https://rbgeomorfologia.org.br/
ISSN 2236-5664

Revista Brasileira de Geomorfologia

v. 24, nº 4 (2023)

http://dx.doi.org/10.20502/rbg.v24i4.2314

Research Article

Binary Logistic Regression applied to erosion susceptibility mapping in the Southern Amazon

Regressão Logística Binária aplicada ao mapeamento da suscetibilidade à erosão no Sul da Amazônia

Elaine Lima da Fonseca [1], and Eliomar Pereira da Silva Filho [2]

[1] Federal Institute of Education Science and Technology of Rondônia, Colorado do Oeste (RO), Brazil. E-mail. elaine.fonseca@ifro.edu.br
ORCID: https://orcid.org/0000-0002-6872-8204

[2] Federal University of Rondônia, Porto Velho (RO), Brazil. E-mail. eliomar@unir.br
ORCID: https://orcid.org/0000-0002-6271-8435

Received: 28/10/2022; Accepted: 03/10/2023; Published: 28/10/2023

Abstract: Problems with soil erosion by water wind in the Brazilian Amazon are intensifying as the forest is replaced by agricultural production. Deforestation, burning, logging, and advancing the agricultural frontier have altered the soil-vegetation balance. In this context, the analysis of soil erosion susceptibility is one of the most significant challenges to developing long-term sustainability strategies and policies. Based on the above, the present study used the principles of Statistical Modeling - Logistic Regression - to develop and validate a model for analysis of susceptibility to soil erosion using 14 environmental factors. The study was carried out in a hydrographic sub-basin with 330 km^2, located in the south of the State of Rondônia in the western Amazon, which combines characteristics of intense anthropic activity, loss of fertile soil, gullies, and silting of rivers. The study area has rainfall above 2000 mm $year^{-1}$, they are transcurrent shear zones, predominant relief forms are flat to slightly convex tops, drainage networks are dendritic in an exorheic system, vegetation cover is composed of areas of forests and natural or regenerated forest fragments, agriculture is destined to annual crops. Livestock is extensive, with a predominance of small rural properties. The logistic regression model showed satisfactory results with an AUC of 0.888 and global accuracy was 0.77. The variables with the most significant effect on the equation were NDVI, erosivity, and TST. The mapping found that 57.71% of the study area is in places susceptible to soil loss due to water events.

Keywords: Susceptibility; Logistic regression; Soil erosion on Amazon.

Resumo: Os problemas com erosão solo por vento hídrico na Amazônia Brasileira estão se intensificando à medida que a floresta é substituída pela produção agropecuária. Atividades como o desmatamento, queimada, exploração madeireira e o avanço da fronteira agrícola alteraram o equilíbrio solo-vegetação e, nesse contexto, a análise da suscetibilidade à erosão do solo é um dos maiores desafios para desenvolver estratégias e políticas de sustentabilidade de longo prazo. Com base no exposto, o presente estudo fez uso dos princípios da Modelagem Estatística - Regressão Logística Binária - para desenvolver e validar um modelo de análise da suscetibilidade à erosão do solo, empregando 14 fatores ambientais. O estudo foi realizado em uma sub-bacia hidrográfica com 330 km^2, localizado no Sul do Estado de Rondônia na Amazônia ocidental, que reúne características de intensa atividade antrópica, perde solo férteis, voçorocamentos e assoreamentos dos rios. A área de estudo apresenta pluviosidade acima de 2000 mm ano^{-1}, são zonas de cisalhamento transcorrentes, as formas predominantes do relevo são topo planos a ligeiramente convexizados, as redes de drenagens são dendrítica em um sistema exorréico, a cobertura vegetal é composta por áreas de florestas e fragmentos florestais naturais ou regenerados, a agricultura é destinada a culturas anuais e a pecuária é extensiva com predomínio de pequenas propriedades rurais. O modelo de regressão logística apresentou resultados satisfatório com AUC de 0,888 e a exatidão global foi de 0,77, as variáveis de maior efeito na equação foram NDVI, erosividade e TST, e o mapeamento constatou que 57,71% da área de estudo encontra-se em locais suscetíveis a perda de solo por evento hídrico.

Palavras-chave: Modelagem**;** Amazônia; Erosão de solo

1. Introduction

Water-related soil erosion is one of the most important natural hazards, bringing on land degradation, loss of soil resources, reduction of soil fertility, desertification, and destruction of human infrastructure (ARABAMERI et al., 2018). It is recognized as the leading cause of land degradation worldwide (VAN LEEUWEN et al., 2019). Den Biggelaar et al., (2003) suggest that the global average rates of entrepreneurs are between 12 and 15 tons ha^{-1} yr^{-1}. Zuazo and Pleguezuelo (2009) claim that approximately 75 billion tons of soil are eroded from terrestrial ecosystems. According to FAO (2015), approximately 33% of the world's soils are degraded, and soil erosion is responsible for losing 25 to 40 billion tons of soil annually.

In the Brazilian Amazon, due to the occupation process, mainly from the 1970s onwards, a series of adverse environmental events such as deforestation, burning, logging, and advances in the agricultural frontier resulted in the breaking of the soil-vegetation balance and have been A considerable increase in erosion processes was observed, marked mainly by large incisions such as ravines and gullies (ALBUQUERQUE; VIEIRA, 2014). Among the factors of anthropogenic modifications with the most significant impact are agricultural practices with inadequate driving techniques, extensive animal production, and the low zootechnical indexes of the herd, in addition to the implantation and development of urban centers without prior planning (FONSECA, 2017).

Analysis of soil erosion susceptibility is one of the significant challenges to developing appropriate long-term sustainability strategies and policies. Several analyses of soil erosion have been conducted to understand, measure, and propose mitigation strategies, ranging from purely quantitative evaluations to studies using complex qualitative methods (SARKAR; MISHRA, 2018). Modeling is a tool that allows spatial extrapolating of the results obtained to new regions to characterize future scenarios. According to Fernandes (2016), they are like simplified structures of reality that supposedly present (or preserve) its most important characteristics or relationships. Several erosion susceptibility modeling studies have been conducted (CHEN et al., 2017; ARABAMERI et al., 2018; SHOLAGBERU et al., 2019; BRAGAGNOLO et al., 2020).

Although these models were developed based on specific geoenvironmental conditions, in general, mapping based on multivariate analysis is a relevant approach because it seeks to understand the phenomenon from a set of direct or indirect factors, considering the relationships between them and with the event. In this sense, modeling complex nonlinear systems, such as the triggering factors of erosive processes, requires integrating models techniques, GIS, and data-mining methods. Modeling techniques are diverse and can be classified into multicriteria decision-making, statistical, and machine-learning models (ARABEMERI et al., 2020). Multicriteria decision-making models involve techniques such as the analytic hierarchy method (AHP) (CALDAS et al., 2019); multicriteria decision-making, in which a scale is applied at a pre-established interval to indicate the range of maximum effect to no effect on the manifestation of erosive processes (FANTINEL; BENEDETTI, 2016); and hierarchization of values according to the importance of factors concerning erosion (VALLE; FRANCELINO and PINHEIRO, 2016).

Statistical models include logistic regression (RAJA et al., 2016); information value (SARKAR; ROY and MARTHA, 2013); conditional probability (RAHMATI et al., 2017); frequency ratio (MELIHO et al., 2018); entropy index (JAAFARI et al., 2014); certainty factor (SOMA; KUBOTA, 2018); and weights of evidence (WoE) (GOYES-PEÑAFIEL; HERNANDEZ-ROJAS, 2021). The most advanced research techniques based on Machine Learning methods involve models such as fuzzy logic and neuro-fuzzy systems (YAVARI et al., 2018); artificial neural networks (ANN) (SHAHRI et al., 2019); support vector machine (SVM) (LEE; HONG; JUNG, 2017); generative adversarial networks (GAN) (AL-NAJJAR; PRADHAN, 2021); convolutional neural network (CNN) (MEENA et al., 2021). Recently, hybrid integration approaches have been gradually introduced in erosion susceptibility mapping (BRAGAGNOLO et al., 2020; ARABAMERI et al., 2021).

Logistic regression, the model chosen in the present study, is a well-recognized and widely applied nonlinear system to predict the probability of presence or absence. Binary logistic regression (BLR), therefore, is an appropriate type of regression preferred when treating dichotomous categorical dependent variables, that is, nominal or non-metric variables with only two groups or classifications (NARDI et al., 2019). In general, they enable the intuitive selection of relevant attributes and are distinguished from other regression techniques because the response variable may be categorical.

