IPG 2023
INTERNATIONAL PIPELINE GEOTECHNICAL CONFERENCE
6ª CONFERENCIA INTERNACIONAL GEOTECNIA DE DUCTOS

Proceedings of the IPG 2023
International Pipeline Geotechnical Conference
IPG2023
23-24 November 2023, Bogotá, D.C., COLOMBIA

IPG2023-0029

UPDATE OF LANDSLIDE SUSCEPTIBILITY MAP FOR THE OCENSA CORRIDOR

Oscar Correa-Calle[1], Francisco J. García-Orozco[1], Joan S. Quintero-Londoño[1], Jorge J. Vélez-Upegui[1], Manuel A. Botía-Díaz[2]

[1]Universidad Nacional de Colombia, Manizales
[2]Oleoducto Central de Colombia, Bogotá

ABSTRACT

The regional analysis of landslide susceptibility is vital for understanding and monitoring threats in hydrocarbon transportation system integrity. This intricate process connects various conditioning factors, including geological materials, terrain geometry, and human activity. A notable challenge is validating robust methods to merge categorical variables (geology, geomorphology, land use, and cover) with continuous variables from the digital elevation model. Another hurdle is obtaining updated land use data, high-resolution digital elevation models, and a substantial number of landslide records over time. 2009 IDEAM created a susceptibility map for the Ocensa corridor using 181 events and a 90-meter resolution DEM. In this work, a new map is introduced, based on 1104 events, 30-meter resolution, updated land use data up to 2021, and more explanatory variables from the Bayesian-weighted evidence model. Validation uses the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve. A satisfactory AUC value of 87.9% was achieved. The results are compared to IDEAM (2009) and SGC (2019) findings.

Keywords: Landslide susceptibility, weights of evidence, principal component analysis, pipeline risk.

NOMENCLATURE

Nomenclature used in equations:

S	landslide presence.
$\bar{S}$	landslide absence.
B_i	presence of class i of factor B.
$\bar{B}_i$	absence of class i of factor B.
A_s	area with landslides.
A_t	total area.
$P\{S\}$	prior probability of landslide in the area.
$P\{B_i\}$	presence probability for class i of factor B.
$P\{S\|B_i\}$	landslide presence conditional probability for presence of class i of factor B.
$P\{\bar{S}\|B_i\}$	landslide absence conditional probability for presence of class i of factor B.
$P\{S\|\bar{B}_i\}$	landslide presence conditional probability for absence of class i of factor B.
$P\{\bar{S}\|\bar{B}_i\}$	landslide absence conditional probability for absence of class i of factor B.
$P\{S \cap B_i\}$	probability of presence of landslide within class i of factor B.
$N_{pix}\{S\}$	number of pixels with presence of landslides.
$N_{pix}\{\bar{S}\}$	number of pixels with absence of landslides.
$N_{pix}\{B_i\}$	number of pixels inside the class i of factor B.
$N_{pix}\{\bar{B}_i\}$	number of pixels outside the class i of factor B.
$N_{pix}\{S \cup \bar{S}\}$	total number of pixels.
$N_{pix}\{B_i \cup \bar{B}_i\}$	total number of pixels.
$N_{pix}\{S \cap B_i\}$	number of pixels with presence of landslides inside the class i of factor B.
$N_{pix}1$	number of pixels with presence of landslides and presence of class i of factor B.
$N_{pix}2$	number of pixels with presence of landslides and absence of class i of factor B.
$N_{pix}3$	number of pixels with absence of landslides and presence of class i of factor B.
$N_{pix}4$	number of pixels with absence of landslides and absence of class i of factor B.
W_i^+, W_i^-	Positive and negative weights for class i of factor B, for landslide presence or absence.
W_{f_i}	Final evidence weight for class i of factor B (class contrast factor).
W_{f_B}	weighted final weight of evidence for factor B.

1. INTRODUCTION

From a conceptual standpoint, landslide susceptibility analysis is required to know and monitor the threat on the oil pipeline. Enhancing this susceptibility analysis by employing higher-resolution information, larger working scales, more frequent updates, and an increased volume of data, along with the use of universally accepted methods, advanced variable treatment strategies, and model performance validation, improves the predictive capability of spatial terrain instability

potential. This improvement contributes to the development of a reliable early warning system for rainfall-triggered landslides.

The objective is to compare three landslide susceptibility maps along the Ocensa oil pipeline corridor. These maps have been generated through the spatial and operational overlay of various conditioning factors. One map was produced by the Instituto de Hidrología, Meteorología y Estudios Ambientales [1]. Another map, with national scope and tailored for this study, was developed by the Servicio Geológico Colombiano [2, 3]. The third map corresponds to the authors' creation under the technical-scientific cooperation agreement between Ocensa and the Universidad Nacional de Colombia - Manizales [4].

The IDEAM susceptibility map [1] was developed using the weights of evidence method [3, 5, 6, 7, 8], based on a 1:500,000 scale base cartography [9], and the MDE SRTM NASA digital elevation model (DEM) with a pixel size of 90 m. It classifies susceptibility into high, medium, and low degrees.

