



Landslide Scars Detection using Remote Sensing and Pattern Recognition Techniques: Comparison Among Artificial Neural Networks, Gaussian Maximum Likelihood, Random Forest, and Support Vector Machine Classifiers

Detecção de Cicatrizes de Movimentos de Massa utilizando Sensoriamento Remoto e Técnicas de Reconhecimento de Padrões: Comparação entre os Classificadores Redes Neurais Artificiais, Máxima Verossimilhança Gaussiana, Random Forest e Support Vector Machine

Tatiana Dias Tardelli Uehara ¹, Sabrina Paes Leme Passos Corrêa ², Renata Pacheco Quevedo ³, Thales Sehn Körting ⁴, Luciano Vieira Dutra ⁵ and Camilo Daleles Rennó ⁶

¹ National Institute for Space Research - INPE, Earth Observation and Geoinformatics Division, São José dos Campos - SP, Brasil. tatiana.uehara@inpe.br

ORCID: <https://orcid.org/0000-0003-1861-8848>

² National Institute for Space Research - INPE, Earth Observation and Geoinformatics Division, São José dos Campos - SP, Brasil. sabrina.correa@inpe.br

ORCID: <https://orcid.org/0000-0002-9956-4134>

³ National Institute for Space Research - INPE, Earth Observation and Geoinformatics Division, São José dos Campos - SP, Brasil. renata.quevedo@inpe.br

ORCID: <https://orcid.org/0000-0002-7528-9166>

⁴ National Institute for Space Research - INPE, Earth Observation and Geoinformatics Division, São José dos Campos - SP, Brasil. thales.korting@inpe.br

ORCID: <https://orcid.org/0000-0002-0876-0501>

⁵ National Institute for Space Research - INPE, Earth Observation and Geoinformatics Division, São José dos Campos - SP, Brasil. luciano.dutra@inpe.br

ORCID: <https://orcid.org/0000-0002-7757-039X>

⁶ National Institute for Space Research - INPE, Earth Observation and Geoinformatics Division, São José dos Campos - SP, Brasil. camilo.renno@inpe.br

ORCID: <http://orcid.org/0000-0001-9920-4473>

Recebido: 04.2020 | Aceito: 07.2020

Abstract: Landslide inventory is an essential tool to support disaster risk mitigation. The inventory is usually obtained via conventional methods, as visual interpretation of remote sensing images, or semi-automatic methods, through pattern recognition. In this study, four classification algorithms are compared to detect landslides scars: Artificial Neural Network (ANN), Maximum Likelihood (ML), Random Forest (RF) and Support Vector Machine (SVM). From Sentinel-2A imagery and SRTM's Digital Elevation Model (DEM), vegetation indices and slope features were extracted and selected for two areas at the Rolante River Catchment, in Brazil. The classification products showed that the ML and the RF presented superior results with OA values above 92% for both study areas. These best accuracy's results were identified in classifications using all attributes as input, so without previous feature selection.

Keywords: Mass movement. Hazard. Supervised Classification. Pattern Recognition.

Resumo: Inventários de cicatrizes de deslizamentos são uma ferramenta essencial para apoiar a mitigação de riscos de desastres. Os inventários são geralmente obtidos por métodos convencionais, como interpretação visual de imagens de sensoriamento remoto, ou métodos semiautomáticos, por meio de reconhecimento de padrões. Neste estudo, quatro algoritmos de classificação são comparados para a detecção de deslizamentos de terra: Redes Neurais Artificiais (ANN), Máxima Verossimilhança (ML), *Random Forest* (RF) e *Support Vector Machine* (SVM). A partir das imagens do Sentinel-2A e um Modelo Digital de Elevação (DEM) da SRTM, índices de vegetação e atributos de declividade foram extraídos e selecionados para duas áreas na Bacia Hidrográfica do Rio Rolante. Os resultados da classificação mostraram que os classificadores ML e RF obtiveram valores de Acurácia Global (AO) acima de 92% para ambas áreas de estudos. Os valores de acurácia mais altos foram identificados nas classificações utilizando todos os atributos como dados de entrada, portanto, sem seleção de atributos prévia.

Palavras-chave: Movimentos de massa. Perigo. Classificação Supervisionada. Reconhecimento de Padrão.

1 INTRODUCTION

Landslides are widespread natural geomorphologic processes and represent a gravity-driven component of erosion (DAVIES, 2015). They are downward movements of slope material triggered by earthquakes, snow melting or heavy rain, which can also be caused or intensified by anthropic activities (GUZZETTI et al., 2012). These phenomena cause economic damages and casualties when occurring in occupied areas (HAQUE et al., 2019). For monitoring, landslide inventory maps are used, identifying mass movement scars, providing information about past events, such as location, types and patterns, and assisting to build landslide susceptibility models (RAMOS-BERNAL et al., 2018). Thus, landslide inventory maps are crucial to support urban planning and disaster risk reduction (LUPIANO et al., 2019).