In Brazil, the modeling of areas susceptible due to hydric events are mainly addressed using multicriteria decision-making models (ADAMI et al., 2012; SILVA; MACHADO, 2014; CORTE et al., 2015; AGRA; ANDRADE, 2021). However, surveys of regions susceptible to erosion processes suggest a need to update both the method and the variables used, improve the quality and scale of products, and significantly reduce subjectivity (RAJA et al., 2017; ARABAMERI et al., 2018). Considering the context described above, the present study used a BLR statistical and validated a model of susceptibility to soil erosion due to hydric events in Rondônia, southern Amazon, Brazil.

2. Materials and Methods

2.1. Location and characterization of the study area

The study area consists of the Sete Voltas River sub-basin, located in the municipality of Colorado do Oeste, Rondônia State, in the southern Amazon (13°04'45,329" S / 60°30'42,942" W) (Figure 1). The area is occupied by cultivated pasture with more than 30 years of formation, implanted after the felling of the native vegetation. Extensive livestock farming is practiced without pasture management and management techniques. The animal stocking rate varies according to forage availability and is seasonal due to rainfall (FONSECA, 2017). The grass is slightly degraded according to the scale (NOGUEIRA, 2016) developed for the municipality. Historically, these are areas with inexpressive economic activity until the 1970s, based on extractivism (SILVA, 2012). The effective occupation of the region occurred after the implantation of the Integrated Settlement Projects (PICs) and Colonization and Directed Settlement Projects (PADs), as part of the national policy for the de facto occupation of the Brazilian territory, during the military government, making a large flow of migrants from various regions of the country came to the State (TEIXEIRA; FONSECA, 2003).

After this migratory flow to the region, Rondônia was incorporated into the national economic system with the exponential growth of agricultural activity in the eighties and nineties of the twentieth century (LINHARES et al., 2014). However, according to Fonseca (2017), as population growth and development of the urban core occurred, there was an increase in the amount of deforested area and the number of forest fragments, with direct consequences in the loss of productive soils through erosion events.

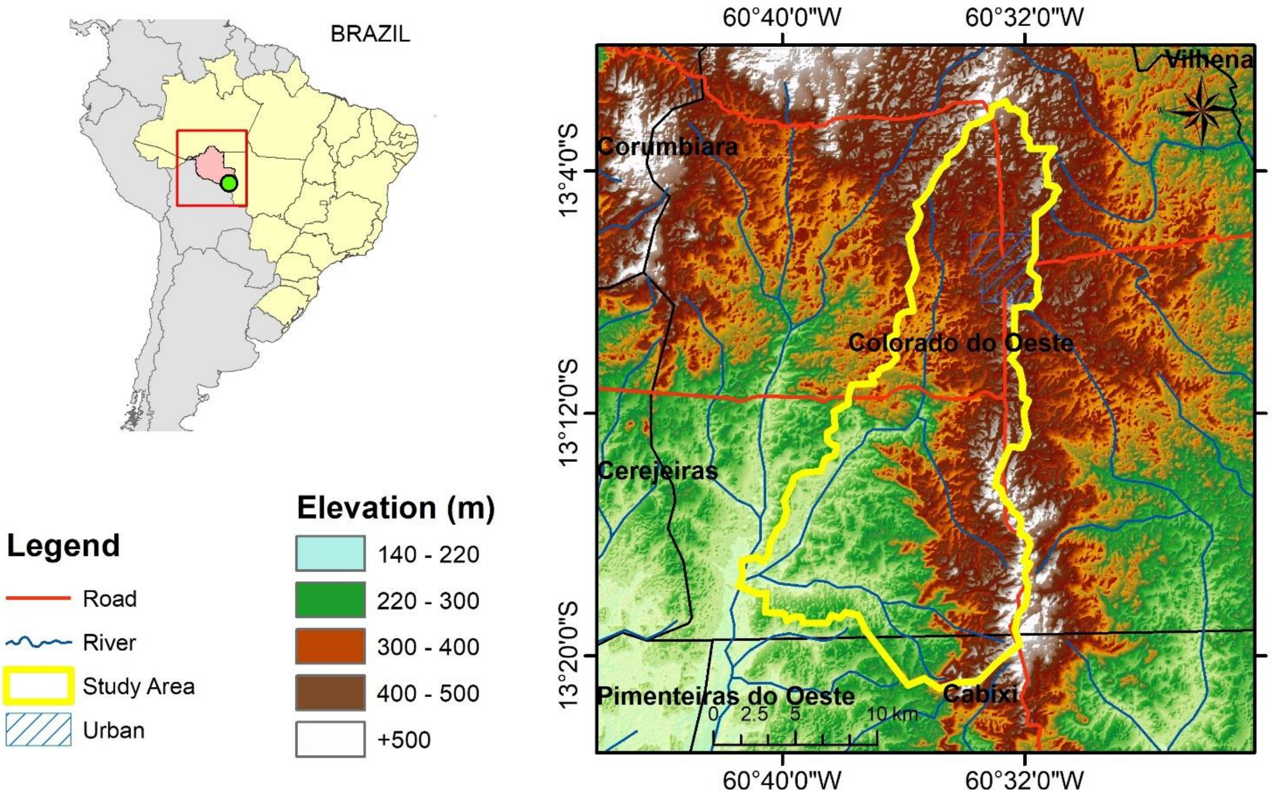

Figure 1. Location of the study area.

Area of study is composed of alluvial plains, from sedimentary processes that commonly occur in river environments, and occasionally from lacustrine processes (PLANAFLORO, 1998), and by denudation units in which the eroded surface and subsurface material are entirely removed from the area due to the processes of physical and chemical weathering of the bedrock. The geological substrate dates different epochs, including Nova Brasilândia Metavolcanic-Sedimentary sequence and minor isolated points formed by the Alto Candeias Intrusive Suite from the Mesoproterozoic Neoproterozoic Eras, the Parecis Formation from the Mesozoic at the northern of the sub-basin, and Cenozoic surface formations (CPRM, 1999).

The average annual rainfall for the municipality of Colorado do Oeste is 1,900 mm year^{-1}. The months with the highest rainfall intensity correspond to the Amazonian winter period (December to March), with rainfall more significant than 300 mm month^{-1}. Dry period is three months, with rainfall below 50 mm (June, July, and August) (FONSECA, 2017). The prevailing climate is Tropical Rainy, with an average air temperature during the coldest month above 18 °C (mega thermal) (SEDAM, 2010). The average annual air temperature is high and uniform throughout the year, with an average of 26 °C (VIEIRA et al., 2014). The vegetation cover comprises Semideciduous Ombrophilous Forest, Cerrado, and transition zones between the two in natural forest fragments (FONSECA; LOCATELLI; SILVA FILHO, 2018).

Drainage networks are tributaries of the Guaporé River and have springs on the edge of Chapada dos Parecis (RADAMBRASIL, 1979). According to Fonseca and Silva Filho (2017), the sub-basins in the region are dendritic in an exorheic system. The vegetation cover mainly comprises forest areas and natural or regenerated arboreal forest fragments, in which Ombrophylous Forest, Cerrado, and transition areas are observed (FONSECA; LOCATELLI; SILVA FILHO, 2018).

2.2. Methodology

This research involved several steps and processes (Figure 2), presented below:

