

Landslide Modeling in a Tropical Mountain Basin Using Machine Learning Algorithms and Shapley Additive Explanations

Authors: Vega, Johnny, Sepúlveda-Murillo, Fabio Humberto, and Parra, Melissa

Source: Air, Soil and Water Research, 16(1)

Published By: SAGE Publishing

URL: <https://doi.org/10.1177/11786221231195824>

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Landslide Modeling in a Tropical Mountain Basin Using Machine Learning Algorithms and Shapley Additive Explanations

Johnny Vega¹, Fabio Humberto Sepúlveda-Murillo² and Melissa Parra¹

¹Facultad de Ingeniería, Universidad de Medellín, Medellín, Colombia, and ²Facultad de Ciencias Básicas, Universidad de Medellín, Medellín, Colombia

Air, Soil and Water Research
Volume 16: 1–20
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/11786221231195824



ABSTRACT: Landslides are a geological hazard commonly induced by rainfall, earthquakes, deforestation, or human activity causing loss of human life every year specially on highlands or mountain slopes with serious impacts that threaten communities and its infrastructure. The incidence and recurrence of landslides are conditioned by several aspects related to soil properties, geological structure, climatic conditions, soil cover, and water flow. Precisely, Colombia is one of the most affected by this type of natural hazard, as well as by floods, since they are the natural phenomena that bring with them the most severe risks for communities. In this work, we articulated the statistical approach of the landslide conditioning factors, Machine Learning Algorithms (MLA), and Geographic Information System (GIS), evaluating a flexible and agile methodology to estimate the landslide susceptibility defining areas prone to the landslide occurrence. The MLA were validated in a case study in the “La Liboriana” River basin, located in the Municipality of Salgar in the Colombian mountains Andes where Landslide Susceptibility Maps (LSMs) were obtained. The obtained MLA results hold immense potential in the field of regional landslide mapping, facilitating the development of effective strategies aimed at minimizing the devastating impacts on human lives, infrastructure, and the natural environment. By leveraging these findings, proactive measures can be devised to safeguard vulnerable areas, mitigate risks, and ensure the safety and well-being of communities. Seven supervised MLA were employed, two regression algorithms (Logistic) and five decision tree algorithms (Recursive Partitioning and Regression Trees [RPART], Conditional Inference Trees [CTREE], Random Forest [RF], Ranger, and Extreme Gradient Boosting Algorithm [XGBoost]). The LSMs were produced for each MLA. Considering different performance metrics, the RF model yields the best classification accuracy with an area under receiver operating characteristic (ROC) curve of 95% and 90% of accuracy, providing the most representative results. Finally, the contribution of each landslide conditioning factor on predictions with RF model is explained using the SHAP method.

KEYWORDS: Colombian Andes, landslides, machine learning, SHAP, statistical methods, susceptibility

RECEIVED: October 12, 2022. **ACCEPTED:** July 3, 2023.

TYPE: Original Research Article

CORRESPONDING AUTHOR: Johnny Vega, Facultad de Ingeniería, Universidad de Medellín, Carrera 87 30 – 65, Medellín, 050026, Colombia.
Email: javega@udemedinellin.edu.co

Introduction

Landslides are a complex hazard that occur on highlands or mountain slopes, conditioned by several topographical aspects related to soil properties, geological structure, lithology and climatic conditions, slope morphology, soil cover, and water flow (Margottini et al., 2013). The landslide triggers are commonly rainfall (Y. Liu, Xu, et al., 2021), earthquakes (Pang et al., 2022), deforestation (García-Ruiz et al., 2017), and human activity that changes the effect of topography (Li et al., 2020).

Landslides cause serious impacts that threaten humans (Hakim et al., 2022; Panahi et al., 2020), and damages to natural resources. Thus, landslide hazard assessment has become a task of interest for decision-making by government entities and municipal and/or urban planning departments.

In general, in regions where urban developments, residential areas, and service infrastructure coincide with mountainous terrain, the risk tends to be high for the population and the economic costs may include relocation of communities, reconstruction of structures, and restoration of the quality of water sources. In many developing countries, where land occupation has generally been carried out without adequate planning and in a disorderly manner, the growth of urban areas occurs in landslide-prone zones.

Susceptibility mapping for landslide prediction is a GIS-based method involving correlation of previous landslides with possible driving factors to identify areas at risk of landslides (Hakim et al., 2022; Hakim & Lee, 2020). In recent years, studies of landslide susceptibility mapping have employed various probabilistic and statistical methods (A. Saha & Saha, 2020; Silalahi et al., 2019). Given the complexity of landslide prediction, many researchers have turned their attention to using hybrid ensemble approaches that combine machine learning methods with metaheuristic algorithms (Jaafari et al., 2019) or ensemble learning techniques (Bui et al., 2019; B. T. Pham et al., 2019).

Currently, landslide susceptibility maps allow the identification of areas prone to the occurrence of a mass removal event where the potential damage to people and infrastructure must be reduced or controlled, however, the accuracy of these approaches varies according to the quality of the data, the model approaches used, and the landslide inventories.

The purpose of this study is to articulate the statistical approach of the landslide conditioning factors, Machine Learning Algorithms (MLA) and GIS, evaluating a flexible and agile methodology to estimate the landslide susceptibility defining areas prone to the landslide occurrence incorporating



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

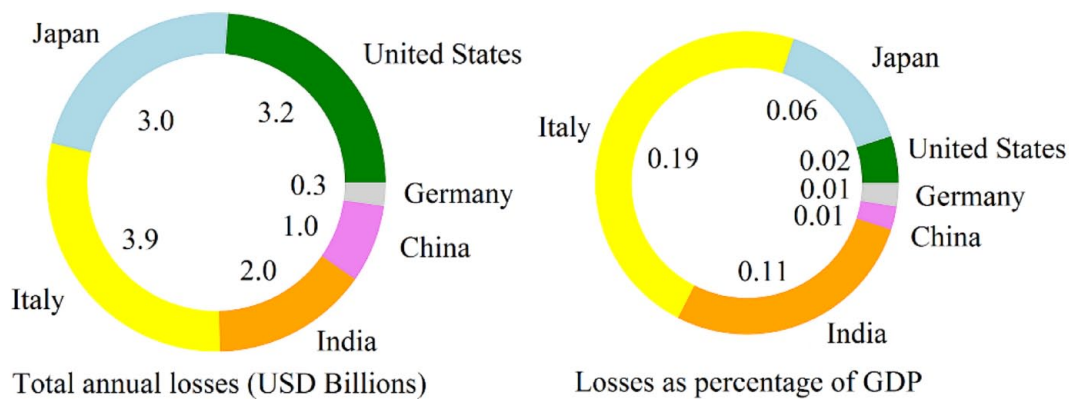


Figure 1. Impacts of landslides in terms of gross domestic product (GDP). Source: Hidalgo and Vega (2021).

the interpretability criteria by mean of SHAP values approach. The MLA were validated in a case study in the “La Liboriana” River basin, located in the Municipality of Salgar (Antioquia) in the Colombian mountains Andes where Landslide Susceptibility Maps (LSMs) were obtained. The results can be used for mapping regional landslides to develop strategies to minimize the loss of human lives, infrastructure, and natural environment.

Antecedents

Due to the high levels of landslides, a great dynamic has been generated worldwide in the study of the associated phenomena in an attempt to understand the physical and economic aspects related to mass movements (Hidalgo & Vega, 2021). The United Nations Office for Disaster Risk Reduction (CRED and UNDRR, 2021) reports that in 2020 the average annual economic losses were higher than those of the last two decades, which amount to US\$ 151.6 billion, in addition to an increase in the phenomena triggered by weather conditions.

In comparison to the previous two decades, 2020 was higher than the annual average in terms of number of recorded events and the annual average of economic losses, which is US\$ 151.6 billion. There were considerably fewer deaths compared to the annual average of 61,709 and fewer people directly affected compared to the annual average of 201.3 million people. However, in 2020 there were 26% more storms than the annual average of 102 events, 23% more floods than the annual average of 163 events, and 18% more flood deaths than the annual average of 5,233 deaths (CRED and UNDRR, 2021).

Landslides cause loss of human life every year. Laccase and Nadim (2009) report that at least 17% of all natural hazard deaths worldwide are caused by landslides. Human losses derived by landslides occur predominantly in developing countries. In contrast, developed countries such as the United States and Japan report few human losses, but high annual economic losses, estimated between 1 and 6 billion dollars (Ospina-Gutiérrez & Aristizábal-Giraldo, 2021). It is estimated that the direct and indirect costs of mass movements

can be significant in terms of gross domestic product (GDP), even in developed countries (Figure 1).

In the case of Colombia, in the period from 2006 to 2014, 21,594 emergencies due to natural events were reported in the country, an average of 2,399 events per year. Of these, 14,641 (67.8%) were concentrated in the period from 2011 to 2014. As a result of the events that occurred, 3,181 people were reported dead in Antioquia, with 586 deaths (414 of them due to floods and landslides).

According to data from the Information System of Mass Movements (SIMMA, in Spanish) of the Colombian Geological Service (SGC, in Spanish), in the period between 1900 and 2018 at least 30,730 landslides have occurred in Colombia, which have left a balance of 31,198 fatalities and economic losses of USD\$ 654 million (Ospina-Gutiérrez & Aristizábal-Giraldo, 2021). According to the DESINVENTAR database, 10,438 landslides have been recorded in Colombia between 1921 and 2020, leaving almost 7,313 dead and disastrous results for the country's economic system.

Figure 2 shows a summary of records of the DESINVENTAR database up to 2019 in the Department of Antioquia. According to the data reported, landslides occur more frequently between the months of May to July and September to November, which coincide with the hydrogeological characterization of the study area. In total, 1,566 events were reported, which show that the areas with the highest occurrence are located in the southwest (462 records) and east (358 records) zones excluding the Aburrá Valley.

Landslide Susceptibility Assessment (LSA)

Landslide susceptibility is the landslide probability of occurrence in a specific area, based on local terrain conditions by interaction of landslide conditioning factors (LCF). Usually, information about landslide magnitude is not available. LSA permits to identify areas potentially affected without considering the time lapse which a landslide might occur, or its magnitude. Commonly, LSA is based on statistical relationships between past landslides and LCF. Future landslides will occur under the conditions that led to past landslides.

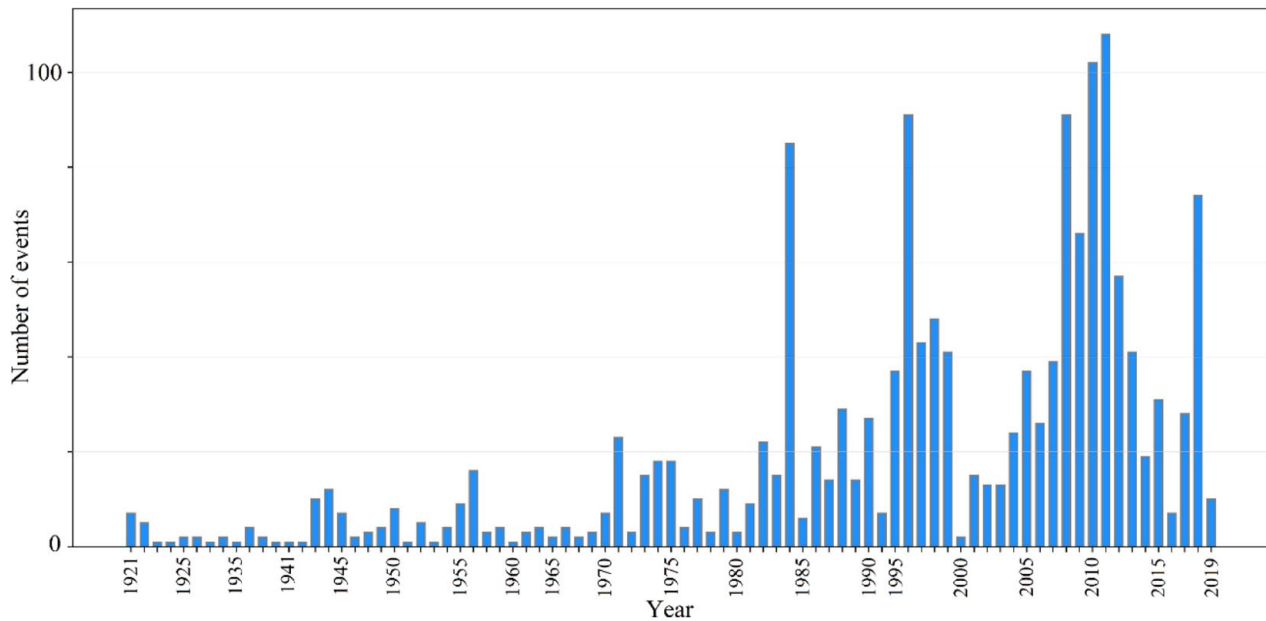


Figure 2. Landslides records for the Department of Antioquia.