Conversely, the SGC susceptibility map [1] considers the following conditioning factors: geomorphological and surface geological units, land use and land cover, slope derived from the DEM, and an inventory of landslides or morphodynamical processes. The map was created using the bivariate statistical method known as the weights of evidence (WoE), which employs Bayes' theorem to determine the conditional and unconditional probabilities of landslide occurrence [3]. This method determines the weights of each conditioning factor based on the frequencies of instability processes' occurrence or absence, using the pixel as the spatial unit. For validation of the calculated SGC landslide susceptibility map, the hypothesis of failure was tested using an acceptable data fitting threshold of 70%, employing the area under the success curve metric.

Subsequently, the process and results of the quantitative landslide susceptibility analysis developed as part of the technical-scientific cooperation agreement between Ocensa and Unal FIA (2023) for a 50 km-wide corridor, are reported.

2. MATERIALS AND METHODS

2.1 Weights of Evidence

The Weights of Evidence (WoE) method, a conditional probability approach [10] and rooted in Bayes' theorem, finds application in landslide susceptibility analysis. It enables the computation of contrast factors or relative weights for the classes of various conditioning factors concerning the absence or presence of instability events. This weight relies on both prior probability and the conditional probability of landslide occurrence.

The prior probability signifies the unconditional probability or density of landslides across the entire analysis area for a specified recording period. This period is determined by the available data in the landslide inventory. It is exclusively calculated based on areas with instability in the inventory, independent of their causality relationships. In other words, it is not contingent on the factor being analyzed or its classes.

The prior probability represents the likelihood of a randomly selected pixel within the analyzed area experiencing a landslide. Accordingly, the calculation of the prior probability involves dividing the area with landslides by the total analysis area [10]. This is equivalent to dividing the number of pixels associated with instability by the total number of pixels in the area, as per equation (1):

$$P_{prior} = P\{S\} = \frac{A_s}{A_t} = \frac{N_{pix}\{S\}}{N_{pix}\{S \cup \bar{S}\}} \quad (1)$$

The conditioning factors, also known as causal factors (independent variables), are divided into classes. The conditional probability of landslide occurrence must be computed for one of the classes of each factor (for instance, a certain geomorphological unit, a slope class, a given land cover). Considering the relationship between a class i (with i = 1, 2, …, m), associated with a causal factor B, hereinafter denoted as B_i, and the landslide map (S), the conditional probability of landslide occurrence due to the presence of the class can be calculated using equation (2):

$$P\{S/B_i\} = \frac{P\{S \cap B_i\}}{P\{B_i\}} = \frac{N_{pix}\{S \cap B_i\}}{N_{pix}\{B_i\}} \quad (2)$$

In other words, the conditional probability of landslide occurrence at a randomly selected pixel within a class i of factor B is the density of landslides within that class. It is calculated as the number of pixels with landslides in the class, divided by the total number of pixels within it. The Weights of Evidence method enables an indirect assessment of landslide susceptibility [5]. To achieve this, it necessitates assigning positive and negative weights (W_i+ and W_i-) to each unit of the map, with which the final weight of the class (referred to as the contrast factor, W_f) is computed concerning landslide occurrence or absence. Using the final weights of all classes, the pixels in the raster corresponding to each of the conditioning factors B are reclassified.

Consequently, based on Bayes' theorem, equation (3) can be formulated as the natural logarithm of the conditional probability ratio for landslide presence or absence, due to the presence of class i of factor B. Similarly, equation (4) represents the natural logarithm of the conditional probability ratio for landslide presence or absence, due to the absence of class i of factor B [5]:

$$W_i^+ = log_e\left(\frac{P\{S|B_i\}}{P\{\bar{S}|B_i\}}\right) \quad (3)$$

$$W_i^- = log_e\left(\frac{P\{S|\bar{B}_i\}}{P\{\bar{S}|\bar{B}_i\}}\right) \quad (4)$$

The calculation of the final weight of evidence or contrast factor, W_{fi}, for class i of factor B, is defined as the difference between the logarithms of the probability ratios computed in equations (3) and (4). The former explains landslide presence or absence due to the presence of class i of factor B. The latter explains landslide presence or absence due to the absence of class i of factor B. The final weight of evidence is obtained using equation (5):

$$W_{f_i} = W_i^+ - W_i^- \quad (5)$$

That is to say, positive weights (W_i^+) and negative weights (W_i^-) are assigned to the various classes into which each conditioning factor has been divided. The final evidence weight for class i of factor B (contrast factor W_{fi} is derived by subtracting the positive evidence weight from the negative evidence weight. These contrast factors signify the statistical association or causal relationship between the classes of each conditioning factor and the occurrence of landslides [11].

Van Westen formulates the problem of evidence weight computation in terms of pixels [5]. Hereinafter, N_{pix} denotes the number of pixels. The universe of potential combinations between each landslide and each class i of conditioning factor B is represented in the matrix of Table 1.