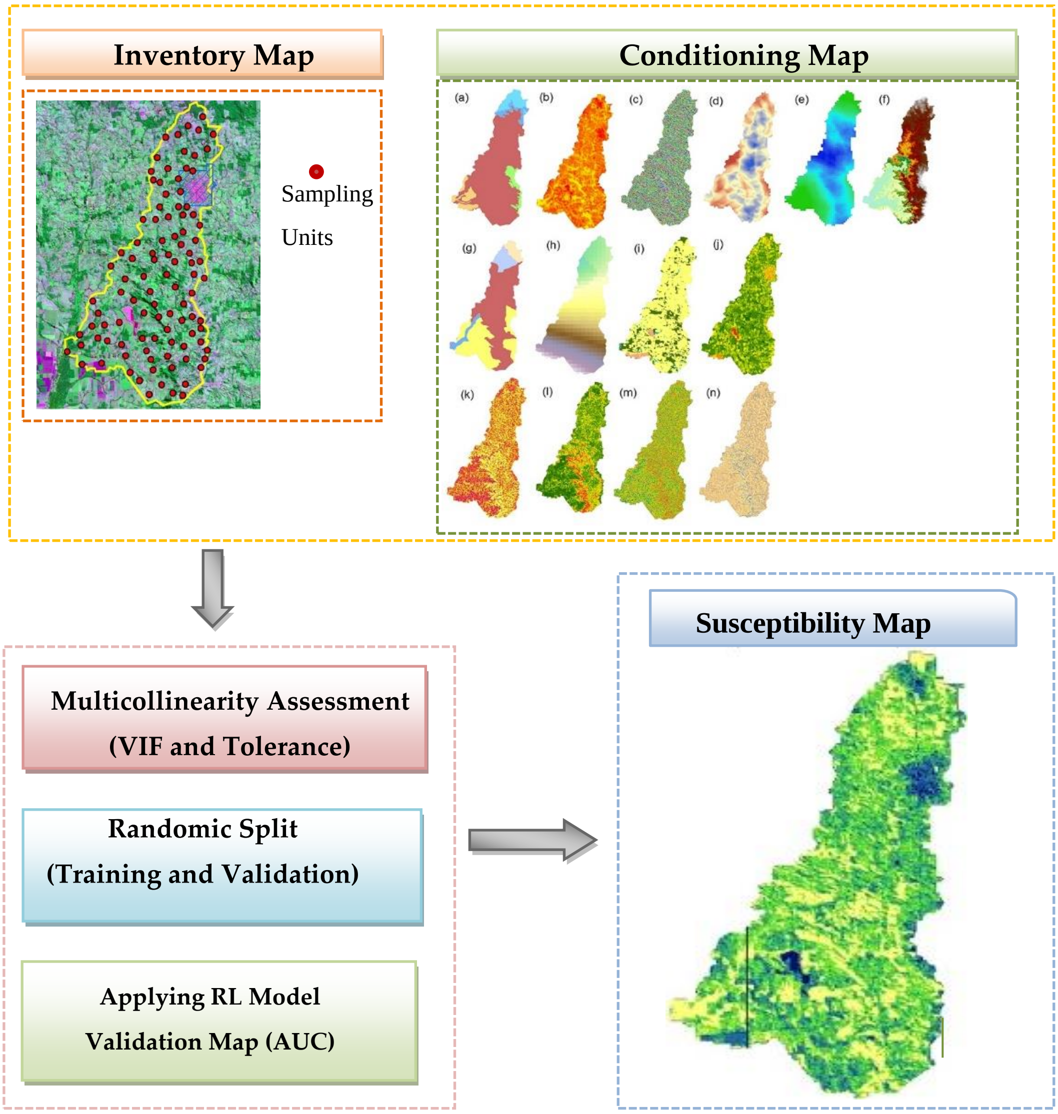

Figure 2. Flowchart of the methodology used in the study. a) Lithology; b) Land Surface Temperature; c) Aspect; d) Drainage Density; e) Elevation; f) Pedology; g) Erosivity; h) Land Use and Occupation; i) NDVI, j) Topographic Wetness Index; k) Slope; l) Stream Power Index; m) Curvature.

2.2.1. Conditioning factors

The selection ensures the maximum accuracy in forecasting, minimization of complexity, and absence of redundancy. The conditioning parameters are under three main categories: 1. topographic parameters derived from the digital elevation model (elevation, slope, aspect, slope curvature, composite topographic index, and stream power index); 2. geological parameters (lithology, drainage density, and lineament density); 3. environmental parameters (normalized difference vegetation index (NDVI), land surface temperature, land use and occupation, erosivity, and soil type). The parameters were chosen based on the researcher's experience, the current specialized literature, the scale of analysis, the data availability, and the study's purpose.

The predictors were developed using the following databases and software: 1. Geomorphometric Database of Brazil (TOPODATA) with a spatial resolution of 30 meters from the Brazilian Institute of Geography and Statistics (IBGE), which offers a digital elevation model (DEM) and its basic local derivations across the country from SRTM data provided by the United States Geological Survey (USGS); 2. Socio-Ecological Zoning of the State of Rondônia (PLANAFLORO) from the Secretary of State for Environmental Development (SEDAM), which provides information on pedology and mapping of ravines and gullies; 3. Geological Service of Brazil (CPRM) through the Carta Geológica Folha Pimenteiras - SD 20-X-D with information on outcrops, structures, and lithologies; 4. LANDSAT-8 with data from scene 230/69 on May 22, 2020. The bands used were B4, B5, B10, and B11, with spatial resolution of 30 meters.

2.2.2. Erosion inventory

The survey was carried out in July 2020 with the aid of a PHANTOM 4 (DJI) Unmanned Aircraft System (UAS), using a regular sampling pattern of 2 km x 2 km, the spatial resolution of around 10 centimeters, aiming at uniformity in the selection of data and the most extensive possible coverage. The flights met Brazilian Civil Aviation Regulations (RBAC-e 94) regarding the safe operation of UAS. Images were complemented with information from PLANAFLORO/RO and Sentinel-2B orbital satellite images (July/2020) (Spatial resolution of 10 meters).

They were submitted to binomial distribution analysis. The selection criteria of the sampling points were visual and based on the indicators described by Guerra et al., (2014) and Guerra and Jorge (2014) for areas with erosive processes in Brazil, including widespread exposed soil patches, the appearance of roots on the surface; removal of organic surface horizons; stretches of laminar erosion; the presence of terracettes; and formation of ravines and gullies.

Sites that widely or fully of these indicators were considered susceptible to erosion (coded as 1 or success); otherwise, susceptibility to erosion was considered minimal or negligible (encoded as 0 or failure). Thus, 50 points with erosion were selected, and, to complete the dichotomous analysis, 50 points with no erosion were randomly created in ArcMap 10.5.

2.2.3. Considerations of categorical variables

For variables with more than two categories, it is essential to encode the explanatory variable, indicating it as a reference for the model. The categorical variables of the present analysis are ordinal, presenting ordering within each predictor, following the guidelines expressed in the literature regarding susceptibility to erosion. Classes with a value weight of 0 are the reference classes for the model and are significant when the null hypothesis (H0) is met, that is when there is no erosion.

2.2.4. Susceptibility map

For applying Binary Logistic Regression, the response variable (Y) assumed an erosional process's dichotomous condition (presence or absence). Thus, 70% of the data set was randomly selected for the training of the BLR model and 30% for the test. The training derived the coefficients (β) for each variable and the equation's constant (α). Finally, the coefficients and constant of the training data set were applied to the predictors within the test data set (corresponding to the remaining 30% of the sample units) to calculate the output value.

The cutoff limit was 0.5 and served as the model's orientation between outputs 0 and 1. If the soil loss (Y=1) is more significant than the cutoff point (P(Y=1) > 0.5), the sample unit is classified as the presence of erosion and receives a code of 1; sample units with lower probability (P(y=1) < 0.5) were coded as 0, indicating an absence of erosion. Variables were introduced, ensuring all independent variables were placed in the same regression model in a single block and that parameter estimates were calculated for each block.

The equations used to obtain the susceptibility map followed the proposal of Sarkar and Mishra (2018). According to the authors, the values of the variables for a given pixel {column c, row r} in the input datasets produce the value of the corresponding output pixel of susceptibility to soil erosion (S) from equation (1).

$$S_{\{\mathrm{c,r}\}} = \frac{1}{1 + e^{-\left(\alpha + \varphi_{cont\{c,r\}} + \varphi_{LIT\{c,r\}} + \varphi_{SOIL\{c,r\}} + \varphi_{USE\{c,r\}}\right)}} \tag{1}$$

α = constant; φ_cont = contribution of continuous variables in the equation; φ_LIT = contribution of the categorical variable 'LIT' in the equation; φ_SOIL = contribution of the categorical variable ‘SOIL’ in the equation; φ_USE = contribution of the categorical variable 'USE' in the equation; {c, r} = the pixel in the scan at column 'c' and row 'r'.

Adequacy of the model was determined with the Cox & Snell R2 and Nagelkerke R2 values (HAIR JÚNIOR, 2009), in addition to the Hosmer-Lemeshow, Log Likelihood Value, and Wald tests (HAIR et al., 2005). The Cox & Snell R2 is comparable to a linear regression R2; Nagelkerke R2 resembles the Cox & Snell R2 as it is a version adapted to provide results between 0 and 1. In general, Cox & Snell and Nagelkerke's measure of R2 are analogs, as both evaluate the variation found in BLR. The difference is that Cox & Snell R2 is limited as it cannot reach the maximum value of 1 (HAIR et al., 2005).

Log-Likelihood Value test ($-2LL$) is an indicator that verifies the adjustment of the estimate with the value −2 times the logarithm of the likelihood value; the lower the value, the better the fit of the model, with 0 (zero) being the perfect estimate. Hosmer-Lemeshow test aims to evaluate the predictive validity of the logistic regression model. It is not based on the likelihood value but rather the actual value of the dependent variable. The function is to divide the number of observations into approximately ten classes and compare the expected with the observed frequencies to assess the significant differences between the classifications of the model and the observed reality.

