



Analysis of landslide explicative factors and susceptibility mapping in an andean context: The case of Azuay province (Ecuador)

Sandra Lucia Cobos-Mora^{a,b,*}, Victor Rodriguez-Galiano^b, Aracely Lima^{c,d}

^a Centro de Investigación, Innovación y Transferencia de Tecnología (CIITT), Universidad Católica de Cuenca, Cuenca, Ecuador

^b Departamento de Geografía Física y Análisis Geográfico Regional, Universidad de Sevilla, Sevilla, Spain

^c Universidad Politécnica de Madrid, Madrid, 28031, Spain

^d Instituto de Investigación Geológico y Energético, Quito, 170518, Ecuador

ARTICLE INFO

Keywords:

Landslide susceptibility
Logistic regression
Exploratory factor analysis
Time-series analysis
Quantitative modeling

ABSTRACT

Landslides are one of the natural phenomena with more negative impacts on landscape, natural resources, and human health worldwide. Andean geomorphology, urbanization, poverty, and inequality make it more vulnerable to landslides. This research focuses on understanding explanatory landslide factors and promoting quantitative susceptibility mapping. Both tasks supply valuable knowledge for the Andean region, focusing on territorial planning and risk management support. This work addresses the following questions using the province of Azuay-Ecuador as a study area: (i) How do EFA and LR assess the significance of landslide occurrence factors? (ii) Which are the most significant landslide occurrence factors for susceptibility analysis in an Andean context? (iii) What is the landslide susceptibility map for the study area? The methodological framework uses quantitative techniques to describe landslide behavior. EFA and LR models are based on a historical inventory of 665 records. Both identified NDVI, NDWI, altitude, fault density, road density, and PC2 as the most significant factors. The latter factor represents the standard deviation, maximum value of precipitation, and rainfall in the wet season (January, February, and March). The EFA model was built from 7 latent factors, which explained 55% of the accumulated variance, with a medium item complexity of 1.5, a RMSR of 0.02, and a TLI of 0.89. This technique also identified TWI, fault distance, plane curvature, and road distance as important factors. LR's model, with AIC of 964.63, residual deviance of 924.63, AUC of 0.92, accuracy of 0.84, and Kappa of 0.68, also shows statistical significance for slope, roads density, geology, and land cover factors. This research encompasses a time-series analysis of NDVI, NDWI, and precipitation, including vegetation and weather dynamism for landslide occurrence. Finally, this methodological framework replaces traditional qualitative models based on expert knowledge, for quantitative approaches for the study area and the Andean region.

* Corresponding author. Universidad Católica de Cuenca, Centro de Investigación, Innovación y Transferencia de Tecnológica (CIITT), Cuenca, Ecuador.

E-mail addresses: scobosm@ucacue.edu.ec, zhandry_cobos@yahoo.com, sancobmor@alum.us.es (S.L. Cobos-Mora).

<https://doi.org/10.1016/j.heliyon.2023.e20170>

Received 17 June 2023; Received in revised form 12 September 2023; Accepted 13 September 2023

Available online 15 September 2023

2405-8440/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Landslides caused 4.9% of all-natural disasters and 1.3% of fatalities worldwide between 1990 and 2015 [1]. According to the World Meteorological Organization [2], landslides are responsible for 10% of disasters in South America and 12% of human deaths. This percentage increases to 23% if the entire South Pacific region is considered [3].

The Andean sub-region in South America lies on Nazca, South American, and Caribbean tectonic plates. It extends from Chile in the south to Colombia and Venezuela in the north, in western South America. This sub-region is marked by the Andes mountain range, the most relevant mountain massif in the geomorphological context. It is also located along the Pacific Ring of Fire, known for its high seismic and volcanic activity [4]. The stability of the high slope of the Andean mountains is reduced in areas with active faults and high precipitations. These areas occur at tropical and subtropical latitudes due to extreme rainfalls [5,6] (above the 90th, 95th, and 99th percentiles) and changes in weather patterns resulting from climate change [7]. Moreover, anthropic activities caused by inadequate urbanization processes, high poverty levels, inequality, and environmental degradation put the Andean sub-region at high-risk for landslides and other hazards [8,9].

The province of Azuay in Ecuador includes areas below 100 m. a.s.l. (Camilo Ponce Enriquez) to areas exceeding 2500 m. a.s.l. (Cuenca) with a wide heterogeneity of climates and vegetation. It is a representative sample of the Andean subregion. Landslides are significant for their human and material implications in this province. On March 29, 1993, approximately 72 people were killed by a rockslide in La Josefina [10]. In Guarumales, an active landslide has also endangered the Molinos Hydroelectric Plant, where approximately 250 persons live [11]. Multiple landslides have occurred on the Cuenca-Molleturo-Naranjal road. Consequently, it has been temporarily closed on several occasions. This road connects Cuenca—the central city in Azuay—with Guayaquil—the main port of Ecuador— [12]. Lastly, alluvium occurred in Sayausí on March 27, 2022, when streams clogged after a heavy rainfall of approximately 30 mm/h. Over 400 people were left homeless, and a minimum of four died [13].

Landslide susceptibility analysis is considered an essential tool for territorial zoning, as it assists land use planners and policy-makers in achieving territorial sustainability and resilience [14,15]. Predicting future landslide areas will help to design safe human settlements and monitor high-susceptibility places [16]. Landslide susceptibility modeling assumes that future landslides will be more likely to occur under similar conditions than previous events. These conditions will vary according to the scale of analysis, characteristics of the sampling site, landslide shape and location, and mechanisms contributing to the fault. The model's accuracy depends on the amount and quality of information (number and quality of factors), among other aspects, such as the reliability of the landslide inventory [17]. Redundant and irrelevant factors will reduce predictive accuracy [18]. In this context, identifying relevant factors for landslide susceptibility modeling is one of the most critical and challenging tasks due to the lack of universal guidelines for its selection [19], the wide variety of landslides, and the particular characteristics of the study area. Previous to the application of the modeling algorithm, the identification of explanatory factors is relevant. It is based on literature review, expert knowledge, availability of information, and data exploration [20]. Some examples of data exploration techniques are Logistic Regression (LR) [21–23], heuristics based on expert opinion [24, 25], Discriminant Analysis [26], Markov Chain [18], machine learning techniques [27], and Exploratory Factor Analysis (EFA). EFA is a statistical technique that aggregates interrelated variables to form fundamental factors for landslide processes. Some of the listed techniques to analyze adequate explicative factors have also been used as modeling algorithms.

LR has attractive features for landslide susceptibility modeling, such as (i) allowing to predict affected and non-affected areas; (ii) the use of independent input factors of any type (nominal, numerical, or categorical); (iii) lack of assumptions on the statistical distribution of factors (i.e., normal distribution); (iv) low computational demand. Furthermore, Nhu et al. [28] pointed out that LR outperformed other algorithms such as Naïve Bayes Tree, Artificial Neural Network, and Support Vector Machine; while Pham and Prakash [29] place LR above Fisher's Linear Discriminate Analysis.

The Andean region has deficiencies in susceptibility analyses due to a lack of detailed technical information from landslide inventories despite having risk management policies [8]. These shortcomings come from insufficient budgets for information acquisition, resulting in outdated equipment, poor maintenance, and a lack of training for technical staff [30]. Furthermore, most studies of landslide susceptibility on a regional scale in Ecuador have used qualitative (semi-quantitative or heuristic) methodologies such as Analytical Hierarchy Process [31]. Although this methodology is widely used worldwide [32–36], it relies on experts' opinions to assign weight to each landslide factor, which influences its selection, comparison, and prioritization. The experts' judgment is based on their professional experience and knowledge of the area, among other aspects [37]. The uncertainty derived from these opinions constitutes the most critical limitation of qualitative modeling. Therefore, its results are somewhat subjective. As a response, quantitative frameworks with statistical techniques are objective. According to Shano et al. [26], these methods include LR [16, 38]; deterministic [39,40]; probabilistic, and artificial intelligence algorithms [41,42].

The authors pretend to answer the following questions based on the described scientific and practical issues: (i) How do LR and EFA evaluate the significance of landslide occurrence factors? LR has been widely used for landslide susceptibility mapping, as it allows identifying significant factors related to landslides [43–45]. However, EFA, as a multicriteria technique, has not been exploited for Landslide factor analysis as in other areas [46,47]. According to the literature review, only a study was found done by Barančoková et al. [48] in the same line. (ii) Which are the most significant landslide occurrence factors for susceptibility analysis in an Andean context? A time series analysis of precipitation, Normalized Difference Vegetation Index (NDVI), and Normalized Difference Water Index (NDWI) will introduce the dynamic development characteristics of landslide susceptibility. Besides, this procedure helps to eliminate atmospheric effects such as disturbance caused by clouds and outliers or missing data from single-date images (iii) What is the landslide susceptibility map for the study area? According to the authors' knowledge, few local quantitative approaches to landslide susceptibility maps in Ecuador [49,50]. Most landslide susceptibility studies have used qualitative methodologies on a regional scale despite the subjectivity limitation discussed above. The present study contemplates the province of Azuay in Ecuador. It

will provide technical inputs and assist municipal governments and planners to better manage the territory by reducing fatal landslide consequences.

2. Data and methodology

This research evaluates landslide explanatory factors. It was done using a multivariate analysis through EFA and LR. Both methods allow for obtaining a statistical model that best describes the landslide behavior in the Andean region. The primary input for the analysis is a training set with landslide occurrence and non-occurrence locations characterized according to 16 factors. These factors are consistent with previous research in similar areas for susceptibility mapping using data-driven techniques such as those developed by Vorpahl et al. [22]. Finally, we obtained a landslide susceptibility map to identify the spatial probability of occurrence from the LR-optimized model. The summarized procedure is shown in Fig. 1.

2.1. Case study

The province of Azuay is located in the south of Ecuador’s inter-Andean region (Fig. 2 (a)). Around 50% of its territory comprises cold peaks of mountain ranges with inherited paleo-glacial forms and external water runoff covered by pyroclastic material, ash, and lapilli. The study area consists primarily of steep slopes between 25% and 50% and precipices with slopes greater than 70% in Cañar and Naranjal-Pagua subbasins. Since 2015, the study area has had 723 mining concessions, including metallic, non-metallic, combinations of both, and construction materials. Landslides represent 39,545 ha; creeping landslides, subsidence, and settlement 9798 ha; debris and mudflows 676.57 ha; while alluvial material next to rivers represents 16,855 ha, regarding mass movements [51].

The geology is complex in the study area, as shown in Fig. 2(b). The lithology comprises metasediments dating back to the late Paleozoic and was affected by a Triassic orogeny. The Pallatanga Unit is part of an ophiolitic sequence accreted against the margin of the South American continent during the middle to late Cretaceous. The suture line for this accretion in the area is marked by the Bulubulu fault, which separates a belt of metamorphic rocks southeast from the Pallatanga Unit to the northwest. The Yunguilla formation of the Maestrichtian age was partially deposited upon the Pallatanga Unit—the Macuchi Formation, which represents an island arc developed in the Early Tertiary. At the same time, the siliciclastic sequence of the Angamarca Group laid down in a forearc-marginal sea which separated the island arc from the continental margin. The Macuchi Arc deformed the Angamarca Group and Macuchi Formation in the Late Eocene. Continental volcanism of the Saraguro Group began in the ocean, deposited on metamorphic rocks and basalts of the Pallatanga Unit. Volcanic activity and periods of erosion and deformation continued during the Early Miocene. These events formed the Ocaña Formation, Chulo Unit, Filo Cajas Unit, Tomebamba Unit, Chanlud Formation, and Río Blanco Formation. During the Miocene, east-west tensional forces formed intramountainous basins to the east, where fluvial sediments of the Ayancay Group and the Turi Formation were deposited. A large-scale pyroclastic event produced the Tarqui Formation.

2.2. Landslide occurrence factors

2.2.1. Geology - lithology

Geology provided information on the diversity of outcropping rock units (lithology) covered by fine soil. Depending on its type, each rock will belong to a different petrogenetic environment and will vary in composition, engineering properties, and stability characteristics. These variations impact the landslide’s frequency and size [52]. The geology map was generalized to a scale of 1:100,000. It comes from regional and Ecuadorian geological maps (1:1’000,000) and maps of the Cordillera Occidental 2–3° S and 3–4° S (1:200,000, with an in-detail survey at 1:50,000 scale). This factor was classified into five categories according to its favorability of causing landslides (Table 1).

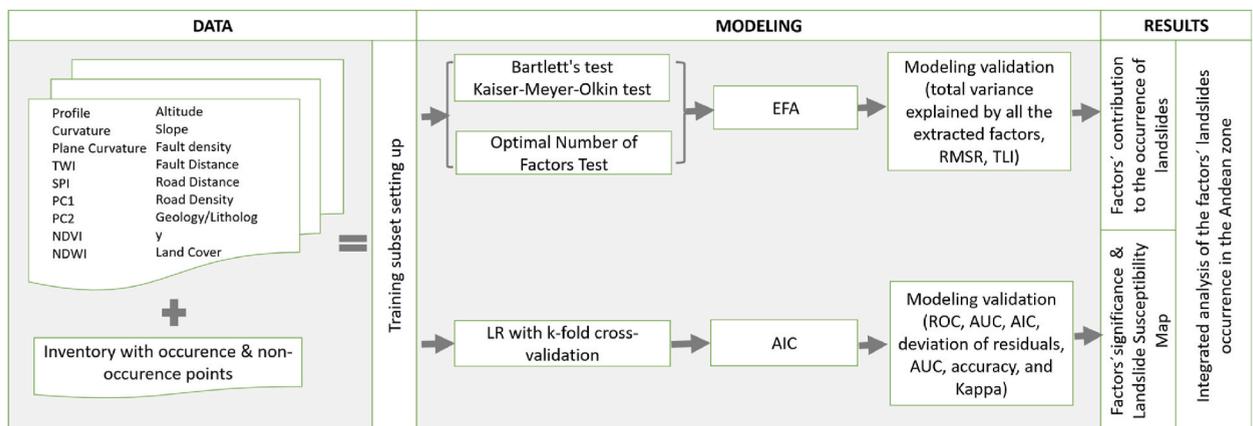


Fig. 1. Methodology flowchart.