In recent years, GIS and remote sensing data have been used to conduct many studies of disasters in mountain regions. Several researchers have built their methodology analyzing data of past landslides, and tested it through unknown landslides events. Different methodologies have been applied to spatially assess landslide susceptibility assessment (LSA). The main methods can be divided into qualitative (known as knowledge-driven or heuristic) or quantitative (data-driven and physically-based) methods (Lima et al., 2022). Their applicability and limitations can be found in literature. The first methods are based on the expert judgment, and usually involve qualitative terms to represent the susceptibility zoning. The quantitative methods establish numerical relationships between LCF and landslide occurrence (Marjanović et al., 2019).

The knowledge-driven or heuristic approach considers a direct mapping methodology establishing a direct relationship between the occurrence of landslides and the LCF using a landslide inventory at regional scale. This category may include the subjective geomorphological method, Analytic Hierarchy Process (AHP), Fuzzy Logic, Weighted Overlay, among others (Ali et al., 2021; Kaur et al., 2023; Q. B. Pham et al., 2021; Sahana & Sajjad, 2017; Sur & Singh, 2019).

The deterministic approach is based on slope stability methods. It is generally only applicable in relatively homogeneous terrain conditions throughout the study area and the types of landslides are known. It requires a high degree of simplification of the intrinsic variables to be used at local scale. This category includes geotechnical methods such as the Newmark's Method (Infinite Slope), Bishop's Method, Morgenstern Price Method, among others. Various studies have used deterministic models, as the Transient Rainfall Infiltration and Grid-Based Regional Slope-Stability (TRIGRS) model (Ma et al., 2021; Marin et al., 2021),

Shallow Slope Stability (SHALSTAB) model (Pradhan & Kim, 2015), Stability Index MAPPING (SINMAP) model (Michel et al., 2014), and Steady-State Infinite Slope Method (SSIS) (Si et al., 2020).

Finally, the data-driven or statistical approach is an indirect susceptibility methodology. It involves statistical analysis of the combinations of variables that led to landslide occurrence in the past. All possible intrinsic variables or LCF are entered and crossed into a GIS for analysis with a landslide inventory in a bivariate or multivariate way at regional scale (Dahal et al., 2012). This category also includes evaluation methods such as Frequency Ratio, Evidence of Weights, Linear and Logistic Regressions, among others. These statistical models have their own advantages and disadvantages and they have been widely used for LSA, but no agreement has reached to select the best method for landslide susceptibility analysis (Sahana & Sajjad, 2017; Zhang et al., 2020). Machine Learning Algorithms (MLA) are included in this category, but this is yet a topic of debate (Merghadi et al., 2020).

Machine Learning (ML) corresponds to a subset of a large discipline called artificial intelligence, which seeks to emulate human behavior through computer algorithms. ML uses statistical methods to train machines from data, that is, from experiences. Specifically on the topic of LSA using MLA, progress has been made since the early 2000s (Merghadi et al., 2020). MLA have been applied to solve geotechnical engineering problems widely in recent years, mainly because MLA are based on historical data, and they are more objective than the expert system methods. Moreover, they do not need more detailed mechanical parameters compared with the mechanical model methods (Y. Liu et al., 2019).

The MLA with the earliest development corresponds to Logistic Regression (LR), Artificial Neural Networks (ANN), Support Vector Machine (SVM), and Decision Trees (DT)

(Bragagnolo et al., 2020; Dou et al., 2019; Nhu et al., 2020; Sahin, 2020; Sun et al., 2020; Wang et al., 2015). Moreover, in recent years, development has focused on bagging methods, including the very popular Random Forest (RF), boosting methods such as AdaBoost and XGBoost (Bui et al., 2019; Chen et al., 2020; Huang et al., 2020; Q. B. Pham et al., 2021; Sahin, 2020).

MLA have been increasingly used in LSA as a result of the fact that they can learn the association between landslide occurrences and LCF without the requirements and assumptions for a statistical model. It has been observed that the predictive power of conventional statistical methods is relatively low. The conventional statistical methods cannot accurately analyze the complex interrelationships between different causative factors. Due to accuracy and very high predictive capability, MLA are getting more importance and attention in spatial analysis of landslides.

With the development of ensemble learning, bagging, and boosting methods are increasingly being used for classification and regression (Chen et al., 2020). In this study, seven supervised MLA were employed. Two regression algorithms (Logistic Regressions) and five decision tree algorithms: RPART, CTREE, RF, Ranger, XGBoost, briefly described below.

The binary logistic regression algorithm seeks to study the relationship between a dichotomous response variable (i.e., it takes only two possible outcomes presence/absence) and one or more explanatory variables, which can be of both qualitative and/or quantitative nature (Hosmer et al., 2013). Two logistic regression algorithms were used, the conventional stepwise selection method and the other where the Least Absolute Shrinkage Selector Operator (LASSO) method is used which imposes a penalty on the regression coefficients and selects variables (Tibshirani, 1996).

RPART is a supervised learning algorithm where a classification tree is obtained if the response variable is dichotomous (discrete) or a regression tree if the response variable is continuous. Initially this algorithm finds the explanatory variable that best divides the data into groups based on a rule. Then, for each of the partitions, the process is repeated. This process is done recursively until it is impossible to find a better partition. A relevant feature of this algorithm is that a variable used to separate the data is not used afterwards (Therneau & Atkinson, 1997). Since the resulting tree is very large and becomes tedious to interpret, pruning techniques are used to reduce its size (Strobl et al., 2009).

CTREE are a special type of decision trees where the choice of variables is made by assessing whether there is an association between the response variable and each of the explanatory variables. If the null hypothesis of independence is not rejected for the whole set of hypotheses, the recursive process is stopped. Otherwise, the level of association of each set of significant tests is quantified, allowing new splits in the tree to be

generated sequentially (Hothorn et al., 2006). One difference of this algorithm with the other decision tree algorithms is that no pruning is done for its statistical support.

Random Forest (RF) is an ensemble of decision trees, which are then combined into a single robust model (Breiman, 2001). RF uses a technique called bagging to build and train the ensemble of decision trees, allowing to reduce variance problems, prediction bias and overfitting when working with large amounts of data. One of the advantages of this classification algorithm is that it can handle many input variables and identify the most significant ones (dimensionality reduction) (Liaw & Wiener, 2002).

The Ranger algorithm is a fast implementation of the RF algorithm, particularly for handling high dimensional data (Wright & Ziegler, 2017). The XGBoost is a supervised MLA (Chen & Guestrin, 2016). The main idea of this algorithm is to generate multiple decision trees sequentially (boosting) where each one takes the results of the previous one and thus to generate an increasingly robust model with better predictive power. This process is repeated until the best possible model is obtained (Y. C. Chang et al., 2018).

Some of the aforementioned MLA have an internal structure that causes difficulty in explaining and interpreting their results, except those of a linear nature known as glass-box models. To address and solve the black-box issue of some MLA, as tree-based ensembles, kernel-based models, and neural networks, at both global and local levels, it is necessary to build eXplainable Artificial Intelligence (XAI) based solution. This approach provides an identification of the LCF influencing mainly effective classification of landslide susceptibility. Interpretable MLA can overcome the limitations of complex MLA in interpreting landslide susceptibility, and in fact, currently, few studies use the SHAP method to interpret the susceptibility of rainfall-induced shallow landslides (Zhou et al., 2022).

Methodology

Study area

Figure 3 shows the area corresponding to the case study in the “La Liboriana” River basin, located in the municipality of Salgar, in the southwest of the Department of Antioquia, western branch of the Colombian Andes. This basin joins El Barroso River basin, and both drain water into Cauca River, one of the most important rivers in the country (Hidalgo & Vega, 2021).

The study basin presents geomorphological, geological, and weather conditions that make it particularly susceptible to landslides and flash floods. The area with slope gradients exceeding 30° accounts for 67% of the total area. It has a humid tropical climate with a mean annual temperature of 22 °C. The rainfall regime is dominated by high interannual and intraseasonal variability with a mean annual rainfall of 3,073 mm; and monthly rainfall distributions show evident seasonal patterns

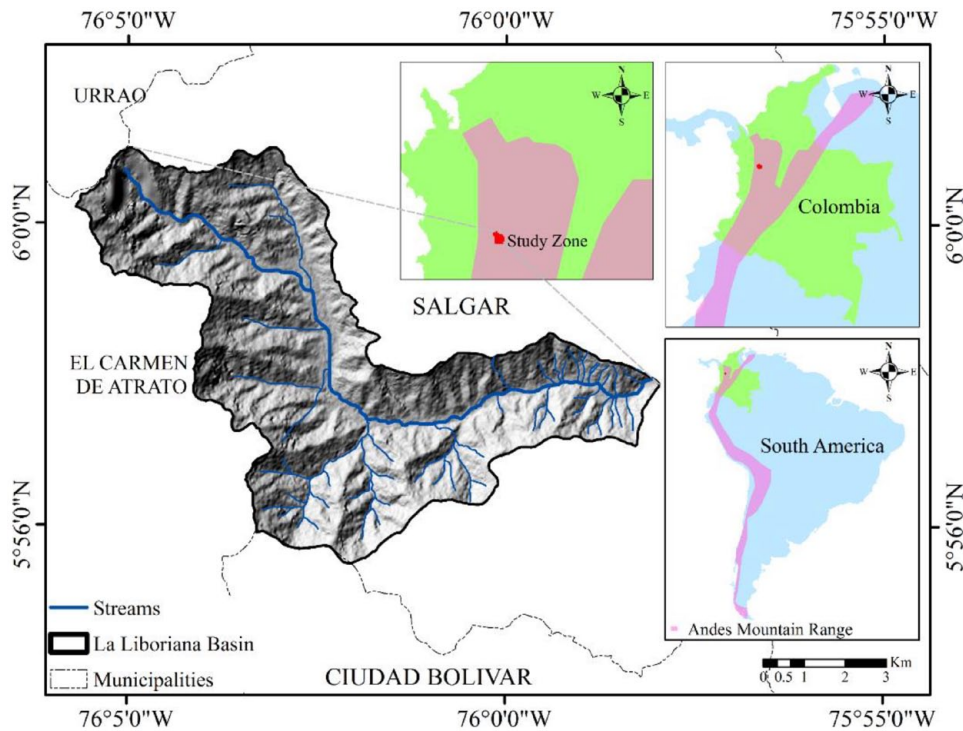


Figure 3. Location of the study area.

with two rainy seasons (peaks in May and October) (Ruiz Vásquez & Aristizábal, 2018).

The geomorphology of the basin exhibits in the upper part a mountain region with a rugged morphology and very steep forested hillslopes. In the middle and lower zones, grasslands (pastures) and coffee plantations have already substituted forest. The basin exhibits grazing areas and urban development near the riverbanks. Geologically, it is composed predominantly by a Cretaceous sedimentary rock formation and an intrusive Miocene body. These rocks have been severely weathered in situ under the humid tropical climate forming well-developed saprolite and residual soils (Ruiz Vásquez & Aristizábal, 2018).