Finally, Wald test assesses the degree of significance of each coefficient of the model and verifies whether each estimated parameter is significantly different from zero. The test follows the Chi-square distribution and is obtained by comparing the maximum likelihood estimator βi with the estimation of its error. In addition, the Odds ratio (OR) was initially proposed to determine an event's probability. Range of OR is from 0 to infinity: A value of 1 = no association with the specified risk (that is, the event or disease is equally likely in the high- and low-risk groups); as the value of OR increases or decreases away from 1, the association grows increasingly more robust (CHEN et al., 2010).

Was obtained using global accuracy calculated according to the following equations (2, 3, 4, and 5), according to Bragagnolo et al. (2020):

$$Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \tag{2}$$

$$Precision = \frac{T_p}{T_p + F_p} \tag{3}$$

$$Sensibility = \frac{T_p}{T_p + F_n} \tag{4}$$

$$Specificity = \frac{V_p}{V_p + F_p} \tag{5}$$

Tp is the true positive; Tn is the true negative; Fp is the false positive; and Fn is the false negative. Values close to 1 indicate better performance.

Sensitivity and specificity of the model are expressed in the curve (Figure 3). The ROC curve is a graph of sensitivity versus specificity for the possible cutoff probability values (cutoff = 0.5), which signaled a satisfactory level of BLR performance with an AUC = 0.888.

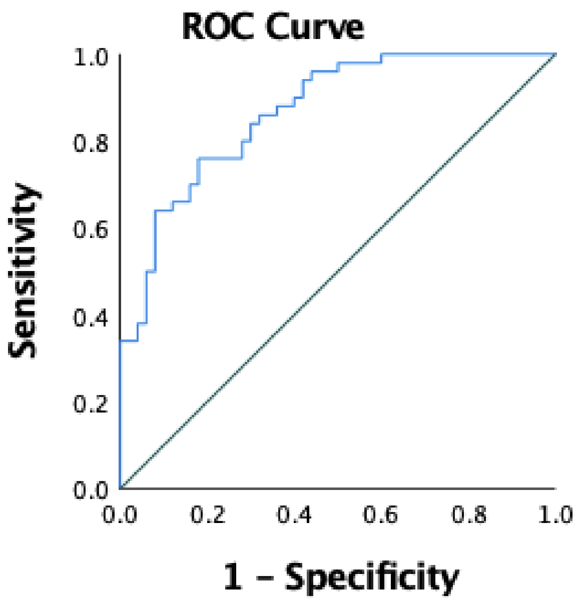

Figure 3. ROC curve for the Logistic Regression Model.

The software used in the mapping were ArcGIS 10.5 and SPSS Statistic 26.0.

3. Results

The parameters analyzed did not present multicollinearity. Values for maximum variance inflation factor (VIF) and minimum tolerance were 2.644 and 0.378, respectively, and meet the requirements for the linear VIF ≤ 10 and tolerance coefficient ≥ 0.1 (CHEN et al., 2019; ARABAMERI et al., 2018) (Table 1).

Table 1. Multicollinearity test between conditioning factors.

Predictors	Multicollinearity test		Predictors	Multicollinearity test	
	Tolerance	VIF		Tolerance	VIF
Elevation	0.688	1.520	USE	0.701	1.426
Slope	0.401	2.494	NDVI	0.602	1.662
ASP	0.686	1.457	DD	0.793	1.261
CV	0.574	1.742	LIT	0.599	1.670
CTI	0.378	2.644	SOIL	0.694	1.440
SPI	0.526	1.901	ERO	0.730	1.540
LD	0.534	1.872	LST	0.546	1.830

ASP: aspect or orientation of the slope; CV: curvature of the slope; CTI: composite topographic index; SPI: stream power index; LD: lineament density; USE: Land Use and Occupation; NDVI: normalized difference vegetation index; DD: drainage density; LIT: lithology; SOIL: soil type; ERO: erosivity; LST: land surface temperature.

Except for the predictors NDVI, erosivity, and LST, the Wald test (Table 2) found that the variables influencing erosion were insignificant for the regression model. Significance is a prerequisite for selecting the variables that must make up using the Stepwise method.

Table 2. Wald test of the Logistic Regression model.

Preditor	SD	Wald	df	p-value	Exp (β)	95% C.I. Exp(β)	
						Inferior	Superior
ELV	0.005	0.028	1	0.868	0.999	0.990	1.008
SLP	0.052	0.298	1	0.585	1.02	0.930	1.138
ASP	0.003	0.344	1	0.558	0.998	0.992	1.004
CV	20.685	0.047	1	0.829	86.99	0.000	-
CTI	0.243	0.001	1	0.977	0.993	0.616	1.600
SPI	0.103	2.973	1	0.085	1.194	0.976	1.461
LD	0.023	0.394	1	0.530	0.986	0.943	1.031
NDVI	24.748	5.122	1	0.024	0.00	0.00	.001
DD	0.014	2.065	1	0.151	1.020	0.993	1.048
LIT		7.475	4	0.113			
LIT(1)	17898.4	0.00	1	0.999	-	0.00	-
LIT(2)	34892.7	0.00	1	1.000	-	0.00	-
LIT(3)	1.214	3.315	1	0.069	9.113	0.845	-
LIT(4)	2.431	5.017	1	0.025	0.004	0.00	0.506
USE		1.148	3	0.766			
USE(1)	1.297	1.148	1	0.284	4.014	0.316	51.024
USE(2)	24442.1	0.00	1	0.999	-	0.00	-
USE(3)	50126	0.00	1	1.000	-	0.00	-
SOIL		1.713	4	0.788			
SOIL(1)	1.858	1.713	1	0.191	11.38	0.298	434.601
SOIL(2)	25312.1	0.00	1	0.999	0.00	0.00	-
SOIL(3)	0.937	0.019	1	0.891	1.138	0.181	7.138
SOIL(4)	34892.7	0.00	1	0.999	0.00	0.00	-
ERO	0.002	8.356	1	0.004	1.005	1.001	1.008
LST	0.523	6.576	1	0.010	3.82	1.372	10.641
Const.	20.134	10.891	1	<.001	0.00		

SD: standard deviation; df: degrees of freedom; p-value < 0.05; CI: confidence interval; Odds ratio values were excessively high depending on the nature of the calculation and were therefore not displayed in the table; ELV: elevation; SLP: slope; ASP: aspect or orientation of the slope; CV: curvature of the slope; CTI: composite topographic index; SPI: stream power index; DL: lineament density; USE: use and occupation of soil; NDVI: normalized difference vegetation index; DD: drainage density; LIT: lithology; SOIL: soil type; ERO: erosivity; LST: land surface temperature.

Pseudo-R2 parameters (Cox & Snell and Nagelkerke), Log-likelihood -2 (-2LL), and the Hosmer-Lemeshow test verified the model. Table 3 presents the fitness measures of the regression model.

Table 3. Statistical summary of the Logistic Regression model.

Log likelihood -2 (-2LL)	Pseudo R^2		Hosmer-Lemeshow Test		
	Cox & Snell R^2	Nagelkerke R^2	Chi-square	df	p-value
81.783a	0.434	0.578	22.474	8	0.004

df: degrees of freedom; p-value: significance $p < 0.05$; [a] - Final estimate with 20 iterations.

Accuracy of the model given by the confusion matrix reached an overall accuracy of 0.77 (Table 4). Matrix indicates the hit-and-miss percentage of the model for the two possible answers. Thus, the model correctly classified 0.80 of the 50 sample units with erosion and 0.74 of the 50 samples without erosion.

Table 4. Logistic Regression Confusion Matrix.

	Observed		Predicted		
			Response		Classification correct
			0	1	
Step 1	Response	0	37	13	0.74
		1	10	40	0.80
	Global				0.77

Table 5 shows the values of the coefficients for the selected predictors, including the value of the constant (α) of the equation.

Table 5. Coefficients (β) of the BLR model equation

Pred.(x)	β(x)	Categorical variables					
		SOIL	$\beta_{(SOIL)}$	USE	$\beta_{(USE)}$	LIT	$\beta_{(LIT)}$
α	-66.446	0	-	0	-	0	-
ELV	-0.001	1	2.432	1	1.390	1	22.628
SLP	0.028	2	-44.430	2	28.093	2	21.369
ASP	-0.002	3	0.129	3	21.859	3	2.210
CV	4.466	4	-23.007			4	-5.445
CTI	-0.007						
SPI	0.177						
LD	-0.014						
NDVI	-56.011						
DD	0.020						
ERO	0.005						
LST	1.340						

ELV: elevation; SLP: slope; ASP: aspect or orientation of the slope; CV: curvature of the slope; CTI: composite topographic index; SPI: stream power index; LD: lineament density; USE: use and occupation of soil; NDVI: normalized difference vegetation index; DD: drainage density; LIT: lithology; SOIL: soil type; ERO: erosivity; LST: land surface temperature.