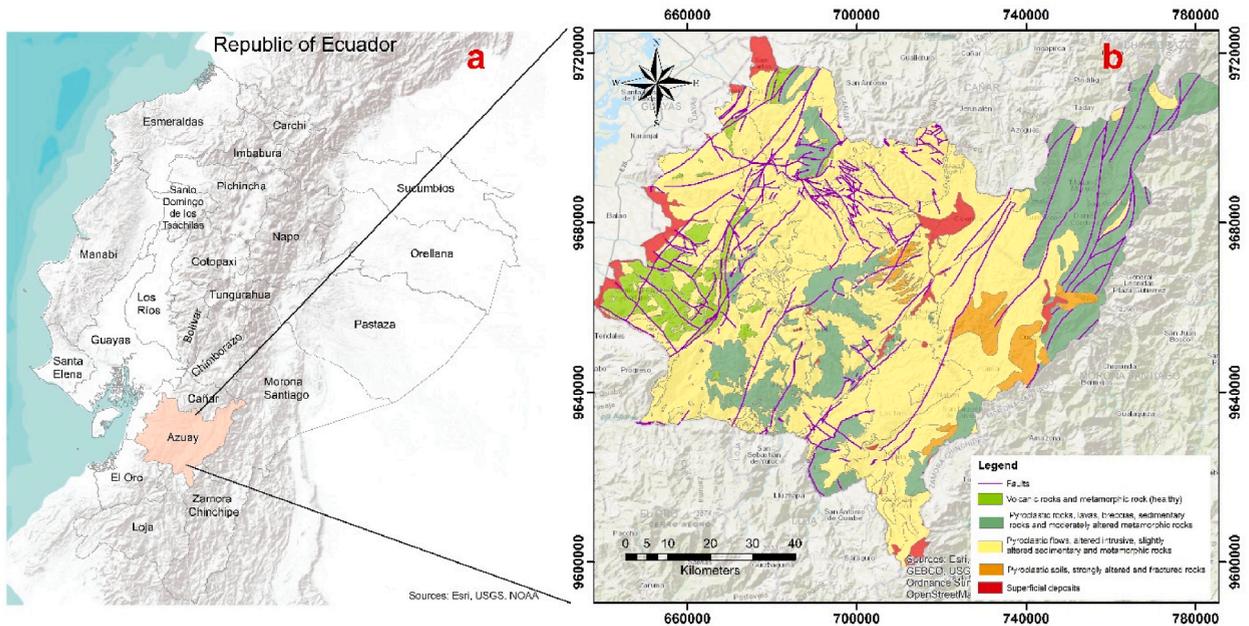


Fig. 2. Study area. (a) Location Map; (b) Geological Map.

Table 1
Favorability of geology for landslide occurrence.

Lithology	Categories	Favorability
Volcanic rocks (basalts, gabbros, andesite) healthy metamorphic rocks.	Low	1
Pyroclastic rocks, lavas, breccias, sedimentary rocks, and moderately altered metamorphic rocks.	Moderated	2
Pyroclastic flows, tuffs, altered intrusives, slightly altered sedimentary rocks, and altered metamorphic rocks.	Medium	3
Pyroclastic soils, strongly altered and fractured rocks.	High	4
Superficial deposits, colluvial, alluvial, alluvial fan, porphyritic rocks (subvolcanic), andesitic porphyries.	Very high	5

Source: Based on the classification of [53].

2.2.2. Faults

The existing geological lineaments, structures, and faults were derived from regional geological maps at a 1:100,000 scale. Two factors are derived: distance and density. The former factor was computed through the Euclidean Distance to the nearest fault and the latter was obtained by computing a Kernel Density function with a search radius according to Eq. (1). It represents two standard deviations (σ) from the median of the distances between inventory points and fault lines (μ).

$$dk = 2\sigma + \mu \tag{1}$$

2.2.3. Morphology

The morphology characterization was based on a digital elevation model (DEM) originally at 3 m spatial resolution (raster format). The slope, profile, and plane curvature, Topographic Wetness Index (TWI), and Stream Power Index (SPI) were derived from it. The strength of shear stresses associated with gravity and driving forces increases as the slope angle does [54]. According to Fanos and Pradhan (2019) [55], rockfall is related to slopes from 45° to 75°, surface landslides with angles from 23° to 43°, and debris flows with angles from 15° to 25°.

Profile and plane curvatures are slope changes based on second-degree altitude derivatives [56]. The profile curvature contributes to erosion and sedimentation. Negative values indicate convex curvature and water flow acceleration, while positive values indicate concave curvature and water flow deceleration [57]. The plane curvature allows for predicting whether a fluid will converge (negative values), diverge (positive values), or behave linearly (values near zero) [58]. Soil erosion will be more significant on concave slopes because water accumulates as it flows down the slope [59]. Curve analysis has been widely applied for landslide zoning [60, 61].

TWI has been used as a landslide explicative factor to assess the structural connectivity of the water network. Besides marking potential surface runoff pathways, TWI high values are also helpful for identifying hydrological sinks [62]. Water seeps into the subsoil or evaporates within hydrological sinks. It causes accumulation processes that alter soil material conditions and promote the occurrence of landslides. Eq. (2) shows how TWI is computed, where SCA is the flow accumulation area per unit contour width. This factor shows the tendency to receive water. β is the slope in radians, representing the tendency to drain water [63].

$$TWI = \ln \left(\frac{SCA}{\tan \beta} \right) \quad (2)$$

SPI can be used to determine the amount of soil loss or erosive power caused by streamflow. This index can predict net erosion in areas of convergence and flow acceleration (convex profile curvature and tangential concavity), as well as net disposal in areas of decreasing flow velocity (concave profile curvature) [64]. Eq. (3) was used for its calculation, where A_s^* is the catchment area and β is the local slope gradient in degrees. The calculation is based on the assumption that discharge is proportional to the catchment area [65]. Generally, low or flat slopes facilitate flooding processes, therefore, have the topographic potential for sediment accumulation. These slopes have a negative SPI value. Positive values indicate steep slopes that contribute significantly to soil erosion and degradation.

$$SPI = A_s^* \tan \beta \quad (3)$$

2.2.4. Land cover

This study obtained the land cover map from secondary sources (Provincial Government of Azuay and Universidad del Azuay). It is a vector layer at a scale of 1:5000. This factor has been categorized into eight classes: water, forest, wasteland, other covers, no information, agricultural land, shrub and herbaceous vegetation, and anthropic zones.

2.2.5. Spectral indexes

The NDVI facilitates the differentiation of vegetation types (e.g., landslide terraces) and provides information of the photosynthetically active biomass. The index was computed using Eq. (4). Higher positive values indicate more vigorous or denser vegetation cover. In contrast, negative values indicate water and snow, while values close to zero indicate bare soil [66]. NDWI, on the other hand, is sensitive to water changes in land cover and is a soil condition indicator [67]. This index was computed using Eq. (5), where negative values indicate dry areas and positive values indicate wet areas and vegetation with higher moisture levels. Both equations are calculated based on satellite images and their spectral bands, where NIR is the abbreviation for Near Infrared Band. RED and GREEN refer to bands in the visible spectrum that represent the colors of their name.

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (4)$$

$$NDWI = \frac{GREEN - NIR}{GREEN + NIR} \quad (5)$$

This research calculates these spectral indices using Sentinel-2 (A and B) time series at a spatial and temporal resolution of 10 m and 5 days, respectively. Sentinel-2's band 8 is NIR with a central wavelength of 842 nm; RED (band 4) and GREEN (band 3) have their central wavelength at 665 nm and 560 nm, respectively.

In total, 1649 Sentinel 2 A and B images were automatically downloaded using the Sen2r package. The study area is covered by 4 tiles (17 MPS, 17 MQS, 17 MQT, 17 MPT) from the five-year period from 2016 to 2020. The images were preferentially acquired at level 2 A, while those only available at level 1-C were atmospherically corrected. These indices values range between -1 and 1 . However, for this research, it was calculated in bytes. These values are re-scaled from 0 to 200; the non-data values take 255. Then, the values are changed to 100, multiplied by 100, and truncated to 200. All steps of this process were carried out using the "sen2cor" module in R from the European Space Agency [68]. A time series analysis was performed using 1649 NDVI and NDWI images to calculate a composite image for each index [69]. Composite images are more explanatory than a single image from a specific date. It highlights vegetation growth information by eliminating background noise caused by weather and other factors [70]. Authors such as Behling et al. [71] used the mean as the base statistic to build the composite, while Cao et al. [72] and Lan et al. [73] used the maximum. This research used the median (to avoid outliers' influence) to have an index for each month. Subsequently, it computed the average of all of them [74], resulting in one NDVI and one NDWI for the entire five-year period.

2.2.6. Roads

The Provincial Government of Azuay provided the road network at a scale of 1:25,000. Two factors were derived from roads, distance, and density, using analogous procedures to those of section 2.2.2.

2.2.7. Precipitation

Daily time-series data between 1985 and 2020 were obtained from 45 weather stations located in the Azuay province and surrounding watersheds. This information was provided by the National Institute of Meteorology and Hydrology, and the Municipal Public Company of Telecommunications, Drinking Water, Sewerage and Sanitation of Cuenca. Climatological data must be accurate and complete for temporal analysis. However, these data have errors due to the technique and type of instrument used for their acquisition. It implies the need for a homogenization process. It was done through the R package "climatol", which implements Standard Normal Homogeneity Test (SNHT) for its homogenization [75].

Precipitation statistics such as means, standard deviations, and extreme data (maxima, 99th percentile, and 95th percentile) were derived from the homogenized daily and monthly data series. These statistics were computed for the complete series and the dry and rainy seasons, representing 31 precipitation variables. According to Tang et al. [23], the principal component analysis (PCA) reduced this large number of variables. This statistical technique, extensively described by Bro and Smilde [76], linearly transforms an original

set of variables into another set of uncorrelated variables. The amount of variance explained in each principal component (PC) defines the order of these new variables. In the framework of this research, Kaiser's rule [77] was followed to choose the number of significant PCs, according to which PCs with an eigenvalue greater than one were chosen [78]. The selected PCs were added as new variables to the landslide inventory, replacing the 31 precipitation statistics computed from the homogenized series.

Homogenous spatial resolution for all raster layers is required for map algebra and modeling. All factors were rasterized and re-sampled to 10 m spatial resolution considering the grid of Sentinel-2 images. Geological and precipitation factors were upsampling to have higher resolution images than its inputs. On the other hand, the inverse process was done with downsampling for DEM and land cover layers. The consequent data loss is not representative due to the factor's spatial variability.

2.3. Inventory

Data-driven susceptibility models start with an inventory of occurrences. These historical records can be obtained through (i) visual photo interpretation of aerial photographs or satellite images [79], (ii) secondary sources [52], (iii) object-based image analysis [80], (iv) damage assessment through field inspection [81], and (v) synthetic aperture radar interferometry [82,83].

National Risk and Emergency Management Service and the Geological and Energy Research Institute provided an inventory of the landslide locations for the study area. The former institution reported 173 points from 2016 to 2020. This dataset was compiled based on reported landslide emergencies, being relatively new. The latter institution elaborated an inventory considering studies, consultancies, and field reports from the '70s, '80s, and '90s, reporting 492 points (e.g., PRECUPA, DINAGE). Both datasets were merged into a unique database to build the occurrence sample.

Samples representing positive and negative occurrences are needed to analyze the significance of the factor related to landslides and to build predictive models. The study area was divided into affected and non-affected areas to obtain the non-occurrence samples. The affected area was defined based on a buffer from occurrence points. The buffer influence distance corresponded to two standard deviations of the median of the minimum distances between each occurrence point. The median was used as a central trend measure because of the non-normality of the distance distribution. The remaining area corresponded to the non-affected zone. 589 random points were extracted from the latter area, maintaining the proportionality between occurrence points versus the affected area and non-occurrence points versus the non-affected area. Occurrence points were labeled with 1, and non-occurrence points were labeled with 0.

2.4. Exploratory factorial analysis (EFA)

EFA probes correlative relationships between manifest factors and models these relationships with one or more latent factors [84]. Its purpose is to generate hypotheses rather than to validate any existing one a-priori [85]. This method can be further described in Watkins (2018) [86] EFA considers essential parameters such as: (i) the optimal number of factors to retain, computed through Kaiser [77], scree test [87], and parallel analysis [88], among other techniques; (ii) the rotation of factors, which can be orthogonal (e.g., varimax and quartimax), or oblique (e.g., Direct Oblimin, Promax). Orthogonal rotations are used with uncorrelated factors or when a general factor is expected. In contrast, oblique ones allow some correlation [89]. Based on the above, this research used the Promax rotation as the most used within the oblique rotations.

The conformation of latent variables is given by grouping observed variables with the highest Factor Loading (FL). According to this value, there is a minor contribution with $FL > |0.3|$, a significant contribution with $FL > |0.5|$, and a relevant contribution with $FL > |0.7|$. All those $FL < |0.3|$ will be non-significant. On the other hand, the total variance of each variable will be explained through communality, uniqueness, and error. The communality (h^2) is the variance explained by the total of the computed factors. An h^2 greater than 0.6 is recommended with sample sizes less than 100 observations, and an h^2 greater than 0.5 with sample sizes greater than 100 observations. Uniqueness refers to the variance explained by the same variable, which does not depend on other variables. Finally, there is the random error.