LSA using machine learning algorithms

Seven supervised MLA were employed, two regression algorithms (Logistic) and five decision tree algorithms (Recursive Partitioning and Regression Trees [RPART], Conditional Inference Trees [CTREE], Random Forest [RF], Ranger, and Extreme Gradient Boosting Algorithm [XGBoost]). The LSMs were produced for each MLA. The LSA study was developed considering five phases: data processing, feature selection, dataset splitting and resampling, evaluation of the susceptibility, and comparison of the performance of the algorithms, as shown in Figure 4.

Data processing and landslide conditioning factors. Several Landslide Conditioning Factors (LCF) have been employed in literature to produce the landslide susceptibility maps. Slope,

aspect, lithology, plan curvature, and drainage density are most extensively used. Initially, data was processed in SAGA GIS 7.7 (<http://www.saga-gis.org>) software and then collected on 16 LCF which can be grouped into categories as terrain, geological, hydrological, and coverage factors. The factors considered in the study are shown in Table 1 and then briefly described. All LCF layers were stacked into a geographic database in raster format, using ArcGIS 10.8 version (<https://desktop.arcgis.com>).

Additionally, a summary of the descriptive statistics for the chosen LCF is shown in Table 2. For the qualitative LCF, the absolute frequency and percentage of each of the categories are presented while for the quantitative, six descriptive statistics are presented, minimum (Min), quartile 1 (Q1), median (Median), mean (Mean), quartile 3 (Q3), and maximum (Max).

Frequency Ratio (FR) which considers the effects of conditioning factors on landslide occurrence (Z. Chang et al., 2020) was determined. FR values were calculated as the ratio between the percentage of landslides and the percentage of class of each LCF. An FR value greater than 1.0, indicates a higher correlation between landslide and conditioning factors; whereas an FR value that is lower than 1.0, suggests a lower effect on landslide.

A landslide inventory documented in Ruiz Vásquez and Aristizábal (2018) for the study basin was used. This landslide inventory was obtained from a multi-temporal analysis of satellite images and aerial photographs, showing a landslide area covering approximately 0.6 km² corresponding to 1% of the basin.

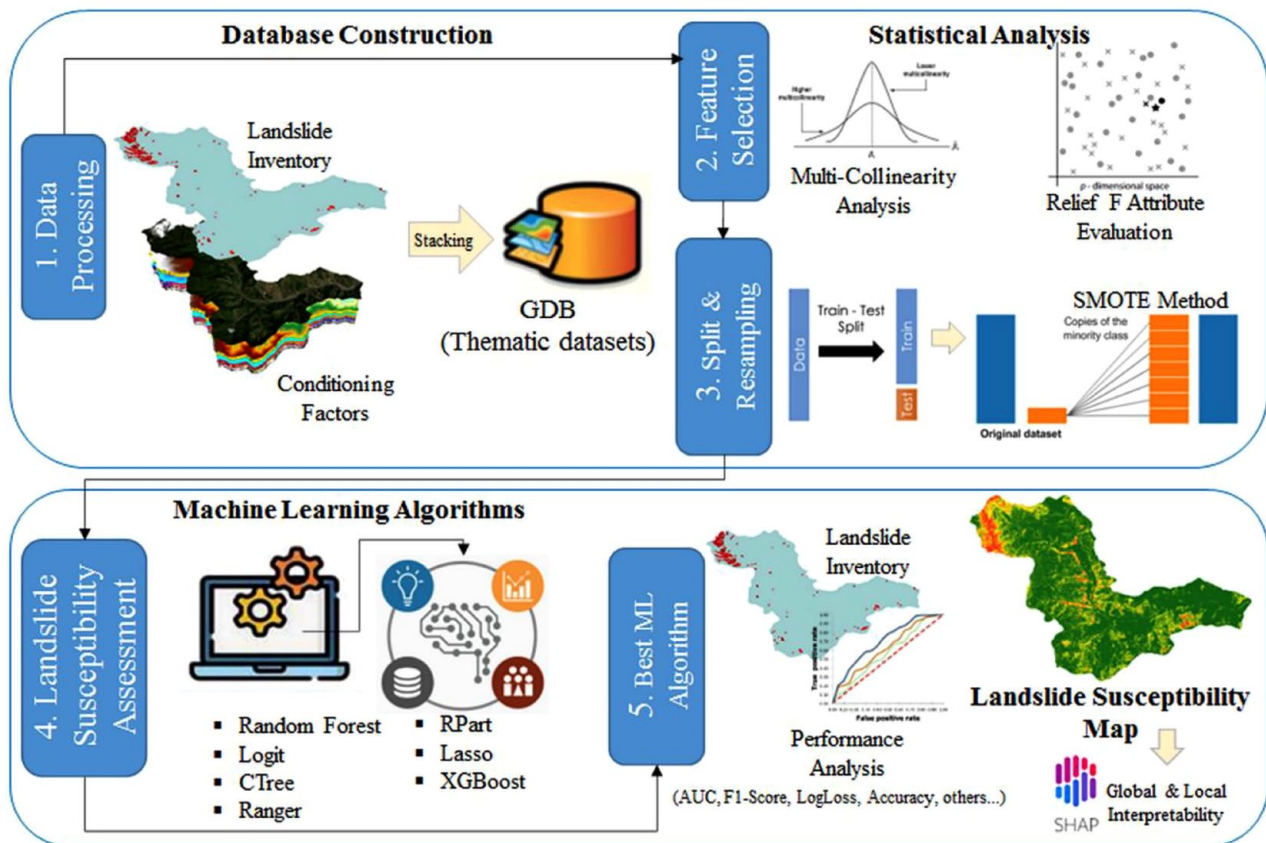


Figure 4. Adopted methodology for landslide susceptibility assessment (LSA).

Feature selection. Statistical methods of correlation and multicollinearity analysis were used. The variance inflation index (VIF) and the tolerance (TOL) were analyzed. VIF focuses on the standard error variations of LCF, which implies the lower the standard error, the lower the multicollinearity risk, which in turn means the increased likelihood that robust results should be obtained using this LCF (Merghadi et al., 2020). LCF with a VIF value less than 5 were used in modeling. LCF with high correlation or multicollinearity were excluded from the training phase of MLA, as they may generate noise in the modeling through erroneous system analysis.

For LCF selection according to relative importance, a Relief F Method provided in WEKA open source (<https://www.cs.waikato.ac.nz/ml/weka>) was used. This method calculates a weight value (average merit) for each variable to quantify its relevance (Dang et al., 2020). Features with the weights exceeding a certain threshold are selected for analysis. Factors assigned zero weight have no contribution to landslide occurrence and therefore, must be removed from further analysis (Hong et al., 2018). This study applied a cross-validation strategy, a standard random 5-fold cross-validation for Relief-F attribute selection method.

Split and resampling dataset. For the purposes of the training and validation stages of the MLA, a usual 70/30 split was

adopted (Achour & Pourghasemi, 2020; S. Saha et al., 2021). In order to balance the input dataset, over sampling techniques can be used to increase the minority class, that is, the landslide class, and for this purpose, the method SMOTE (Synthetic Minority Over sampling TEchnique) (Qing et al., 2020) was implemented under the R Studio environment (<http://www.rstudio.com>). This method is an oversampling technique that creates synthetic minority class data points to balance the dataset using a K-nearest neighbor algorithm.

Selection of the best performance MLA. An alternative method for evaluating and determining the statistical significance of systematic pairwise differences between the MLA is the Wilcoxon signed-rank test. This non-parametric test, as described by Dou et al. (2019) and Merghadi et al. (2020), was employed with a significance level of $\alpha = 5\%$. The test is based on a null hypothesis that assumes an equality between models only if the p-value is greater than 0.05. Rejecting the null hypothesis indicates statistically significant difference in performance between a pair of models, thereby establishing its reliability.

Once it has been identified that the models used showed statistically significant differences, seven statistical metrics were used to measure performance and validate the trained MLA. These metrics are widely used in the machine learning literature and are briefly described below. Five metrics were

Table 1. Dataset of Landslide Conditioning Factors.

DATA TYPE	CONDITIONING FACTOR	DATA SOURCE	CLASS	CLASS %	LANDSLIDE %	FREQUENCY RATIO		
1. Terrain	Elevation (m)	12.5m DTM obtained from ALOS PALSAR data (https://asf.alaska.edu)	0–1,300	0.55	0.70	1.27		
			1,300–1,500	6.05	2.32	0.38		
			1,500–1,800	20.51	9.37	0.46		
			1,800–2,000	19.15	17.57	0.92		
			2,000–2,500	37.02	26.63	0.72		
			2,500–3,000	12.50	20.90	1.67		
			3,000–3,500	3.72	22.37	6.01		
			>3,500	0.50	0.15	0.30		
			Slope (°)	Obtained from DTM	0–5	1.54	0.08	0.05
					5–15	9.54	3.48	0.36
15–30	40.65	34.44			0.85			
30–45	43.15	50.70			1.17			
>45	5.12	11.30			2.21			
Aspect	Obtained from DTM	Flat			0.18	0.00	0.00	
		North	15.59	16.02	1.03			
		Northeast	14.61	15.79	1.08			
		East	16.94	21.21	1.25			
		Southeast	16.61	9.44	0.57			
		South	13.23	9.60	0.73			
		Southwest	9.04	4.64	0.51			
		West	5.55	7.66	1.38			
		Northwest	8.25	15.63	1.89			
		Curvature	Obtained from DTM	<–5	1.15	1.16	1.01	
–5 to 0	56.86			54.72	0.96			
0–1	14.45			16.49	1.14			
>1	27.54			27.63	1.00			

(Continued)

Table 1. (Continued)

DATA TYPE	CONDITIONING FACTOR	DATA SOURCE	CLASS	CLASS %	LANDSLIDE %	FREQUENCY RATIO
Landforms	Obtained from DTM		Canyons, deeply incised streams	6.08	3.64	0.60
			Midslope drainages, shallow valleys	7.96	5.03	0.63
			Upland drainages, headwaters	0.57	0.77	1.35
			U-shape valleys	14.07	7.43	0.53
			Plains	0.29	0.00	0.00
			Open slopes	43.95	48.22	1.10
			Upper slopes, mesas	11.49	17.96	1.56
			Local ridges/hills in valleys	0.58	0.39	0.67
			Midslope ridges, small hills in plains	8.47	10.84	1.28
			Mt. Tops, high ridges	6.53	5.73	0.88
Terrain ruggedness index	Obtained from DTM		<3	20.65	9.83	0.48
			3–6	57.27	55.88	0.98
			6–9	19.17	26.24	1.37
			9–12	2.07	4.57	2.21
			>12	0.83	3.48	4.19
			0–2	65.67	69.86	1.06
2. Drainage	Drainage density (km/km ²)	Extracted stream network (Regional Autonomous Corporation of Central Antioquia)	2–4	15.85	17.04	1.08
			4–6	10.93	10.45	0.96
			6–8	7.55	2.65	0.35
			>8	9.61	3.86	0.40
3. Soil	Soil depth (m)	Calculated based on (Botero et al., 2015)	0.2–0.3	9.99	13.62	1.36
			0.3–0.5	3.44	5.11	1.49

(Continued)

Table 1. (Continued)