	Potential landslide conditioning factor Bi		
	Presence *Bi*	Absence $\bar{B}i$	
Presence *S*	Npix1	Npix2	Total area with presence of landslides
Absence $\bar{S}$	Npix3	Npix4	Total area with absence of landslides
	Total area of *Bi*	Total area of $\bar{B}i$	Total area (A) of the map

TABLE 1: Matrix of potential combinations between each landslide and each class i of a conditioning factor B (modified from [2, 5, 8]).

Alternatively, considering: A, the study area or universal set; Bi, the set of areas occupied by class i of the conditioning factor B; S, the set of areas with landslides, then Figure 1. Representation of the relationship between the occurrence or absence of landslides and the presence or absence of class i of factor B (adapted from [2, 5, 7]), depicts the relationship between the occurrence or absence of landslides and the presence or absence of class i of factor B within the study area A.

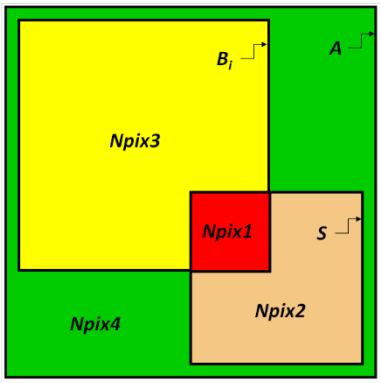


FIGURE 1: Representation of the relationship between the occurrence or absence of landslides and the presence or absence of class i of factor B (modified from [2, 5, 8]).

From Table 1 and Figure 1, illustrating the relationship between the occurrence or absence of landslides and the presence or absence of class i of factor B (adapted from [2, 5, 8]), it is feasible to derive the conclusions provided by equations (6) to (14):

$$N_{pix}1 = N_{pix}\{S \cap B_i\} \quad (6)$$
$$N_{pix}2 = N_{pix}\{S \cap \bar{B}_i\} \quad (7)$$
$$N_{pix}3 = N_{pix}\{\bar{S} \cap B_i\} \quad (8)$$
$$N_{pix}4 = N_{pix}\{\bar{S} \cap \bar{B}_i\} \quad (9)$$
$$N_{pix}\{S\} = N_{pix}1 + N_{pix}2 \quad (10)$$
$$N_{pix}\{\bar{S}\} = N_{pix}3 + N_{pix}4 \quad (11)$$
$$N_{pix}\{B_i\} = N_{pix}1 + N_{pix}3 \quad (12)$$
$$N_{pix}\{\bar{B}_i\} = N_{pix}2 + N_{pix}4 \quad (13)$$
$$N_{pix}\{B_i \cup \bar{B}_i\} = N_{pix}\{S \cup \bar{S}\} = N_{pix}1 + N_{pix}2 + N_{pix}3 + N_{pix}4 \quad (14)$$

According to van Westen [5], equations (3) and (4) are equivalent to equations (15) and (16), but in terms of pixels:

$$W_i^+ = log_e\left[\frac{\frac{N_{pix}1}{N_{pix}1+N_{pix}2}}{\frac{N_{pix}3}{N_{pix}3+N_{pix}4}}\right] \quad (15)$$

$$W_i^- = log_e\left[\frac{\frac{N_{pix}2}{N_{pix}1+N_{pix}2}}{\frac{N_{pix}4}{N_{pix}3+N_{pix}4}}\right] \quad (16)$$

The final evidence weight for class i of factor B is obtained through the expression (5), which for clarity is reiterated below as equation (17):

$$W_{f_i} = W_i^+ - W_i^- \quad (17)$$

The factor (W_i^+) indicates the importance of class i of factor B in explaining the presence or absence of landslides. The factor (W_i^-) indicate the importance of the absence of class i of factor B in explaining the presence or absence of landslides.

Hence, the positive and negative evidence weights corresponding to class i of factor B must be interpreted according to the following rules [7, 8]:

If ($W_i^+ > 0$), the presence of class i of factor B contributes to the occurrence of landslides. Its magnitude indicates the degree of direct correlation or contribution [5].

If ($W_i^+ < 0$), the presence of class i of factor B contributes to the absence of landslides. Its magnitude indicates the degree of inverse correlation [5].

If ($W_i^+ \cong 0$), the presence of class i of factor B shows no relation to landslides. In other words, the factor is not conditioning [5].

If ($W_i^- > 0$), the absence of class i of factor B contributes to the occurrence of landslides.

If ($W_i^- < 0$), the absence of class i of factor B contributes to the absence of landslides.

If ($W_i^- \cong 0$), the absence of class i of factor B shows no relation to landslides.

In conclusion, “the final class weight W_{fi} will be zero (or very close to zero) when the spatial distribution of mass movements is independent of the considered factor, positive when a positive association exists (the presence of the factor contributes to the occurrence of mass movement), and negative when a negative association exists (the absence of the factor contributes to the occurrence of mass movement).” [12].

