, and 0.026 for LR) shows how much random folds overstate performance when training and test points are spatially close; the conclusions of this study rest on the spatial-block results.

The lower performance of LR did not depend on a particular pseudo-absence dataset or on how pseudo-absences were placed (Supplementary Table S5). Across the 1,000 random pseudo-absence datasets (100 per block assignment), LR reached AUC 0.654 ± 0.011 (range 0.618 to 0.687) and TSS 0.227 ± 0.021 (range 0.152 to 0.283). Restricting pseudo-absences to at least 500 m from the training presences changed these values little (AUC 0.657 ± 0.010; TSS 0.232 ± 0.020), as

did a 1,000 m restriction (AUC 0.661 ± 0.010; TSS 0.237 ± 0.020). None of the 3,000 datasets reached the AUC of FR in the same block assignment, and 2, 4, and 14 of the 1,000 datasets of the three schemes reached that of MaxEnt. Under random folds the same checks gave the same picture, and the pseudo-absence ratio had no significant effect (AUC 0.679, 0.679, and 0.673 for 1:1, 1:2, and 1:5; DeLong p = 0.165 to 0.959).

Thinning the inventory more strictly lowered the AUC of all three models slightly but changed neither their order nor the conclusions (Supplementary Table S6). With a minimum distance of 100 m between points (401 points), the mean AUC under spatial blocks was 0.681 for FR, 0.678 for MaxEnt, and 0.646 for LR, and MaxEnt had a significantly higher pAUC than FR in all ten block assignments (0.601 against 0.571). With 250 m (347 points), the corresponding AUC values were 0.671, 0.664, and 0.641; MaxEnt had a higher pAUC than FR in all ten assignments (0.591 against 0.568), significantly in eight, and LR remained below both other models in all ten, although with fewer points its difference from MaxEnt on the full AUC was significant in six. The full-AUC difference between FR and MaxEnt was significant in at most one assignment at either distance.

The comparison of variable importance was made between the MaxEnt and LR models (Table 4). The FR model was not included in this comparison because of its bivariate nature. Both models place slope in first rank, and in the LR model that factor consistently remains in first rank across all pseudo-absence draws. However, the LR model concentrates variable importance on slope far more sharply than MaxEnt does (71.1% against 48.9%), while MaxEnt distributes the remaining contribution across elevation, rainfall, and distance to river. The Spearman rank correlation between MaxEnt and LR has a median of 0.390 across 50 pseudo-absence draws, with an interquartile range of 0.320 to 0.490 and a full range of 0.198 to 0.830. Only 24% of the draws produce a correlation above 0.5. A distribution this wide indicates that the agreement in variable ranking beyond slope is unstable and depends heavily on the particular pseudo-absence draw used. The most consistent differences occur for elevation and rainfall, both of which systematically occupy lower ranks in LR across all pseudo-absence draws, in line with the high correlation between those variables (r = 0.741; Supplementary Table S1) and the instability of the sign of their coefficients across folds.

Susceptibility maps

The five-class landslide susceptibility maps from the three models (Figure 5) show broadly consistent spatial patterns, although with variation in the extent of each susceptibility class. Areas of high susceptibility tend to cluster on the western side of Semarang Regency, which forms part of the slopes of Mount Ungaran. That area is characterised by steep slopes and thick soils (Juhadi et al., 2022). The five-class classification scheme follows the percentile threshold approach described in the Methods section, so the extent of each class can be compared directly between models.

Table 4. Variable importance

Variable	MaxEnt importance (%)	MaxEnt rank	LR importance (%)	LR rank	LR rank range
Slope	48.9	1	71.1	1	1-1
Elevation	13.1	2	2.3	5	2-13
Rainfall	9.7	3	1.4	7	2-13
Distance to river	8.1	4	0.7	8	2-12
NDBI	7.4	5	6.7	3	3-11
HAND	3.3	6	1.9	6	4-13
TWI	3.0	7	0.1	12	2-13
Northness	2.8	8	0.7	9	3-13
NDVI	1.6	9	10.0	2	2-10
Profile curvature	0.9	10	0.4	11	3-13
TPI	0.8	11	4.2	4	2-13
Plan curvature	0.3	12	0.1	13	5-13
Eastness	0.1	13	0.6	10	8-13