DATA TYPE	CONDITIONING FACTOR	DATA SOURCE	CLASS	CLASS %	LANDSLIDE %	FREQUENCY RATIO
			0.5–0.8	9.41	11.61	1.23
			0.8–1.0	9.7	11.84	1.22
			1.0–1.5	31.96	29.49	0.92
			1.5–2.0	23.32	21.05	0.90
			2.0–2.5	12.17	7.28	0.60
	Soil type	Generated from lithological units	Silty clay	94.36	98.45	1.04
			Silty sand	5.62	1.55	0.28
			Sandy silt	0.02	0.00	0.00
4. Geological	Lithology ^b	Digitized from 1:100,000 geology maps (Colombian Institute of Geology and Mining)	Kaa	92.01	85.45	0.93
			Qar	5.62	1.55	0.28
			Tcf	0.02	0.00	0.00
			Tdt	2.35	13.00	5.53
5. Coverage	Landcover	Extracted from 1:100,000 land cover map 2005–2009 (Colombian Institute of Hydrology, Meteorology and Environmental Studies)	Weedy pastures	9.85	3.33	0.34
			Continuous urban zone	0.56	0.00	0.00
			Mosaic of pastures and crops	7.18	1.16	0.16
			Clean pastures	9.36	14.09	1.51
			Coffee crop	3.48	6.19	1.78
			Mosaic of crops with natural spaces	8.28	3.56	0.43
			Fragmented forest with pastures and crops	0.09	0.00	0.00
			Secondary vegetation or in transition	27.59	39.63	1.44
			Gallery forest and riparian	1.29	1.32	1.02
			Pasture mosaic with natural spaces	6.08	1.93	0.32
			Crop mosaic	2.09	0.85	0.41
			Mosaic of crops, Pastures and natural spaces	2.37	0.00	0.00
			Dense forest under the mainland	4.76	5.73	1.20
			Dense highland forest	17.00	22.21	1.31

(Continued)

Table 1. (Continued)

DATA TYPE	CONDITIONING FACTOR	DATA SOURCE	CLASS	CLASS %	LANDSLIDE %	FREQUENCY RATIO
NDVI		Extracted from a Landsat 8 OLI sensor image (11/01/2015, path 9 row 56) (https://earthexplorer.usgs.gov)	<0.2	3.39	13.00	3.83
			0.2–0.3	5.88	24.85	4.23
			0.3–0.4	27.98	28.79	1.03
			0.4–0.5	50.09	30.88	0.62
			>0.5	12.66	2.48	0.20
6. Hydrological	Stream power index ($\times 10^6$)	Calculated using DTM	0–10	99.73	99.92	1.00
			100–30	0.16	0.08	0.50
			30–70	0.09	0.00	0.00
			>70	0.02	0.00	0.00
Sediment transport index ($\times 10^3$)		Calculated using DTM	0–1	86.90	86.15	0.99
			1–9	12.40	13.16	1.06
			9–18	0.49	0.54	1.10
			18–40	0.18	0.08	0.44
			>40	0.03	0.08	2.67
Topographic wetness index		Calculated using DTM	>10	83.60	63.08	0.75
			10–12	12.27	30.03	2.45
			12–15	3.26	6.11	1.87
			>15	0.88	0.77	0.88
Antecedent rainfall index (mm)		Calculated using IMERG rainfall data based on (Kirschbaum & Stanley, 2018)	>9	13.31	16.72	1.26
			9–13	14.40	26.86	1.87
			>13	72.28	56.42	0.78

bKaa: Interstratifications of the shales, siltstones, mudstones, sandstones, cherts, and polymictic conglomerates; Qar: Alluvial deposits; Tcf: Granodiorite batholith; Tdt: Pyroxene dioritic stock. [AQ:3]

Table 2. Statistical Description of Most Significant LCF.

RESPONSE VARIABLE	CLASSES	FREQUENCY (%)					VIF	TOL
Landslide occurrence	(0) No	374,682 (99.66%)						
	(1) Yes	1,292 (0.34%)						
LCF								
Landforms	(1) Canyons, deeply, incised streams						1.22	0.81
	(2) Midslope drainages, shallow valleys	29,951 (7.97%)						
	(3) Upland drainages, headwaters	2,143 (0.57%)						
	(4) U-shape valleys	52,952 (14.08%)						
	(5) Plains	1,104 (0.29%)						
	(6) Open slopes	165,337 (43.98%)						
	(7) Upper slopes, mesas	43,169 (11.48%)						
	(8) Local ridges/hills in valleys	2,200 (0.59%)						
	(9) Midslope ridges, small hills in plains	31,831 (8.47%)						
	(10) Mountain tops, high ridges	24,395 (6.49%)						
Soil type	(1) Silty clay	354,790 (94.37%)					2.41	0.41
	(2) Silty sand	21,184 (5.63%)						
Landcover	(1) Weedy pastures	37,082 (9.86%)					1.13	0.88
	(2) Continuous urban zone	2,124 (0.56%)						
	(3) Mosaic of pastures and crops	26,991 (7.18%)						
	(4) Clean pastures	35,199 (9.36%)						
	(5) Coffee crop	13,108 (3.49%)						
	(6) Mosaic of crops with natural spaces	31,115 (8.28%)						
	(7) Fragmented forest with pastures and crops	319 (0.08%)						
	(8) Secondary vegetation or in transition	103,777 (27.60%)						
	(9) Gallery forest and riparian	4,870 (1.30%)						
	(10) Pasture mosaic with natural spaces	22,890 (6.09%)						
	(11) Crop mosaic	7,848 (2.09%)						
	(12) Mosaic of crops, pastures, and natural spaces	8,929 (2.37%)						
	(13) Dense forest under the mainland	17,855 (4.75%)						
	(14) Dense highland forest	63,867 (16.99%)						
	MIN	Q1	Q2	MEAN	Q3	MAX		
TWI	5.85	8.91	9.83	10.17	10.89	33.60	1.21	0.82
NDVI	0.04	0.37	0.43	0.41	0.47	0.61	1.29	0.77
TRI	0.00	3.29	4.54	4.68	5.81	42.04	2.16	0.46
Curvature	-46.08	-1.28	0.00	-0.25	0.64	31.36	4.39	0.22
Elevation	1,238	1,776	2,036	2,093	2,351	3,724	1.44	0.69

Note. The value between parenthesis in the class column corresponds to coding of categorical variables.

calculated based on the binary confusion matrix (Agarwal, 2020; Ahamad et al., 2020). To do this, first the following measures were calculated, True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Then, the following performance metrics were calculated:

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (1)$$

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

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

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (4)$$

$$\text{F1-Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (5)$$

Another statistical metric used was Area Under Curve (AUC) of the Receiver Operating Characteristic (ROC). The AUC ranges from 0 to 1 and represents the predictive performance of the model (Fawcett, 2006). There are five categories to understand this value with respect to the level of accuracy: excellent (0.9–1.0), good (0.8–0.9), fair (0.7–0.8), poor (0.6–0.7), and fail (0.5–0.6) (Das & Lepcha, 2019; Rasyid et al., 2016). In addition, the probability-based log loss classification metric was also calculated and used to compare the performance of MLA.

These metrics were calculated for each of the fitted algorithms. The choice of the best algorithm to explain the occurrence of a possible landslide was carried out by comparing and counting the number of metrics in favor of each algorithm. The MLA used in this study were implemented using R software (R Core Team, 2022). They are all part of the “alookr” package. For description of this package see <https://CRAN.R-project.org/package=alookr>.

Interpretation of MLA output with SHAP method. In this work, the SHapley Additive exPlanations (SHAP) approach was adopted to understand the MLA output. This method is a game theoretic approach to explain the output of any MLA. It provides a unified framework to interpret predictions through calculating the Shapley values that provide the coherence of the explanations (Inan & Rahman, 2023; Kavzoglu et al., 2021). The SHAP method allows local interpretation using the Shapley value for each feature in a single sample to show the contribution of each feature to the predicted value. The SHAP method can be consistent by aggregating local explanations into global explanations and by separating single-factor effects from interaction effects (Zhou et al., 2022).

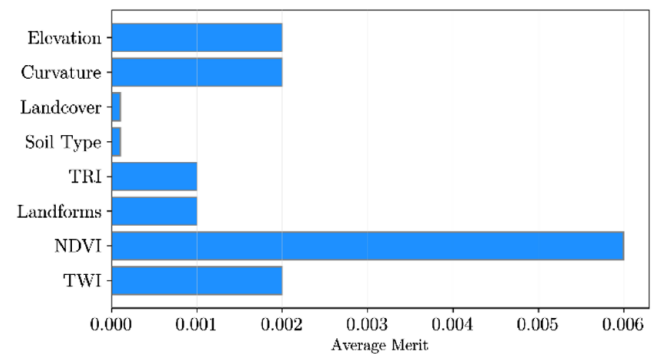


Figure 5. Average merit of landslide conditioning factors considered.

Results and Discussion

Statistical description of most significant LCF

A multicollinearity analysis was performed. Then, the Relief-F attribute selection method for most significance LCF was used. LCF with average merit (AM) values less than or equal to zero were not considered in the MLA training process since they were not statistically significant in explaining the response variable. Figure 5 shows the significant LCF and their respective AM, indicating their relative importance for explaining the landslide occurrence. These results were obtained using a standard random 5-fold cross-validation with Relief-F method. All these values are positive, and indicate that NDVI has the highest AM, followed by the factors Elevation, Curvature, and TWI. Landcover and Soil Type have the lowest AM values. Table 2 shows the values of the tolerance and variance inflation factor for multicollinearity diagnosis. All the conditioning factors previously selected by the Relief-F method have a tolerance greater than 0.2 and a VIF less than 5, indicating that there are no multicollinearity problems among the LCF.

Spatial distribution of most significant LCF and frequency ratio

Figure 6 shows the spatial distribution of the chosen LCF for modeling and its value ranges. In general terms it can be noticed that the spatial distribution of the elevation factor shows the altimetric variation of the study area, showing a definite pattern with a marked trend of increasing low values from the east to higher values in the west. In addition, it is observed that most of the landslides have occurred at higher elevations. In fact, it can be noticed that the greatest occurrence of landslides in the basin corresponds to the elevation range between 3,000 and 3,500 m.a.s.l., where there are landforms with gradients of slopes greater than 45°.

The curvature convex and straight (plan) present the greater frequency ratio (FR), due to this formation is favorable for landslide initiation according to used landslide inventory. High values of Terrain Ruggedness Index (TRI) were present in the landslide zone where predominantly silty clay soils exist formed from pyroxene dioritic stocks. The greater frequency ratio in

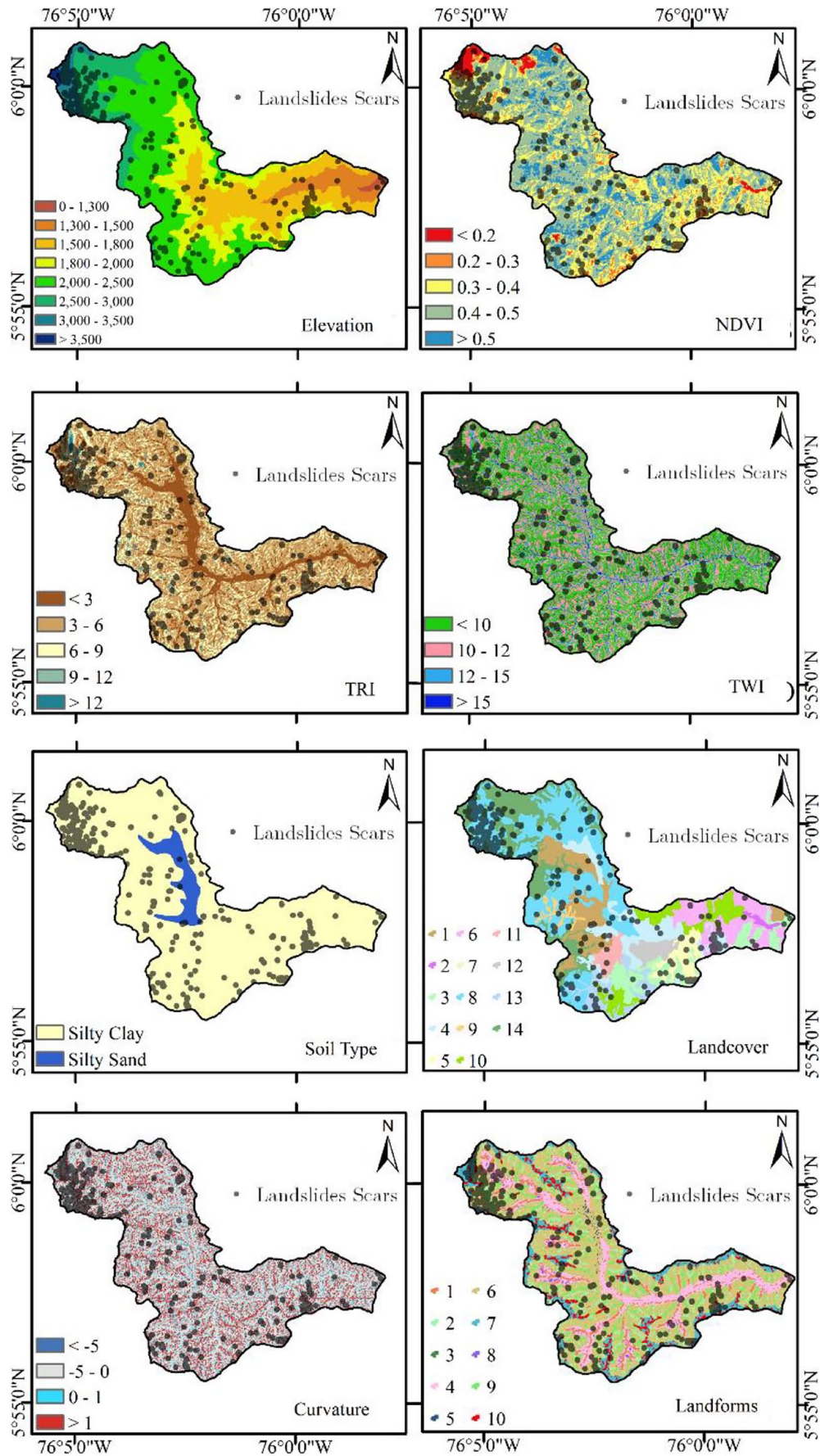


Figure 6. Spatial distribution of landslide conditioning factors used in modeling.