In other words, the final contrast factor or evidence weight obtained from equation (17) constitutes a measure of the correlation between class i of conditioning factor B and mass movements S. If the final evidence weight exhibits extreme values (positive or negative), this indicates that the studied class is useful for landslide susceptibility analysis and mapping. However, if the contrast factor value centers around zero, it suggests that the class has no relationship with the occurrence of landslides.

Finally, this study proposes obtaining the weighted factor weight for each conditioning factor based on the individual final weights of its component classes, proportional to the area they occupy in the study area, using equation (18).

$$W_{f_B} = \frac{\sum_{i=1}^{m} N_{pix\{B_i\}} \times W_{f_i}}{\sum_{i=1}^{m} N_{pix\{B_i\}}} \quad (17)$$

Similarly, following van Westen [5], it is possible to interpret, for the weighted final evidence weight of the factor, that "Weights with extreme values indicate that the factor is useful for susceptibility mapping, whereas factors with a weight around zero have no relation to the occurrence of landslides." In this latter case, factor B could be excluded from the susceptibility map calculation model.

2.2 ROC Curve in Landslide Susceptibility Analysis

For model validation, a true positive is referred to when the forecast for an analysis unit (or pixel) indicates terrain instability and indeed corresponds to an unstable area in the landslide inventory. Conversely, a true negative exists when the forecast for an analysis unit (or pixel) indicates terrain stability, and indeed no landslides have occurred in the analyzed area, according to the landslide inventory.

Quantitative models for landslide susceptibility provide continuous values of dependent variables. In this case, results can be classified in a binary form: occurrence (1) or absence (0) of landslides. The classification of spatial probability of instability or susceptibility depends on the distributions of data with and without landslides, as well as the selected threshold value to predict positive and negative cases of phenomenon occurrence.

Table 2 presents the universe of possible cases that may arise when evaluating a landslide occurrence prediction model in a matrix form. This table allows contrasting the evidence of occurrence or non-occurrence of real events against the forecast of the presence or absence of instability processes. The match between reality and the occurrence prediction of events is known as the model's Sensitivity, which represents the classification as true positive cases. On the other hand, the match between reality and the prediction of NO occurrence of events is referred to as the Specificity of the model (true negatives). The classification of each case into one of the four categories shown depends, as mentioned earlier, on the chosen discrimination or decision threshold.

Modelo / Eveto	Prediction no landslide	Prediction landslide
No landslide (q = 1 - p = 0.0)	Specificity: True negatives 1 - α	TYPE I ERROR: False positives α
Landslide (p = 1.0)	TYPE II ERROR: False negatives β	Sensitivity: True positives 1 - β

TABLE 2: Confusion matrix for landslides.

2.3 Study Area and Analysis Area

The study area was defined as the corridor contained within a 50 km strip (25 km buffer) along the pipeline (see Figure 2). However, the analysis was conducted for a larger area to avoid model boundary issues and utilize a greater amount of information from the available landslide inventory within the pipeline's influence zone. It encompasses 42 IGAC sheets at a 1:100,000 scale.

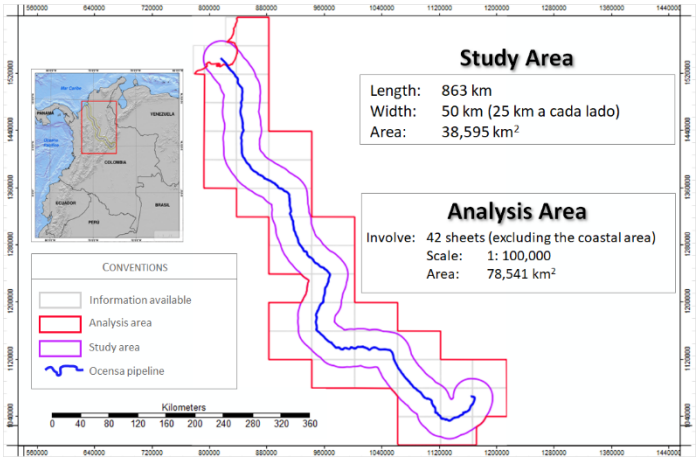


FIGURE 2: Study area for landslide susceptibility analysis along the Ocensa pipeline...

Thus, each raster analysis matrix corresponding to each input or output variable consists of 290,705,464 pixels (14,536 columns x 19,999 rows) across the entire map extent (261,635 km^2 block). However, within the mosaic boundary of the 42 sheets, there are 203,437,265 pixels with no value and 87,268,199 pixels with data within the analysis area. Considering the coastline trim, the information covers a total analysis area of 78,541 km^2. In the end, a trimming of the landslide susceptibility map is performed for the study area. Thus, the analysis area contains 42,882,733 pixels with information, for a total study area of 38,595 km^2.