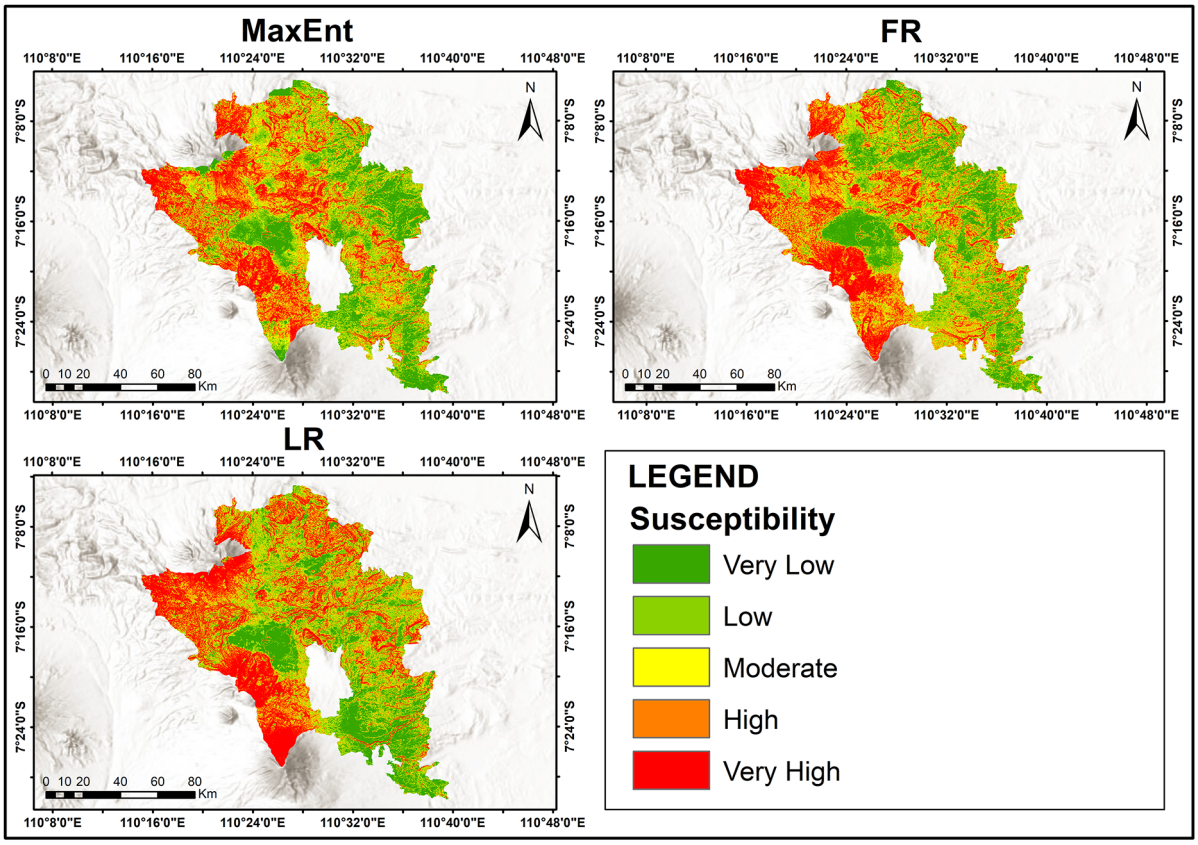

Figure 5. Five-class landslide susceptibility maps (MaxEnt, FR, LR)

The differences in class extent between models were validated further using the frequency ratio (FR) of each class (Figure 6). All three models show a monotonic increasing trend from the Very Low to the Very High class, indicating that the gradation of susceptibility levels is consistent with the landslide phenomenon in the field. In the Very Low to Moderate classes, the frequency ratio of all models is below 1.0. A pronounced jump occurs in the Very High class, where MaxEnt records the highest value of 2.88, followed by FR (2.62) and LR (2.28).

The performance of the three models in isolating landslide events within the Very High class shows a different pattern (Table 5). MaxEnt captures the highest proportion of events, namely 57.5% of all landslide events, followed by FR at 52.4% and LR at 45.6%. The difference between MaxEnt and FR is relatively small (5.2 percentage points). Because the Very High class covers the 20% of the area with the highest scores, it corresponds approximately to the specificity range 0.8-1.0 used for the pAUC, so this result and the pAUC advantage of MaxEnt describe the same behaviour measured in two ways rather than two independent lines of evidence. In contrast, the difference between MaxEnt and LR is far larger, consistent with the significant difference in AUC between those two models. This pattern indicates that MaxEnt is relatively more effective at concentrating landslide events within the highest-susceptibility zone than LR, while its advantage over FR is more limited.

DISCUSSION

Comparison

The comparison between models was made in order to establish the performance of landslide susceptibility models in Semarang Regency. Differing performance can reveal differences in the applicability of a model given the data available. The landslide event data used in this study come from reports of the Regional Disaster Management Agency (BPBD) of Semarang Regency, which are based on community reporting of events that genuinely occurred. This characteristic means the data are available only in presence form, without field-verified non-occurrence points. This condition forms an important basis

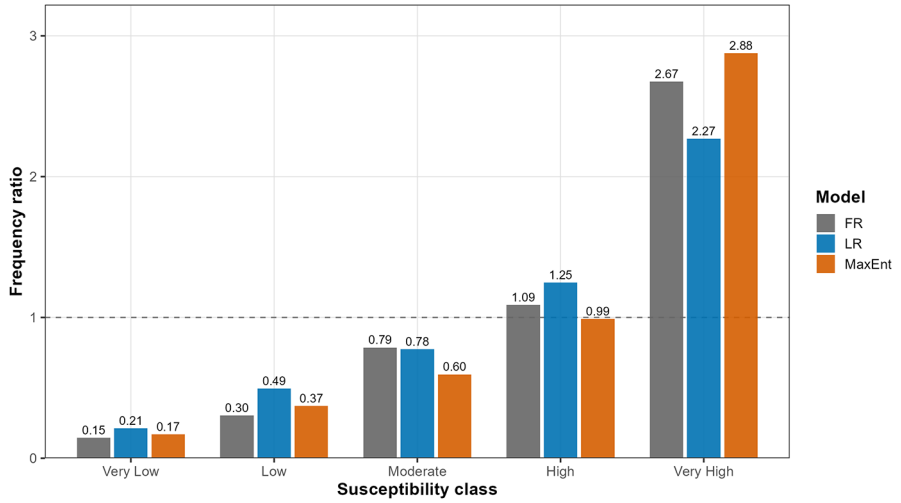

Figure 6. Frequency ratio per susceptibility class for the three models

Table 5. Performance in the very high class (the most susceptible 20% of the area)