The inventory can be achieved either by conventional methods or state-of-the-art techniques. Conventional methods include field mapping and visual interpretation of remote sensing images; although, as a drawback, these methods are time and resource consuming (QIN; LU; LI, 2018). On the other hand, semi-automatic recognition of landslide scars and analysis of changes in the spectral signature of land surface can provide rapid mapping (GUZZETTI et al., 2012). Machine learning techniques such as Artificial Neural Network (ANN), Maximum Likelihood (ML), Random Forest (RF) and Support Vector Machine (SVM) are pattern recognition methods, commonly applied to remote sensing images for landslide scars identification. Manfré et al. (2014) use SVM and ML to identify landslides in São Paulo State coast, in Brazil. The authors claim that SVM presented better performance than ML, especially when associated to the Normalized Difference Vegetation Index (NDVI). Moosava, Talebi and Shirmohammadi (2014) compare ANN and SVM to mapping landslides and the results have shown no significant differences between both methods. Additionally, much research uses ANN for landslide issues, for instance, the results shown by Chen et al. (2017) at Wanyuan area, China and by Kalantar et al. (2018) at Dodangeh watershed, Iran. Moreover, RF algorithms, as shown by Ghorbanzadeh et al. (2019), Chen et al. (2018), Stumpf and Kerle (2011) have also shown satisfactory outcomes for landslide scars detection.

The reduction of time to conduct the process is usually significant in hazard-related issues. Thus, in order to accelerate the decision making, one can employ data mining techniques, such as feature selection. In general, data mining is based on finding valuable structure in large datasets, as stated by Hastie, Tibdhirani and Friedman (2009). Feature selection discriminates the most important features from a dataset, avoiding overfitting (HASTIE; TIBDHIRANI; FRIEDMAN, 2009).

In this context, the aim of this study is to compare different pattern recognition techniques: ANN, ML, RF and SVM in order to identify which of them presents better results concerning landslide scars detection. Likewise, different feature selection methods are also compared aiming to understand the attribute's contribution for this detection.

This paper is an extended version of Uehara et al. (2019), presented in the XX Brazilian Symposium on GeoInformatics (GEOINFO 2019).

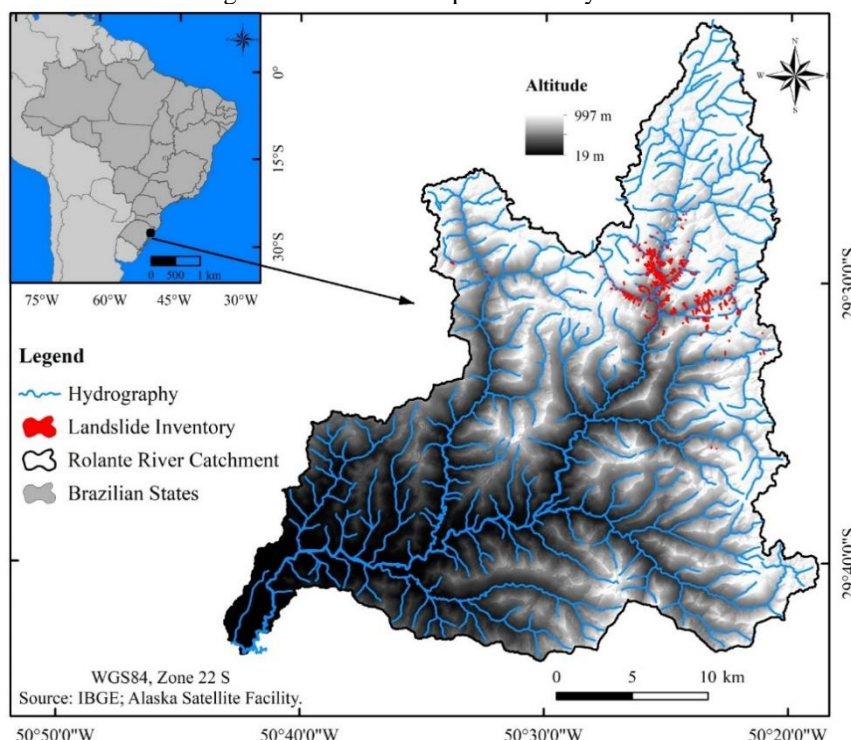
2 STUDY AREA

The Rolante River Catchment is located in the State of Rio Grande do Sul, Brazil (Figure 1), and it embraces three cities: Riozinho, Rolante and São Francisco de Paula. Its drainage area is 828 km², with altitudes varying from 19 to 997 m, according to data extracted from a Digital Elevation Model (DEM) provided by the Shuttle Radar Topography Mission (SRTM). This area is almost entirely located in the Serra Geral geomorphological unit, with a predominance of basaltic rocks and sandstone. According to Rossato (2011), the climate is characterized as very humid subtropical, with precipitation regime distributed throughout the year, with annual averages between 1700 and 2000 mm.

On January 5th, 2017, there was a landslide event in the upstream area of Rolante River Catchment triggered by an extreme precipitation event. The rains lasted for approximately four hours, and local private measurers estimated volumes between 90 and 272 mm (SEMA, 2017). These rains moved a large amount of material from the slopes, generating a natural dam on the Mascarada river, a tributary of the Rolante river, with subsequent rupture of this barrier and consequent flash flood, reaching Rolante city.