2.4 Landslide Inventory Map

A landslide inventory map polygons from various sources is employed, namely the National Mass Movement Inventory - INMM (2021 with 6826 data points), supplemented with data from the Mass Movement Information System – SIMMA [15], both from SGC, totaling 7200 records for the entire country, of which 1929 are within the analysis area. These were further supplemented with 91 records provided by Ocensa. Out of these landslides, 1104 are located within the study area (see Figure 3).

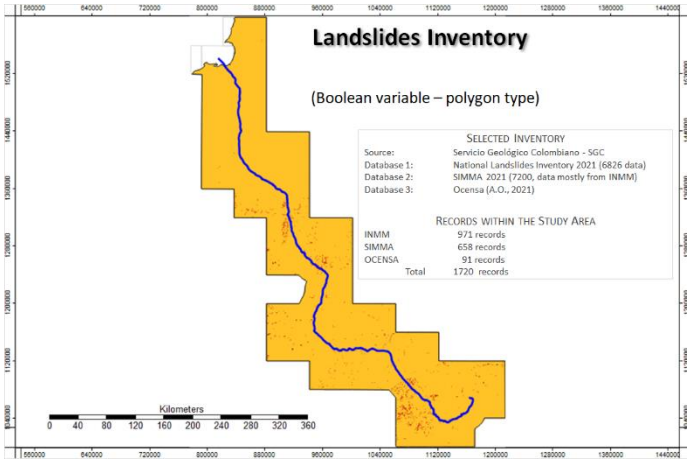


FIGURE 3: Landslide inventory within the analysis area of the Ocensa pipeline.

2.5 Factors Conditioning Landslides

Following recent recommendations [2, 7, 8], input variables of the following types are considered: continuous, categorical, boolean or binary. Continuous variables are derived from the digital elevation model. These include slopes, curvatures, roughness, and humidity indices, as well as flow accumulation and length. Categorical variables encompass superficial geological units, geomorphological units, and land use and land cover units. Additionally, boolean variables consist of the mass movement polygons. Table 3 specifies the coding of variables used in the weights of evidence analysis.

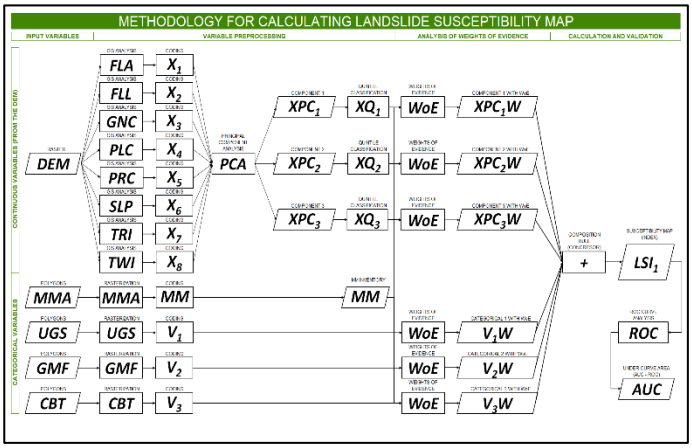
Code	Nomenclature	Variable Name	Data Type	Variable Type
X_1	FLA	Flow Accumulation	raster	continuous
X_2	FLL	Flow Length	raster	continuous
X_3	GNC	General Curvature	raster	continuous
X_4	PLC	Planar Curvature	raster	continuous
X_5	PRC	Profile Curvature	raster	continuous
X_6	SLP	Slope	raster	continuous
X_7	TRI	Terrain Roughness Index	raster	continuous
X_8	TWI	Topographic Wetness Index	raster	continuous
V_1	UGS	Shallow Geological Unit	polygon	categorical
V_2	GMF	Geomorphological Unit	polygon	categorical
V_3	CBT	Land use and cover	polygon	categorical
P	MMA	Landslide Inventory	polygon	boolean
Y	SampPoint	Training and Validation Points	point	binary

TABLE 3: Nomenclature of susceptibility analysis variables.

2.6 Methodology

The methodology for landslide susceptibility analysis consists of four phases: 1. Acquisition of input variables; 2. Preprocessing of variables; 3. Weights of evidence analysis; 4. Calculation and validation of susceptibility map. Figure 4 depicts a diagram of the methodological process for conducting quantitative landslide susceptibility analysis based on the weights of the evidence method.

FIGURE 4: Methodology for the calculation and validation of landslide susceptibility scenario, LSI1.



This analysis strategy is referred to as the LSI1 scenario. It involves performing conventional weights of evidence (WoE) analysis while preceding it with the preprocessing of continuous variables derived from the digital elevation model (DEM) as well as categorical variables. The preprocessing of continuous variables involves transforming the eight continuous variables into three (dimensionality reduction) using the principal component analysis method, followed by classifying the resulting variables into five quintiles.

Moreover, the preprocessing of categorical variables involves converting polygons into raster. Subsequently, the WoE method is applied to each of the resulting variables, and the susceptibility index is obtained as the sum of contrast factors of explanatory variables, pixel by pixel (as suggested by the Colombian Geological Survey [2, 3]). Validation is carried out using the area under the curve (AUC) of the ROC curve.