Model	FR of very high class	% of events in very high
MaxEnt	2.88	57.5%
FR	2.62	52.4%
LR	2.28	45.6%

for the comparison of the three models tested, namely frequency ratio (FR) and MaxEnt, which can work directly from presence data, and logistic regression (LR), which requires the additional construction of pseudo-absence data.

The comparison among the three models shows that, under spatial-block validation, MaxEnt and FR perform equivalently on the full AUC and TSS, while LR performs significantly lower. MaxEnt records the highest mean values under random folds, but its margin over FR is not significant there and is not preserved under spatial blocks, where FR has the higher full AUC in six of the ten block assignments. Even so, the advantage of MaxEnt appears over the high-specificity range through a significantly higher partial AUC in all ten block assignments. The partial AUC is used to examine the range relevant to the assignment of the highest susceptibility class on the final map.

The magnitude of the advantage of MaxEnt over FR appears to vary between studies, whether measured through AUC or through other validation metrics. Hong et al., (2016) reported nearly identical AUC values for FR (0.779) and MaxEnt (0.790) in landslide susceptibility mapping in Ganzhou, China, with a difference of only 0.011. Mokhtari & Abedian, (2019) found a larger difference in a landslide study in the Taleghan basin, Iran, with the AUC of MaxEnt (0.812) exceeding that of FR (0.732) by about 0.08 points, although statistical significance for this particular pair was not reported. A far larger difference was found by Qasimi et al., (2024) in flood susceptibility mapping for the Kokcha River, Afghanistan, with MaxEnt reaching a success rate of 97% compared with 86% for FR. This variation indicates that the size of the advantage of MaxEnt over FR probably depends on the characteristics of the data and the type of hazard studied, rather than being a fixed property of the method itself. Consistent with that pattern, the present study finds that the advantage of MaxEnt over FR only becomes apparent over the high-specificity range through the partial AUC, and not in the full AUC.

The comparison between MaxEnt and LR shows a different pattern from the comparison with FR. Although LR is a multivariate model that explicitly accounts for correlation among conditioning factors, a characteristic that is theoretically superior to bivariate approaches such as FR (Yalcin et al., 2011), its performance is lower than that of MaxEnt on all three metrics, significantly so for the AUC in nine of the ten block assignments (median difference 0.037) and for the pAUC in all ten. This pattern is consistent with the finding that the ranking of variable importance in LR is unstable across different pseudo-absence draws. The dependence of LR on non-landslide points drawn at random from the study area, which cannot be guaranteed to be free of landslide susceptibility, is a plausible contributor,

but it does not explain the whole gap: keeping pseudo-absences at least 500 m or 1,000 m from known landslides raised the AUC of LR by less than 0.01. The remaining gap may reflect the smaller amount of background information available to LR, which is fitted to the training presences of a fold (294 to 418) and an equal number of pseudo-absences, whereas MaxEnt and FR use all of the about 8,000 training background points, and the linear additive form of LR, which cannot represent the non-linear responses that MaxEnt fits through quadratic and hinge features. This condition indicates that presence-based modelling such as MaxEnt is more suitable under data conditions where only event reports are available, without field-verified non-occurrence points.

The magnitude of the advantage of MaxEnt over LR also varies between studies. Felicisimo et al., (2013) found a narrow difference between MaxEnt (AUC 0.78) and multiple logistic regression (AUC 0.76) in landslide susceptibility mapping at several sites in Spain. The opposite pattern was found by Boussouf et al., (2023), who reported that geographically weighted logistic regression (GWLR), a variant of LR that accounts for spatial variation in its coefficients, in fact outperformed MaxEnt on AUC (0.972 against 0.927) in the Rio Aguas catchment, Spain, although MaxEnt still performed better on the degree-of-fit validation of the rupture zones. This variation again underlines that the advantage of MaxEnt over presence-absence models is not a fixed property, but depends on how the presence-absence model is constructed and validated.

The comparison between the FR and LR models shows a statistically significant difference. FR is a model that relies on bivariate statistical comparison and does not assume correlation among the conditioning factors, yet it still yields better results than LR. The DeLong test indicates a median AUC difference of 0.043, significant in all ten block assignments. This condition is thought to be related to the requirement of LR for non-landslide (pseudo-absence) data, which cannot be guaranteed to fully represent areas that are genuinely free of landslide susceptibility, although excluding pseudo-absences near known landslides did not close the gap.