Previous works identified approximately 300 landslide scars in this region (GAMEIRO et al., 2019; QUEVEDO et al., 2019a; QUEVEDO et al., 2019b).

Figure 1 – Location map of the study area.

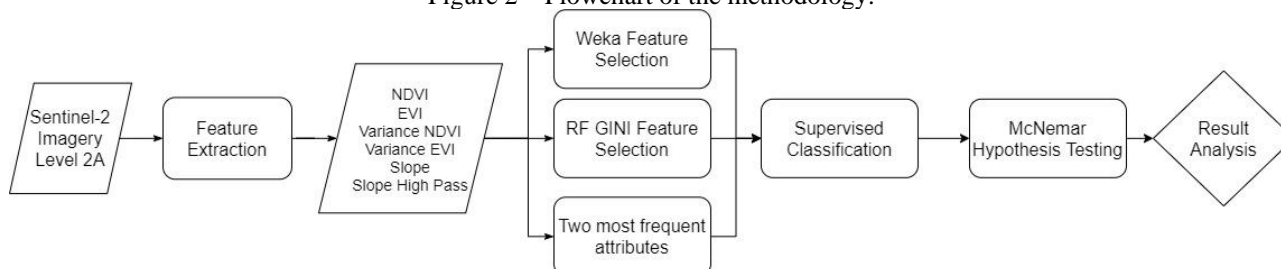


Source: The authors (2020).

3 METHODOLOGY

To fulfill the proposed objective, 11 features were used. From Sentinel: band 02 – Blue (B02), band 03 – Green (B03), band 04 – Red (B04), and band 08 – Near Infra-Red (B08); NDVI; Enhance Vegetation Index (EVI); NDVI Variance Texture; EVI Variance Texture. From SRTM: Slope (degrees) and High Pass Filter applied to Slope, with a 3x3 kernel size. All features were ranked in order of importance using Weka software and its functions Correlation Attribute Evaluation and RF GINI regressor. Then, four image classifiers were applied: a parametric classifier, the ML and three non-parametric: ANN, RF and SVM. The methodological process of this study is depicted in Figure 2.

Figure 2 – Flowchart of the methodology.

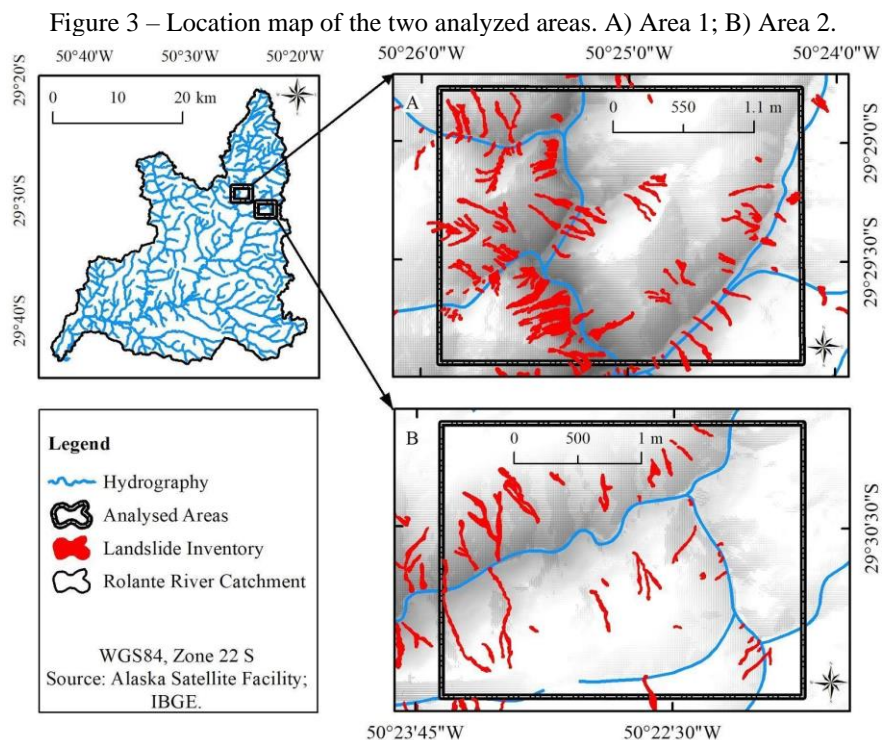


Source: The authors (2020).

The identification of landslide scars usually presents better results when high spatial resolution images are used (JOYCE et al., 2009). Considering that it provides orthorectified reflectance products of Bottom-of-Atmosphere (BOA), the imagery from the MultiSpectral Imager (MSI) sensor on board Sentinel-2 and product level 2A was used. For the purpose of this study, among all products available for Level-2A, only 10m spatial resolution data was used. The original images were clipped in for the two areas of interest, each one containing 6 km² (Figure 3). The choosing criteria for these areas considered that both contained a significant amount of

landslide scars presented in the landslide inventory (QUEVEDO et al., 2019a): Area 1 contains 91 landslide scars with 39 ha whilst Area 2 contains 34 landslide scars with approximately 16 ha. The Sentinel scenes selected are from March 30th, 2020 for Area 1 and February 09th, 2019 for Area 2.