Table 3. Performances Indicators (Metrics) for MLA Considered.

PERFORMANCE INDICATOR	LOGISTIC	RPART	CTREE	RANDOM FOREST	RANGER	XGBOOST	LAGSSO
Accuracy	0.83	0.89	0.86	0.90	0.91	0.88	0.88
Precision	0.01	0.02	0.01	0.03	0.03	0.02	0.01
Recall	0.53	0.51	0.50	0.89	0.81	0.58	0.46
Sensitivity	0.53	0.51	0.50	0.89	0.81	0.58	0.46
Specificity	0.84	0.89	0.86	0.90	0.91	0.88	0.88
F1 Score	0.02	0.03	0.02	0.06	0.06	0.03	0.03
LogLoss	0.45	0.42	0.39	0.29	0.30	0.33	0.47
AUC	0.77	0.72	0.80	0.95	0.93	0.82	0.77

NDVI is presented in the range values of 0.2 to 0.3 indicating that the landslides occurred in areas where bare soil or non-healthy low vegetation were present. Moderate values of Topographic Wetness Index (TWI) were present in the area affected by landslides. Areas of coffee crops and secondary vegetation were commonly affected by the landslide history in the study area which was reflected in its FR values.

Performance of MLA

Since the original used dataset presents balancing problems in terms of the variable landslide occurrence (0.34% landslide records, 99.66% non-landslide records), before applying the seven MLA to the training dataset, the SMOTE resampling method was applied to them, creating synthetic samples of the minority class using a k-nearest neighbor algorithm.

On the other hand, according to the results of the Wilcoxon test, the null hypothesis is rejected, concluding the existence of statistically significant differences between all the MLA pairs considered. Once it has been identified, different metrics for each algorithm used to measure their performance were calculated using the test dataset. The validation metrics used, and their results are shown in Table 3. The values of the validation metric highlighted in bold in Table 3, identify the best performing algorithm on that metric. According to these results the best performing algorithm for predicting landslide susceptibility in the study area considering the adopted LCF is the RF algorithm (AUC=0.95).

The AUC values for each model are shown in Table 3. This assessment method was used previously by several authors (Hong et al., 2019; Ozer et al., 2020; Rahali, 2019) to check the model performance, and in this work, the ROC curves were also used for the same purpose. The RF method yields the best classification accuracy, and RPART yields the worst performance (AUC=0.72). All MLA employed in the present study provide acceptable results.

Regarding the results of the other models (Figure 7), the Logit (Figure 7d) and Lasso (Figure 7e) models, which are based on regressions, present similarities in terms of their landslide

prediction capacity, showing differences of the order of 10% in all the validation metrics. Regarding the models based on decision trees, XGBoost (Figure 7a), Ranger (Figure 7c), and CTREE (Figure 7f), the spatial distribution of the predicted values of landslide susceptibility do not differ much and show similarities in terms of the spatial patterns of each landslide susceptibility class, except for the RPART model (Figure 7b), in which there is a clear predominance of the “Low” landslide susceptibility class in almost the entire basin that differs spatially with the other models, even in terms of predictive ability using the AUC indicator (Table 3). Compared to the other models, RPART model performed very poorly, indicating that this model is unsuitable for landslide susceptibility mapping in the study area.

According to Huang and Zhao (2018), two main steps should be followed in order to create a landslide susceptibility map: Generate the landslide susceptibility values, and then reclassify them. The method used to reclassify these values depends on the histogram values (Natural breaks, equal intervals, standard deviation, among others). The landslides maps were obtained using the results of the models, which were categorized in five classes: very low, low, moderate, high, and very high classes, using natural break values distribution.

Regarding other studies in the same area with other kinds of methodologies and LCF used, it can be mentioned that Ruiz Vásquez and Aristizábal (2018) obtained an AUC=0.69 using a multivariate statistical logistic regression analysis. Marin et al. (2021) obtained a performance of the models close to AUC=0.80 for the total basin, and an AUC=0.56 for the upper part of the basin using deterministic models (TRIGRS model). Hidalgo and Vega (2021) obtained an AUC=0.56 using the EPADYM model to calculate the failure probability and factor of safety under seismic and static conditions. Finally, Vega and Hidalgo, (2023) obtained AUC values of 0.95, 0.86 and 0.60 using SVM model, fuzzy gamma model and TRIGRS model respectively, for the landslide-event hazard mapping using the records of May 18, 2015. In other regions with similar topographic conditions with intense rainfall have been carried out studies where the efficiency of MLA was better than the deterministic models, as the case reported in Z. Liu, Gilbert,

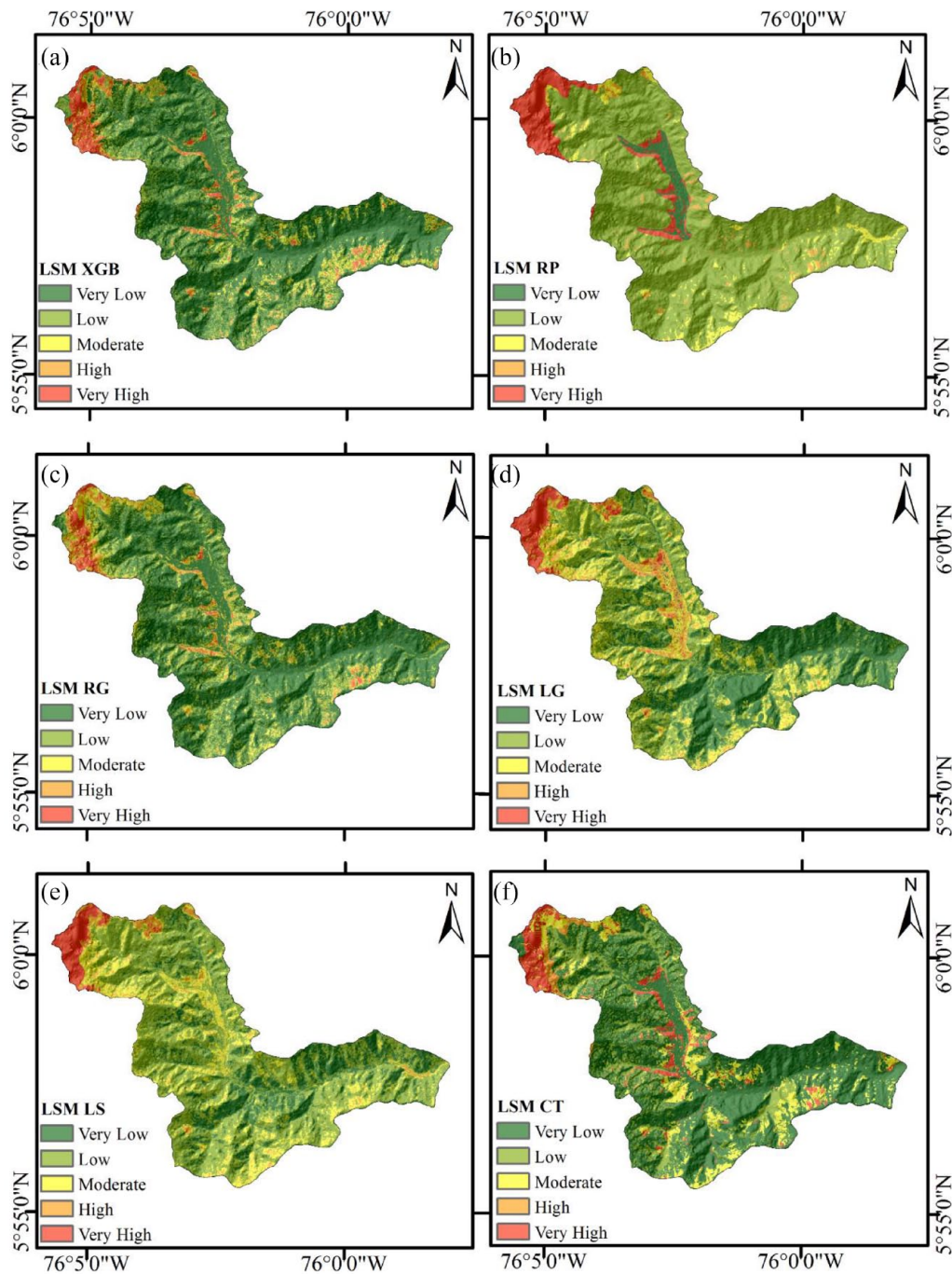


Figure 7. Landslide susceptibility maps using: (a) XGBoost, (b) RPART, (c) Ranger, (d) Logit, (e) Lasso, and (f) CTREE.

et al. (2021), where the accuracy was about 82% against accuracies above 90% using RF algorithm.

RF algorithm has better accuracy in prediction compared to linear models. But it cannot be interpreted, so it is often considered as a black-box model. In this work, an interpretable algorithm, SHAP method, is explored for the interpretation of LSA models and the determination of predominant LCF. Figure 8a shows the LSM of the study basin obtained using the RF model. The global prediction results are reasonable to the statistical analysis of different landslide susceptibility classes.

The very low susceptibility class covers about more than half of the study area (62.4%; 36.7km²) and low class 23.2%; (13.6km²). The areas with moderate and high susceptibility classes cover 7.6% (4.5 km²) and 4.3% (2.5km²) of the study area respectively. Finally, very high landslide susceptibility class covers 2.4% (1.4km²) of the basin, almost doubling the area of landslide registered in the basin according to landslide inventory used.