2.5. Logistic Regression (LR)

LR is one of the most reliable approaches for evaluating landslide susceptibility and optimizing the model using the most significant factors [45, 90]. The factor labeling with 1 for occurrence and 0 for non-occurrence locations is considered the categorical-binary dependent variable. The independent variables are morphological, environmental, and anthropic factors. They can be numerical or categorical [91]. The LR model was calculated through Eq. (6), where y represents the dependent variable, x_i is the i -th explanatory variable, β_0 is a constant, β_i is the i -th coefficient of the regression, and e is the error [92]. The probability of occurrence of the event instead was represented through Eq. (7) using maximum likelihood estimation.

$$\text{logit}(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i + e \quad (6)$$

$$p = \frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i)}} \quad (7)$$

2.6. Models validation

The statistical validation for EFA was based on: (i) the cumulative variance explained by the extracted factors, which must be greater than half of the total variance (0.5) [93]; (ii) the Root Mean Square of the Residuals (RMSR), where the lowest value was sought (<0.08: reasonable fit and <0.05: perfect fit) [94]; (iii) the Tucker Lewis Index (TLI), which looks for values closest to 1. Values greater than 0.90 represent acceptable fit [95], while values greater than 0.95 represent excellent fits [96]. These threshold values refer to confirmatory factor analysis, being more flexible for EFA.

The selection and optimization of factors are usually made stepwise in LR [97]. It adds factors one by one and tests their significance by the *f* or *t*-test. The Akaike Information Criterion (AIC) developed by Akaike [98] was applied in this research. This criterion evaluates how well the model fits the observed data by penalizing the complexity of the model. The most suitable model will be the one with the smallest AIC value. A detailed description of AIC can be found in Liao et al. [99]. The stepAIC function in R tries various combinations of the explanatory variables and stops when the AIC no longer decreases [100].

Spatially-distributed data could present spatial autocorrelation or lack of statistical independence. It causes an overestimation of the predictive power of models. K-fold cross-validation reduces the predictive model's bias and thus helps avoid overfitting [101]. This method consists of splitting the data into *k* subsets of nearly equal size. The model is estimated from *k*-1 subsets because the subset left out is used to validate the model. This process is repeated *k* times [102]. *K* will depend on the amount of data, which in practice uses 5 or 10 [97, 102, 103]. In this study, 10 folds were used. Thus, the total efficiency of the model corresponds to the average of each iteration error [104].

Accuracy, Kappa, Sensitivity, and Specificity statistics were calculated through the confusion matrix. It is a fundamental term to compare predicted values with real values to calculate Accuracy [105]. Accuracy represents the ratio of correctly classified samples in the landslide inventory according to Eq. (8). Kappa coefficient is interpreted as a measure of agreement beyond chance compared to the maximum possible beyond chance agreement [106]. It is calculated with Eq. (9). Values lower than 0 represent a poor agreement, from 0 to 0.2, a slight agreement, from 0.21 to 0.4 fair, 0.41 to 0.6 moderate, 0.61 to 0.8 substantial and 0.81 to 1 almost perfect [107]. Sensitivity is the True Positive rate or the probability that a positive observation is classified as positive. It is calculated according to Eq. (10), and Specificity is the True Negative Rate, or the probability that a negative observation is classified as negative, calculated with Eq. (11) [108]. These four metrics are based on the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). A good model must contain a high percentage of TP and TN because they represent the number of hits between predicted and actual values. In contrast, the model error is presented in the FP and FN percentage, representing overestimations and underestimations, respectively [109]. In this context, all four metrics must be closer to one.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (8)$$

$$Kappa = \frac{(TP + TN) - [(TP + FN)(TP + FP) + (FP + FN)(FN + TN)] / (TP + FP + FN + TN)}{(TP + FP + FN + TN) - [(TP + FN)(TP + FP) + (FP + FN)(FN + TN)] / (TP + FP + FN + TN)} \quad (9)$$

$$Sensitivity = \frac{(TP)}{(TP + FN)} \quad (10)$$

$$Specificity = \frac{(TN)}{(TN + FP)} \quad (11)$$

On the other hand, curve ROC is one of the most efficient methods to evaluate the performance of predictive models. The area under the ROC curve (AUC) describes its ability to accurately predict the occurrences and non-occurrence of the landslide [110,111], considering TPR and False Positive Rate (FPR) [112]. AUC is calculated from Eq. (12), where *P* is the number of records with landslides presence, and *N* is the number of records with landslides absence. The result is expressed in values ranging from 0 to 1. The range 0.9–1 represents an excellent performance; 0.8–0.9 is an excellent performance; 0.7–0.8 is a good performance; 0.6–0.7 is an average performance; 0.5–0.6 is a fair performance; and below 0.5 represents a poor performance [113].

$$AAUC = \sum TP + \sum \frac{TP}{P} + N \quad (12)$$

3. Results

3.1. EFA

PC1 explained (above a |0.20|) the monthly and daily means of 36 years and the rainy season, the daily maximum of the dry season, the monthly 95th percentile, and the precipitation of April and May. PC2 (also with eigenvalues above |0.20|) explained the monthly precipitation of January, February, and March, the daily and monthly maximum values of the rainy season, the maximum of the complete time series, and the daily and monthly standard deviations in the rainy season.

Bartlett's test of sphericity ($X^2 = 7517.19$, p -value = 0, $df = 136$) and the Kaiser-Meyer-Olkin test ($KMO = 0.54$) were significant for the training set, allowing EFA analysis. Two of the 23 methods used to identify the number of relevant factors (Parallel Analysis,

Kaiser Criterion) determined seven factors for reliable EFA modeling. It is based on the premise of getting the model that best fits reality rather than the perfect one (because it does not exist). In this framework and according to Table 2, the latent factor ML1 explains the factors NDVI and NDWI, so it was named the Vegetation Factor. ML4 has higher factor loadings for road distance, density, and for the landslide classifier, so it was named the Anthropropic Factor. TWI, slope, and SPI represent ML3; it was named Morphological Factor. ML2 represents only altitude. ML5 is represented by fault density and distance, called the Geological Factor. ML6 is represented by plane and profile curvatures, called the Terrain Curvature Factor. Finally, ML7 represents PC1 and PC2 of the precipitation, denominated the Climatic Factor.

In summary, the EFA model for landslide occurrence identified a relevant contribution of the NDVI, NDWI, TWI, Altitude, and PC2 with h^2 greater than 0.7; a significant contribution of roads (distance and density), faults (distance and density), and plane curvature, with h^2 higher than 0.5; a minimal contribution from the categorical factor landslide classifier, slope, and profile curvature, with h^2 greater than 0.3. Finally, SPI, land cover, PC1, and geology do not significantly contribute to the model because they present h^2 lower than 0.3. This seven-factor model explains 55% of the accumulated variance, with a mean element complexity of 1.5, an RMSR of 0.02 (<0.08), and a TLI equal to 0.89.

3.2. LR

The complete LR model with all the factors (Model A) is expressed in Eq. (13), and the optimized model (Model B) under AIC stepwise procedure is fully described in Table 3. The latter model dropped Profile and plane curvatures, SPI, TWI, Road Distance, and PC1 factors to simplify it. The significance of each controlling factors was given using the Z statistic and its corresponding p-value. Wald Statistic has the same purpose, and its P-value $P(X^2)$ is closely similar to $Pr(z)$. Model B directly correlates the probability of landslide occurrence with intercept, altitude, slope, faults distance, and density. On the other hand, there is an inverse relationship with road density, NDVI, NDWI, and PC2. Concerning the land cover with agricultural land as the control group, forests, areas without information, shrub, and herbaceous vegetation could be more prone to landslides than agricultural land. Regarding lithology, igneous, sedimentary, and moderately altered metamorphic rocks are more significant for landslides than superficial deposits, colluvial, alluvial, alluvial fan, porphyritic rocks (subvolcanic), and andesitic porphyries.

Table 2
Factor loadings of the observed factors in the latent factors.

Factors	ML ^a 1	ML ^a 4	ML ^a 3	ML ^a 2	ML ^a 5	ML ^a 6	ML ^a 7	h^2 ^b
NDVI ^c	-0.98	-0.12	-0.02	-0.02	-0.02	0.03	-0.1	1
NDWI ^d	0.95	0.19	0.05	-0.11	-0.02	-0.02	0.07	0.95
Road distance	-0.08	-0.7	-0.02	0.06	-0.06	-0.04	0.05	0.51
Road density	0.11	0.69	0.04	-0.08	0.07	0.03	0.2	0.54
Landslide classifier	0.08	0.64	0.06	-0.06	0.04	0.02	-0.08	0.44
TWI ^e	0.02	0.08	0.97	-0.03	0	0.21	0.04	1
Slope	-0.16	-0.3	-0.47	-0.03	-0.15	0.16	-0.21	0.43
SPI ^f	-0.01	-0.02	0.35	-0.03	0.01	0.12	-0.06	0.14
Altitude	-0.03	-0.1	-0.01	0.85	0.09	0.01	0.07	0.75
Land Cover	-0.03	-0.13	-0.05	0.28	0	-0.01	0.09	0.11
Fault density	0.03	-0.16	-0.05	0.01	-0.77	0	-0.01	0.62
Fault distance	0.02	-0.03	0.01	0.07	0.7	0.02	0.2	0.54
Plane curvature	0.03	-0.01	-0.26	-0.04	0.03	-0.68	-0.05	0.53
Profile curvature	-0.01	0.05	0.04	-0.03	0.02	0.61	-0.03	0.38
PC2 ^g	0.03	0.33	0.07	0.66	0.04	0	-0.66	1
PC1 ^h	0.08	0.01	0.01	0.09	0.1	0.01	0.41	0.2
Geology	0.03	0.16	0.03	0.18	0.22	-0.04	0.28	0.19

^a Latent Factors.

^b Communality.

^c Normalized Difference Vegetation Index.

^d Normalized Difference Water Index.

^e Topographic Wetness Index.

^f Stream Power Index.

^g Principal Component number 2.

^h Principal Component number 1.

Table 3
Optimized LR model for landslide susceptibility.

	Estimate	Std. Error	z value	Pr(> z)b	WaldTest (X ²)	P(X ²)	
(Intercept)	2.97E+01	6.25E+00	4.754	2.00E-06	2.26E+01		***
Altitude	3.79E-04	1.51E-04	2.519	0.011767	6.35E+00	0.011892	*
Slope	2.85E-02	7.70E-03	3.706	0.000211	1.37E+01	0.000220	***
Fault distance	6.19E-05	4.23E-05	1.463	0.143389	2.14E+00	0.143640	
Fault density	1.56E+03	4.82E+02	3.23	0.001237	1.04E+01	0.001269	**
Road density	-2.45E+03	2.28E+02	-10.745	<2E-16	1.15E+02	<2.22E-16	***
NDV1c	-1.23E-01	2.72E-02	-4.518	6.25E-06	2.04E+01	6.85E-06	***
NDW1d	-1.92E-01	3.99E-02	-4.797	1.61E-06	2.30E+01	1.81E-06	***
PC2e	-1.13E-01	4.56E-02	-2.466	0.01365	6.08E+00	0.013786	*
Geology: low	-8.33E-01	9.24E-01	-0.901	0.367705	8.11E-01	0.367880	
Geology:moderated	-3.06E+00	7.17E-01	-4.263	2.01E-05	1.82E+01	2.17E-05	***
Geology:medium	-2.40E+00	6.98E-01	-3.432	0.000599	1.18E+01	0.000618	***
Geology:high	1.67E+00	1.04E+00	1.602	0.109049	2.57E+00	0.109300	
LandCover:water	-1.27E+01	8.86E+02	-0.014	0.988607	2.04E-04	0.988610	
LandCover:forest	8.63E-01	3.10E-01	2.783	0.00538	7.75E+00	0.005462	**
LandCover:wasteland	3.93E-02	4.72E-01	0.083	0.933608	6.94E-03	0.933620	
LandCover:other-covers	-1.57E+01	2.40E+03	-0.007	0.99477	4.30E-05	0.994770	
LandCover:no-information	3.46E+00	4.55E-01	7.613	2.69E-14	5.80E+01	5.32E-14	***
LandCover:shrub-and-herbaceous-vegetation	8.60E-01	2.10E-01	4.099	4.16E-05	1.68E+01	4.43E-05	***
LandCover:anthropic zones	-1.06E+01	5.69E+02	-0.019	0.985136	3.47E-04	0.98514	

*Significancia codes: 0 '***' 0.001 '***' 0.01 '**' 0.05 '*' 0.1 '.' 1.

^a Standard Error, ^b probability, ^c Normalized Difference Vegetation Index, ^d Normalized Difference Water Index, ^ePrincipal Component number 2.