From the original images, a feature extraction process was performed, using features based on Gerente et al. (2017a), Gerente et al. (2017b) and Joyce et al (2009). The EVI and NDVI, which consider NIR and red wavelengths for their computation, were directly downloaded from Sentinel-2 products. NDVI values are used to detect varying densities of vegetation coverage which could be used for natural disasters (BHANDARI; KUMAR; SINGH, 2012). Moreover, textures can characterize the heterogeneity of classes. This concept is related to the spatial distribution of intensity values; hence, it contains information regarding rugosity, regularity, contrast, and others (RUIZ; FDEZ-SARRÍA; RECIO, 2004). Among the statistical features, variance extracted from both vegetation indices was used to characterize texture. This step was performed using ENVI 4.7.



The original 30m spatial resolution DEM, from SRTM, was resampled to 10m using the nearest neighbor method available at SAGA toolbox resampling function in QGIS 3.4. SRTM data was chosen because no other data with higher spatial resolution was available free of charge. Tests were not realized to prove whether the interpolation presented an increase in the quality of the DEM data. However, this procedure was necessary in order to assure the same pixel size for all attributes in the stack, a requisite for the classification process performed. The slope was extracted from the DEM. After that, in ENVI 4.7, a high pass filter was applied to the slope to identify the valley regions and aid the differentiation between landslide scars and bare soil *per se*.

Hall (1999) defines feature selection as a learning step that focuses on the most useful data aspects for analysis and feature prediction. This step was conducted using the Correlation Attribute Evaluation Function available on Weka, which ranks the importance of data based on the Pearson's correlation. All the features were ranked in order of importance and the first five were selected. According to Raykov and Marcoulides (1999), the principle of parsimony reflects the notion that researchers should strive for simple measurement models that use the minimum number of parameters needed to explain a given phenomenon, so these five features were chosen to test whether half of the variables were able to map landslides scars and, consequently, make the model more parsimonious. Furthermore, another method used to rank the importance of variables

was the GINI importance in the RF model. Moreover, the features that were present among the first five in the importance rank from RF and Weka were selected for another round of classifications.

With all selected attributes, the classification assessment was performed via holdout method where testing samples are given independently of training samples (KIM, 2009). The image size was approximately 282x212 pixels, in which 200 training and 60 testing samples were used for each class. In order to assure our decision about the classes, Google Earth's high spatial resolution imagery from different dates were consulted. Visual interpretation of this imagery allowed the identification of the main classes present in the study area. Hence, the classes were: Forest, Shadowed Forest, Grass, Bare Soil and Landslide.

3.1 Classification methods

The analysis of different classifiers and features for detecting landslides scars aims to present the best performance available in order to attend risk assessments in urgent situations. Considering that, it is important to take into account the computational efforts, time and feasibility of such methods presented hereafter.

Artificial Neural Networks (ANN) is a supervised, non-parametric classification method, which is inspired on human brain functioning, composed of a variety of processing units, called neurons, that work in parallel classifying input data in output classes. Generally, a feed-forward multi-layer network is adopted. It typically consists of three layers—input, output, and a hidden layer between the first two—with a sufficient number of neurons in each layer (AURORA; DAS GUPTA; GUPTA, 2004). This method uses the error backpropagation algorithm (RUMELHART; HINTON; WILLIAMS, 1986), which consists of minimizing the output errors.

Maximum Likelihood (ML) is a supervised classification method determined by the Bayes theorem and employs a discriminant function to assign pixels to user-defined classes with the maximum likelihood (PAWLUSZEK; BORKOWSKI; TAROLLI, 2018). According to the author, ML continues to be the most widely used parametric classification algorithm. Hastie, Tibdhirani and Friedman (2009) state that the feature is assigned to a class with the highest probability of belonging, considering their correlation and variance. A probability function describes the distribution of reflectance values and evaluates the possibility of a pixel to belong to a certain category. It is important to emphasize that, for this study, a gaussian ML was used.

Random Forest (RF) is an ensemble-learning non-parametric method which uses the Bootstrap Aggregation technique and performs as classifier and regressor, combining decision trees. In the classification, the input data is randomly selected and separated into subsets. To each sub-sample, a decision tree is constructed, then the RF model selects the more frequent result. The RF predictor model also uses the GINI importance ranking the most important features used in the model, hence being also a feature selection tool. Another highlight is that the RF model measures the model statistical quality by the Out-Of-Bag (OOB) with samples not used for the classification. Two parameters are to be defined by the user: the number of trees and the number of the features, this allows the algorithm to create trees that have high variance and low bias (BREIMAN, 2001; HASTIE et al., 2009; SOTHE, 2019).

Support Vector Machine (SVM) is a machine learning technique, based on statistical learning theory which transforms original input space into a higher-dimensional feature space to find an optimal separating hyperplane (VAPNIK, 1998; KAVZOGLU; COLKESEN, 2009; ABE, 2005). The goal of the optimal separating hyperplane is a correct discrimination between two sorts of samples while maximizing the classification margin (HUANG; ZHAO, 2018) avoiding under and overfitting. According to Feizizadeh et al. (2017), the resulting SVM classifications are affected by the choice of the kernel function and among different possibilities of available kernels, the Radial Basis Function (RBF) has been found the most feasible and reliable to produce susceptibility maps. Based on that, the authors chose RBF kernel for SVM.