3. RESULTS AND DISCUSSION

3.1 Calculation of Susceptibility Map

The methodology described in Figure 12 was employed to calculate the landslide susceptibility map at a 1:100,000 scale, corresponding to the LSI1 scenario, for the available information within the analysis area.

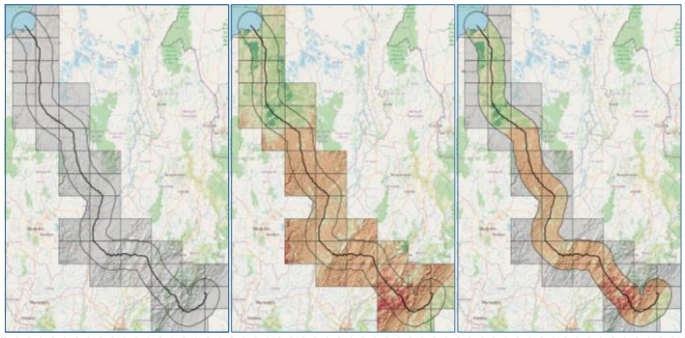


FIGURE 12: Left: Analysis area (gray color); center: Landslide susceptibility map, LSI1; right: Cropped susceptibility map for the study area of the Ocensa pipeline.

Subsequently, the map was cropped for the study area, which corresponds to a 50 km corridor (25 km buffer) along the 836 km length. The final map is presented in Figure 12, both for the analysis area and the study area.

3.2 Validation of Susceptibility Map

For the validation of the landslide susceptibility model, the area under the curve (AUC) metric of the ROC curve was used. A high AUC value (greater than 80% according to Goyes [8], or greater than 70% according to SGC [2, 3]) determines the validity of the hypothesis regarding the existence of a causal relationship between the analyzed conditioning factors and the occurrence of a specific type of landslide (in this case, the slides). Figure 13 shows the ROC curve for the LSI1 scenario. It was

constructed by testing different thresholds on the distributions of data with the presence and absence of landslides, and for each threshold, obtaining a coordinate point (false positives, true positives).

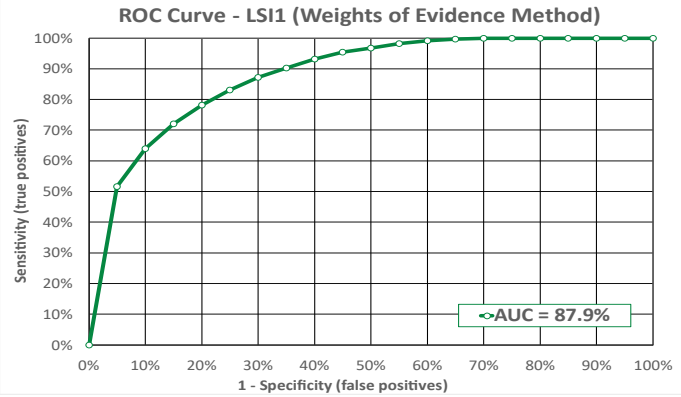


FIGURE 13: ROC Curve of the landslide susceptibility map using weights of evidence, LSI1 scenario.

The ROC curve was constructed using the algorithm by Professor Vakhshoori from Shiraz University (Iran) [13]. As seen in Figure 13, the area under the curve was 87.9%, confirming the reliability of the resulting map.

3.3 Analysis of Landslide Susceptibility Map Results

The map was classified into five categories. Table 4 presents the information derived from the susceptibility map. The category percentages for the analysis area were very low (3.0%), low (14.6%), medium (25.6%), high (52.1%), and very high (4.7%). For the study area, these percentages are not significantly different.

Susceptibility Class	Analysis Area			Study Area (buffer of 25 km)		
	# of Pixels per Class	Km²	%	# of Pixels per Class	Km²	%
Very Low	2,597,487	2,338	3.0	1,576,067	1,418	3.7
Low	12,760,950	11,485	14.6	6,168,583	5,552	14.4
Moderate	22,327,782	20,095	25.6	9,516,696	8,565	22.2
High	45,455,961	40,910	52.1	23,470,659	21,124	54.7
Very High	4,126,019	3,713	4.7	2,150,728	1,936	5.0
Total	87,268,199	78,541	100	42,882,733	38,594	100
	Number of Pixels without Data					
	203,437,265			207,357,257		
	Total Number of Pixels on the map					
	290,705,464			250,239,990		

TABLE 4: Susceptibility category for analysis and study areas.

3.4 Results of Model Comparison

The landslide susceptibility map provided in this report differs from the IDEAM [1] and SGC [3] maps in the following aspects:

The number of landslide events used for creating the map increased from 181 [1] to 1104 events. The number of events in the SGC map [3] cannot be determined as it is a subset of the result. However, it is expected to be lower since it uses the same database but with a smaller time window than the one used in the Unal FIA [4] map. These events correspond to 15% of the official records available in the country, contained in SIMMA [15]. This is a highly reliable database, as around 94% of the information was collected directly in the field and verified by the landslides group of SGC (National Landslide Inventory of SGC, updated as of March 5, 2020). The observation window was also extended from 30 years of records (1979 - 2009) to 111 years (1909 - 2020).