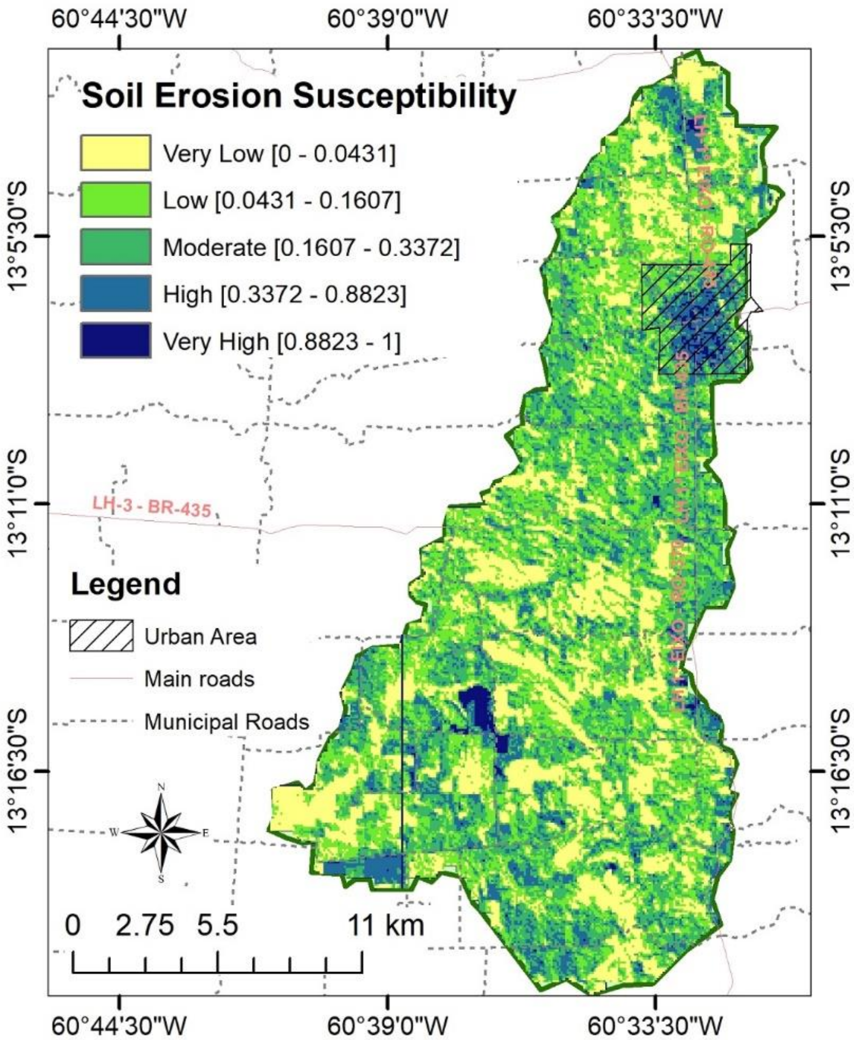

Figure 4. Soil erosion susceptibility map of the Sete Voltas sub-basin - Regression Model.

The susceptible areas are those classified as moderate, high, and very high, representing 57.69% of the area (Table 7).

Table 7. Erosion susceptibility class and area based on the BLR model

Erosion Susceptibility Class	BLR Model	
	Num. of pixels	**% of area**
Very Low	67085	19.44
Low	78779	22.83
Moderate	121053	35.09
High	73153	21.20
Very High	4856	1.40
Total	**344926**	**100**

Figure 5. Aerial photographs taken with the UAS: a) areas with terracettes and trails formed by cattle trampling; b) presence of erosion in first-order rivers; c) erosion near the rivers. Source: Author (July, 2020).

4. Discussion

Choice of input parameters constitutes the subjective criterion with the most significant impact on modeling studies, as they are directly or indirectly involved in the manifestation of the erosion event and have the power to impact the mapping results. In the literature, several studies point out that the number of predictive parameters in modeling erosion events does not follow a specific pattern. Instead, they are determined following several other aspects guided by the perspective of the researcher and the selected landscape. However, it is possible to observe that some parameters are considered typical in modeling processes, such as NDVI and terrain slope (CAN et al., 2017; SHAHRI et al., 2019; GOES-PEÑAFIEL; HERNANDEZ-ROJAS, 2021).

The NDVI index is crucial because it reflects changes in the Earth's surface (DING et al., 2017). Removing vegetation reduces soil moisture through exposure to radiant energy, making it dry, affecting the stability of aggregates, shear strength (XUE et al., 2011), and the decomposition of organic matter. The NDVI analysis reveals that areas with a more significant presence of vegetation present lower susceptibilities, and areas with less vegetation had a higher frequency indicating areas susceptible to soil loss. A similar observation was described by Rahmati et al., (2017). Terrain slope affects with increasing slope gradient, where the shear forces applied by the runoff flow velocity increase, and these factors tend to increase erosion, by increasing the detachment of soil particles. (GUERRA et al. 2014).

4.1. Application of Logistic Regression

The first step of the analysis was the multicollinearity test. This analysis has no multicollinearity among the conditioning factors as they meet the required VIF and tolerance requirements. Thus, there is a strong indication that the selected set is coherent and capable of providing robustness to the model.

Wald tests should be taken with caution because when the regression coefficient (β) is significant (Table 2), the standard error tends to become inflamed, resulting in an underestimation of the Wald statistic (MENARD, 1995). Furthermore, the inflation of the standard error increases the probability that the predictor will be rejected when it contributes to the model. According to Field (2009), this increases the possibility of making Type II errors, which occur when the result shows no effect on the sample while there is an effect.

The odds ratio (OR) of the event of interest presented values with different behaviors for the pre-established set. The predictors as elevation, slope orientation, CTI, and lineament density have OR < 1 (Exp (β) < 1.000 - Table 2). Thus, for the RLB model, an increase in value for this set of variables would decrease the chances of the output resulting in 'success.' In contrast, increasing the other variables' value would increase the probability of soil erosion susceptibility.

A reduction in the probabilities of the predictors causing a positive event - '1' - would imply possible rejection in the composition of the predictor table. This occurs because the lowest p-value is compared with the previously chosen p-value (p-value = 0.05). If the lowest p-value is less than the chosen p-value, the covariate is included in the model; otherwise, the final is the null model.

Variables curvature of the slope, NDVI, and LST showed the highest OR values. An increment of one unit in the degree of slope curvature would increase the chances of erosion by four times. Meanwhile, an increase of 1°C in the LST would increase the erosive possibility by at least two times.

The NDVI was high due to the nature of the calculation, so its value is not shown in Table 3. High values imply significance within the RLB model. When increased by one unit, the NDVI values considerably increase the probabilities of erosion in the experimental area. NDVI predictor is included in most homologous models across various climatic contexts.

Some variables have no effect on the model; an increase of one unit in the variable's value or a reduction of one unit would not result in a greater or lesser chance of erosion occurring. However, in general, the further the coefficient is from one (1), regardless of direction, the more significant the impact of a given independent variable on the probability of the event occurring.

This applies to the variables slope, erosivity, and drainage density with a value of Exp (β) = 1.000. The absence of effects in the model would also exclude the predictor from the final set of independent variables. However, covariates indicated as not significant can become significant in the presence of other covariates, requiring further analysis of the impact of their inclusion or exclusion from the RLB model, as discussed by Chowdhury and Turin (2020).

Regarding the categorical variables, OR values are compared concerning the reference class. In the case of lithology, the reference class was the lithostratigraphic unit Granitoids, which have more excellent resistance to weathering. For the independent variable soil type, the reference class was Latosols because they present greater resistance to the movement of particles as a function of their physicochemical characteristics. Finally, for the variable land use and occupancy, the reference class was defined as native vegetation, which offers protection against the transportation of particles.

Statistically, the OR for lithology indicates that in the LIT(1) class, composed predominantly of quartz-sandstone rocks, the chance of erosion is significantly high, similar to the LIT(2) class with bimodal sandstones. LIT (3), immature polymictic conglomerates, in turn, have a nine (9) times greater chance of erosion. However, the class of undifferentiated sedimentary cover and alluvial deposits, LIT(4), with a predominance of sand, silt, and clay, presented OR < 1, indicating a decreased chance of erosion when compared to the reference class.

Regarding soil type, the OR indicates that the class SOIL(1), Gleysols, is 11 times more likely to experience erosion with Latosols; SOIL(3), Ultisols, are 1.3 times more prone to erosion, while SOIL(2), Cambisols, and SOIL(4), Neosols, have an OR < 1 indicating a lower probability of erosion as the OR for the event to occur is lower with the reference class.