Finally, the distribution of landslides in the different susceptibility classes was statistically considered and an analysis of

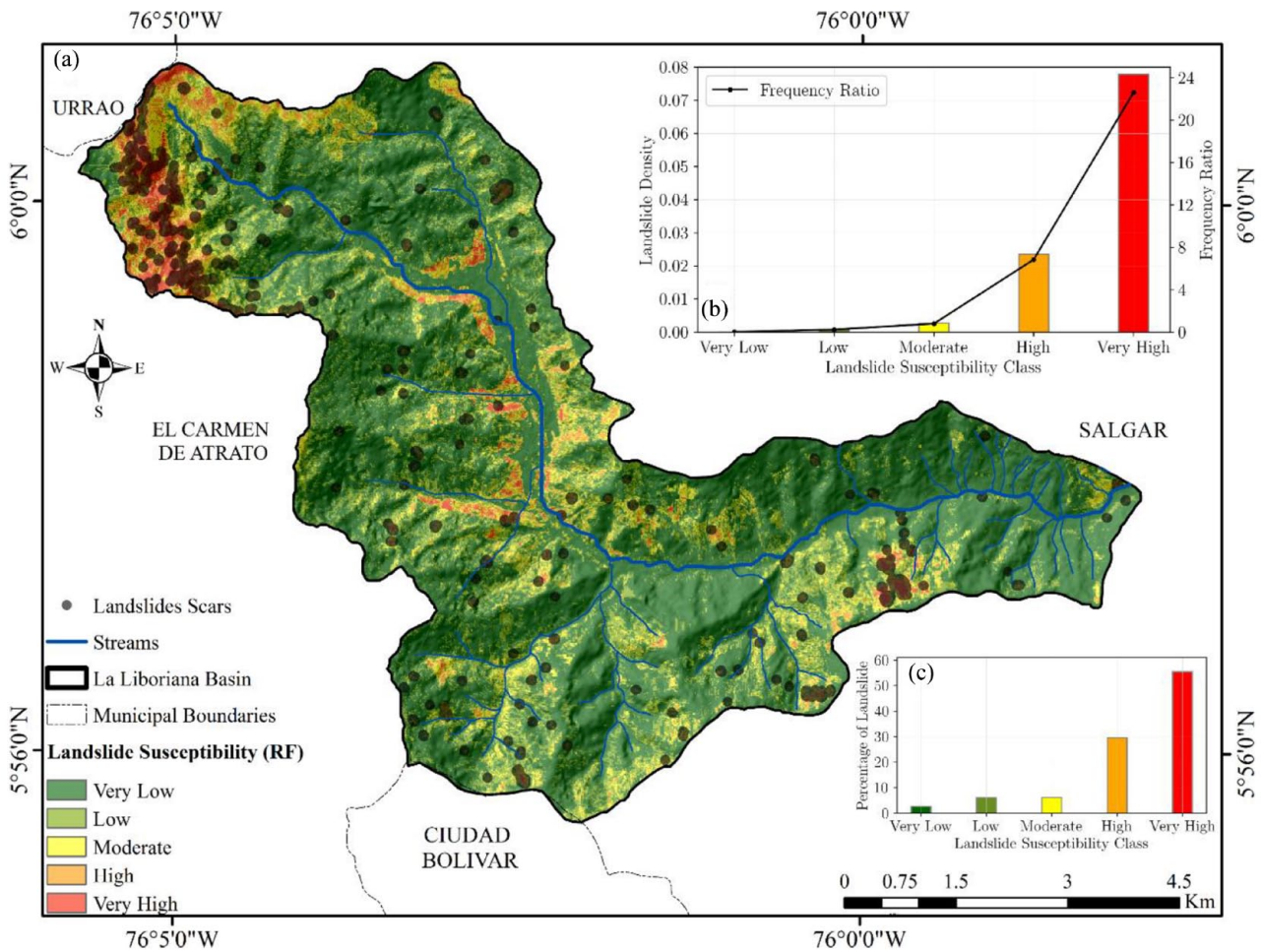


Figure 8. Landslide susceptibility map using random forest (RF).

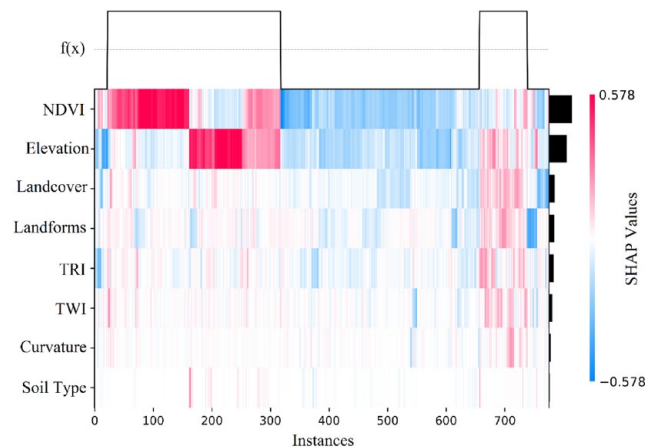


Figure 9. SHAP values heatmap for RF model output.

landslide density in the five susceptibility classes was performed. In order to assess whether the generated landslide susceptibility map meets the requirements, two principles are used: First, an ideal landslide susceptibility map has landslide density values that increase from lower to higher susceptibility class (Pradhan & Kim, 2016). Finally, the high-risk areas should account for a small percentage of the total area (Huang & Zhao, 2018). Figure 8b shows the frequency ratio and landslide densities for all classes. Indeed, the very high class has the

highest landslide density (LD=0.08) and landslide frequency ratio (FR=22.6). Additionally, according to Figure 8c, more than 85% of the events recorded in the inventory occur in this class, according to the values determined with the RF model.

In order to provide global interpretability to the output of RF model, SHAP method was conducted. According to the SHAP values obtained for the test dataset, Elevation and NDVI are the most significant factors for predicting the occurrence of landslides based on SHAP values heatmap (Figure 9).

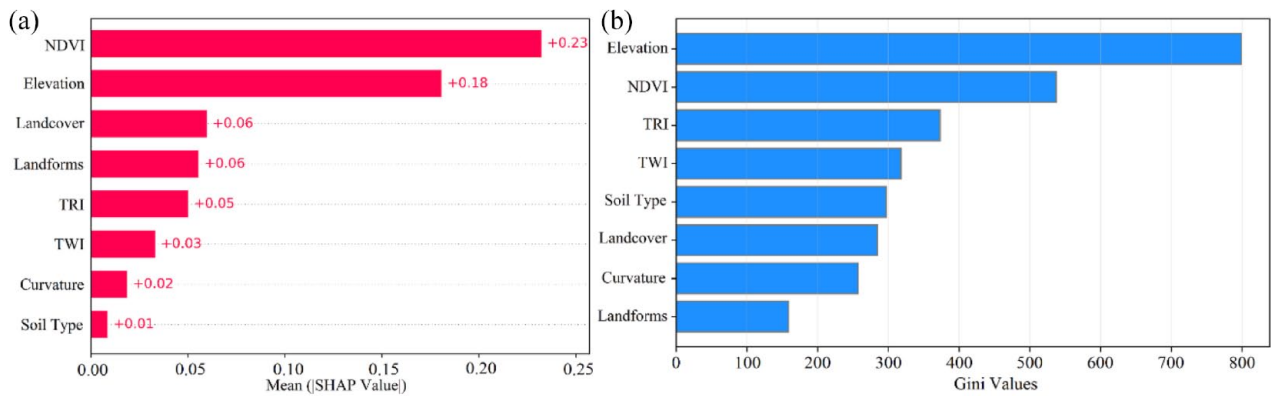


Figure 10. Contribution of LCF in RF model output: (a) mean SHAP values and (b) Gini values.

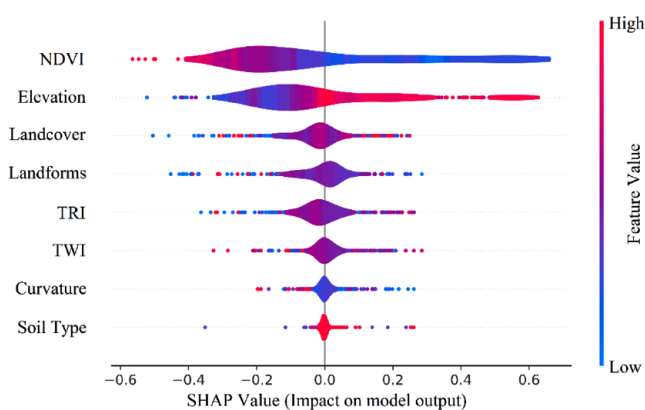


Figure 11. SHAP values plot for RF model output.

In fact, almost 40% of the dataset has high SHAP values (red color) in both conditioning factors, associated to landslide occurrences while approximately in a 50%, the combination of low SHAP values (blue color) in all LCF is associated to non-occurrence of landslide. The remaining 10%, corresponds to a combination of medium SHAP values of all LCF which lead to landslide occurrence. The obtained SHAP values are consequents with the results of the mean decrease Gini values calculated previously (Figure 10), where both Elevation and NDVI has the higher importance.

Regarding to the magnitude of each LCF and its influence in RF model output, in Figure 11, it can be noticed that the higher impact on RF model predictions is related to high values of Elevation and low values of NDVI, which is very consistent with the analysis carried out in section 5.2 about spatial distribution of most significant LCF and frequency ratio. High Elevation values in the basin are related to steepest zones prone to landslide occurrence. Low NDVI values are related to a bare soil and poor vegetation areas, very susceptible to infiltration process with the consequent loss of slope stability. The remaining LCF do not present a remarkable impact on RF model predictions.

On the other hand, and as illustration of an example of local interpretability, in Figure 12, it is shown the SHAP values waterfall of a landslide scar pixel (cyan color) in the upper part

of study basin, which was correctly classified as landslide with RF model. Again, in this case Elevation and NDVI have the most effect on prediction, whose values are in the conditioning factor class with higher frequency ratio (Table 2). Additionally, the values of Soil Type and Landform in this pixel, correspond to silty clay soils and open steep slopes prone to occurrence of landslides. The remain LCF have no major relevance on RF model output.

Conclusions

Landslides are one of the most widespread and complex natural geodynamics phenomena. In frequently affected regions by landslide, such as tropical mountainous regions, an increase in the number of landslides studies have been shown with the impulse of researchers and regional and local planners. Nowadays, MLA have been applied for landslide susceptibility mapping, making valuable progress. Model algorithm accuracy and diversification have improved susceptibility mapping, particularly with the rapid growth of computer technology and the popularization of GIS techniques.

Relief-F method enabled the detection of potential LCF that could adversely impact the MLA performance, as well as the understanding of LCF that could contribute to the model performance. The LCF representing the terrain characteristics of the study basin were considered very important. Conversely, geologic factors were found to be the least effective factor.

In this study, the landslide susceptibility maps were produced by applying seven different MLA (RF, XGBoost, RPART, Ranger, Logit, Lasso, CTREE). According to some performance metrics, all MLA showed good performance. However, the RF method yields the best classification accuracy with an AUC= 0.95, RPART yields the worst performance with an AUC = 0.72. Based on surface comparison analysis, the most representative results were provided by RF and consequently, RF is the most appropriate approach for landslide susceptibility assessment.

SHAP method provide both global and local interpretability to MLA. According to this method, most influent LCF on landslide susceptibility probabilities were determined. Results

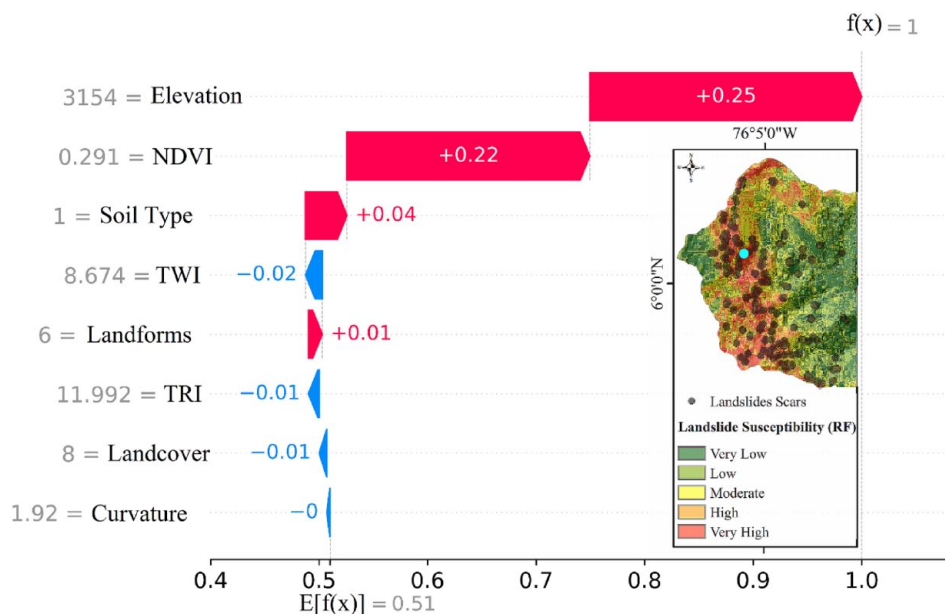


Figure 12. SHAP values waterfall for the pixel N° 8669 classified as landslide for RF model.

revealed that the Elevation and NDVI had the highest positive contribution, while Soil Type and Curvature had no major implications. In addition, it is worth mentioning that terrain and coverage factors had a significant contribution to the landslide occurrences in the study basin, considering that four out of the eight LCF modeled, were the most influential factors on landslide phenomena (Elevation, NDVI, Landcover, and Landforms).