Furthermore, the scale of the new map is 1:100,000, which is 5 times larger than the previous maps. Additionally, the spatial resolution of the digital elevation model (SRTM - NASA) was improved from 90 m to 30 m (30x30 m pixel), providing 9 times more definition for map areas and consequently for morphometric variables derived from DEM. The new map also enhances the scale of the base cartography by IGAC, going from 1:500,000 to 1:100,000.

In the IDEAM report [1], the methodology used for constructing the landslide susceptibility map and the validation results were not evident. The 2009 map likely lacks validation, given that the application of validation methods for susceptibility analysis has gained prominence over the last decade. The validation of the SGC [3] map and the new map was performed using the area under the curve (AUC) metric of the Receiver Operating Characteristic (ROC) curve. This curve represents the model's performance in terms of accuracy rates (true positives) and errors (false positives) in predicting landslide occurrences.

The conditioning factors considered by IDEAM [1] could not be determined either, representing another aspect that generates uncertainty. At the very least, it is known that geological units and the digital elevation model were used.

It should be noted that for the new map, rigorous preliminary work was done to integrate the 42 sheets of the cartographic mosaic containing the studied corridor for each categorical variable (geological surface units, geomorphological units, and land cover), addressing existing issues in the base information.

In the previous IDEAM map, 3 classes were used for categorizing landslide susceptibility: High, Medium, and Low. The SGC [3] and Unal FIA [4] maps use 5 categories: Very High, High, Medium, Low, and Very Low, which enhances spatial differentiation of susceptibility.

In summary, the new landslide susceptibility map was calculated using more and better information, recent and more widely accepted methodologies within the academic and scientific community of the country, and it undergoes validation or performance evaluation of the susceptibility model calculated using the weights of evidence method.

Table 5 presents a comparison of the most relevant parameters among the three available susceptibility maps for the Ocensa pipeline corridor, according to IDEAM [1], SGC [3], and Unal FIA [4].

Item	IDEAM 2009	SGC 2019	UNAL FIA 2021
Source of base cartography	IGAC	IGAC	IGAC
Scale of base cartography	1:500,000	1:500,000	1:100,000
Digital Elevation Model	MDE SRTM	MDE SRTM	SRTM - NASA
Spatial resolution (meters)	90	90	30
Number of landslides for analysis	181	ND	1104
Event observation window	1979 - 2006	1909 - 2014	1909 - 2020
Work scale	1:500,000	1:100,000	1:100,000
Analysis method	ND	WoE	WoE
Input variables	UGS, SLP, DEM?	UGS, GMF, USU, CBT	UGS, GMF, CBT, DEM, SLP, TRI, TWI, FLA, FLL, GNC, PLC, PRC
Validation method	ND	area under the curve (AUC), of the ROC curve	area under the curve (AUC), of the ROC curve
Performance	ND	> 70.0%	87.9%
Susceptibility categories	Low, Medium, High	Very low, Low, Medium, High, Very high	Very low, Low, Medium, High, Very high

TABLE 5: Comparison of susceptibility models between IDEAM [1], SGC [3], and Ocensa Unal FIA [4].

In Figure 14, it is evident that the Ocensa - Unal FIA [4] map more finely discriminates susceptibility differences between neighboring analysis units.

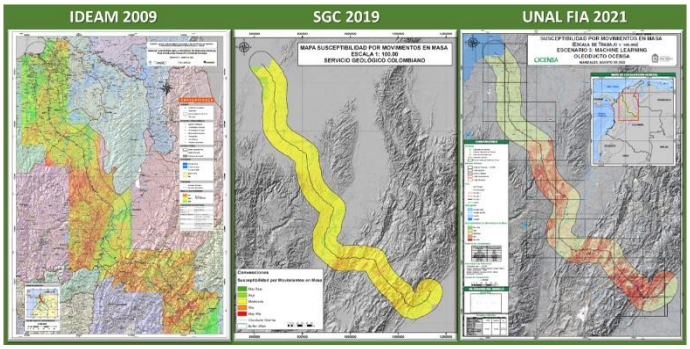


FIGURE 14: Comparison between landslide susceptibility maps of IDEAM [1], SGC [3], and Ocensa-Unal FIA [4].

4. CONCLUSION

Landslide susceptibility, reflecting the likelihood of slope instability, primarily depends on intrinsic terrain factors related to morphological characteristics and soil/rock resistance. However, it is also influenced by external factors like land use/cover changes and variations in natural moisture conditions. Susceptibility maps provide snapshots of geological processes tied to slope erosion and instability, which are intrinsically dynamic phenomena.