For the predictor of land use and occupancy, the OR suggests that agriculture, USE(2), and water resources, USE(3), are critical for the probability of erosion due to the high values of Exp (β). Pasture, USE(1), in turn, is four times more likely to experience erosive events when compared to the reference class.

Values obtained for log -2 in soil erosion susceptibility modeling studies should be mentioned in the model fitness analysis. The absence of information about regression models, such as the procedure for selecting variables, the type of model used, and how the model's fit is adjusted, are all issues indicated by Bagley et al., (2001). Nevertheless, these flaws persist in many scientific studies.

Nagelkerke coefficient of determination indicates that 57.8% of the variation that occurred in the log of the OR can be explained by the set of independent variables. Cox & Snell R2 predicts that 43.8% of the variation in the dependent variables of susceptibility to erosion is explained by this model. Pseudo values R2 (Cox & Snell and Nagelkerke) presented the same results for the coefficient of determination but were based on different calculations.

The ROC curve indicates that the greater the area under the curve (AUC), the more capable the model is of identifying true positives while minimizing the number or percentage of false positives. The AUC describes the performance of the classification model and illustrates the diagnostic capacity of a binary classifier whose acceptable value varies between 0.5 and 1.0 (RAHMATI et al., 2016). The AUC of the model was 0.888 (88.8%), which is satisfactory to accept the model and its subsequent application in mapping susceptibility to soil erosion.

Several authors have used AUC estimates as parameters to accept model performance. For example, Sholagberu et al., (2019) obtained an estimated AUC value of 0.887 for multivariate logistic regression analysis, indicating the model's excellent performance in evaluating susceptibility to erosion. Same evaluation was applied by Arabameri et al., (2018) with an AUC of 0.802 and Chen, Pourghasemi and Zhao (2016) with an AUC of 0.836.

When analyzing the causal factors in a soil erosion susceptibility study in the state of Pahang, Malaysia, the map produced by Sholagberu et al., (2019) obtained a predictive accuracy of the model in the order of 81.9%, indicating that on average, the map was accurate and acceptable and could be used by planners and decision-makers for sustainable development. However, the author suggested that additional studies should be implemented because the accuracy of the mapping can be affected by the quality and number of variables.

RLB models were also applied to identify areas of ravine erosion. For example, Arabameri et al., (2018) verified soil loss in northeastern Semnan province, Iran. They obtained an index of 80.2%, indicating that the results can contribute to sustainable development in this area, helping minimize soil loss's economic and environmental impacts. Similarly, Arabameri et al., (2019) in Yazd province, Iran, obtained an index of 83.4% with logistic regression. Results show that model, slope, TPI, and elevation were the key factors generating gullying in the study area.

Chen et al., (2017), when mapping susceptibility in Shaanxi province, China, achieved more than 77% accuracy. Raja et al., (2016) used logistic regression to explore the probable future distribution of landslides in the region in Turkey, reaching accuracy values close to 80%. Authors concluded that logistic regression could be an essential means by which landslides in the Black Sea region can be assessed. The parameters selected included slope angle, aspect, lithology, TWI, river proximity, and relative relief. According to the authors, the study emphasized the importance of aspect and lithology, which leads to the conclusion that slope is a significant factor in determining landslide activity.

Conoscenti et al., (2013) employed RLB with an accuracy of approximately 80% to explore susceptibility in the north-central region of Sicily, Italy. The attributes described the variability of lithology, land use, topography, and road position. In addition, they were able to determine the spatial distribution of the gullies and develop a reliable map of susceptibility to ravine erosion.

4.2. Mapping of susceptibility

When analyzing the susceptibility map developed herein (Figure 3), we can see a belt of low susceptibility to erosion in the southeast region of the sub-basin, which is associated with the maintenance of native forest cover. Low erosion risk is also observed in the extreme southwest of the sub-basin. Lowland areas experience flooding at certain times of the year and are unsuitable for agricultural activities. In adjacent regions, agricultural practices increase susceptibility to soil loss, especially in the first year of cultivation, due to the exposure of soil to erosion from rainfall.

At the northern end of the sub-basin, susceptibility to erosion is related to the sites where Cerrado vegetation is being converted to other land uses or grazing areas in transition zones between the Ombrophilous forest and Cerrado vegetation. These sandy soils, classified as Neosols, are located on the edge of the Parecis Plateau, strongly influenced by weathering agents.

Neosols are naturally susceptible to soil loss due to hydric events. Fonseca (2017), when evaluating erosion conditions in Quartzarenic Neosols in the municipality of Colorado do Oeste, found that the predisposition to sediment loss stems from the predominance of very coarse and coarse sand, high values of soil density, low porosity, compaction related to trampling by cattle, and an increase in the average penetration resistance, particularly in the first 10 centimeters.

The central area of the sub-basin and the region close to the urban core have moderate susceptibility. The greatest mapped susceptibilities are in areas with degraded pastures and lowland areas close to water bodies (Figure 4). Fonseca, Locatelli, and Silva Filho (2018), when classifying the stages of degradation of pastures in the municipality of Colorado do Oeste, observed that areas that present some degree of degradation have similar characteristics, such as a variety of the type of forage (braquiarão; *Brachiaria brizantha*), low forage height, limited plant population per m^2, presence of weeds, high grazing pressure, and poor pasture formation and management.

Erosion surrounding drainage networks is predominant in the study area. The rivers in the sub-basin are mostly first-order channels that are fragile, intermittent, and highly susceptible to anthropogenic pressure

(Figure 4). In Figure 4, we can see an absence of vegetation in lowland areas and along rivers. The immediate consequence is a disequilibrium in the soil-landscape relationship and the disappearance of some waterways.

When analyzing the morphometry of the sub-basins in the municipality, Fonseca and Silva Filho (2017) identified that the first-order channels almost always flow directly to the main river, have limited flow, with little or no riparian forest, and may cease to exist due to inadequate management of agricultural activities. Therefore, mapping susceptibility to erosion is the first and most crucial step in managing the effects of soil loss to achieve sustainable development (ARABEMARI et al., 2019). Sustainable management practices, considering the challenges of land use and occupation, require indexes of quantification of soil erosion, spatial distribution, and identification of critical areas for sediment transportation.

Land use and occupation is an essential indicator of erosion potential, as it is related to resistance to particle removal due to the levels of organic matter, protection from the direct impact of rainfall, and more significant water infiltration into the soil (BERTONI; LOMBARDI NETO, 2008). In addition, the soil cover is highly significant in preventing and controlling erosion and the aggravation of erosive processes. It can compromise water infiltration, damage the soil structure, and lead to sediment loss (PERUSI; OAK, 2008).

Damage to the soil includes animal trails near watercourses that increase soil density, compacting the surface and intensifying the speed of surface water runoff. Animal trampling has direct and indirect effects and can impact several aspects of the geomorphological system in both the high and lowlands (TOMAS; DAYS, 2009). However, a reduction in the area available for productive activities (Figure 4) can have an economic impact on producers because most of the rural properties in the sub-basin are small-scale (up to four fiscal modules), that practice intensive livestock production and do not use pasture maintenance and management techniques.

5. Conclusions

The soil erosion susceptibility map is essential to predict areas susceptible. Soil erosion in the Amazon generates significant negative impacts on the environment with direct consequences on local production due to the loss of productive soil and reduced areas suitable for cultivation, affecting ecological stability and reducing soil quality and quantity. Therefore, in the present study, the statistical principle of Logistic Regression was applied in the mapping of areas susceptible to erosion using 14 conditions, including Altitude, Slope, Aspect or Orientation of the Slope, Curvature of the Slope, Composite Topographic Index, Index of Potency of the Flow, Lineament Density, Vegetation Index by Normalized Difference, Drainage Density, Lithology, Soil Type, Land Use and Land Coverage, Erosivity, and Land Surface Temperature. Attached is an erosion inventory map of the entire basin prepared using high-resolution orbital imagery and field survey with remotely piloted aircraft.

Results show that the model could identify the existing relationships between the conditioning factors and generate susceptibility maps consistent with the local reality. The factors NDVI, erosivity, and land surface temperature showed the most significant effects on the RLB model to identify potential areas for soil loss. The areas susceptible to erosion were classified as moderate, high, and very high susceptibility and represented 57.71% of the study area.

The application of spatial modeling can promote more adequate and effective land use planning, providing a preliminary assessment of large areas. Strategies such as vegetation revitalization and management of agricultural activities should be implemented to prevent and reduce the destructive effects of erosion rates. Further studies should be conducted that carefully evaluate the choice of input parameters, verifying their influence on triggering erosion, the associated factors, and their impacts on modeling processes. An excess or lack of parameters may compromise the techniques' quality.