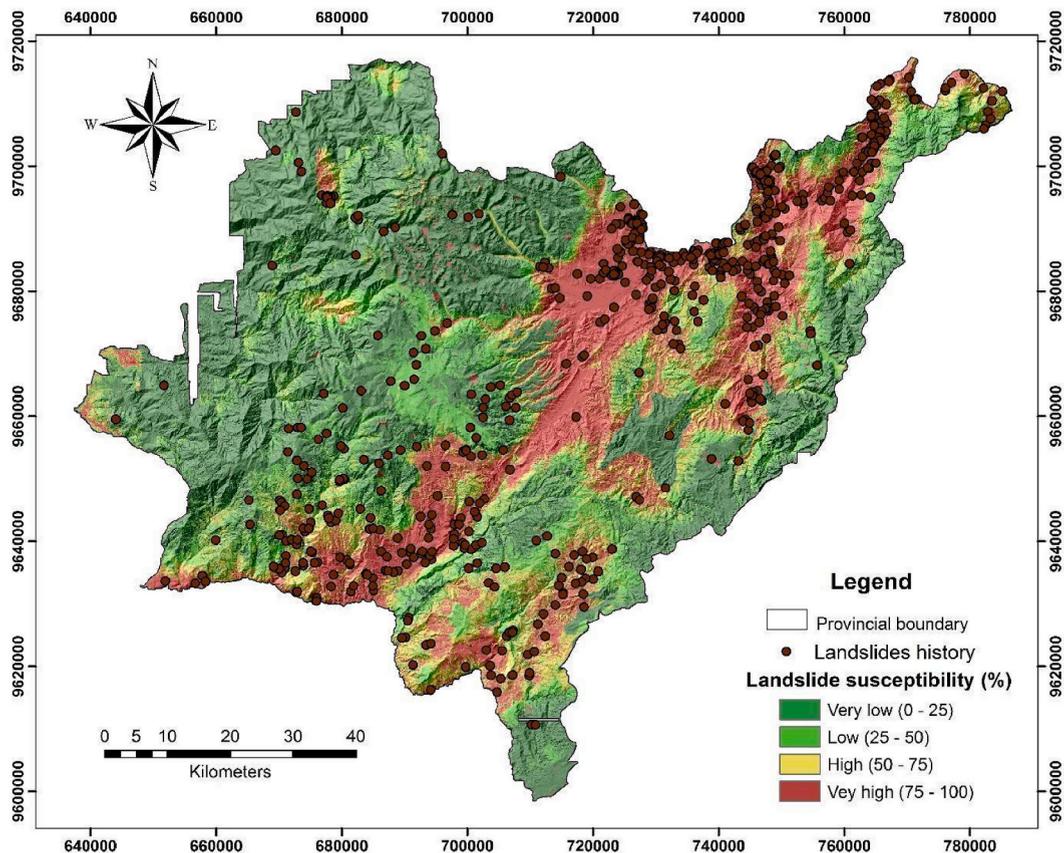


Fig. 3. Landslide susceptibility map from LR modeling.

$$\begin{aligned}
 &2.92E + 01 + 3.49E - 04 \text{ altitude} + 2.68E - 02 \text{ slope} - 3.06E - 02 \text{ profile curvature} - 2.14E - 02 \text{ plane curvature} + 6.54E \\
 &- 05 \text{ fault distance} + 1.60E + 03 \text{ fault density} + 9.11E - 06 \text{ SPI} - 3.86E - 02 \text{ TWI} + 4.07E - 05 \text{ road distance} - 2.30E \\
 &+ 03 \text{ road density} - 1.23E - 01 \text{ NDVI} - 1.92E - 01 \text{ NDWI} + 1.11E - 02 \text{ PC1} - 9.38E - 02 \text{ PC2} - 7.29E - 01 \text{ geology} \\
 &: \text{ low} - 3.01E + 00 \text{ geology} : \text{ moderated} - 2.34E + 00 \text{ geology} : \text{ medium} + 1.72E + 00 \text{ geology} : \text{ high} - 1.22E + 01 \text{ LandCover} \\
 &: \text{ water} + 8.55E - 01 \text{ LandCover} : \text{ forest} + 9.66E - 02 \text{ LandCover} : \text{ wasteland} - 1.58E + 01 \text{ LandCover} \\
 &: \text{ other} - \text{ covers} + 3.53E + 00 \text{ LandCover} : \text{ no} - \text{ information} + 8.47E - 01 \text{ LandCover} \\
 &: \text{ shrub} - \text{ and} - \text{ herbaceous} - \text{ vegetation} - 1.06E + 01 \text{ LandCover} : \text{ anthropic} - \text{ zones}
 \end{aligned}
 \tag{13}$$

The landslide susceptibility map shown in Fig. 3 was extracted through Model B. It is classified into 4 categories, divided into equal ranges from very low susceptibility for those areas with a probability of 0%–25%, low susceptibility for areas from 25% to 50%, high susceptibility from 50% to 75%, and very high from 75% to 100%. It can be observed that the probability increases in the glacial valley, in the foothills of the Andes mountain range, with a trend that goes from southwest to northeast. It means that the most active landslides are found to the east of the study area, and the activity decreases to the west. This pattern corresponds to lithology and geological faults. The most susceptible geological units are found along the Andes mountain range, such as metamorphic and sedimentary rocks, pyroclastic flows, and intrusive, altered, and fractured rocks belonging to the Turi, La Paz, Saraguro, and Ayancaya geological formations. The areas mostly affected by this phenomenon are located in the cantons of Cuenca (Turi, Santa María del Vergel, Asunción, Sigsig, Paccha, Lourdes, San Pedro de Llacao, La Unión de Llacao, La Merced, Los Trigales, Gusho, Nulti, Chaulabamba, Pacay, Ricaurte, Sinincay, Octavio Cordero, Santa Ana, San José de Balzay), Santa Isabel, San Fernando, Pucará, Girón, Oña, Nabón (Rosas, Tamboloma, Bellavista sectors), Sigsig, Gualaceo, Paute (Guaraynac sector), Sevilla de Oro, and Gapal (Jadán and Pumayunga).

3.3. Validation LR models

The confusion matrices of Models A and B are summarized in Table 4. The difference in Accuracy, Kappa, Sensitivity, and Specificity is small, as the AUC. The latter has a difference of 0.1%, as shown in Fig. 4, where Fig. 4(a) represents the ROC curve for Model A, and Fig. 4(b) illustrates the ROC curve for Model B. However, Model A has an AIC of 974.56 and a Residual Deviance of 922.56, while Model B has an AIC of 964.63 and Residual deviance of 924.63. The higher the AIC, the more complex the model. With slight differences in performance, Model B is more parsimonious than Model A, and the dropped factors (Profile and plane curvatures, SPI, TWI, Road Distance, and PC1) do not demonstrate to be significant.

4. Discussion

The knowledge about the spatial and temporal distribution of landslides, captured through landslide inventories, is essential to identifying the evolution of the landscape and its relationships with human activities and land management. It becomes a critical tool for estimating the probability of occurrence through susceptibility maps, necessary for establishing land use policies [9] and risk assessment. Volume, area extent, velocity, material, and kinematics often allow the identification of landslide behavior and the associated hazard [114]. A weak risk management system represents a significant limitation in developing countries. It is caused by insufficient scientific knowledge about the phenomenon, inefficient communication and information dissemination processes, and political-cultural barriers [115]. In addition, other studies have reported a lack of baseline information, which has been replaced by remote sensing techniques [116].

In this research, two models have been used to determine the importance of morphological, environmental, and anthropic factors in landslide occurrence. Thus, EFA and LR have coincided in the importance of NDVI, NDWI, fault and road density, altitude, and PC2. This last factor reflects precipitation in the wet season months of January, February, and March, extreme precipitation, and the dispersion of precipitation values around the mean. This result agrees with Zhang et al. [117], who stated that mountainous regions with complex geology, factors such as: elevation, precipitation, and vegetation are the most important factors in the occurrence of landslides. In this context, it is also important to note that, unlike LR, EFA did not select categorical factors (geology and land cover) as relevant.

Table 4
Confusion matrix of LR models.

			Prediction		
			Landslides	Non-landslides	
Modelo A	Observation	Landslides	547	95	Sensitivity = 0.86 Specificity = 0.84
		Non-landslides	93	495	
Accuracy = 0.85; Kappa = 0.69					
Modelo B	Observation	Landslides	545	97	Sensitivity = 0.85 Specificity = 0.84
		Non-landslides	93	495	
Accuracy = 0.85; Kappa = 0.69					

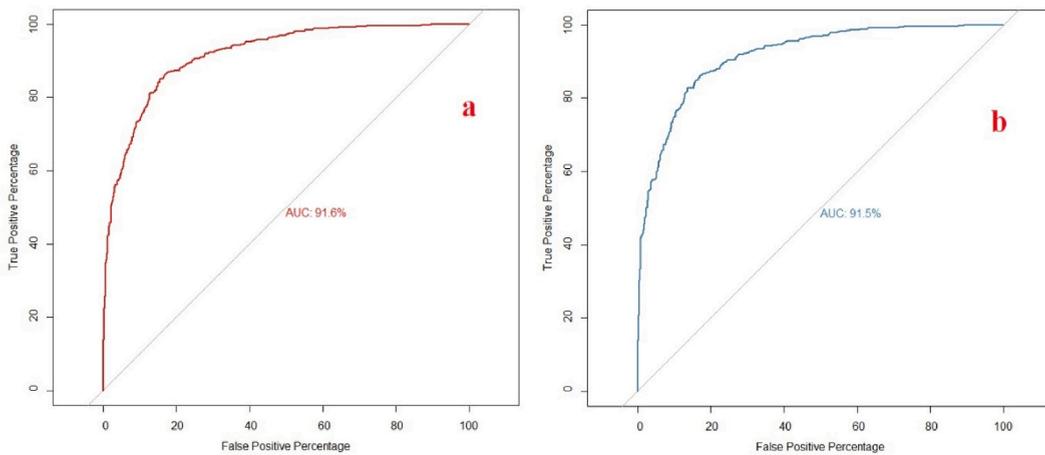


Fig. 4. ROC curves of LR models. (a) Model A with all factors; (b) Model B optimized.

NDVI, as a vegetation indicator, is one of the most significant indices for identifying landslide footprints or susceptibility maps [118,119]. The lower the NDVI value representing bare soil, anthropic zones, water bodies, or sparse vegetation (e.g., wasteland), the higher the probability of landslide for the study area. The vegetation canopy can act as a natural umbrella, reducing the amount of water for infiltration and evapotranspiration processes. However, under extreme precipitation events, the umbrella effect disappears due to canopy saturation, impacting slope stabilization [120]. Spiekermann [121] developed a model to represent trees' hydrological and mechanical influence on slopes, understanding that each vegetation type has different stabilizing effects. For instance, Guo et al. [122] show that grass favors landslides under rainy conditions. Despite this, Miandad et al. [123] state that using NDVI coupled with morphological factors helps to improve the overall model accuracy. In addition, this index becomes a key factor for the analysis of land recovery after a landslide [124].

This study shows the importance of NDWI for the characterization of landslide occurrence. It is consistent with Zhang et al. [125], Maqsoom et al. [126], and Nguyen et al. [20]. The water infiltrates when soil moisture uptake exceeds its water-holding capacity, resulting in subsurface runoff. In addition, soil particles expand with long-duration rainfall and swell, which reduces the inter-particle gap. This result—regarding the implications of water fluctuations in their study area—agrees with Hua et al. [127]. High NDWI values may also result from areas with higher precipitation regimes. However, the results of the present investigation show that the lower the NDWI value, the higher the probability of landslide occurrence. It is attributed to the fact that the lowest NDWI values for the study area are reflected in agricultural lands, shrubs, and herbaceous vegetation, which are sensitive to this type of phenomenon. It is observed that some researchers, such as Wang et al. [128], eliminate this factor from their study due to its high correlation with NDVI.

In mountainous areas with contrasting climatic patterns, such as the Azuay province in Ecuador, landslide history is highly related to rainfall storms [129,130], especially with large-scale fluctuations and extreme rainfall [3]. The study results determined a higher landslide probability when there is a higher rainfall standard deviation, higher precipitation in the wet season (January to March), and a higher maximum rainfall (due to their inverse relationship in the PC technique). Therefore, Orejuela and Toulkeridis [31] give greater weight to precipitation in their study of susceptibility in central Ecuador. However, according to Jesus et al. [131], groundwater rather than precipitation is more critical to triggering a landslide. Groundwater destabilizes the slope, reduces strength, changes density, and generates pore pressures. Therefore, monitoring and hydrometeorological forecasting conditions would allow timely risk management by cross-checking with information on landslides in vulnerable areas. Short-duration rainfall causes more significant water infiltration into the soil, while more extended rainfalls cause soil saturation; both events favor landslide occurrence. Chousianitis et al. [132] showed that slope failure scenario change in different seasons, highlighting that rainfall's effect on slope stability depends on the season of their occurrence. Studies conducted by Soto et al. [133] in the Loja-Ecuador basin show that rainfall with values of 90%–145% above average cause the studied landslides. This area is also located in the Ecuadorian Andes, sharing similar characteristics with the province of Azuay.

On the other hand, the geological factor is represented by three variables, faults distance, faults density, and geology. However, EFA does not indicate the importance of geology because the explained h^2 is minimal (0.2). Despite this, geology is well known as one of the most important factors for landslides in Andean areas [4,133] since it is directly related to landslide frequency and size [52]. Soils formed by young sediments containing poorly consolidated rocks, exposed sheared rocks, and soft or weak rocks covered by hard rocks are susceptible to this phenomenon [92,134]. LR confirms that igneous, sedimentary, moderately altered metamorphic rocks greatly influence the occurrence of landslides (p -value <0.001), both by the type of existing material and the geological structures. However, geology is not always the most significant factor in other mountainous areas, Pham et al. [111] showed that slope, road distance, and rainfall are the most important factors.

EFA and LR agree that the faults' presence could be an important factor in the occurrence of landslides. They produce stresses and induce deformations due to water flowing through the fractured rocks, producing chemical weathering, scouring, and thus, erosion. In this way, they cause a variation in stability, rock strength, and soil texture [135]. Faults or geological structures appearing in the

province of Azuay have preferential northeast-southwest and, to a lesser extent, northwest-southeast directions, as shown in Fig. 2(b). The three-dimensional geometry of the rock units and the geological faults' orientation allows for quantifying the formations' depth and volume. Therefore, the higher the density of faults, the greater the probability of landslide occurrence; the exact relationship is found with the fault's distance. However, the latter does not show statistical significance in LR. This behavior may occur because geological faults cross the entire study area.