Landslide susceptibility analysis is vital for identifying zones of future landslide occurrence and the proper estimation of landslide-induced risk. In this sense, MLA are efficient and can help to predict disaster risk and decrease disaster costs. The results of a landslide susceptibility analysis using MLA hold immense potential in the field of regional landslide mapping, facilitating the development of effective strategies aimed at minimizing the devastating impacts on human lives, infrastructure, and the natural environment. By leveraging these findings, proactive measures can be devised to safeguard vulnerable areas, mitigate risks, and ensure the safety and well-being of communities. Nevertheless, to fulfill this purpose, interpretation, and explanation of the results of such MLA is required for its correct implementation. Therefore, methods such as SHAP values are used to address this issue in decision making processes based on learning algorithms, providing more abundant and relevant information for landslide risk management.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Sincere thanks to “High Level Training Program for

Full-time Professors in their own Doctorates” of the Academic and Research Vice Rector's Offices of the University of Medellín and National Doctorate Program for Teachers of Higher Education Institutions of the Ministry of Science, Technology and Innovation. Data was provided by Research Program “Vulnerability, resilience and risk of communities and supplying basins affected by landslides and avalanches,” code 1118-852-71251, project “Functions for vulnerability assessment due to water shortages by landslides and avalanches: micro-basins of southwest Antioquia,” contract 80740-492-2020 held between Fiduprevisora and the Universidad de Medellín, with resources from the National Financing Fund for science, technology, and innovation, “Francisco José de Caldas.”

REFERENCES

- Achour, Y., & Pourghasemi, H. R. (2020). How do machine learning techniques help in increasing accuracy of landslide susceptibility maps? *Geoscience Frontiers*, 11, 871–883. <https://doi.org/10.1016/j.gsf.2019.10.001>
- Agarwal, R. (2020). *The 5 classification evaluation metrics every data scientist must know* [www Document]. Towards Data Scientist.
- Ahamad, M. M., Aktar, S., Rashed-Al-Mahfuz, M., Uddin, S., Liò, P., Xu, H., Summers, M. A., Quinn, J. M. W., & Moni, M. A. (2020). A machine learning model to identify early-stage symptoms of SARS-Cov-2 infected patients. *Expert Systems with Applications*, 160, 113661. <https://doi.org/10.1016/j.eswa.2020.113661>
- Ali, S. A., Parvin, F., Vojteková, J., Costache, R., Linh, N. T. T., Pham, Q. B., Vojtek, M., Gigović, L., Ahmad, A., & Ghorbani, M. A. (2021). GIS-based landslide susceptibility modeling: A comparison between fuzzy multi-criteria and machine learning algorithms. *Geoscience Frontiers*, 12, 857–876. <https://doi.org/10.1016/j.gsf.2020.09.004>
- Botero, E. M., Azevedo, G. F., Souza, H. E. M. C., De Souza, N. M., & Aristizabal, E. F. G. (2015). Estimativa da profundidade do solo pelo uso de técnicas de geoprocessamento, estudo de caso: Setor Pajarito, Colômbia. In *Anais XVII Simpósio Brasileiro Sensoriamento Remoto - SBR, João Pessoa-PB, Brasil, 25 a 29 abril 2015*, INPE 4551–4558.
- Bragagnolo, L., Silva, R. V. D., & Grzybowski, J. M. V. (2020). Artificial neural network ensembles applied to the mapping of landslide susceptibility. *Catena*, 184, 104240. <https://doi.org/10.1016/j.catena.2019.104240>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>

- Bui, D. T., Shirzadi, A., Shahabi, H., Geertsema, M., Omidvar, E., Clague, J. J., Pham, B. T., Dou, J., Asl, D. T., Ahmad, B. B., & Lee, S. (2019). New ensemble models for shallow landslide susceptibility modeling in a semi-arid watershed. *Forests*, 10(9), 743. <https://doi.org/10.3390/f10090743>
- Centre for Research on the Epidemiology of Disasters (CRED) and United Nations Office for Disaster Risk Reduction (UNDRR). (2021). Disaster year in review 2020: Global trends and perspectives. *Cred Crunch*, 62. <https://www.prevention-web.net/quick/52005>
- Chang, Y. C., Chang, K. H., & Wu, G. J. (2018). Application of eXtreme gradient boosting trees in the construction of credit risk assessment models for financial institutions. *Applied Soft Computing*, 73, 914–920. <https://doi.org/10.1016/j.asoc.2018.09.029>
- Chang, Z., Du, Z., Zhang, F., Huang, F., Chen, J., Li, W., & Guo, Z. (2020). Landslide susceptibility prediction based on remote sensing images and GIS: Comparisons of supervised and unsupervised machine learning models. *Remote Sensing*, 12, 502. <https://doi.org/10.3390/rs12030502>
- Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system* [Conference session]. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 13–17 August 2016, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Chen, T., Zhu, L., Niu, R. Q., Trinder, C. J., Peng, L., & Lei, T. (2020). Mapping landslide susceptibility at the Three Gorges Reservoir, China, using gradient boosting decision tree, random forest and information value models. *Journal of Mountain Science*, 17, 670–685. <https://doi.org/10.1007/s11629-019-5839-3>
- Dahal, R. K., Hasegawa, S., Bhandary, N. P., Poudel, P. P., Nonomura, A., & Yatabe, R. (2012). A replication of landslide hazard mapping at catchment scale. *Geomatics, Natural Hazards and Risk*, 3, 161–192. <https://doi.org/10.1080/19475705.2011.629007>
- Dang, V. H., Hoang, N. D., Nguyen, L. M. D., Bui, D. T., & Samui, P. (2020). A novel GIS-Based random forest machine algorithm for the spatial prediction of shallow landslide susceptibility. *Forests*, 11, 118. <https://doi.org/10.3390/f11010118>
- Das, G., & Lepcha, K. (2019). Application of logistic regression (LR) and frequency ratio (FR) models for landslide susceptibility mapping in Relli Khola river basin of Darjeeling Himalaya, India. *SN Applied Sciences*, 1, 1453. <https://doi.org/10.1007/s42452-019-1499-8>
- Dou, J., Yunus, A. P., Tien Bui, D., Merghadi, A., Sahana, M., Zhu, Z., Chen, C. W., Khosravi, K., Yang, Y., & Pham, B. T. (2019). Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. *Science of The Total Environment*, 662, 332–346. <https://doi.org/10.1016/j.scitotenv.2019.01.221>
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27, 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- García-Ruiz, J. M., Beguería, S., Arnáez, J., Sanjuán, Y., Lana-Renault, N., Gómez-Villar, A., Álvarez-Martínez, J., & Coba-Pérez, P. (2017). Deforestation induces shallow landsliding in the montane and subalpine belts of the Urbión Mountains, Iberian Range, Northern Spain. *Geomorphology*, 296, 31–44. <https://doi.org/10.1016/j.geomorph.2017.08.016>
- Hakim, W. L., & Lee, C. W. (2020). A review on remote sensing and GIS applications to monitor natural disasters in Indonesia. *Korean Journal of Remote Sensing*, 36, 1303–1322. <https://doi.org/10.7780/kjrs.2020.36.6.1.3>
- Hakim, W. L., Rezaie, F., Nur, A. S., Panahi, M., Khosravi, K., Lee, C. W., & Lee, S. (2022). Convolutional neural network (CNN) with metaheuristic optimization algorithms for landslide susceptibility mapping in Icheon, South Korea. *Journal of Environmental Management*, 305, 114367. <https://doi.org/10.1016/j.jenvman.2021.114367>
- Hidalgo, C. A., & Vega, J. A. (2021). Probabilistic landslide risk assessment in water supply basins: La Liboriana River Basin (Salgar-Colombia). *Natural Hazards*, 109, 273–301. <https://doi.org/10.1007/s11069-021-04836-0>
- Hong, H., Liu, J., Bui, D. T., Pradhan, B., Acharya, T. D., Pham, B. T., Zhu, A. X., Chen, W., & Ahmad, B. B. (2018). Landslide susceptibility mapping using J48 Decision Tree with AdaBoost, Bagging and Rotation Forest ensembles in the Guangchang area (China). *Catena*, 163, 399–413. <https://doi.org/10.1016/j.catena.2018.01.005>
- Hong, H., Shahabi, H., Shirzadi, A., Chen, W., Chapi, K., Ahmad, B. B., Roodposhti, M. S., Yari Hesar, A., Tian, Y., & Tien Bui, D. (2019). Landslide susceptibility assessment at the Wuning area, China: A comparison between multi-criteria decision making, bivariate statistical and machine learning methods. *Natural Hazards*, 96, 173–212. <https://doi.org/10.1007/s11069-018-3536-0>
- Hosmer, D. W. Jr., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). John Wiley and Sons. <https://doi.org/10.1002/9781118548387>
- Hothorn, T., Hornik, K., Van De Wiel, M. A., & Zeileis, A. (2006). A lego system for conditional inference. *The American Statistician*, 60, 257–263. <https://doi.org/10.1198/000313006X118430>
- Huang, J., Hales, T. C., Huang, R., Ju, N., Li, Q., & Huang, Y. (2020). A hybrid machine-learning model to estimate potential debris-flow volumes. *Geomorphology*, 367, 107333. <https://doi.org/10.1016/j.geomorph.2020.107333>
- Huang, Y., & Zhao, L. (2018). Review on landslide susceptibility mapping using support vector machines. *Catena*, 165, 520–529. <https://doi.org/10.1016/j.catena.2018.03.003>
- Inan, M. S. K., & Rahman, I. (2023). Explainable AI integrated feature selection for landslide susceptibility mapping using TreeSHAP. *SN Computer Science*, 4, 482. <https://doi.org/10.1007/s42979-023-01960-5>
- Jaafari, A., Panahi, M., Pham, B. T., Shahabi, H., Bui, D. T., Rezaie, F., & Lee, S. (2019). Meta optimization of an adaptive neuro-fuzzy inference system with grey wolf optimizer and biogeography-based optimization algorithms for spatial prediction of landslide susceptibility. *Catena*, 175, 430–445. <https://doi.org/10.1016/j.catena.2018.12.033>
- Kaur, H., Gupta, S., Parkash, S., & Thapa, R. (2023). Knowledge-driven method: a tool for landslide susceptibility zonation (LSZ). *Geology, Ecology, and Landscapes*, 7(1), 1–15. <https://doi.org/10.1080/24749508.2018.1558024>
- Kavzoglu, T., Teke, A., & Yilmaz, E. O. (2021). Shared blocks-based ensemble deep learning for shallow landslide susceptibility mapping. *Remote Sensing*, 13, 4776. <https://doi.org/10.3390/rs13234776>
- Kirschbaum, D., & Stanley, T. (2018). Satellite-based assessment of rainfall-triggered landslide hazard for situational awareness. *Earth's Future*, 6, 505–523. <https://doi.org/10.1002/2017EF000715>
- Laccase, S., & Nadim, F. (2009). Landslide risk assessment and mitigation strategy. In K. Sassa & P. Canuti (Eds.), *Landslide disaster risk reduction* (pp. 31–61). Springer. https://doi.org/10.1007/978-3-540-69970-5_3
- Li, Y., Wang, X., & Mao, H. (2020). Influence of human activity on landslide susceptibility development in the Three Gorges area. *Natural Hazards*, 104, 2115–2151. <https://doi.org/10.1007/s11069-020-04264-6>
- Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, 2, 18–22.
- Lima, P., Steger, S., & Glade, T. (2022). Literature review and bibliometric analysis on data-driven assessment of landslide susceptibility landslide susceptibility. *Journal of Mountain Science*, 19, 1670–1698. <https://doi.org/10.1007/s11629-021-7254-9>
- Liu, Y., Zhang, J. J., Zhu, C. H., Xiang, B., & Wang, D. (2019). Fuzzy-support vector machine geotechnical risk analysis method based on Bayesian network. *Journal of Mountain Science*, 16, 1975–1985. <https://doi.org/10.1007/s11629-018-5358-7>
- Liu, Y., Xu, P., Cao, C., Shan, B., Zhu, K., Ma, Q., Zhang, Z., & Yin, H. (2021). A comparative evaluation of machine learning algorithms and an improved optimal model for landslide susceptibility: A case study. *Geomatics, Natural Hazards and Risk*, 12, 1973–2001. <https://doi.org/10.1080/19475705.2021.1955018>
- Liu, Z., Gilbert, G., Cepeda, J. M., Lysdahl, A. O. K., Piciullo, L., Hefre, H., & Lacasse, S. (2021). Modelling of shallow landslides with machine learning algorithms. *Geoscience Frontiers*, 12, 385–393. <https://doi.org/10.1016/j.gsf.2020.04.014>
- Ma, S., Shao, X., Xu, C., He, X., & Zhang, P. (2021). MAT.TRIGRS (V1.0): A new open-source tool for predicting spatiotemporal distribution of rainfall-induced landslides. *Natural Hazards Research*, 1, 161–170. <https://doi.org/10.1016/j.nhres.2021.11.001>
- Margottini, C., Canuti, P., & Sassa, K. (2013). *Landslide science and practice*. Springer. <https://doi.org/10.1007/978-3-642-31445-2>
- Marin, R. J., Velásquez, M. F., & Sánchez, O. (2021). Applicability and performance of deterministic and probabilistic physically based landslide modeling in a data-scarce environment of the Colombian Andes. *Journal of South American Earth Sciences*, 108, 103175. <https://doi.org/10.1016/j.jsames.2021.103175>
- Marjanović, M., Samardžić-Petrović, M., Abolmasov, B., & Đurić, U. (2019). Concepts for improving machine learning based landslide assessment. In H. Pourghasemi, & M. Rossi (Eds.), *Natural hazards GIS-based spatial modeling using data mining techniques. Advances in natural and technological hazards research*, (Vol. 48, pp. 27–58). Springer. https://doi.org/10.1007/978-3-319-73383-8_2
- Merghadi, A., Yunus, A. P., Dou, J., Whiteley, J., ThaiPham, B., Bui, D. T., Avtar, R., & Abderrahmane, B. (2020). Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Science Reviews*, 207, 103225. <https://doi.org/10.1016/j.earscirev.2020.103225>
- Michel, G. P., Kobiyama, M., & Goerl, R. F. (2014). Comparative analysis of SHAL-STAB and SINMAP for landslide susceptibility mapping in the Cunha River basin, southern Brazil. *Journal of Soils and Sediments*, 14, 1266–1277. <https://doi.org/10.1007/s11368-014-0886-4>
- Nhu, V. H., Shirzadi, A., Shahabi, H., Singh, S. K., Al-Ansari, N., Clague, J. J., Jaafari, A., Chen, W., Miraki, S., Dou, J., Luu, C., Górski, K., Pham, B. T., Nguyen, H. D., & Ahmad, B. B. (2020). Shallow landslide susceptibility mapping: A comparison between logistic model tree, logistic regression, naïve bayes tree, artificial neural network, and support vector machine algorithms. *International Journal of Environmental Research and Public Health*, 17, 2749. <https://doi.org/10.3390/ijerph17082749>
- Ospina-Gutiérrez, J. P., & Aristizábal-Giraldo, E. V. (2021). Application of Artificial Intelligence and machine learning techniques for landslide susceptibility assessment. *Revista Mexicana de Ciencias Geológicas*, 38, 43–54. <https://doi.org/10.22201/cgeo.20072902e.2021.1.1605>