In addition, studies with a smaller scope should be applied to explore spatial modeling techniques better. For example, analyses could evaluate the performance of modeling techniques in contexts with variable climate, geomorphology, and hydrology; compare models employed in the same context; and introduce new approaches, such as hybrid models, to overcome possible flaws detected previously studied models or expand the accuracy of classifications.

Author Contributions: Elaine Lima da Fonseca - conception, methodology, validation, formal analysis, investigation, and writing; Eliomar Pereira da Silva Filho - data curation, visualization, supervision, and project administration.

Funding: This research did not receive any external funding.

Acknowledgments: To the Federal Institute of Education, Science and Technology of Rondônia, Campus Colorado do Oeste. To the research group "Technologies for the management of tropical agro-ecosystems". And the Federal University of Rondônia and the GEOGRAPHY AND CARTOGRAPHY LABORATORY (LabCART).

Conflict of Interest: The authors declare that there is no conflict of interest in publishing this article.

References

1. ADAMI, S.F. et al. Environmental fragility and susceptibility mapping using geographic information systems: applications on Ribeirão do Pinhal watershed (Limeira, State of São Paulo). **Acta Scientiarum**. Technology Maringá, v. 34, n. 4, p. 433-440, 2012. DOI: 10.4025/actascitechnol.v34i4.10005
2. AGRA, R.M.R.T.; ANDRADE, C.D. Análise Multicritério da Suscetibilidade à Erosão na Bacia Hidrográfica do Rio da Batateira, CE. **Rev. Bras. Cartografia**, vol. 73, n. 1, 2021. DOI:10.14393/rbcv73n1-50066
3. ALBUQUERQUE, A.R.C.; VIEIRA, A.F.S.G. **Erosão dos solos na Amazônia**. In: Degradação dos solos no Brasil. Guerra, A.J.T.; Jorge, M.C.O.(Org). Rio de Janeiro. Ed. Bertrand Brasil, 230p., 2014.
4. AL-NAJJAR, H.H.; PRADHAN, B. Spatial landslide susceptibility assessment using machine learning techniques assisted by additional data created with generative adversarial networks. **Geoscience Frontiers** 12, 625–637, 2021. DOI : 10.1016/j.gsf.2020.09.002
5. ARABAMERI A. et al. Spatial modelling of gully erosion using evidential belief function, logistic regression, and a new ensemble of evidential belief function–logistic regression algorithm. **Land Degrad Dev**.; 2018, 29:4035–4049. DOI: 10.1002/ldr.3151
6. ARABAMERI, A. et al. Evaluation of Recent Advanced Soft Computing Techniques for Gully Erosion Susceptibility Mapping: A Comparative Study. **Sensors,** 20, 2020, 335. DOI:10.3390/s20020335
7. ARABAMERI, A. et al. Prediction of gully erosion susceptibility mapping using novel ensemble machine learning algorithms. **Geomatics, Natural Hazards and Risk**, v. 12, n. 1, p. 469–498, 2021. DOI: 10.1080/19475705.2021.1880977
8. ARABAMERI, A.; PRADHANB, B.; LOMBARDOD, L. Comparative assessment using boosted regression trees, binary logistic regression, frequency ratio and numerical risk factor for gully erosion susceptibility modelling. **CATENA** 183, 2019, 104223. DOI: 10.1016/j.catena.2019.104223
9. BAGLEY, S.C. et al. Logistic regression in the medical literature: standards for use and reporting, with particular attention to one medical domain. **Journal of Clinical Epidemiology**, 54:979-985, 2001. http://10.1016/s0895-4356(01)00372-9
10. BERTONI, J.; LOMBARDI NETO, F. **Conservação do solo**. 6.ed. São Paulo: Ícone, 2008. 355p.
11. BRAGAGNOLO, L.; SILVA, R.V.A.; GRZYBOWSKI, J.M.V. Artificial neural network ensembles applied to the mapping of landslide susceptibility. **CATENA** 184, 2020, 104240. DOI: 10.1016/j.catena.2019.104240
12. CAN, A.; et al. Landslide susceptibility mapping at Ovacık-Karabu¨k (Turkey) using different artificial neural network models: comparison of training algorithms. **Bull Eng Geol Environ**. 2017. https://doi.org/10.1007/s10064-017-1034-3
13. CALDAS, V.I.P.; SILVA, A.S.S.; SANTO, J.P.C.S. Suscetibilidade a Erosão dos Solos da Bacia Hidrográfica Lagos – São João, no Estado do Rio de Janeiro – Brasil, a partir do Método AHP e Análise Multicritério. **Revista Brasileira de Geografia Física** v.12, n.04 (2019) 1415-1430.
14. CHEN, W. et al. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. **CATENA** 151, 2017, 147–160. DOI: 10.1016/j.catena.2016.11.032
15. CHEN, H.; COHEN, P.; CHEN, S. How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies, Communications in Statistics - Simulation and Computation, 39:4, 860-864, 2010. DOI: 10.1080/03610911003650383
16. CHEN, W.; et al. Novel hybrid integration approach of bagging-based fisher's linear discriminant function for groundwater potential analysis. **Nat. Resour. Res**. 1–20. 2019. DOI: 10.1007/s11053-019-09465-w
17. CHEN, W.; POURGHASEMI, H.R.; ZHAO, Z. **A GIS based comparative study of Dempster-Shafer, logistic regression, and artificial neural network models for landslide susceptibility mapping**, Geocarto International. 2016.

18. CHOWDHURY, M.Z.I.; TURIN, T.C. Variable selection strategies and its importance in clinical prediction modelling. Fam Med Community Health. 2020 Feb 16;8(1):e000262. DOI: 10.1136/fmch-2019-000262.
19. CONOSCENTI, C. **Gully erosion susceptibility assessment by means of GIS-based logistic regression**: A case of Sicily (Italy). Geomorphology. 2013. DOI: 10.1016/j.geomorph.2013.08.021
20. CORTE, A.P.D. et al. Environmental fragility of Iguaçu River watershed, Paraná, Brazil. **BOSQUE** 36(2): 287-297, 2015. DOI: 10.4067/S0717-92002015000200014
21. CPRM – SERVIÇO GEOLÓGICO DO BRASIL. **Geologia e Recursos Minerais do Estado de Rondônia**: texto explicativo e mapa geológico do Estado de Rondônia, escala 1.1.000.000, Brasília: CPRM. 1999.
22. DEN BIGGELAAR, C. et al. **Crop losses to soil erosion at regional and global scales**: evidence from plot-level and GIS data. In K. Wiebe, ed. Land Quality, Agricultural Productivity, and Food Security. pp. 262-279. UK, Cheltenham, Edward Elgar. (2003).
23. DING, Q.; CHEN, W.; HONG, H. Application of frequency ratio, weights of evidence and evidential belief function models in landslide susceptibility mapping. **Geocarto Int**. 32 (6), 619–639, 2017. DOI: 10.1080/10106049.2016.1165294
24. FANTINEL, R.A.; BENEDETTI, A.C.P. Avaliação dos fatores influentes na vulnerabilidade à erosão do solo por meio de decisão multicritério e de técnicas de geoprocessamento no município de Piratini-RS. **Ciência e Natura**, Santa Maria v.38 n.1, 2009, p. 156 – 163. DOI: 10.5902/2179460X19783
25. FERNANDES, N.F. (2016). **Modelagem em Geografia Física**: Teoria, Potencialidades e Desafios. Espaço Aberto, V. 6, N.1, p. 209-247.
26. FIELD, A. **Descobrindo a estatística usando o SPSS** [recurso eletrônico] Tradução Lorí Viali. – 2. ed. Porto Alegre: Artmed. 2009.
27. FONSECA, E.L. **Processos erosivos em superfícies tabulares com evolução de voçorocamento em áreas de pastagens cultivadas (*Braquiária brizantha* cv. marandu) no município de Colorado do Oeste – Rondônia**. Dissertação. Fundação Universidade Federal de Rondônia – UNIR. Porto Velho. 2017.
28. FONSECA, E.L.; LOCATELLI, M.; SILVA FILHO, E.P. NDVI aplicado na detecção de degradação de pastagens cultivadas. **Confins [Online],** 35, 2018 | DOI: 10.4000/confins.13180.
29. FONSECA, E.L.; SILVA FILHO, E.P. Análise fisiográfica como subsídio ao estudo da suscetibilidade erosiva em bacias hidrográficas. **ACTA Geográfica**, Boa Vista, v.11, n.25. 2017, pp. 137-158. DOI: 10.5654/acta.v11i25.4029
30. FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). (2015). **Status of the World's Soil Resources**. Roma: ITPS.
31. GOYES-PEÑAFIEL, P.; HERNANDEZ-ROJAS, A. Doble evaluación de la susceptibilidad por movimientos en masa basada en redes neuronales artificiales y pesos de evidencia**. Boletín de Geología**, 43(1), 2021, 173-191. DOI: 10.18273/revbol.v43n1-2021009
32. GUERRA, A.J.T. O início do processo erosivo. In: **Erosão e conservação do solo**: conceitos, temas e aplicações. Org. GUERRA, A.J.T.; SILVA, A.S.; BOTELHO, R.G.M. 9ª ed., Rio de Janeiro: Berthrand Brasil, 340p. 2014.
33. GUERRA, A.J.T.; JORGE, M.C.O. **Degradação dos solos no Brasil**. Rio de Janeiro. Ed. Bertrand Brasil, 230p. 2014.
34. HAIR JÚNIOR, J.F. et al. **Análise multivariada de dados**. 6 ed. Porto Alegre: Bookman. 2009
35. HAIR, J.F.; ANDERSON, R.E.; TATHAM, R.L.; BLACK, W.C. **Análise multivariada de dados**. Porto Alegre: Bookman. 2005.
36. JAAFARI, A. et al. GIS-based frequency ratio and index of entropy models for landslide susceptibility assessment in the Caspian Forest, northern Iran. **Int. J. Environ. Sci. Technol**. 11:909–926, 2014. DOI: 10.1007/s13762-013-0464-0
37. LEE, S.; HONG, S.M.; JUNG, H.S. A Support Vector Machine for Landslide Susceptibility Mapping in Gangwon Province, Korea. **Sustainability**, 9, 2017, 48. DOI: 10.3390/su9010048
38. LINHARES, M.S. et al. **Geoprocessamento aplicado ao diagnóstico de uso e cobertura da terra na microbacia hidrográfica do córrego Hermes – Rondônia**. HOLOS, Ano 30, Vol. 01. 2014. DOI: 10.15628/holos.2014.752
39. MEENA, S.R., et al. **Rapid mapping of landslides in the Western Ghats (India) triggered by 2018 extreme monsoon rainfall using a deep learning approach**. Landslides. 2021.