3.2 Feature Selection

According to Yu and Liu (2004) and Chen et al. (2018), the role of feature selection is to reduce the number of features used aiming the elimination of redundant information in addition to finding sensitive features optimal combination. Chen et al. (2018) adds that feature selection algorithms can be split into filter,

wrapper and embedded methods. The filter method, more suitable for large scale dataset, eliminates the classifier training steps, ranking the subset (CHEN et al., 2018). Moreover, Chen et al. (2018), state that the wrapper method depends on the machine learning algorithm performance evaluating merits of the feature subset and Tang, Alelyani and Liu (2014) add that this approach uses the predictive accuracy of a learning algorithm predefined for determining the feature selection quality. Lastly, the embedded method is used as component of the machine learning algorithm, considering the bias of the classifier, as stated by Tang, Alelyani and Liu (2014) and Chen et al. (2018), being decision tree the most typical one. This study concerns a filter method, entitled Correlation Attribute Evaluation (CAE) and an embedded method, which is the GINI importance inside the RF classification model.

According to Hall (2009) and Aldehim and Wang (2017), CAE is a filtering feature selection tool which ranks features considering a correlation-based feature heuristic evaluation function, assessing, therefore, the efficacy of individual features taking into account their predicting strength and how strong they are inter-correlated. On the other hand, Gini is a splitting metric between two classes measuring the weighted average between the trees (BREIMAN, 2001; BREIMAN, 2017; ALER et al., 2020). Moreover, Maretto et al. (2015), have proven that feature selection can reduce both amount of data and computational time, being therefore an asset for detection of land change and land use.

3.3 Classification Accuracy and Comparison

For comparing classification results, both elements of Spatial Data Quality are considered: thematic accuracy and completeness. For thematic accuracy, both overall accuracy and *kappa* index are used, and for completeness, Commission and Omission Errors are computed, as stated by ISO (2013).

In addition to that, the McNemar hypothesis test is carried out. This test was proposed by Bradley (1968) and Agresti (1996) and is explained with examples by Foody (2004). In a nutshell, this hypothesis test statistically compares two classifications with related samples by evaluating their confusion matrices using a Chi-squared distribution.

4 RESULTS AND DISCUSSION

Feature selection results are shown in Table 1, revealing that each area and method presented different rankings. None of them ranked the same top-5 features, however, the B03 and the NDVI variance were present in all cases. Considering that, we used four datasets for each area and each classifier: all the attributes; the five more important by the Weka ranking; the five more important by the RF regressor ranking; B03 and the NDVI variance. In this sense, we obtained 32 classifications, 16 for each study area.

As mentioned above, among the top-5 features for both selection methods and areas, B03 and the NDVI variance texture are present. B03 corresponds to the green band in the visible range of the electromagnetic spectrum, presenting high values for vegetated areas. Since the study areas are in dense forested regions with Bare Soil and Landslide, this feature could be useful to describe these classes, presenting high values for the former, and low for the two latter. In that sense, the fact that the variance of the NDVI feature was also selected can be explained by the relevance of the abrupt NDVI change in Landslide scars compared to their location site. In this case, most of the scars have occurred in forested areas at the slopes.

Table 1 – Feature Selection and Selection for both studied areas.

Rank	Weka CAE		GINI RF	
	Area1	Area2	Area1	Area2
1	NDVI	Texture: NDVI variance	Texture: NDVI variance	Texture: NDVI variance
2	EVI	Texture: NDVI variance	NDVI	Sentinel B03 (Green)
3	Sentinel B03 (Green)	Sentinel B03 (Green)	DEM	Slope
4	Texture: NDVI variance	Slope	Sentinel B3 (Green)	Sentinel B04 (Red)
5	DEM	Sentinel B04 (Red)	Slope	DEM
6	Sentinel B02 (Blue)	Sentinel B02 (Blue)	Sentinel B04 (Red)	Texture: EVI variance
7	Sentinel B04 (Red)	Texture: EVI variance	Sentinel B02 (Blue)	Sentinel B02 (Blue)
8	NIR	Texture: EVI variance	Texture: EVI variance	Sentinel B08 (NIR)
9	Texture: EVI variance	Sentinel B08 (NIR)	Sentinel B08 (NIR)	Texture: NDVI variance
10	Slope	DEM	EVI	EVI
11	Slope High Pass Filter	Slope High Pass Filter	Slope High Pass Filter	Slope High Pass Filter

Source: The authors (2020).