Another highly significant factor (according to LR but not to EFA) is the quantitative description of the slope relief. It is commonly observed that a slope becomes progressively unstable as its gradient increases [134]. Therefore, LR states that the steeper the slope, the greater the probability of a landslide. This behavior agrees with Aditian et al. [136], Çellek et al. [137], and Guo et al. [138]. The study area presents slopes between 25° and 50°, which according to Van Zuidam classification [139] are steep and very steep slopes. However, the landslide susceptibility does not increase only with the slope degree increment. It must be coupled with other factors such as gravity, water content, soil structure, potential erosion, hydrological and geomorphological processes, vegetation type and density, and geological formation discontinuity [137,140]. The slope is not the most influential factor in all territorial contexts [141]. It is shown by Tang et al. [23], who conclude a slope relevance only for rockfall and not for all mass movement types. Altitude is another analyzed factor within the geomorphology, considered an essential piece for landslide prediction and recognition [142], as it is affected by geological tectonics. This research shows that the higher the altitude, the greater the probability of landslide occurrence in the study area. It is due to the strong association between altitude, vegetation cover, and geological formations. However, it contrasts with Sun et al. [143] and Das et al. [144], where a negative correlation with altitude is stated. They focus on the fact that high areas are dominated by resistant to weathering rocks.

In addition to the factors mentioned above, there is also anthropic action, mainly due to road construction, especially when roads are built on steep slopes using weak constructive standards [145]. The statistical significance of road distance and density provided by this research is consistent with the realities of developing countries. Such is the case of Brenning et al. [146], who established an increment of more than one order of magnitude in landslide probability for areas closer to paved interurban roads in the Andes of southern Ecuador. Moreover, McAdoo et al. [147] show that in Nepal, the probability is twice as much as in areas with poorly constructed road infrastructure. Road construction destabilizes slopes, modifies the original topography of the terrain, physically alters the landscape, changes the slope hydrology, and adds unstable materials along the roadside [148]. This infrastructure allows sediment transport and water accumulation in areas with abundant precipitation. Therefore, the drainage system and proper management become indispensable to ensuring road network sustainability.

Land cover is a highly dynamic factor due to human intervention and climate [149]. There is a slope modification (cuts in the slopes), changes in drainage systems, and deforestation (product of civil works construction), among other effects, when the land cover is altered. This alteration causes landslides [150], especially in Ecuadorian and Andean realities shown by Zimmermann and Elsenbeer [151] and Guns and Vanacker [152]. Landslide susceptibility can decrease with changes in land cover [153]. It would be the case from bare soil or grassland to cultivated land or from shrubland to forest, which favors slope stability [154]. A study applied to an Andean region in Colombia also reported that the existence of a forest decreases the probability of landslide occurrence [155,156].

5. Conclusion

Landslide occurrence and its relationship with the environment (factors) is a complex multilevel system that must be understood before taking any action in territorial planning and risk management. Altitude, fault density, vegetation indexes (NDVI and NDWI), and precipitation were relevant for Azuay's landslide susceptibility models built by both EFA and LR methodologies. Regarding precipitation, maximum rainfall in the wet season was highlighted, as it quickly saturates the soil. Further, EFA highlighted the importance of terrain morphology through a significant contribution of TWI and plane curvature. The anthropic influence is represented by road density and distance and geological faults distance. On the other hand, the LR model disagrees with EFA regarding terrain morphology by selecting slope as a significant factor. Unlike EFA, LR was more prone to select categorical factors such as geology and land cover.

The EFA model used 7 latent factors, which explained 55% of the accumulated variance, with a medium item complexity of 1.5, an RMSR of 0.02, and a TLI equal to 0.89. An AIC of 964.63, Residual deviance of 924.63, AUC of 0.92, an accuracy of 0.84, and a Kappa index of 0.68 were obtained in the case of LR.

NDVI, NDWI, and precipitation were used within a time-series analysis, which provides more comprehensive information than data from a single date. Those factors include vegetation and weather dynamism, which are directly related to landslide occurrence as triggering factors. This multi-temporal analysis is archived thanks to remote sensing information based on satellite images of medium spatial resolution, and daily data from weather stations.

The importance of the provided results resides in the identification of critical factors in the occurrence of landslides under an Andean context and the generation of a landslide susceptibility model through quantitative analysis on a regional scale for the study area. This methodological framework replaces traditional heuristic models based on expert knowledge. From our point of view, very few studies have been carried out in Ecuador for landslide susceptibility, and none of them focus on factor selection. In this context, this study contributes to state-of-the-art, covering geographical areas rarely explored but widely affected by landslides. The factors selected in the study can be extrapolated to other Andean cities in South America, such as La Paz, Cusco, Quito, and Bogota. However, they need to be calibrated with local inventories.

Some limitations that could be addressed in future research are (i) considering the typology of landslide (e.g., flow-like, rotational, planar, soil creep, rockfalls): since each landslide type could have different conditioning and triggering factors associated with them. (ii) Exploring new algorithms for landslide susceptibility models, such as machine learning techniques.

Author contribution statement

Sandra Lucia Cobos-Mora: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper. Victor Rodríguez-Galiano: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper. Aracely Lima: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Data availability statement

Data will be made available on request.

Funding

This work was supported by: Universidad Católica de Cuenca (UC) and the Instituto de Investigación Geológico y Energético (IIGE), who have co-executed the research project “Landslides: characterization of their morphological and environmental factors”, PICCIITT19-8 [grant numbers Oficio Nro.: UCACUE-JIEI-2020-85-OF].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to acknowledge the Universidad Católica de Cuenca (UC) and the Instituto de Investigación Geológico y Energético (IIGE) for funding this research through project number PICCIITT19-8 [grant numbers Oficio Nro.: UCACUE-JIEI-2020-85-OF]. The UC dependencies that were part of this research were laboratories DataAcademics/SigData, and Computational Calculation, Modeling and Data Analytics (C2MAD); the research groups Urban and Earth Data Science; and Faculties of Postgraduate and Agricultural Sciences to whom we are thankful. Finally, we also want to thank Jose Guamán for the graphic support.

References

- [1] M.J. Froude, D.N. Petley, Global fatal landslide occurrence from 2004 to 2016, *Nat. Hazards Earth Syst. Sci.* 18 (8) (2018 Aug 23) 2161–2181.
- [2] World Meteorological Organization, WMO Atlas of Mortality and Economic Losses from Weather, Climate and Water Extremes (1970–2019) [Internet]. Geneva - Switzerland, 2021 [cited 2022 Sep 26]. Available from: https://library.wmo.int/doc_num.php?explnum_id=10989.
- [3] U. Haque, P.F. da Silva, G. Devoli, J. Pilz, B. Zhao, A. Khaloua, W. Wilopo, P. Andersen, P. Lu, J. Lee, T. Yamamoto, D. Keellings, W. Jian-Hong, G.E. Glass, The human cost of global warming: deadly landslides and their triggers (1995–2014), *Sci. Total Environ.* 682 (2019 Sep 10) 673–684.
- [4] K. Chunga, F.A. Livio, C. Martillo, H. Lara-Saavedra, M.F. Ferrario, I. Zevallos, A.M. Michetti, Landslides triggered by the 2016 Mw 7.8 pedernales, Ecuador earthquake: correlations with ESI-07 intensity, lithology, slope and PGA-h, *Geosciences* 9 (9) (2019 Sep 1).
- [5] K.E. Clark, A.J. West, R.G. Hilton, G.P. Asner, C.A. Quesada, M.R. Silman, S.S. Saatchi, W. Farfan-Rios, R.E. Martin, A.B. Horwath, K. Halladay, M. New, Y. Malhi, Storm-triggered landslides in the Peruvian Andes and implications for topography, carbon cycles, and biodiversity, *Earth Surf. Dyn.* 4 (1) (2016) 47–70.
- [6] E. Aristizábal, H. Martínez-Carvajal, E. García-Aristizábal, Modelling shallow landslides triggered by rainfall in tropical and mountainous basins [Internet], *Adv Cult Living with Landslides* (2017) [cited 2022 Sep 24];207–12. Available from: https://link.springer.com/chapter/10.1007/978-3-319-53485-5_23.
- [7] U. Ozturk, E. Bozzolan, E.A. Holcombe, R. Shukla, F. Pianosi, T. Wagener, How climate change and unplanned urban sprawl bring more landslides [Internet], *Nature* (2022 Aug 8) [cited 2023 Jan 21];608(7922):262–5. Available from: <https://www.nature.com/articles/d41586-022-02141-9>.
- [8] Comunidad Andina [Internet]. Lima, Estrategia Andina para la Gestión del Riesgo de Desastres - EAGRD- Decisión 819 (2017 May) [cited 2022 Sep 24]. Available from: https://www.comunidadandina.org/StaticFiles/2017522151956ESTRATEGIA_ANDINA.pdf.
- [9] F. Puente-Sotomayor, A. Egas, J. Teller, Land policies for landslide risk reduction in Andean cities, *Habitat Int.* 107 (2021 Jan 1), 102298.
- [10] S.G. Evans, J.V. Degraff, Catastrophic landslide: effects, occurrence, and mechanisms [Internet], in: *Reviews in Engineering Geology*, vol. XV, The Geological Society of America, Inc. (GSA), Colorado, 2002 [cited 2022 Sep 24]. Available from: http://pubs.geoscienceworld.org/gsa/books/book/795/chapter-pdf/3742686/9780813758152_frontmatter.pdf.
- [11] A.U. Vinueza, J. Robles, M. Bakker, P. Guzman, T. Bogaard, Characterization and hydrological analysis of the guarumales deep-seated landslide in the tropical Ecuadorian Andes [Internet], *Geosciences* (2020 Jul 10) [cited 2022 Sep 24];10(7):267. Available from: <https://www.mdpi.com/2076-3263/10/7/267/htm>.
- [12] C.M. Sevilla, J. Idrovo, Use of lightweight fill for landslide repair: media via stabilization project, the cuenca-molleturo roadway, Ecuador [Internet], *Geotech Front* (2017). Mar 30 [cited 2022 Sep 24];126–35. Available from: <https://ascelibrary.org/doi/10.1061/9780784480458.013>.
- [13] El Comercio, Cifra de fallecidos asciende, en Cuenca, debido a aluvión registrado en Sayausí - El Comercio, 2022. Mar 28 [cited 2022 Sep 24]; Available from: <https://www.elcomercio.com/actualidad/cuenca/cofra-fallecidos-afectadas-aluvion-cuenca.html>.
- [14] H. Petschko, A. Brenning, R. Bell, J. Goetz, T. Glade, Assessing the quality of landslide susceptibility maps - case study Lower Austria, *Nat. Hazards Earth Syst. Sci.* 14 (1) (2014 Jan 16) 95–118.
- [15] A. Roccati, G. Paliaga, F. Luino, F. Faccini, L. Turconi, GIS-based landslide susceptibility mapping for land use planning and risk assessment [Internet], *Land* (2021). Feb 5 [cited 2023 Jan 21];10(2):162. Available from: <https://www.mdpi.com/2073-445X/10/2/162/htm>.
- [16] S. Steger, V. Mair, C. Kofler, M. Pittore, M. Zebisch, S. Schneiderbauer, Correlation does not imply geomorphic causation in data-driven landslide susceptibility modelling - benefits of exploring landslide data collection effects, *Sci. Total Environ.* 776 (2021), 145935.
- [17] K. Gaidzik, M.T. Ramírez-Herrera, The importance of input data on landslide susceptibility mapping [Internet], *Sci. Rep.* (2021 Sep 29) [cited 2023 Jan 16];11 (1):1–14. Available from: <https://www.nature.com/articles/s41598-021-98830-y>.
- [18] T. Kavzoglu, A. Teke, Ensemble Conditioning Factor Selection with Markov Chain Framework for Shallow Landslide Susceptibility Mapping in Lake Sapanca Basin and its Vicinity, Turkey [Internet], *Balt J Mod Comput*, 2022, <https://doi.org/10.22364/bjmc.2022.10.2.09> [cited 2023 Jan 22];10(2):224–40.
- [19] T. Kavzoglu, E. Kutlug Sahin, I. Colkesen, Selecting optimal conditioning factors in shallow translational landslide susceptibility mapping using genetic algorithm, *Eng. Geol.* 192 (2015 Jun 18) 101–112.