Several studies show considerable variation in the performance of FR and LR. Yalcin et al., (2011) reported that FR captured 60.98% of active landslide zones in the high and very high susceptibility classes, far exceeding LR, which captured only 42.58%. (Hailu et al., 2024) likewise found FR to be superior, with a validation accuracy of 81.45% against 75% for LR. Conversely, (Hamal et al., 2026) reported LR to be higher than FR (AUC 0.816 against 0.767), although without any formal significance testing. This variation in direction indicates that the superiority of either model is not universal, but depends on the characteristics of the study area and the scheme used to determine non-landslide data. In the present study, FR was consistently superior to LR, in line with the findings of Yalcin et al., (2011) and Hailu et al., (2024). From an application standpoint, FR is adequate for rapid and transparent mapping because its computation is simple and its performance is equivalent on the full AUC and TSS, whereas MaxEnt is the more appropriate choice when the priority is to sharpen the delineation of the most susceptible zone.

Main parameters

The main parameter contributing the largest value in both the MaxEnt and the LR model is slope. Slope is the strongest predictor in both models that produce variable importance scores. Slope is also the parameter most frequently used and tends to be consistently high-ranked (Chicas et al., 2024; Kerekes et al., 2026). In MaxEnt, slope contributes 48.9% of the total permutation importance, followed by elevation (13.1%), rainfall (9.7%), distance to river (8.1%), and NDBI (7.4%), so that the five leading factors account for 87.2% of the total importance of the model. In LR, the dominance of slope is even more pronounced, with a contribution of 71.1%, and it is the only variable whose rank is entirely stable in that model, always occupying first position across all 50 pseudo-absence draws. This finding is in line with Chicas et al., (2024) and Kerekes et al., (2026), Importance values measure predictive contribution, not causal control. Slope is correlated with elevation (r = 0.391), NDVI (0.391), and TWI (-0.418), so part of its importance may be shared with these factors, and the values do not establish slope as an independent physical cause. The FR model was not included in this comparison because the method produces ratio values per class within each variable rather than cross-variable importance scores that can be ranked on equal terms.

This dominance of slope is consistent with the character of the shallow landsliding commonly encountered in the regolith of volcanic slopes in

tropical regions, where slope governs the balance between the gravitational force driving material downslope and the shear strength of the residual soil resisting it. This pattern has been reported in several other tropical volcanic settings. Noviyanto et al., (2020), on Mount Sumbing in Central Java, documented slope stability decreasing as steepness increases, with shallow slip surfaces forming in soil layers of high clay content in the crown zone and in former landslide scars. Nakileza & Nedala, (2020), on the volcanic slopes of Mount Elgon, Uganda, likewise found landslides concentrated in the 15-20 degree slope class. These field studies provide the physical basis for treating slope as a control on shallow landsliding in tropical volcanic settings; the importance values of the present study are consistent with that interpretation but do not test it.

Limitation

This study has several limitations, most of which stem from the characteristics of the available data. First, the landslide inventory is dominated by community reports collected through BPBD, and is therefore vulnerable to detection bias, in that events in accessible and densely populated areas have a greater chance of being recorded than events in remote areas. The additional landslides recorded during the field survey were found on or near roads and share this bias. The quality of the inventory thus sets an upper bound on the quality of the resulting map. Second, restricting the predictors to continuous data meant that lithology and land cover were not included in the modelling. The influence of land cover is still partly represented through NDVI and NDBI as proxies, whereas lithology has no suitable proxy among the variables used, even though lithology is one of the conditioning factors most frequently used in landslide susceptibility mapping (Chicas et al., 2024). Third, the DeLong and bootstrap tests assume independent observations. With spatially autocorrelated points their p-values are optimistic, which is why spatial-block folds were used as the primary validation; blocks reduce this dependence but do not remove it, and the block folds were unbalanced, holding 27 to 151 presence points.

Factors relating to the data used and to the models as run also carry limitations. The rainfall data come from CHIRPS, with a native spatial resolution of 0.05 degrees or about 5.5 km, far coarser than the 30 m analysis resolution, so rainfall values in neighbouring cells are practically uniform even after resampling. This variable is therefore able to distinguish rainfall conditions only at the regional scale and does not represent rainfall variation at the slope scale, so its contribution within the model must be interpreted with caution. MaxEnt was run without interaction terms between variables, so the possible presence of interaction effects among conditioning factors was not explored. Further research may strengthen these results through the inclusion of lithology as a conditioning factor, independent field validation of the resulting map, systematic inventory work to correct detection bias, and development from susceptibility assessment towards hazard assessment by integrating triggering rainfall data.

CONCLUSIONS

This study compared the performance of MaxEnt, frequency ratio, and logistic regression in mapping landslide susceptibility in Semarang Regency using an identical set of conditioning factors and an identical validation scheme. Under spatial-block validation, MaxEnt and FR performed equivalently on the full AUC and TSS, both with moderate discrimination. The advantage of MaxEnt becomes apparent only over the high-specificity range and within the highest susceptibility class, the part that most determines the assignment of priority zones on the final map, and it appeared in all ten spatial-block assignments. Logistic regression occupies the lowest position on all metrics, with differences that are significant relative to both other models and that held across pseudo-absence datasets, ratios, and distance restrictions, and when the inventory was thinned to 100 m or 250 m.

This pattern indicates that a landslide inventory based on community reporting, which provides presence data only, can be modelled adequately without constructing pseudo-absence data. The requirement of logistic regression for artificial non-landslide points adds a source of uncertainty that presence-background models avoid, although restricting pseudo-absences away from known landslides did not close its performance gap. The practical consequence is that FR is adequate for rapid and transparent mapping by regional agencies because its computation is simple and its performance is equivalent on the full AUC and TSS, whereas MaxEnt is the more appropriate choice when the priority is to sharpen the delineation of the most susceptible zone.