One should notice that the best ranked features in at least one method or area were: B03, Texture NDVI variance, NDVI, EVI, B04, DEM and slope. Both NDVI and EVI are vegetation indices, so their contribution is assembled with the B03 explanation, regarding the identification of (non-)vegetated areas. Nonetheless, as the feature selection differentiates correlations between attributes, it is expected that the red band is not close to its derived products (NDVI, EVI, Textures). Besides, they can help with the mapping of different vegetation types, as Grass and Forest. According to Fox et al. (2004), the soil line is a linear relationship between the NIR and red reflectance of Bare Soil. In that sense, since B04 corresponds to the red band, it presents relevant information to Bare Soil detection. When it comes to geomorphometric features, as the DEM and slope, they stand out due to the spatial concentration of landslide occurrences. In other words, there is a pattern between the altitude and the slope angle where landslides happen. This pattern can help the model to recognize the landslide scars, considering that elevation and slope values are similar among them. Related results can be seen in Quevedo et al. (2020) and Oliveira et al. (2019). The High Pass Slope Filter achieved the lowest position for all the rankings. This could be explained by the fact that it could present similar information when compared to the slope and DEM.

After performing the classification including all 11 features, feature selection was performed, and three other datasets were used as input. Tables 2 and 3 show the Overall Accuracy (OA) and *kappa* index for each classifier related to a feature selection method for Areas 1 and 2.

For Area 1, concerning the results with all features as input, SVM and ML presented the highest values, with 96.33% and 95.33% of OA, respectively, as can be seen on Table 2. Since their values are similar, the McNemar statistical test was applied, showing that they are not different, considering 5% of significance, which, based on that, cannot lead to any conclusion concerning their performance. The datasets built from the RF GINI and Weka feature selection showed a general decrease in the accuracy values, especially for the RF classifier. The effects on the RF classifier were expected, since it is not recommended to perform a feature selection before this type of classification, considering that this procedure is already implemented by the algorithm. However, tests after the feature selection were made in order to standardize the methodology for all cases. On the other hand, the other classification algorithms maintained a relative stability in the results, presenting less than 3% of changes in their accuracy values.

Besides, we can verify that the use of only two variables, B03 and NDVI variance, can present an acceptable accuracy for ML, SVM, and RF with OA values above 77%. The reduction to a couple of features as input for ML caused a decrease of 11% on the OA. Nevertheless, comparing to the others, this classifier still presented the highest accuracy values with OA of 84%. Based on that, one can affirm that ML was the most parsimonious model with two variables for Area 1.

Table 2 – Classification results for Area 1. Overall Accuracy (%) and kappa index values for the classification algorithms in each feature selection method. All: 11 attributes are considered in the classification. RF: only top-5 features selected by RF. Weka: only top-5 features selected by Weka. B03 + NDVI variance: only these two features.

Area 1	Feature Selection	ANN		ML		SVM		RF	
		OA (%)	kappa	OA (%)	kappa	OA (%)	kappa	OA (%)	kappa
	All	86.33%	0.8292	95.33%	0.9417	96.33%	0.9542	90.57%	0.8821
	RF GINI	84.67%	0.803	93.33%	0.9167	93.67%	0.9208	84.85%	0.8106
	Weka CAE	85.67%	0.8208	94.33%	0.9292	90.33%	0.8792	84.18%	0.8022
	B03 + NDVI Variance	57.67%	0.4708	84.00%	0.8	77.67%	0.7208	77.10%	0.7138

Source: The authors (2020).

For Area 2, the OA and kappa index for the classifications are shown in Table 3. The best performance with all features included was achieved by RF (92.33%) and SVM (92.00%). In order to compare these values, the McNemar test was applied, presenting a negative result, meaning that, on 5% of significance, one cannot deny the hypothesis that the two accuracy values are different, concluding that the RF indeed presented a better performance than the SVM. On this approach, the ANN presented the lowest result (74.33%), being this value decreased to 59% when using the first five features selected by Weka. However, the combination of ANN and the two most frequent attributes (B03 + NDVI variance) increased the OA value in 7% when compared to all features, reaching an acceptable OA value (81.67%). The same value is verified for the SVM with two features, which decreased only 10% of the accuracy when reducing about 80% of its input data.

Table 3 – Classification results for Area 2. Overall Accuracy (%) and kappa index values for the classification algorithms in each feature selection method. All: the 11 attributes are considered in the classification. RF: only top-5 features selected by RF. Weka: only top-5 features selected by Weka. B03 + NDVI variance: only these two features.

Area 2	Feature Selection	ANN		ML		SVM		RF	
		OA (%)	kappa	OA (%)	kappa	OA (%)	kappa	OA (%)	kappa
	All	74.33%	0.6792	84.67%	0.8083	92.00%	0.9000	92.33%	0.9042
	RF GINI	81.00%	0.7625	90.33%	0.8792	92.00%	0.9000	89.33%	0.8667
	Weka CAE	59.00%	0.4875	88.33%	0.8542	92.00%	0.9000	89.67%	0.8708
	B03 + NDVI Variance	81.57%	0.7708	78.88%	0.725	81.67%	0.7708	72.00%	0.6500

Source: The authors (2020).

Considering all of the above, for the two-features approach, ANN and SVM revealed to be the most parsimonious models applied to Area 2. Moreover, it is interesting to point out that RF presented the most accurate results for the input considering all features, however, the worst for the two-features approach. This reveals the classifier’s high level of sensibility at the classification’s accuracy with previously applied feature selection procedure.