40. MELIHO, M.; et al. A GIS-based approach for gully erosion susceptibility modelling using bivariate statistics methods in the Ourika watershed, Morocco. **Environ. Earth Sci**. 77, 2018, 655. DOI: 10.1007/s12665-018-7844-1
41. MENARD, S. **Applied Logistic Regression Analysis**: Sage University Series on Quantitative Applications in the Social Sciences. Thousand Oaks, CA: Sage, 1995.
42. NARDI, I.R. (2019). O desenvolvimento de um modelo matemático para a previsão da aprovação da disciplina de cálculo 1 utilizando regressão logística. **Braz. J. of Develop**., Curitiba, v. 5, n. 10, p. 22245-22256. DOI: 10.34117/bjdv5n10-352
43. PERUSI, M. C.; CARVALHO, W. A. Comparação de Métodos para Determinação da Estabilidade de Agregados por Vias Seca e Úmida em Diferentes Sistemas de Uso e Manejo do Solo. **Geociências**, São Paulo, v. 27, n. 2, 2008, p. 197-206.
44. NOGUEIRA, W.K.S. Aplicação do NDVI na detecção de degradação de pastagens cultivadas. **Monografia** (Curso de Agronomia). IFRO, 2016.
45. RAHMATI, O. et al. Evaluating the influence of geo-environmental factors on gully erosion in a semi-arid region of Iran: An integrated framework. **Sci. Total Environ**. 579, 2017, 913–927. DOI : 10.1016/j.scitotenv.2016.10.176
46. RAHMATI, O.; HAGHIZADEH, A.; POURGHASEMI, H.R. et al. Gully erosion susceptibility mapping: the role of GIS-based bivariate statistical models and their comparison. **Nat Hazards** 82, 2016, 1231–1258. DOI : 10.1007/s11069-016-2239-7
47. RAJA, N.B. et al. Landslide susceptibility mapping of the Sera River Basin using logistic regression model. **Nat Hazards** 85, 2016. DOI: 10.1007/s11069-016-2591-7
48. RONDONIA, Secretaria de Estado do Planejamento. (2002). **Plano agroflorestal e Pecuária de Rondônia – PLANAFLORO** (bando de dados geográfico). Porto Velho.
49. RONDÔNIA. Secretaria de Estado do Desenvolvimento Ambiental (SEDAM). **Boletim Climatológico de Rondônia**, ano 2008, Porto Velho, 2010, 36p.
50. SARKAR, S.; ROY. A.K.; MARTHA, T.R. Landslide Susceptibility Assessment using Information Value Method in Parts of the Darjeeling Himalayas. **Journal Geological Society of India**. Vol.82, 2013, pp.351-362. DOI: 10.1007/s11069-016-2591-7
51. SARKAR, T.; MISHRA, M. Soil Erosion Susceptibility Mapping with the Application of Logistic Regression and Artificial Neural Network. **Journal of Geovisualization and Spatial Analysis**, 2, 8, 2018. DOI: 10.1007/s41651-018-0015-9
52. SHAHRI, A.A.; SPROSS, J.; JOHANSSON, F.; LARSSON, S. Landslide susceptibility hazard map in southwest Sweden using artificial neural network. **CATENA**, Volume 183, 2019. DOI: 10.1016/j.catena.2019.104225
53. SHOLAGBERU, A.T. Multivariate Logistic Regression Model for soil erosion susceptibility assessment under static and dynamic causative factors. **Pol. J. Environ. Stud**. Vol. 28, 2019, No. 5, 3419-3429. DOI: 10.15244/pjoes/91943
54. SILVA, V.C.B.; MACHADO, P.S. SIG na análise ambiental: susceptibilidade erosiva da bacia hidrográfica do córrego Mutuca, Nova Lima – Minas Gerais. **Revista de Geografia** (UFPE) V. 31, No. 2., 2014.
55. SILVA, A.G.S. Vale do Guaporé: território das espacialidades das culturas desviantes. **Revista Labirinto**, Ano XII, nº 16, 2012.
56. SOMA, S.A.; KUBOTA, T. Landslide susceptibility map using certainty factor for hazard mitigation in mountainous areas of Ujung-loe watershed in South Sulawesi. **Forest and Society**. Vol. 2(1): 79-91, 2018. DOI: 10.24259/fs.v2i1.3594
57. TEIXEIRA, M.A.D.; FONSECA, D.R. **História Regional**: Rondônia-4ª Edição; Porto Velho: Editora Rondoniana, 2003;
58. STUDENMUND, A.H.; CASSIDY, H.J. **Using econometrics**: a practical guide. Boston, Little, Brown. 1987.
59. TOMAS, E.; DIAS W.B. Evolução do rebanho bovino brasileiro e implicações nos processos geomorfológicos, **Revista Brasileira de Geomorfologia** vol.10, 2009, N° 2, 9p. DOI: 10.20502/rbg.v10i2.125
60. VALLE, I.C.; FRANCELINO, M.R.; PINHEIRO, H.S.K. Mapeamento da Fragilidade Ambiental na Bacia do Rio Aldeia Velha, RJ. **Floresta e Ambiente**; 23(2): 295-308, 2016.

61. VAN LEEUWEN, C.C.E. et al. The evolution of soil conservation policies targeting land abandonment and soil erosion in Spain: A review, **Land Use Policy**, Volume 83, 2019, DOI: 10.1016/j.landusepol.2019.01.018.
62. VIEIRA, I.C.G.; JARDIM, M.A.G.; ROCHA, E.J.P. **Amazônia em tempo**: estudos climáticos e socioambientais. Belém: Universidade Federal do Pará. Embrapa Amazônia Oriental, 462 p., 2014.
63. XUE, X.; LUO, Y.; ZHOU, X.; SHERRY, R.; JIA, X. Climate warming increases soil erosion, carbon and nitrogen loss with biofuel feedstock harvest in tallgrass prairie. **Gcb Bioenergy**, 3 (3), 198., 2011. DOI: 10.1111/j.1757-1707.2010.01071.
64. YAVARI S.; MAROUFPOOR, S.; SHIRI, J. Modeling soil erosion by data-driven methods using limited input variables. **Hydrology Research** | 49.5. 2018. DOI: 10.2166/nh.2017.041
65. ZUAZO, V.H.D.; PLEGUEZUELO, C.R.R. (2009). **Soil-erosion and runoff prevention by plant covers**: a review. Sustainable agriculture. Springer. p 785.