Periodic updating of these maps captures changes in conditioning factors such as land use/cover and the occurrence of new erosion and landslide processes, altering the spatial probability of terrain instability. Sensitivity, reflected by the true positive rate (on the ordinate axis), defines the model's sensitivity, while the false positive rate defines the complement of specificity (1 - true negative rate). For the susceptibility map obtained via the weights of evidence method, an AUC (Area Under the Curve) value of 0.879 was achieved, indicating a highly satisfactory result.

The present study's generated map demonstrates reduced uncertainty and relative variation in landslide susceptibility. This improvement is attributed to a larger scale (5 times), higher spatial resolution (9 times), utilization of finer pixel size, an extended time window, increased dataset size, and inclusion of more explanatory variables. Notably, the proposed map excels by achieving highly satisfactory validation metrics compared to the SGC map [3] and acceptable literature-based standards. This achievement is noteworthy given that the IDEAM map [1] lacks reported validation results.

ACKNOWLEDGEMENTS

The authors would like to thank the Colombian Geological Survey (Servicio Geológico Colombiano – SGC in Spanish) for providing the information used to prepare this research.

REFERENCES

[1] Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) – Datum Ingeniería Ltda. (2009). Mapa de susceptibilidad a procesos de remoción en masa en la zona de influencia del Oleoducto Central de Colombia. Proyecto: Revisión, ajuste y mejoramiento del modelo de pronóstico de amenaza a procesos de remoción en masa a escala 1:500,000. Bogotá (Colombia).

[2] Servicio Geológico Colombiano - SGC (2017). Guía Metodológica para la Zonificación de Amenaza por Movimientos en Masa, Escala 1:25.000. Colección de Guías y Manuales. Bogotá: Imprenta Nacional de Colombia.

[3] Servicio Geológico Colombiano - SGC (2019). Mapa Integrado de Susceptibilidad a Movimientos en Masa, Escala 1:100.000. v1 2017. Bogotá, Colombia.

[4] Universidad Nacional de Colombia Sede Manizales – Facultad de Ingeniería y Arquitectura – Unal FIA (2021). Susceptibilidad por Movimientos en Masa, Escenario LSI1: Pesos de Evidencia – Oleoducto Ocensa, Escala 1:100,000, Convenio de cooperación técnico – científica Ocensa – Unal FIA, 2020. Manizales, Colombia.

[5] van Westen, C.J. (2002). Weights of evidence modeling for landslide susceptibility mapping. International Institute for Geoinformation Science and Earth Observation (ITC), Netherlands.

[6] Aristizábal, E., López, S., Sánchez, O., Vásquez, M., Rincón, F., Ruiz-Vásquez, D., Restrepo, S., y Valencia, J.S. (2019). Evaluación de la amenaza por movimientos en masa detonados por lluvias para una región de los Andes colombianos estimando la probabilidad espacial, temporal, y magnitud. Boletín de Geología, 41(3), 85-105. doi: 10.18273/revbol.v41n3-2019004.

[7] Goyes-Peñafiel, P., & Hernandez-Rojas, A. (2020). Landslide susceptibility index based on the integration of logistic regression and weights of evidence: A case study in Popayan, Colombia. Engineering Geology, 105958.

[8] Goyes-Peñafiel, P.; Hernández-Rojas, A. (2021). Doble evaluación de la susceptibilidad por movimientos en masa basada en redes neuronales artificiales y pesos de evidencia. Boletín de Geología, 43(1), 173-191. https://doi.org/10.18273/revbol.v43n1-2021009

[9] Instituto Geográfico Agustín Codazzi – IGAC. (2005). Cartografía vectorial a escala 1:500.000 con cobertura total de la República de Colombia. Sistema de coordenadas MAGNA-SIRGAS. Bogotá (Colombia).

[10] Bonham-Carter, G. F., & Bonham-Carter, G. (1994). Geographic information systems for geoscientists: modeling with GIS. Oxford: Pergamon Press 32(11):302–334.

[11] Neuhäuser B, Bodo D, Birgit T (2012) GIS-based assessment of landslide susceptibility on the base of the weights-of-evidence model. Landslides 9: 511–528. https://doi.org/10.1007/s10346-011-0305-5

[12] van Westen, C. J., Rengers, N., & Soeters, R. (2003). Use of Geomorphological Information in Indirect Landslide Susceptibility Assessment. Natural Hazards, 30(3), 339-419. https://doi.org/10.1023/B:NHAZ.0000007097.42735.9e

[13] Vakhshoori, V., Pourghasemi, H. R., Zare, M., & Blaschke, T. (2019). Landslide susceptibility mapping using GIS-based data mining algorithms. Water, 11(11), 2292.

[14] Servicio Geológico Colombiano - SGC (2022). Sistema de Información de movimientos en masa. De http://simma. sgc. gov. co.

[15] Servicio Geológico Colombiano - SGC (2021). Sistema de Información de movimientos en masa - SIMMA. De http://simma. sgc. gov. co.