Acknowledgements

This research was funded by Universitas Negeri Semarang (UNNES) through the Lecturer Research Grant Scheme (Faculty) Dana DPA FISIP, 2026, under contract/decree number 13.27.4/UN37/PPK.03/2026. The authors thank the Regional Disaster Management Agency (Badan Penanggulangan Bencana Daerah, BPBD) of Semarang Regency for providing the landslide records used in this study.

REFERENCES

1. Ado, M., Amitab, K., Maji, A. K., Jasińska, E., Gono, R., Leonowicz, Z., & Jasiński, M. (2022). Landslide Susceptibility Mapping Using Machine Learning: A Literature Survey. *Remote Sensing*, *14*(13). https://doi.org/10.3390/rs14133029
2. Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, *43*(6), 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x
3. Amarasinghe, M. P., Kulathilaka, S. A. S., Robert, D. J., Zhou, A., & Jayathissa, H. A. G. (2024). Risk assessment and management of rainfall-induced landslides in tropical regions: a review. In *Natural Hazards* (Vol. 120, Number 3). Springer Netherlands. https://doi.org/10.1007/s11069-023-06277-3
4. Azarafza, M., Azarafza, M., Akgün, H., Atkinson, P. M., & Derakhshani, R. (2021). Deep learning-based landslide susceptibility mapping. *Scientific Reports*, *11*(1), 1–16. https://doi.org/10.1038/s41598-021-03585-1
5. Badan Nasional Penanggulangan Bencana. (2025). *Data Informasi Bencana Indonesia (DIBI)*. Badan Nasional Penanggulangan Bencana.
6. Barman, J., Ali, S. S., Biswas, B., & Das, J. (2023). Application of index of entropy and Geospatial techniques for landslide prediction in Lunglei district, Mizoram, India. *Natural Hazards Research*, *3*(3), 508–521. https://doi.org/10.1016/j.nhres.2023.06.006
7. Boussouf, S., Fernández, T., & Hart, A. B. (2023). Landslide susceptibility mapping using maximum entropy (MaxEnt) and geographically weighted logistic regression (GWLR) models in the Río Aguas catchment (Almería, SE Spain). *Natural Hazards*, *117*(1), 207–235. https://doi.org/10.1007/s11069-023-05857-7
8. Budimir, M. E. A., Atkinson, P. M., & Lewis, H. G. (2015). A systematic review of landslide probability mapping using logistic regression. *Landslides*, *12*(3), 419–436. https://doi.org/10.1007/s10346-014-0550-5
9. Çağan, Ç., Kurnaz, T. F., & Erden, C. (2026). Frequency ratio and machine learning-based classification of landslide susceptibility in the west of Mersin province, Türkiye: An integrated approach using SMOTE. *Journal of African Earth Sciences*, *238*(February). https://doi.org/10.1016/j.jafrearsci.2026.106085
10. Cellek, S. (2020). Morphological parameters causing landslides: A case study of elevation. *Bulletin of the Mineral Research and Exploration*, *162*, 197–224. https://doi.org/https://doi.org/10.19111/bulletinofmre.649758
11. Cemiloglu, A., Zhu, L., Mohammednour, A. B., Azarafza, M., & Nanehkaran, Y. A. (2023). Landslide Susceptibility Assessment for Maragheh County, Iran, Using the Logistic Regression Algorithm. *Land*, *12*(7), 1–20. https://doi.org/10.3390/land12071397
12. Chen, W., Pourghasemi, H. R., Kornejady, A., & Zhang, N. (2017). Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. *Geoderma*, *305*(May), 314–327. https://doi.org/10.1016/j.geoderma.2017.06.020
13. 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*(14), 12573–12593. https://doi.org/10.1007/s11069-024-06697-9
14. Conoscenti, C., Di Maggio, C., & Rotigliano, E. (2008). GIS analysis to assess landslide susceptibility in a fluvial basin of NW Sicily (Italy). *Geomorphology*, *94*(3–4), 325–339. https://doi.org/10.1016/j.geomorph.2006.10.039
15. DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, *44*(3), 837–845. https://doi.org/10.2307/2531595
16. Dou, H., He, J., Huang, S., Jian, W., & Guo, C. (2023). Influences of non-landslide sample selection strategies on landslide susceptibility mapping by machine learning. *Geomatics, Natural Hazards and Risk*, *14*(1). https://doi.org/10.1080/19475705.2023.2285719
17. Felicisimo, A. M., Cuartero, A., Remondo, J., & Quiros, E. (2013). Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. *Landslides*, *10*(2), 175–189. https://doi.org/10.1007/s10346-012-0320-1
18. Gentilucci, M., Pelagagge, N., Rossi, A., Domenico, A., & Pambianchi, G. (2023). Landslide Susceptibility Using Climatic–Environmental Factors Using

the Weight-of-Evidence Method—A Study Area in Central Italy. *Applied Sciences (Switzerland)*, *13*(15). https://doi.org/10.3390/app13158617