- [20] H.D. Nguyen, V.D. Pham, Q.H. Nguyen, V.M. Pham, M.H. Pham, V.M. Vu, Q.T. Bui, An optimal search for neural network parameters using the Salp swarm optimization algorithm: a landslide application, *Remote Sens Lett* [Internet] 11 (4) (2020) 353–362, <https://doi.org/10.1080/2150704X.2020.1716409>.
- [21] M. Fressard, Y. Thiery, O. Maquaire, Which data for quantitative landslide susceptibility mapping at operational scale case study of the pays d'auge plateau hillslopes (Normandy, France), *Nat. Hazards Earth Syst. Sci.* 14 (3) (2014 Mar 13) 569–588.
- [22] P. Vorpahl, H. Elsenbeer, M. Märker, B. Schröder, How can statistical models help to determine driving factors of landslides? *Ecol. Model.* 239 (2012 Jul 24) 27–39.
- [23] Y. Tang, F. Feng, Z. Guo, W. Feng, Z. Li, J. Wang, Q. Sun, H. Ma, Y. Li, Integrating Principal Component Analysis with Statistically-Based Models for Analysis of Causal Factors and Landslide Susceptibility Mapping: A Comparative Study from the Loess Plateau Area in Shanxi (China), vol. 277, *J Clean Prod* [Internet, 2020, 124159, <https://doi.org/10.1016/j.jclepro.2020.124159>.
- [24] B. Mandal, S. Mandal, Analytical hierarchy process (AHP) based landslide susceptibility mapping of Lish river basin of eastern Darjeeling Himalaya, India, *Adv. Space Res.* 62 (11) (2018 Dec 1) 3114–3132.
- [25] M. Mokarram, A.R. Zarei, Landslide susceptibility mapping using fuzzy-AHP [Internet], *Geotech. Geol. Eng.* (2018 May 28) [cited 2022 Sep 24];36(6): 3931–43. Available from: <https://link.springer.com/article/10.1007/s10706-018-0583-y>.
- [26] L. Shano, T.K. Raghuvanshi, M. Meten, Landslide susceptibility evaluation and hazard zonation techniques – a review [Internet], *Geoenvironmental Disasters* (2020 May 20) [cited 2022 Sep 24];7(1):1–19. Available from: <https://geoenvironmental-disasters.springeropen.com/articles/10.1186/s40677-020-00152-0>.
- [27] S. Alqadhi, J. Mallick, S. Talukdar, A.A. Bindajam, N. Van Hong, T.K. Saha, Selecting optimal conditioning parameters for landslide susceptibility: an experimental research on Aqabat Al-Sulbat, Saudi Arabia [Internet], *Environ. Sci. Pollut. Res.* (2022 Jan 1) [cited 2023 Jan 22];29(3):3743–62. Available from: <https://link.springer.com/article/10.1007/s11356-021-15886-z>.
- [28] V.H. Nhu, A. Shirzadi, H. Shahabi, S.K. Singh, N. Al-Ansari, J.J. Clague, A. Jaafari, W. Chen, S. Miraki, J. Dou, C. Luu, K. Górski, B.T. Pham, H.D. Nguyen, B. Bin Ahmad, Shallow landslide susceptibility mapping: a comparison between logistic model tree, logistic regression, Naïve Bayes tree, artificial neural network, and support vector machine algorithms [Internet], *Int. J. Environ. Res. Publ. Health* (2020 Apr 16) [cited 2023 Jan 23];17(8):2749. Available from: <https://www.mdpi.com/1660-4601/17/8/2749>.
- [29] B.T. Pham, I. Prakash, Evaluation and comparison of LogitBoost Ensemble, Fisher's Linear Discriminant Analysis, logistic regression and support vector machines methods for landslide susceptibility mapping [Internet], *Geocarto Int.* (2017 Feb 23) [cited 2023 Jan 23];34(3):316–33. Available from: <https://www.tandfonline.com.vpn.ucacue.edu.ec/doi/abs/10.1080/10106049.2017.1404141>.
- [30] J.A. Palenzuela Baena, J. Soto Luzuriaga, C. Irigaray Fernández, Characteristics of rainfall events triggering landslides in two climatologically different areas: southern Ecuador and southern Spain [Internet], *Hydrology* (2020 Jul 21) [cited 2022 Sep 25];7(3):45. Available from: <https://www.mdpi.com/2306-5338/7/3/45>.
- [31] I.P. Orejuela, T. Toulkeridis, Evaluation of the susceptibility to landslides through diffuse logic and analytical hierarchy process (AHP) between Macas and Riobamba in Central Ecuador, ICEDEG 2020 [Internet], in: 7th International Conference on eDemocracy and eGovernment, Institute of Electrical and Electronics Engineers Inc., Buenos Aires - Argentina, 2020 [cited 2022 Sep 24]. pp. 201–7. Available from: <https://ieeexplore.ieee.org/abstract/document/9096879>.
- [32] G. Zhang, Y. Cai, Z. Zheng, J. Zhen, Y. Liu, K. Huang, Integration of the statistical index method and the analytic hierarchy process technique for the assessment of landslide susceptibility in huizhou, China, *Catena* 142 (2016 Jul 1) 233–244.
- [33] F. Yan, Q. Zhang, S. Ye, B. Ren, A novel hybrid approach for landslide susceptibility mapping integrating analytical hierarchy process and normalized frequency ratio methods with the cloud model, *Geomorphology* 327 (2019 Feb 15) 170–187.
- [34] S. Das, S. Sarkar, D.P. Kanungo, GIS-based landslide susceptibility zonation mapping using the analytic hierarchy process (AHP) method in parts of Kalimpong Region of Darjeeling Himalaya [Internet], *Environ. Monit. Assess.* (2022). Mar 1 [cited 2023 Jan 16];194(3):1–28. Available from: <https://link.springer.com/article/10.1007/s10661-022-09851-7>.
- [35] D. Bălăteanu, M. Micu, M. Jurchescu, J.P. Malet, M. Sima, G. Kucsicsa, C. Dumitrică, D. Petrea, M.C. Mărgărint, Bilașco Ș. C.F. Dobrescu, E.A. Călărășu, E. Olinic, I. Boți, F. Senzaconi, National-scale landslide susceptibility map of Romania in a European methodological framework, *Geomorphology* 371 (2020 Dec 15), 107432.
- [36] I. Bostjančić, M. Filipović, V. Gulam, D. Pollak, Regional-scale landslide susceptibility mapping using limited LiDAR-based landslide inventories for sisak-moslavina county, Croatia [Internet], *Sustainability* (2021 Apr 19) [cited 2023 Jan 21];13(8):4543. Available from: <https://www.mdpi.com/2071-1050/13/8/4543/htm>.
- [37] G.D. Bathrellos, H.D. Skilodimou, K. Chousianitis, A.M. Youssef, B. Pradhan, Suitability estimation for urban development using multi-hazard assessment map, *Sci. Total Environ.* 575 (2017 Jan 1) 119–134.
- [38] Y. Liu, L. Zhao, A. Bao, J. Li, X. Yan, Chinese high resolution satellite data and GIS-based assessment of landslide susceptibility along highway G30 in guozigou valley using logistic regression and MaxEnt model [Internet], *Rem. Sens.* (2022 Jul 28) [cited 2023 May 28];14(15):3620. Available from: <https://www.mdpi.com/2072-4292/14/15/36200>.
- [39] B. Dashbold, L.S. Bryson, M.M. Crawford, Landslide hazard and susceptibility maps derived from satellite and remote sensing data using limit equilibrium analysis and machine learning model [Internet], *Nat. Hazards* (2022 Oct 23) [cited 2023 Jan 23];1–31. Available from: <https://link.springer.com/article/10.1007/s11069-022-05671-7>.
- [40] B. Rashid, J. Iqbal, L. Jun Su, Landslide susceptibility analysis of Karakoram highway using analytical hierarchy process and scoops 3D [Internet], *J. Mt. Sci.* (2020 Jul 1) [cited 2023 Jan 23];17(7):1596–612. Available from: <https://link.springer.com/article/10.1007/s11629-018-5195-8>.
- [41] W.L. Hakim, F. Rezaie, A.S. Nur, M. Panahi, K. Khosravi, C.W. Lee, S. Lee, Convolutional neural network (CNN) with metaheuristic optimization algorithms for landslide susceptibility mapping in Icheon, South Korea, *J. Environ. Manag.* 305 (2022 Mar 1), 114367.
- [42] J. Dou, A.P. Yumus, D.T. Bui, M. Sahana, C.W. Chen, Z. Zhu, W. Wang, B.T. Pham, Evaluating GIS-based multiple statistical models and data mining for earthquake and rainfall-induced landslide susceptibility using the LiDAR DEM, *Rem. Sens.* 11 (6) (2019 Mar 15) 638.
- [43] A. Ganga, M. Elia, E. D'Ambrosio, S. Tripaldi, G.F. Capra, F. Gentile, G. Sanesi, Assessing landslide susceptibility by coupling spatial data analysis and logistic model [Internet], *Sustainability* (2022 Jul 1) [cited 2023 Jan 22];14(14):8426. Available from: <https://www.mdpi.com/2071-1050/14/14/8426/htm>.
- [44] V.E. Nwazelibe, C.O. Unigwe, J.C. Egbueri, Integration and comparison of algorithmic weight of evidence and logistic regression in landslide susceptibility mapping of the Orumba North erosion-prone region, Nigeria [Internet], *Model Earth Syst Environ* (2022 Oct 2) [cited 2023 Jan 22];1–20. Available from: <https://link.springer.com/article/10.1007/s40808-022-01549-6>.
- [45] T. Van Phong, N.D. Dam, P.T. Trinh, N. Van Dung, N. Hieu, C.Q. Tran, T.D. Van, Q.C. Nguyen, I. Prakash, B.T. Pham, GIS-based logistic regression application for landslide susceptibility mapping in son La hydropower reservoir basin [Internet], in: CIGOS 2021, Emerging Technologies and Applications for Green Infrastructure Lecture Notes in Civil Engineering, Springer Science and Business Media Deutschland GmbH, 2022 [cited 2023 Jan 22]. pp. 1841–9. Available from: https://link.springer.com/chapter/10.1007/978-981-16-7160-9_186.
- [46] O. Genc, Identifying principal risk factors in Turkish construction sector according to their probability of occurrences: a relative importance index (RII) and exploratory factor analysis (EFA) approach [Internet], *Int J Constr Manag* (2021) [cited 2023 Jan 23]; Available from: <https://www.tandfonline.com.vpn.ucacue.edu.ec/doi/abs/10.1080/15623599.2021.1946901>.
- [47] R.D. Ledesma, P.J. Ferrando, M.A. Trógo, F.M. Poó, J.D. Tosi, C. Castro, Exploratory factor analysis in transportation research: current practices and recommendations, *Transport. Res. F Traffic Psychol. Behav.* 78 (2021 Apr 1) 340–352.
- [48] M. Barancoková, Z. Krnáčová, Assessment of landslide susceptibility using statistical modelling in the flysch zone of the Western Carpathians (NW Slovakia), *Phys Geogr* [Internet] 41 (6) (2020) 558–586, <https://doi.org/10.1080/02723646.2020.1770039>.
- [49] F. Puente-Sotomayor, A. Mustafa, J. Teller, Landslide susceptibility mapping of urban areas: logistic regression and sensitivity analysis applied to Quito, Ecuador [Internet], *Geoenvironmental Disasters* (2021 Dec 1) [cited 2023 Jan 22];8(1):1–26. Available from: <https://geoenvironmental-disasters.springeropen.com/articles/10.1186/s40677-021-00184-0>.