For completeness, the commission (CE) and omission (OE) errors regarding each classification product derived from the procedures including all features are shown on Tables 4 and 5. For both areas, the classification algorithm presenting the highest error rates was the ANN. For Area 1, when comparing completeness, SVM presented the best prediction with maximum CE of 7.69% for the class Grass and OE and 8.33% for Bare Soil. For Landslide detection, SVM also presented the greatest outcomes, with 0.00% of OE and 3.33% of CE. On the contrary, ANN presents the highest error rates for landslides, which are 11.86% of CE and 13.33% of OE.

For Area 2, SVM also achieved the best performance either in general or considering only the Landslide class. In the general scenario, the highest error rates were 16.67% of CE for Shadowed Forest and 20.00% of OE for Forest. Similarly, to Area 1, for Area 2 classifications, ANN also presented the highest error rates for both general scenario and Landslide class: 37.78% of CE and 53.33% of OE.

For both areas, apart from the ANN classification, in general, Landslide detection showed greater outcomes when compared to other classes. A possible explanation for that relies on the fact that all extracted attributes were chosen specifically for landslide scar detection. In a nutshell, in all possible studied scenarios, SVM model showed the most satisfactory results while ANN showed the opposite.

Table 4 – Percentage of Commission (CE) and Omission (OE) errors for the four classifiers for Areas 1 without feature selection.

Area 1	ANN		ML		RF		SVM	
Class	CE	OE	CE	OE	CE	OE	CE	OE
Bare Soil	31.82	25.00	10.77	3.33	15.38	25.42	3.51	8.33
Forest	5.26	10.00	0.00	10.00	5.17	8.33	6.25	0.00
Grass	15.15	6.67	1.72	5.00	4.76	0.00	7.69	0.00
Landslide	11.86	13.33	7.81	1.67	5.36	8.62	0.00	3.33
Shadowed Forest	0.00	13.33	1.69	3.33	16.18	5.00	0.00	6.67

Source: The authors (2020).

Table 5 – Percentage of Commission (CE) and Omission (OE) errors for the four classifiers for Area 2 without feature selection.

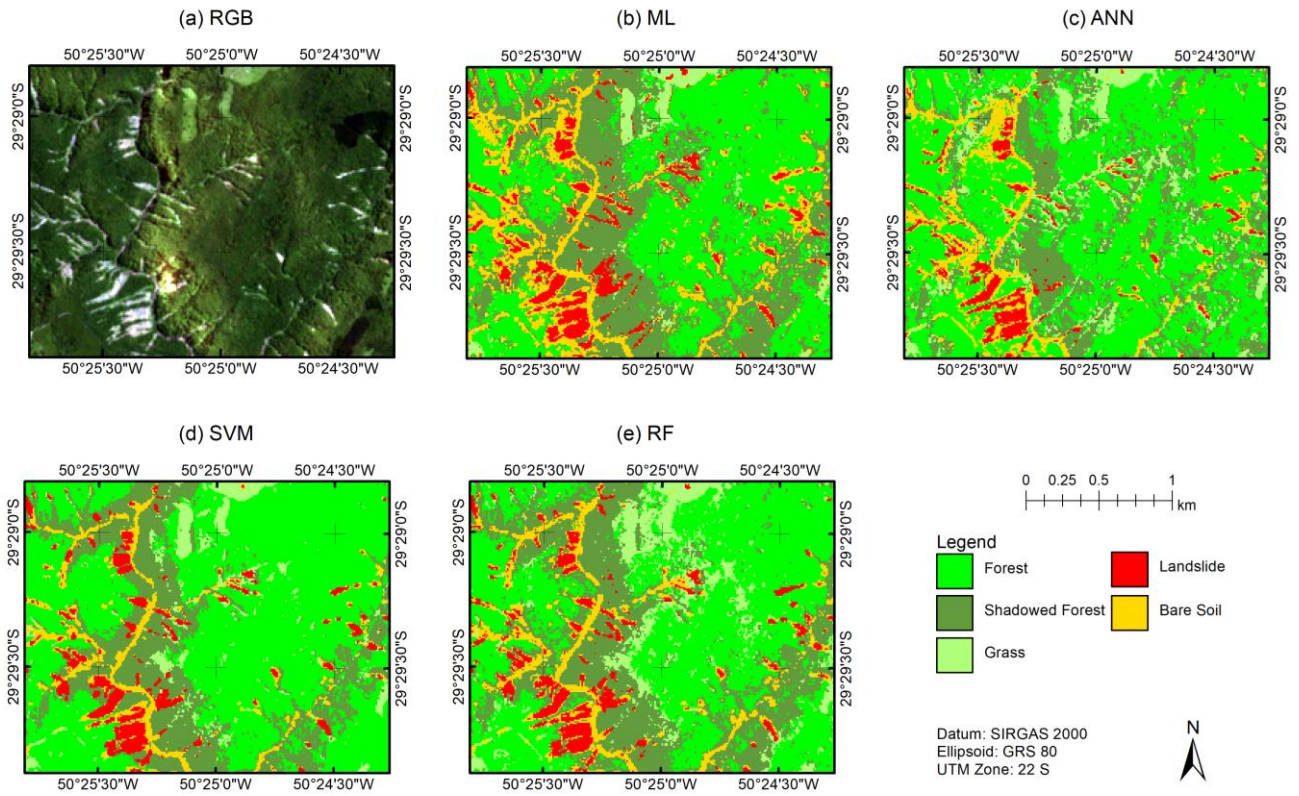
Area 2	ANN		ML		RF		SVM	
Class	CE	OE	CE	OE	CE	OE	CE	OE
Bare Soil	45.95	33.33	26.39	11.67	1.89	13.33	9.84	8.33
Forest	6.82	31.67	19.4	10.00	0.00	23.33	2.04	20.00
Grass	6.67	6.67	0.00	33.33	3.28	1.67	0.00	11.67
Landslide	37.78	53.33	10.61	1.67	10.45	0.00	7.69	0.00
Shadowed Forest	24.68	3.33	12.73	20.00	17.81	0.00	16.67	0.00