19. Gorokhovich, Y., & Vustianiuk, A. (2021). Implications of slope aspect for landslide risk assessment: A case study of Hurricane Maria in Puerto Rico in 2017. *Geomorphology*, *391*, 107874. https://doi.org/10.1016/j.geomorph.2021.107874
20. Habumugisha, J. M., Chen, N., Rahman, M., Islam, M. M., Ahmad, H., Elbeltagi, A., Sharma, G., Liza, S. N., & Dewan, A. (2022). Landslide Susceptibility Mapping with Deep Learning Algorithms. *Sustainability (Switzerland)*, *14*(3), 1–22. https://doi.org/10.3390/su14031734
21. Hadmoko, D. S., Lavigne, F., Sartohadi, J., Gomez, C., & Daryono, D. (2017). Spatio-Temporal Distribution of Landslides in Java and the Triggering Factors. *Forum Geografi*, *31*(1), 1–15. https://doi.org/10.23917/forgeo.v31i1.3790
22. Hailu, S., Deribew, K. T., Teferi, E., Moisa, M. B., Roba, Z. R., Dagne, S. S., & Woldetsadik, M. (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*(1). https://doi.org/10.1186/s40068-024-00382-3
23. Hamal, S., Dhakal, S., Shahi, B., Budha, P. B., Alkhuraiji, W. S., & Zhran, M. (2026). Assessing landslide susceptibility using frequency ratio and logistic regression models in the Khudi Watershed. *Geoscience Letters*, *13*(1). https://doi.org/10.1186/s40562-026-00467-0
24. Hanafi, F., & Pamungkas, D. (2021). Aplikasi Model Rusle untuk Estimasi Kehilangan Tanah Bagian Hulu di Sub Das Garang, Jawa Tengah. *Jurnal Geografi : Media Informasi Pengembangan Dan Profesi Kegeografian*, *18*(1), 30–36. https://doi.org/10.15294/jg.v18i1.28079
25. He, Q., Wang, M., & Liu, K. (2021). Rapidly assessing earthquake-induced landslide susceptibility on a global scale using random forest. *Geomorphology*, *391*, 107889. https://doi.org/10.1016/j.geomorph.2021.107889
26. Hong, H., Naghibi, S. A., Pourghasemi, H. R., & Pradhan, B. (2016). GIS-based landslide spatial modeling in Ganzhou City, China. *Arabian Journal of Geosciences*, *9*(2), 1–26. https://doi.org/10.1007/s12517-015-2094-y
27. Huang, Y., Tyata, U., Liang, D., Gao, Y., & Yang, Q. (2025). Considering tectonic uplift in landslide susceptibility assessment using MaxEnt model: a case study of Trishuli River watershed. *Environmental Earth Sciences*, *84*(7), 1–16. https://doi.org/10.1007/s12665-025-12164-w
28. Ilinca, V., Șandric, I., & Gheuca, I. (2025). Landslide susceptibility analysis using machine learning: Insights from the bend Subcarpathians-Carpathians, Romania. *Geomorphology*, *486*(May). https://doi.org/10.1016/j.geomorph.2025.109872
29. Juhadi, Pratiwi, E. S., Trihatmoko, E., Findayani, A., Sartohadi, J., & Hamid, N. (2022). The Study of Anthropogenic Disturbance on Geopedogenetic Process at the Lower Slope of Mt. Ungaran, Indonesia. *Revista Brasileira de Geomorfologia*, *23*(3), 1735–1752. https://doi.org/10.20502/rbg.v23i3.2142
30. Kamran, K. V., Feizizadeh, B., Khorrami, B., & Ebadi, Y. (2021). A comparative approach of support vector machine kernel functions for GIS-based landslide susceptibility mapping. *Applied Geomatics*, *13*(4), 837–851. https://doi.org/10.1007/s12518-021-00393-0
31. Kerekes, A. H., Baciu, C., & Poszet, S. L. (2026). Integrating MaxEnt and CA–Markov–MLP for Multi-Temporal Landslide Susceptibility Modelling. *Sustainability (Switzerland)*, *18*(12), 1–29. https://doi.org/10.3390/su18126232
32. Koesoemadinata, R. (2020). *An introduction into the geology of Indonesia: Part 2 central Indonesia and part 3 eastern Indonesia*. Ikatan Alumni Geologi, Institut Teknologi Bandung.
33. Li, M., Wang, H., Chen, J., & Zheng, K. (2024). Assessing landslide susceptibility based on the random forest model and multi-source heterogeneous data. *Ecological Indicators*, *158*(January), 111600. https://doi.org/10.1016/j.ecolind.2024.111600
34. Liu, S., Du, J., Yin, K., Zhou, C., Huang, C., Jiang, J., & Yu, J. (2024). Regional early warning model for rainfall induced landslide based on slope unit in Chongqing, China. *Engineering Geology*, *333*(July 2023), 107464. https://doi.org/10.1016/j.enggeo.2024.107464
35. Liu, S., Wang, L., Zhang, W., He, Y., & Pijush, S. (2023). A comprehensive review of machine learning-based methods in landslide susceptibility mapping. *Geological Journal*, *58*(6), 2283–2301. https://doi.org/https://doi.org/10.1002/gj.4666
36. McClish, D. K. (1989). Analyzing a portion of the ROC curve. *Medical Decision Making*, *9*(3), 190–195. https://doi.org/10.1177/0272989X8900900307
37. McColl, S. T. (2022). Landslide causes and triggers. In *Landslide hazards, risks, and disasters* (pp. 13–41). Elsevier.
38. Md.Sharafat, C., Md.Naimur, R., Md.Sujon, S., Md.Abu, S., Mahmud, K. H., & Hafsa, B. (2024). GIS-based landslide susceptibility mapping using logistic regression, random forest and decision and regression tree models in Chattogram District, Bangladesh. *Heliyon*, *10*(1), e23424. https://doi.org/10.1016/j.heliyon.2023.e23424
39. Mokhtari, M., & Abedian, S. (2019). Spatial prediction of landslide susceptibility in Taleghan