- [50] E. Bravo-López, T. Fernández Del Castillo, C. Sellers, J. Delgado-García, Landslide susceptibility mapping of landslides with artificial neural networks: multi-approach analysis of backpropagation algorithm applying the neuralnet package in Cuenca, Ecuador, *Rem. Sens.* 14 (14) (2022 Jul 21) 3495.
- [51] Gobierno Provincial del Azuay, Plan de Desarrollo y Ordenamiento Territorial del Azuay Actualizado 2015 -2030 [Internet], Cuenca (2015) [cited 2022 Sep 26]. Available from: <https://odsterritorioecuador.ec/wp-content/uploads/2019/04/PDOT-PROVINCIA-AZUAY-2015-2030.pdf>.
- [52] S. Junquera-Torradó, S.M. Moreiras, S.A. Sepúlveda, Distribution of landslides along the Andean active orogenic front (Argentinean Precordillera 31–33° S), *Quat Int* [Internet] (2019), <https://doi.org/10.1016/j.quaint.2019.01.030>, 512(August 2018):18–34.
- [53] C.S. Mora, Vahron Wilhelm-Günther, Macrozonation methodology for landslide hazard determination, *Environ. Eng. Geosci.* xxxi (1) (1994 Mar 1) 49–58.
- [54] A. Dehnavi, I.N. Aghdam, B. Pradhan, M.H. Morshed Varzandeh, A new hybrid model using step-wise weight assessment ratio analysis (SWARA) technique and adaptive neuro-fuzzy inference system (ANFIS) for regional landslide hazard assessment in Iran, *Catena* 135 (2015 Dec 1) 122–148.
- [55] A.M. Fanos, B. Pradhan, A novel hybrid machine learning-based model for rockfall source identification in presence of other landslide types using LiDAR and GIS, *Earth Syst Environ* 3 (3) (2019 Dec 1) 491–506.
- [56] Z. Chen, S. Liang, Y. Ke, Z. Yang, H. Zhao, Landslide susceptibility assessment using evidential belief function, certainty factor and frequency ratio model at Baxie River basin, NW China, *Geocarto Int.* 34 (4) (2019 Mar 21) 348–367.
- [57] I. Ahmad, Digital elevation model (DEM) coupled with geographic information system (GIS): an approach towards erosion modeling of Gumara watershed, Ethiopia, *Environ. Monit. Assess.* (10) (2018 Oct 1) 190.
- [58] L. Blaga, Aspects regarding the significance of the curvature types and values in the studies of geomorphometry assisted by GIS [Internet], *Analele Univ din Oradea, Ser Geogr* (2012) [cited 2022 Sep 24]; Available from: <http://istgeoint.utoradea.ro/Reviste/Anale/anale.htm>.
- [59] I. Iliá, P. Tsangaratos, Applying weight of evidence method and sensitivity analysis to produce a landslide susceptibility map [Internet], *Landslides* (2016 Apr 7) [cited 2022 Sep 24];13(2):379–97. Available from: <https://link.springer.com/article/10.1007/s10346-015-0576-3>.
- [60] R. Schlögel, I. Marchesini, M. Alvioli, P. Reichenbach, M. Rossi, J.P. Malet, Optimizing landslide susceptibility zonation: effects of DEM spatial resolution and slope unit delineation on logistic regression models, *Geomorphology* 301 (2018 Jan 15) 10–20.
- [61] Z. Chen, F. Ye, W. Fu, Y. Ke, H. Hong, The influence of DEM spatial resolution on landslide susceptibility mapping in the Baxie River basin, NW China [Internet], *Nat. Hazards* (2020 Apr 1) [cited 2022 Dec 7];101(3):853–77. Available from: <https://link.springer.com/article/10.1007/s11069-020-03899-9>.
- [62] K. Jancewicz, P. Migoń, M. Kasprzak, Connectivity patterns in contrasting types of tableland sandstone relief revealed by Topographic Wetness Index, *Sci. Total Environ.* 656 (2019 Mar 15) 1046–1062.
- [63] S. Panchal, A.K. Shrivastava, Landslide hazard assessment using analytic hierarchy process (AHP): a case study of National Highway 5 in India, *Ain Shams Eng. J.* 13 (3) (2022 May 1), 101626.
- [64] T. Danielson, Utilizing a high resolution digital elevation model (DEM) to develop a stream power index (SPI) for the Gilmore Creek watershed in Winona County, Minnesota [Internet], *Pap Resour Anal* (2013) [cited 2022 Sep 24];15. Available from: <http://www.gis.smumn.edu>.
- [65] A. Sharma, Integrating terrain and vegetation indices for identifying potential soil erosion risk area [Internet], *Geo-Spatial Inf. Sci.* (2010) [cited 2022 Sep 24]; 13(3):201–9. Available from: <https://www.tandfonline.com/doi/abs/10.1007/s11806-010-0342-6>.
- [66] T. Lillesand, R. Kiefer, J. Chipman, Remote sensing and image interpretation, *Photogramm Eng Remote Sens.* 7th Edition 81 (8) (2015 Aug 12) 615–616.
- [67] V.D. Pham, Q.H. Nguyen, H.D. Nguyen, V.M. Pham, V.M. Vu, Q.T. Bui, Convolutional neural network - optimized moth flame algorithm for shallow landslide susceptible analysis, *IEEE Access* 8 (2020) 32727–32736.
- [68] L. Ranghetti, M. Boschetti, F. Nutini, L. Busetto, sen2r™: an R toolbox for automatically downloading and preprocessing Sentinel-2 satellite data, *Comput. Geosci.* 139 (2020 Jun 1), 104473.
- [69] S. Li, L. Xu, Y. Jing, H. Yin, X. Li, X. Guan, High-quality vegetation index product generation: a review of NDVI time series reconstruction techniques, *Int. J. Appl. Earth Obs. Geoinf.* 105 (2021 Dec 25), 102640.
- [70] Z. Ni, Z. Yang, W. Li, Y. Zhao, Z. He, Decreasing trend of geohazards induced by the 2008 Wenchuan earthquake inferred from time series NDVI data [Internet], *Rem. Sens.* (2019 Sep 20) [cited 2023 Jan 23];11(19):2192. Available from: <https://www.mdpi.com/2072-4292/11/19/2192>.
- [71] R. Behling, S. Roessner, D. Golovko, B. Kleinschmit, Derivation of long-term spatiotemporal landslide activity—a multi-sensor time series approach, *Remote Sens. Environ.* 186 (2016 Dec 1) 88–104.
- [72] R. Cao, Y. Chen, M. Shen, J. Chen, J. Zhou, C. Wang, W. Yang, A simple method to improve the quality of NDVI time-series data by integrating spatiotemporal information with the Savitzky-Golay filter, *Remote Sens. Environ.* 217 (2018 Nov 1) 244–257.
- [73] S. Lan, Z. Dong, Incorporating vegetation type transformation with NDVI time-series to study the vegetation dynamics in Xinjiang, *Sustainability* [Internet] (2022 Jan 5) [cited 2023 May 9];14(1):582. Available from: <https://www.mdpi.com/2071-1050/14/1/582>.
- [74] W. De Keersmaecker, S. Lhermitte, M.J. Hill, L. Tits, P. Coppin, B. Somers, Assessment of regional vegetation response to climate anomalies: a case study for Australia using gims NDVI time series between 1982 and 2006 [Internet], *Rem. Sens.* (2017 Jan 4) [cited 2022 Sep 24];9(1):34. Available from: <https://www.mdpi.com/2072-4292/9/1/34/htm>.
- [75] J.A. Guijarro, Homogenization of climatic series with Climatol [Internet], Balearic Islands (2018 May) [cited 2022 Sep 24]. Report No.: Version 3.1.1. Available from: https://www.researchgate.net/publication/325203476_Homogenization_of_climatic_series_with_Climatol?channel=doi&linkId=5afda3fea6fdcc3a5a90bd5b&showFulltext=true.
- [76] R. Bro, A.K. Smilde, Principal component analysis, *Anal. Methods* 6 (9) (2014) 2812–2831.
- [77] H. Kaiser, The application of electronic computers to factor analysis, *Educ. Psychol. Meas.* 20 (1) (1960).
- [78] T. Basu, A. Das, S. Pal, Application of geographically weighted principal component analysis and fuzzy approach for unsupervised landslide susceptibility mapping on Gish River Basin, India, *Geocarto Int.* 37 (5) (2020) 1294–1317.
- [79] F. Bucci, M. Santangelo, F. Fiorucci, F. Ardizzone, D. Giordan, M. Cignetti, D. Notti, P. Allasia, D. Godone, D. Lagomarsino, A. Pozzoli, E. Norelli, M. Cardinali, Geomorphologic landslide inventory by air photo interpretation of the High Agri Valley (Southern Italy), *J. Maps* 17 (2) (2021) 376–388.
- [80] L. Hao, A. R. K.S.S. Van Westen C, T. Ranjan Martha, P. Jaiswal, G. McAdoo B, Constructing a complete landslide inventory dataset for the 2018 monsoon disaster in Kerala, India, for land use change analysis, *Earth Syst. Sci. Data* 12 (4) (2020 Nov 16) 2899–2918.
- [81] C. Xu, X. Xu, J.B.H. Shyu, Database and spatial distribution of landslides triggered by the Lushan, China Mw 6.6 earthquake of 20 April 2013, *Geomorphology* 248 (2015 Nov 1) 77–92.
- [82] E.H. Bouali, T. Oommen, R. Escobar-Wolf, Mapping of slow landslides on the Palos Verdes Peninsula using the California landslide inventory and persistent scatterer interferometry, *Landslides* 15 (3) (2018 Mar 1) 439–452.
- [83] A. Rosi, V. Tofani, L. Tanteri, C. Tacconi Stefanelli, A. Agostini, F. Catani, N. Casagli, The new landslide inventory of Tuscany (Italy) updated with PS-InSAR: geomorphological features and landslide distribution [Internet], *Landslides* (2018 Jan 1) [cited 2022 Sep 24];15(1):5–19. Available from: <https://link.springer.com/article/10.1007/s10346-017-0861-4>.
- [84] D. Goretzko, T.T.H. Pham, M. Bühner, Exploratory factor analysis: current use, methodological developments and recommendations for good practice, *Curr. Psychol.* 40 (7) (2021 Jul 1) 3510–3521.
- [85] L. Luo, C. Arizmendi, K.M. Gates, Exploratory factor analysis (EFA) programs in R [Internet], *Struct Equ Model A Multidiscip J* (2019 Sep 3) [cited 2023 Jan 21];26(5):819–26. Available from: <https://www.tandfonline.com.vpn.ucacue.edu.ec/doi/abs/10.1080/10705511.2019.1615835>.
- [86] M.W. Watkins, Exploratory factor analysis: a guide to best practice [Internet], *J. Black Psychol.* (2018 Apr 27) [cited 2022 Sep 24];44(3):219–46. Available from: <https://journals.sagepub.com/doi/10.1177/0095798418771807>.
- [87] R.B. Cattell, The scree test for the number of factors [Internet], *Multivariate Behav. Res.* (1966 Apr 1) [cited 2022 Sep 24];1(2):245–76. Available from: https://www.tandfonline.com/doi/abs/10.1207/s15327906mbr0102_10.
- [88] J.L. Horn, A rationale and test for the number of factors in factor analysis [Internet], *Psychometrika* (1965 Jun) [cited 2022 Sep 26];30(2):179–85. Available from: <https://link.springer.com/article/10.1007/BF02289447>.
- [89] A. Hefetz, G. Liberman, The factor analysis procedure for exploration: a short guide with examples, *Cult. y Educ.* 29 (3) (2017) 526–562.

- [90] B. Zhao, W. Li, Y. Wang, J. Lu, X. Li, Landslides triggered by the Ms 6.9 Nyingchi earthquake, China (18 November 2017): analysis of the spatial distribution and occurrence factors, *Landslides* 16 (4) (2019) 765–776.
- [91] H. Hemasinghe, R.S.S. Rangali, N.L. Deshapriya, L. Samarakoon, Landslide susceptibility mapping using logistic regression model (a case study in Badulla District, Sri Lanka), *Procedia Eng.* 212 (2018 Jan 1) 1046–1053.
- [92] M.E.A. Budimir, P.M. Atkinson, H.G. Lewis, A systematic review of landslide probability mapping using logistic regression, *Landslides* 12 (3) (2015 Jun 1) 419–436.
- [93] P. Samuels, Advice on exploratory factor analysis [Internet], ResearchGate (2017) [cited 2022 Sep 25]. Available from: <http://bcu-test.eprints-hosting.org/6076/>.
- [94] D. Almaleki, Stability of the data-model fit over increasing levels of factorial invariance for different features of design in factor analysis [Internet], *Eng. Technol. Appl. Sci. Res.* (2021 Apr 11) [cited 2022 Sep 24];11(2):6849–56. Available from: <https://etasr.com/index.php/ETASR/article/view/4047>.
- [95] Y. Yang, D. Montgomery, Exploratory and confirmatory factor analyses of the multicultural teaching scale [Internet], *J. Psychoeduc. Assess.* (2011 Aug 31), <https://doi.org/10.1177/0734282910378983> [cited 2022 Sep 24];29(3):261–72. Available from:
- [96] Y. Xia, Y. Yang, RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: the story they tell depends on the estimation methods [Internet], *Behav. Res. Methods* (2019 Feb 15) [cited 2022 Sep 24];51(1):409–28. Available from: <https://link.springer.com/article/10.3758/s13428-018-1055-2>.
- [97] Y. Zhang, H. Wen, P. Xie, D. Hu, J. Zhang, W. Zhang, Hybrid-optimized logistic regression model of landslide susceptibility along mountain highway [Internet], *Bull. Eng. Geol. Environ.* (2021 Aug 20) [cited 2022 Sep 24];80(10):7385–401. Available from: <https://link.springer.com/article/10.1007/s10064-021-02415-y>.
- [98] H. Akaike, A new look at the statistical model identification, *IEEE Trans. Automat. Control* 19 (6) (1974) 716–723.
- [99] J.G. Liao, J.E. Cavanaugh, T.L. McMurry, Extending AIC to best subset regression [Internet], *Comput. Stat.* (2018 Feb 9) [cited 2022 Sep 24];33(2):787–806. Available from: <https://link.springer.com/article/10.1007/s00180-018-0797-8>.
- [100] Z. Zhang, Variable selection with stepwise and best subset approaches [Internet], *Ann. Transl. Med.* (2016 Apr 1) [cited 2022 Sep 24];4(7):136–136. Available from: <https://atm.amegroups.com/article/view/9706/html>.
- [101] P. López, J. Quiñe, C. Martínez, Applicability of spatial prediction models for landslide susceptibility in land-use zoning instruments: a guideline in a coastal settlement in South-Central Chile [Internet], *Geocarto Int.* (2022) [cited 2023 Feb 5];37(22):6474–93. Available from: <https://www.tandfonline.com/vpn.ucacue.edu/doi/abs/10.1080/10106049.2021.1939440>.
- [102] Y. Jung, Multiple predicting K-fold cross-validation for model selection [Internet], *J. Nonparametric Statistics* (2018 Jan 2) [cited 2022 Sep 24];30(1):197–215. Available from: <https://www.tandfonline.com/doi/abs/10.1080/10485252.2017.1404598>.
- [103] P. Roy, S.C. Pal, A. Arabameri, R. Chakraborty, B. Pradhan, I. Chowdhuri, S. Lee, D.T. Bui, Novel ensemble of multivariate adaptive regression spline with spatial logistic regression and boosted regression tree for gully erosion susceptibility [Internet], *Rem. Sens.* (2020 Oct 10) [cited 2023 Feb 5];12(20):3284. Available from: <https://www.mdpi.com/2072-4292/12/20/3284/html>.
- [104] Z. Xiong, Y. Cui, Z. Liu, Y. Zhao, M. Hu, J. Hu, Evaluating explorative prediction power of machine learning algorithms for materials discovery using k-fold forward cross-validation, *Comput. Mater. Sci.* 171 (2020 Jan 1), 109203.
- [105] G. Zeng, On the confusion matrix in credit scoring and its analytical properties [Internet], *Commun. Stat. Theor. Methods* (2019 May 2), <https://doi.org/10.1080/03610926.2019.1568485> [cited 2023 Aug 17];49(9):2080–93. Available from:
- [106] D. Chicco, M.J. Warrens, G. Jurman, The Matthews correlation coefficient (MCC) is more informative than Cohen's Kappa and brier score in binary classification assessment, *IEEE Access* 9 (2021) 78368–78381.
- [107] N.W.S. Wardhani, M.Y. Rochayani, A. Iriany, A.D. Sulistyono, P. Lestanyo, Cross-validation metrics for evaluating classification performance on imbalanced data, in: *International Conference on Computer, Control, Informatics and its Applications: Emerging Trends in Big Data and Artificial Intelligence, IC3INA 2019*, Institute of Electrical and Electronics Engineers Inc., 2019, pp. 14–18, 2019.
- [108] I.M. De Diego, A.R. Redondo, R.R. Fernández, J. Navarro, J.M. Mogueza, General performance score for classification problems [Internet], *Appl. Intell.* (2022 Aug 1) [cited 2023 Aug 17];52(10):12049–63. Available from: <https://link.springer.com/article/10.1007/s10489-021-03041-7>.
- [109] L.C. Quesada-Ruiz, V.F. Rodríguez-Galiano, R. Zurita-Milla, E. Izquierdo-Verdiguier, Area and Feature Guided Regularised Random Forest: a novel method for predictive modelling of binary phenomena. The case of illegal landfill in Canary Island [Internet], *Int. J. Geogr. Inf. Sci.* (2022) [cited 2023 Aug 11];36(12):2473–95. Available from: <https://www.tandfonline.com/us.debiblio.com/doi/abs/10.1080/13658816.2022.2075879>.
- [110] S. Khezri, A. Ahmadi Dehrashid, B. Nasrollahzadeh, H. Moayed, H. Ahmadi Dehrashid, H. Azadi, J. Scheffran, Prediction of landslides by machine learning algorithms and statistical methods in Iran, *Environ Earth Sci* [Internet] (2022 Jun 1) [cited 2023 Aug 11];81(304):1–22. Available from: <https://link.springer.com/article/10.1007/s12665-022-10388-8>.
- [111] B.T. Pham, I. Prakash, D. Tien Bui, Spatial prediction of landslides using a hybrid machine learning approach based on Random Subspace and Classification and Regression Trees, *Geomorphology* 303 (2018 Feb 15) 256–270.
- [112] V. Rodríguez-Galiano, M.P. Mendes, M.J. García-Soldado, M. Chica-Olmo, L. Ribeiro, Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: a case study in an agricultural setting (Southern Spain), *Sci. Total Environ.* (2014 Apr 1), 476–477:189–206.
- [113] A. Wubalem, M. Meten, Landslide susceptibility mapping using information value and logistic regression models in Goncha Siso Eneses area, northwestern Ethiopia [Internet], *SN Appl. Sci.* (2020 May 1) [cited 2022 Sep 24];2(5):1–19. Available from: <https://link.springer.com/article/10.1007/s42452-020-2563-0>.
- [114] M. Bunn, B. Leshchinsky, M.J. Olsen, Estimates of three-dimensional rupture surface geometry of deep-seated landslides using landslide inventories and high-resolution topographic data, *Geomorphology* 367 (2020 Oct 15), 107332.
- [115] J. Klimeš, A.M. Rosario, R. Vargas, P. Raška, L. Vicuña, C. Jurt, Community participation in landslide risk reduction: a case history from Central Andes, Peru [Internet], *Landslides* (2019 Jul 13) [cited 2022 Sep 24];16(9):1763–77. Available from: <https://link.springer.com/article/10.1007/s10346-019-01203-w>.
- [116] N. Younes Cárdenas, E. Erazo Mera, Landslide susceptibility analysis using remote sensing and GIS in the western Ecuadorian Andes [Internet], *Nat. Hazards* (2016 Jan 16) [cited 2022 Sep 24];81(3):1829–59. Available from: <https://link.springer.com/article/10.1007/s11069-016-2157-8>.
- [117] J. Zhang, X. Ma, J. Zhang, D. Sun, X. Zhou, C. Mi, H. Wen, Insights into geospatial heterogeneity of landslide susceptibility based on the SHAP-XGBoost model, *J. Environ. Manag.* 332 (2023 Apr 15), 117357.
- [118] X. Zheng, G. He, S. Wang, Y. Wang, G. Wang, Z. Yang, J. Yu, N. Wang, Comparison of machine learning methods for potential active landslide hazards identification with multi-source data [Internet], *ISPRS Int. J. Geo-Inf.* (2021 Apr 9) [cited 2022 Sep 24];10(4):253. Available from: <https://www.mdpi.com/2220-9964/10/4/253/html>.
- [119] G. Acosta, A. Rodríguez, P. Euillades, L. Euillades, F. Ruiz, P. Rosell, M. Sanchez, F. Leiva, J. Ariza, H. García, Detection of active landslides by DInSAR in andean precordillera of san juan, Argentina, *J South Am Earth Sci* (2021 Jun 1) 108.
- [120] A. Gonzalez-Ollauri, S.B. Mickovski, Hydrological effect of vegetation against rainfall-induced landslides, *J. Hydrol.* 549 (2017 Jun 1) 374–387.
- [121] R.I. Spiekermann, S. McColl, I. Fuller, J. Dymond, L. Burkitt, H.G. Smith, Quantifying the influence of individual trees on slope stability at landscape scale, *J. Environ. Manag.* (2021 May 15), 286:112194.
- [122] W.Z. Guo, Z.X. Chen, W.L. Wang, W.W. Gao, M.M. Guo, H.L. Kang, P.F. Li, W.X. Wang, M. Zhao, Telling a different story: the promote role of vegetation in the initiation of shallow landslides during rainfall on the Chinese Loess Plateau, *Geomorphology* 350 (2020 Feb 1), 106879.
- [123] J. Miandad, M.M. Darrow, M.D. Hendricks, R.P. Daanen, Landslide mapping using multiscale LiDAR digital elevation models, *Environ. Eng. Geosci.* 26 (4) (2020 Nov 1) 405–425.
- [124] M. Chen, C. Tang, X. Wang, J. Xiong, Q. Shi, X. Zhang, M. Li, Y. Luo, Y. Tie, Q. Feng, Temporal and spatial differentiation in the surface recovery of post-seismic landslides in Wenchuan earthquake-affected areas, *Ecol. Inf.* 64 (2021 Sep 1), 101356.