- Ozer, B. C., Mutlu, B., Nefeslioglu, H. A., Sezer, E. A., Rouai, M., Dekayir, A., & Gokceoglu, C. (2020). Correction to: On the use of hierarchical fuzzy inference systems (HFIS) in expert-based landslide susceptibility mapping: the central part of the Rif Mountains (Morocco). *Bulletin of Engineering Geology and the Environment*, 79(1), 551–568. <https://doi.org/10.1007/s10064-019-01585-0>
- Panahi, M., Gayen, A., Pourghasemi, H. R., Rezaie, F., & Lee, S. (2020). Spatial prediction of landslide susceptibility using hybrid support vector regression (SVR) and the adaptive neuro-fuzzy inference system (ANFIS) with various meta-heuristic algorithms. *Science of The Total Environment*, 741, 139937. <https://doi.org/10.1016/j.scitotenv.2020.139937>
- Pang, Y., Meng, R., Li, C., & Li, C. (2022). A probabilistic approach for performance-based assessment of highway bridges under post-earthquake induced landslides. *Soil Dynamics and Earthquake Engineering*, 155, 107207. <https://doi.org/10.1016/j.soildyn.2022.107207>
- Pham, B. T., Prakash, I., Singh, S. K., Shirzadi, A., Shahabi, H., Tran, T. T. T., & Bui, D. T. (2019). Landslide susceptibility modeling using Reduced Error Pruning Trees and different ensemble techniques: Hybrid machine learning approaches. *Catena*, 175, 203–218. <https://doi.org/10.1016/j.catena.2018.12.018>
- Pham, Q. B., Achour, Y., Ali, S. A., Parvin, F., Vojtek, M., Vojteková, J., Al-Ansari, N., Achu, A. L., Costache, R., Khedher, K. M., & Anh, D. T. (2021). A comparison among fuzzy multi-criteria decision making, bivariate, multivariate and machine learning models in landslide susceptibility mapping. *Geomatics, Natural Hazards and Risk*, 12, 1741–1777. <https://doi.org/10.1080/19475705.2021.1944330>
- Pradhan, A. M. S., & Kim, Y. T. (2015). Application and comparison of shallow landslide susceptibility models in weathered granite soil under extreme rainfall events. *Environmental Earth Sciences*, 73, 5761–5771. <https://doi.org/10.1007/s12665-014-3829-x>
- Pradhan, A. M. S., & Kim, Y. T. (2016). Evaluation of a combined spatial multi-criteria evaluation model and deterministic model for landslide susceptibility mapping. *Catena*, 140, 125–139. <https://doi.org/10.1016/j.catena.2016.01.022>
- Qing, F., Zhao, Y., Meng, X., Su, X., Qi, T., & Yue, D. (2020). Application of machine learning to debris flow susceptibility mapping along the China – Pakistan Karakoram Highway. *Remote Sensing*, 12(18), 2933. <https://doi.org/10.3390/rs12182933>
- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rahali, H. (2019). Improving the reliability of landslide susceptibility mapping through spatial uncertainty analysis: A case study of Al Hoceima, Northern Morocco. *Geocarto International*, 34, 43–77. <https://doi.org/10.1080/10106049.2017.1357767>
- Rasyid, A. R., Bhandary, N. P., & Yatabe, R. (2016). Performance of frequency ratio and logistic regression model in creating GIS based landslides susceptibility map at Lompobattang Mountain, Indonesia. *Geoenvironmental Disasters*, 3, 19. <https://doi.org/10.1186/s40677-016-0053-x>
- Ruiz Vásquez, D., & Aristizábal, E. (2018). Landslide susceptibility assessment in mountainous and tropical scarce-data regions using remote sensing data: A case study in the Colombian Andes. *Geophysical Research Abstracts EGU2018-3408*, EGU General Assembly.
- Saha, A., & Saha, S. (2020). Comparing the efficiency of weight of evidence, support vector machine and their ensemble approaches in landslide susceptibility modelling: A study on Kurseong region of Darjeeling Himalaya, India. *Remote Sensing Applications: Society and Environment*, 19, 100323. <https://doi.org/10.1016/j.rsase.2020.100323>
- Saha, S., Roy, J., Pradhan, B., & Hembram, T. K. (2021). Hybrid ensemble machine learning approaches for landslide susceptibility mapping using different sampling ratios at East Sikkim Himalayan, India. *Advances in Space Research*, 68, 2819–2840. <https://doi.org/10.1016/j.asr.2021.05.018>
- Sahana, M., & Sajjad, H. (2017). Evaluating effectiveness of frequency ratio, fuzzy logic and logistic regression models in assessing landslide susceptibility: A case from Rudraprayag district, India. *Journal of Mountain Science*, 14, 2150–2167. <https://doi.org/10.1007/s11629-017-4404-1>
- Sahin, E. K. (2020). Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using XGBoost, gradient boosting machine, and random forest. *SN Applied Sciences*, 2, 1–17. <https://doi.org/10.1007/s42452-020-3060-1>
- Si, A., Zhang, J., Zhang, Y., Kazuva, E., Dong, Z., Bao, Y., & Rong, G. (2020). Debris flow susceptibility assessment using the integrated random forest based steady-state infinite slope method: A case study in Changbai Mountain, China. *Water*, 12, 2057. <https://doi.org/10.3390/w12072057>
- Silalahi, F. E. S., Arifanti, Y., & Hidayat, F. (2019). Landslide susceptibility assessment using frequency ratio model in Bogor, West Java, Indonesia. *Geoscience Letters*, 6, 1–7. <https://doi.org/10.1186/s40562-019-0140-4>
- Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, 14, 323–348. <https://doi.org/10.1037/a0016973>
- Sun, D., Wen, H., Wang, D., & Xu, J. (2020). A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm. *Geomorphology*, 362, 107201. <https://doi.org/10.1016/j.geomorph.2020.107201>
- Sur, U., & Singh, P. (2019). Landslide susceptibility indexing using geospatial and geostatistical techniques along Chakrata-Kalsi road corridor. *Journal of the Indian National Cartographic Association*, 38, 2018.
- Therneau, T. M., & Atkinson, E. J. (1997). *An introduction to recursive partitioning using the RPART routines* (Technical report). Mayo Foundation.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1), 267–288.
- Vega, J. & Hidalgo, C. (2023). Comparison study of a landslide-event hazard mapping using a multi-approach of fuzzy logic, TRIGRS model, and support vector machine in a data-scarce Andes Mountain region. *Arabian Journal of Geosciences*, 16, 527. <https://doi.org/10.1007/s12517-023-11627-3>
- Wang, Y., Seijmonsberger, A. C., Bouten, W., & Chen, Q. (2015). Using statistical learning in regional landslide susceptibility zonation with limited landslide field data. *Journal of Mountain Science*, 12, 268–288. <https://doi.org/10.1007/s11629-014-3134-x>
- Wright, M. N., & Ziegler, A. (2017). Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, 77, 1–17. <https://doi.org/10.18637/jss.v077.i01>
- Zhang, Y. X., Lan, H. X., Li, L. P., Wu, Y. M., Chen, J. H., & Tian, N. M. (2020). Optimizing the frequency ratio method for landslide susceptibility assessment: A case study of the Caiyuan Basin in the southeast mountainous area of China. *Journal of Mountain Science*, 17, 340–357. <https://doi.org/10.1007/s11629-019-5702-6>
- Zhou, X., Wen, H., Li, Z., Zhang, H., & Zhang, W. (2022). An interpretable model for the susceptibility of rainfall-induced shallow landslides based on SHAP and XGBoost. *Geocarto International*, 37(26), 13419–13450. <https://doi.org/10.1080/10106049.2022.2076928>