basin, Iran. *Stochastic Environmental Research and Risk Assessment*, *33*(7), 1297–1325. https://doi.org/10.1007/s00477-019-01696-w

40. Nakileza, B. R., & Nedala, S. (2020). Topographic influence on landslides characteristics and implication for risk management in upper Manafwa catchment, Mt Elgon Uganda. *Geoenvironmental Disasters*, *7*(1). https://doi.org/10.1186/s40677-020-00160-0
41. Ngadisih, Samodra, G., Bhandary, N. P., & Yatabe, R. (2017). Landslide inventory: challenge for landslide hazard assessment in Indonesia. In *GIS landslide* (pp. 135–159). Springer.
42. Nikoobakht, S., Azarafza, M., Akgün, H., & Derakhshani, R. (2022). Landslide Susceptibility Assessment by Using Convolutional Neural Network. *Applied Sciences (Switzerland)*, *12*(12). https://doi.org/10.3390/app12125992
43. Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A., Waterloo, M., & Saleska, S. (2011). Height Above the Nearest Drainage - a hydrologically relevant new terrain model. *Journal of Hydrology*, *404*(1–2), 13–29. https://doi.org/10.1016/j.jhydrol.2011.03.051
44. Noviyanto, A., Purwanto, B. H., & Sartohadi, J. (2020). The distribution of soil morphological characteristics for landslide-impacted Sumbing Volcano, Central Java - Indonesia. *Geoenvironmental Disasters*, *7*(1), 25. https://doi.org/10.1186/s40677-020-00158-8
45. Nseka, D., Kakembo, V., Bamutaze, Y., & Mugagga, F. (2019). Analysis of topographic parameters underpinning landslide occurrence in Kigezi highlands of southwestern Uganda. *Natural Hazards*, *99*(2), 973–989. https://doi.org/10.1007/s11069-019-03787-x
46. Ohlmacher, G. C. (2007). Plan curvature and landslide probability in regions dominated by earth flows and earth slides. *Engineering Geology*, *91*(2–4), 117–134. https://doi.org/10.1016/j.enggeo.2007.01.005
47. Orefice, S., & Innocenti, C. (2025). Regional assessment of coastal landslide susceptibility in Liguria, Northern Italy, using MaxEnt. *Natural Hazards*, *121*(3), 2613–2639. https://doi.org/10.1007/s11069-024-06833-5
48. Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). *Maximum entropy modeling of species geographic distributions*. *190*, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026
49. Phillips, S. J., Dudík, M., & Schapire, R. E. (2017). *Maxent software for modeling species niches and distributions*.
50. Qasimi, A. B., Isazade, V., & Berndtsson, R. (2024). Flood susceptibility prediction using MaxEnt and frequency ratio modeling for Kokcha River in Afghanistan. *Natural Hazards*, *120*(2), 1367–1394. https://doi.org/10.1007/s11069-023-06232-2
51. Qiu, H., Xu, Y., Tang, B., Su, L., Li, Y., Yang, D., & Ullah, M. (2024). Interpretable Landslide Susceptibility Evaluation Based on Model Optimization. *Land*, *13*(5), 639. https://doi.org/10.3390/land13050639
52. Rajabian, A. (2023). Effect of initial failure geometry on the progress of a retrogressive seepage-induced landslide. *International Journal of Geo-Engineering*, *14*(1). https://doi.org/10.1186/s40703-023-00189-8
53. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*(1), 77. https://doi.org/10.1186/1471-2105-12-77
54. Rózycka, M., Migoń, P., & Michniewicz, A. (2017). Topographic Wetness Index and Terrain Ruggedness Index in geomorphic characterisation of landslide terrains, on examples from the Sudetes, SW Poland. *Zeitschrift Fur Geomorphologie*, *61*(2014), 61–80. https://doi.org/10.1127/zfg_suppl/2016/0328
55. Sahin, E. K. (2020). Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using XGBoost, gradient boosting machine, and random forest. *SN Applied Sciences*, *2*(7), 1–17. https://doi.org/10.1007/s42452-020-3060-1
56. Saleem, N., Enamul Huq, M., Twumasi, N. Y. D., Javed, A., & Sajjad, A. (2019). Parameters derived from and/or used with digital elevation models (DEMs) for landslide susceptibility mapping and landslide risk assessment: A review. *ISPRS International Journal of Geo-Information*, *8*(12). https://doi.org/10.3390/ijgi8120545
57. Samodra, G., Ramadhan, M. F., Sartohadi, J., Setiawan, M. A., Christanto, N., & Sukmawijaya, A. (2020). Characterization of displacement and internal structure of landslides from multitemporal UAV and ERT imaging. *Landslides*, *17*(10), 2455–2468. https://doi.org/10.1007/s10346-020-01428-0
58. Silalahi, F. E. S., Pamela, Arifianti, Y., & Hidayat, F. (2019). Landslide susceptibility assessment using frequency ratio model in Bogor, West Java, Indonesia. *Geoscience Letters*, *6*(1). https://doi.org/10.1186/s40562-019-0140-4
59. Soehaimi, A. (2008). Seismotektonik dan Potensi Kegempaan Wilayah Jawa. *Indonesian Journal on Geoscience*, *3*(4), 227–240.
60. Soma, A. S., Saharudin, N. C. P., & Ramadhan, M. D. R. (2026). Landslide hazard modeling using the artificial neural network approach in the Biang Loe River watershed. *Ecological Engineering and Environmental Technology*, *27*(2), 156–169. https://doi.org/10.12912/27197050/217186
61. Sujatha, E. R., & Sridhar, V. (2021). Landslide Susceptibility Analysis : A Logistic Regression Model. *Hydrology*, *8*(41).