- [125] H. Zhang, Y. Song, S. Xu, Y. He, Z. Li, X. Yu, Y. Liang, W. Wu, Y. Wang, Combining a class-weighted algorithm and machine learning models in landslide susceptibility mapping: a case study of Wanzhou section of the Three Gorges Reservoir, China, in: *Comput Geosci* [Internet], vol. 158, 2022, <https://doi.org/10.1016/j.cageo.2021.104966>. October 2021;104966.
- [126] A. Maqsoom, B. Aslam, U. Khalil, Z.A. Kazmi, S. Azam, T. Mehmood, A. Nawaz, Landslide susceptibility mapping along the China Pakistan Economic Corridor (CPEC) route using multi-criteria decision-making method [Internet], *Model Earth Syst Environ* (2021), 0123456789, <https://doi.org/10.1007/s40808-021-01226-0>.
- [127] Y. Hua, X. Wang, Y. Li, P. Xu, W. Xia, Dynamic development of landslide susceptibility based on slope unit and deep neural networks [Internet], *Landslides* (2021 Jun 30) [cited 2022 Sep 24];18(1):281–302. Available from: <https://link.springer.com/article/10.1007/s10346-020-01444-0>.
- [128] Y. Wang, L. Feng, S. Li, F. Ren, Q. Du, A hybrid model considering spatial heterogeneity for landslide susceptibility mapping in Zhejiang Province, China, *Catena* 188 (2020 May 1), 104425.
- [129] E. Aristizábal, E. García, R. Marin, F. Gómez, J. Guzmán, Rainfall-intensity effect on landslide hazard assessment due to climate change in north-western Colombian Andes [Internet], *Rev Fac Ing Univ Antioquia* (2022) [cited 2022 Sep 24];(103):51–66. Available from: <https://revistas.udea.edu.co/index.php/ingenieria/article/view/343311>.
- [130] A. Beyene, N. Tesema, F. Fufa, D. Tsige, Geophysical and numerical stability analysis of landslide incident [Internet], *Heliyon* (2023), <https://doi.org/10.1016/j.heliyon.2023.e13852> [cited 2023 Jun 17];9.
- [131] C.C. Jesus, S.C. Oliveira, C. Sena, F. Marques, Understanding constraints and triggering factors of landslides: regional and local perspectives on a drainage basin, *Geosci.* 8 (1) (2018).
- [132] K. Chousianitis, V. Del Gaudio, N. Sabatakakis, K. Kavoura, G. Drakatos, G.D. Bathrellos, H.D. Skilodimou, Assessment of earthquake-induced landslide hazard in Greece: from arias intensity to spatial distribution of slope resistance demand, *Bull. Seismol. Soc. Am.* 106 (1) (2016 Feb 1) 174–188.
- [133] J. Soto, J.P. Galve, J.A. Palenzuela, J.M. Azañón, J. Tamay, C. Irigaray, A multi-method approach for the characterization of landslides in an intramontane basin in the Andes (Loja, Ecuador), *Landslides* 14 (6) (2017 Dec 1) 1929–1947.
- [134] J.W. Lin, M.H. Hsieh, Y.J. Li, Factor analysis for the statistical modeling of earthquake-induced landslides [Internet], *Front. Struct. Civ. Eng.* (2020 Nov 19) [cited 2022 Sep 24];14(1):123–6. Available from: <https://link.springer.com/article/10.1007/s11709-019-0582-y>.
- [135] L. Ayalew, H. Yamagishi, The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan, *Geomorphology* 65 (1–2) (2005 Feb 1) 15–31.
- [136] A. Aditian, T. Kubota, Y. Shinohara, Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia, *Geomorphology* [Internet 318 (2018) 101–111, <https://doi.org/10.1016/j.geomorph.2018.06.006>.
- [137] S. Çellek, Effect of the slope angle and its classification on landslide, *Nat. Hazards Earth Syst. Sci.* (May) (2020) 1–23.
- [138] Z. Guo, Y. Shi, F. Huang, X. Fan, J. Huang, Landslide susceptibility zonation method based on C5.0 decision tree and K-means cluster algorithms to improve the efficiency of risk management, *Geosci. Front.* 12 (6) (2021 Nov 1), 101249.
- [139] R.A. Van Zuidam [Internet]. Smits Publishers, editor. Aerial photo-interpretation in terrain analysis and geomorphologic mapping. Smits Publishers, The Hague, Aerial photo-interpretation in terrain analysis and geomorphologic mapping (1986) [cited 2022 Sep 25]. Available from: https://books.google.com.ec/books/about/Aerial_Photo_interpretation_in_Terrain_A.html?id=zGLGQgAACAAJ&redir_esc=y.
- [140] P. Tsangaratos, I. Ilia, Landslide susceptibility mapping using a modified decision tree classifier in the Xanthi Perfection, Greece [Internet], *Landslides* (2016) [cited 2022 Sep 25];13(2):305–20. Available from: <https://link.springer.com/article/10.1007/s10346-015-0565-6>.
- [141] C.J. Van Westen, N. Rengers, R. Soeters, Use of geomorphological information in indirect landslide susceptibility assessment [Internet], *Nat. Hazards* (2003 Nov) [cited 2022 Sep 24];30(3):399–419. Available from: <https://link.springer.com/article/10.1023/B:NHAZ.0000007097.42735.9e>.
- [142] S. Ji, D. Yu, C. Shen, W. Li, Q. Xu, Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks [Internet], *Landslides* (2020 Jun 1) [cited 2022 Sep 24];17(6):1337–52. Available from: <https://link.springer.com/article/10.1007/s10346-020-01353-2>.
- [143] X. Sun, L. Yuan, S. Tao, M. Liu, D. Li, Y. Zhou, H. Shao, A novel landslide susceptibility optimization framework to assess landslide occurrence probability at the regional scale for environmental management, *J. Environ. Manag.* 322 (2022 Nov 15), 116108.
- [144] J. Das, P. Saha, R. Mitra, A. Alam, M. Kamruzzaman, GIS-based data-driven bivariate statistical models for landslide susceptibility prediction in Upper Tista Basin, India [Internet], *Heliyon* (2023), <https://doi.org/10.1016/j.heliyon.2023.e16186> [cited 2023 Jun 17];9.
- [145] S. Saha, A. Arabameri, A. Saha, T. Blaschke, P.T.T. Ngo, V.H. Nhu, S.S. Band, Prediction of landslide susceptibility in Rudraprayag, India using novel ensemble of conditional probability and boosted regression tree-based on cross-validation method, *Sci. Total Environ.* 764 (2021 Apr 10), 142928.
- [146] A. Brenning, M. Schwinn, A.P. Ruiz-Páez, J. Muenchow, Landslide susceptibility near highways is increased by 1 order of magnitude in the Andes of southern Ecuador, Loja province, *Nat. Hazards Earth Syst. Sci.* 15 (1) (2015 Jan 7) 45–57.
- [147] B.G. McAdoo, M. Quak, K.R. Gnyawali, B.R. Adhikari, S. Devkota, P. Lal Rajbhandari, K. Sudmeier-Rieux, Roads and landslides in Nepal: how development affects environmental risk, *Nat. Hazards Earth Syst. Sci.* 18 (12) (2018 Nov 30) 3203–3210.
- [148] C. Vuillez, M. Tonini, K. Sudmeier-Rieux, S. Devkota, M.H. Derron, M. Jaboyedoff, Land use changes, landslides and roads in the Phewa Watershed, Western Nepal from 1979 to 2016, *Appl. Geogr.* 94 (2018 May 1) 30–40.
- [149] B.M. Meneses, S. Pereira, E. Reis, Effects of different land use and land cover data on the landslide susceptibility zonation of road networks, *Nat. Hazards Earth Syst. Sci.* 19 (3) (2019 Mar 11) 471–487.
- [150] L. Chen, Z. Guo, K. Yin, D. Pikha Shrestha, S. Jin, The influence of land use and land cover change on landslide susceptibility: a case study in Zhushan Town, Xuan'en County (Hubei, China), *Nat. Hazards Earth Syst. Sci.* 19 (10) (2019 Oct 10) 2207–2228.
- [151] B. Zimmermann, H. Elsenbeer, The near-surface hydrological consequences of disturbance and recovery: a simulation study, *J. Hydrol.* 364 (1–2) (2009 Jan 15) 115–127.
- [152] M. Guns, V. Vanacker, Shifts in landslide frequency–area distribution after forest conversion in the tropical Andes, *Anthropocene* 6 (2014 Jun 1) 75–85.
- [153] L. Pisano, V. Zumpano, Malek, C.M. Roskopf, M. Parise, Variations in the susceptibility to landslides, as a consequence of land cover changes: a look to the past, and another towards the future [Internet], *Sci. Total Environ.* (2017), <https://doi.org/10.1016/j.scitotenv.2017.05.231>, 601–602:1147–59.
- [154] J. Lopez-Saez, C. Corona, N. Eckert, M. Stoffel, F. Bourrier, F. Berger, Impacts of land-use and land-cover changes on rockfall propagation: insights from the Grenoble conurbation, *Sci. Total Environ.* 547 (2016 Mar 15) 345–355.
- [155] F.F. Ávila, R.C. Alvalá, R.M. Mendes, D.J. Amore, The influence of land use/land cover variability and rainfall intensity in triggering landslides: a back-analysis study via physically based models, *Nat Hazards* [Internet] 105 (1) (2021) 1139–1161, <https://doi.org/10.1007/s11069-020-04324-x>.
- [156] N. Grima, D. Edwards, F. Edwards, D. Petley, B. Fisher, Landslides in the Andes: forests can provide cost-effective landslide regulation services, *Sci. Total Environ.* (2020 Nov 25) 745.