62. Sun, D., Xu, J., Wen, H., & Wang, D. (2021). Assessment of landslide susceptibility mapping based on Bayesian hyperparameter optimization: A comparison between logistic regression and random forest. *Engineering Geology*, *281*(December 2020), 105972. https://doi.org/10.1016/j.enggeo.2020.105972
63. Taloor, A. K., Abraham, A., Jaiswal, A., Katiyar, R., Gupta, A. K., Sharma, K., & Pundir, S. K. (2026). Hybrid modelling of landslide susceptibility using frequency ratio and maximum entropy models in the MPRRU region of Jammu and Kashmir, India. *Physics and Chemistry of the Earth*, *144*(P2), 104460. https://doi.org/10.1016/j.pce.2026.104460
64. Thomas, A. V., Saha, S., Danumah, J. H., Raveendran, S., Prasad, M. K., Ajin, R. S., & Kuriakose, S. L. (2021). Landslide Susceptibility Zonation of Idukki District Using GIS in the Aftermath of 2018 Kerala Floods and Landslides: a Comparison of AHP and Frequency Ratio Methods. *Journal of Geovisualization and Spatial Analysis*, *5*(2). https://doi.org/10.1007/s41651-021-00090-x
65. Turahmansyah, A. I., Ahmad, A., Soma, A. S., & Santi, M. A. (2026). Landslide susceptibility modelling using the weight of evidence and logistic regression approaches in the Balease watershed, Indonesia. *Ecological Engineering and Environmental Technology*, *27*(2), 374–389. https://doi.org/10.12912/27197050/217597
66. Ullah, I., Aslam, B., Shah, S. H. I. A., Tariq, A., Qin, S., Majeed, M., & Havenith, H. B. (2022). An Integrated Approach of Machine Learning, Remote Sensing, and GIS Data for the Landslide Susceptibility Mapping. *Land*, *11*(8), 1–20. https://doi.org/10.3390/land11081265
67. Van Bemmelen, R. W. (1949). *The Geology of Indonesia: IA*. Government Printing Office.
68. Vujović, F., Grozdanić, G., Đurović, R., Valjarević, A., & Milevski, I. (2025). Comparative assessment of GIS-based multi-criteria decision analysis (AHP) and machine learning (MaxEnt) approaches for wildfire susceptibility modeling in Montenegro. *Egyptian Journal of Remote Sensing and Space Science*, *28*(4), 724–736. https://doi.org/10.1016/j.ejrs.2025.11.003
69. Xu, Z., Che, A., & Zhou, H. (2024). Seismic landslide susceptibility assessment using principal component analysis and support vector machine. *Scientific Reports*, *14*(1), 1–16. https://doi.org/10.1038/s41598-023-48196-0
70. Yalcin, A., Reis, S., Aydinoglu, A. C., & Yomralioglu, T. (2011). A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. *Catena*, *85*(3), 274–287. https://doi.org/10.1016/j.catena.2011.01.014
71. Zhang, J., Ma, X., Zhang, J., Sun, D., Zhou, X., Mi, C., & Wen, H. (2023). Insights into geospatial heterogeneity of landslide susceptibility based on the SHAP-XGBoost model. *Journal of Environmental Management*, *332*(August 2022), 117357. https://doi.org/10.1016/j.jenvman.2023.117357
72. Zhou, C., Gan, L., Cao, Y., Wang, Y., Segoni, S., Shi, X., Motagh, M., & Singhc, R. P. (2025). Landslide susceptibility assessment of the Wanzhou district: Merging landslide susceptibility modelling (LSM) with InSAR-derived ground deformation map. *International Journal of Applied Earth Observation and Geoinformation*, *136*(April 2024), 104365. https://doi.org/10.1016/j.jag.2025.104365
73. Zul, A. Z., Massinai, M. A., Arif, S., Karurung, W. S., & Wahyu Laode, A. M. (2026). Random forest modeling for landslide susceptibility assessment in the complex terrain of Tinggimoncong, South Sulawesi, Indonesia. *Ecological Engineering and Environmental Technology*, *27*(2), 46–59. https://doi.org/10.12912/27197050/216355