Source: The authors (2020).

When it comes to visual interpretation, the classification results for Area 1 are shown in Figures 4 and 5 whilst for Area 2, they are shown in Figures 6 and 7. Due to the large number of results, only the classification without feature selection as well as the classification using B03 and NDVI variance as inputs are shown herein. The point is to show the reader the visual effect of reducing attributes on classifications.

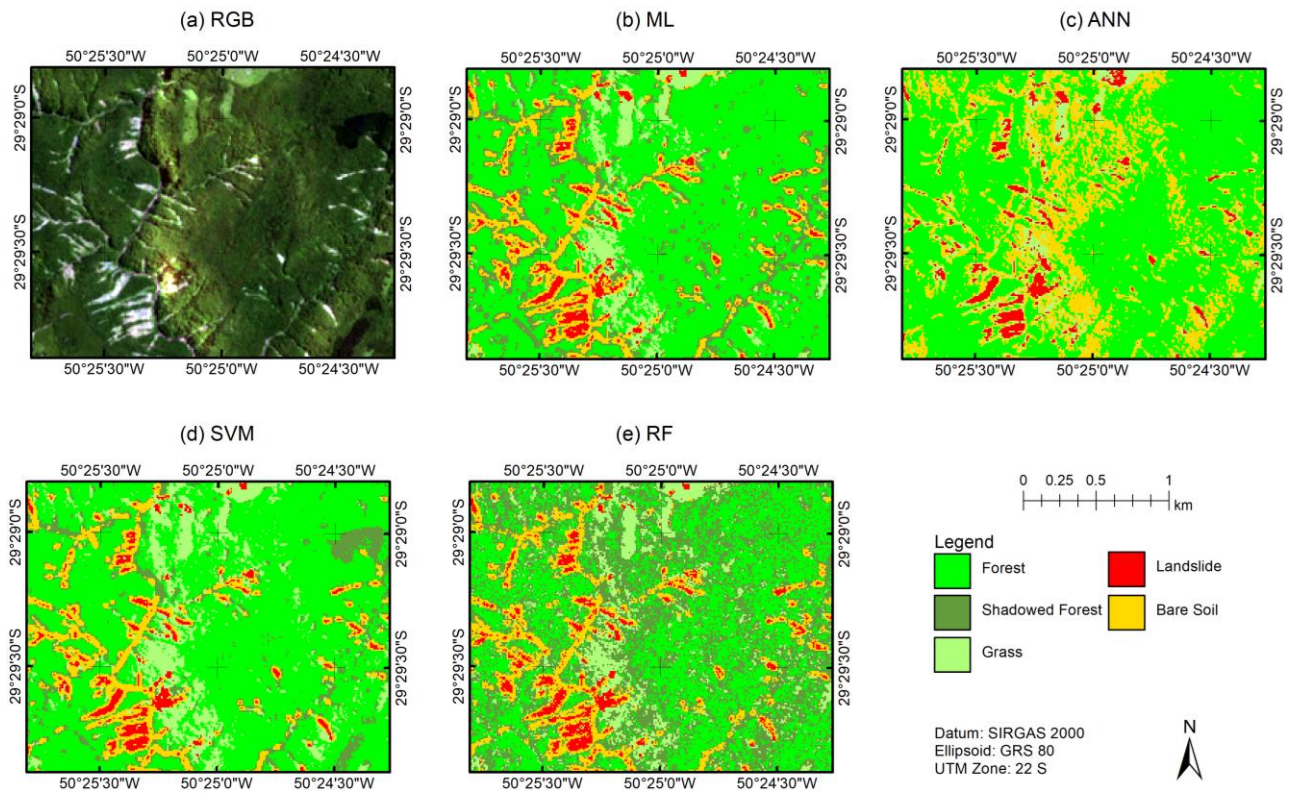
As before mentioned, for Area 1, the classifiers presented better OA values without the realization of previous feature selection. Figure 4 shows that, in general, the four methods presented similar results. All of them could detect an acceptable quantity of landslide scars, which can be identified at the RGB composition by their bright color and long shape. SVM presented the least noisy result, with well delineated forms and a reduced number of spare pixels (salt-and-pepper effect). Simple visual interpretation could conclude that the ANN had more variations on results compared to the others, having classified less Shadowed Forest areas.

Figure 4 – Classification products with all Features for Area 1.



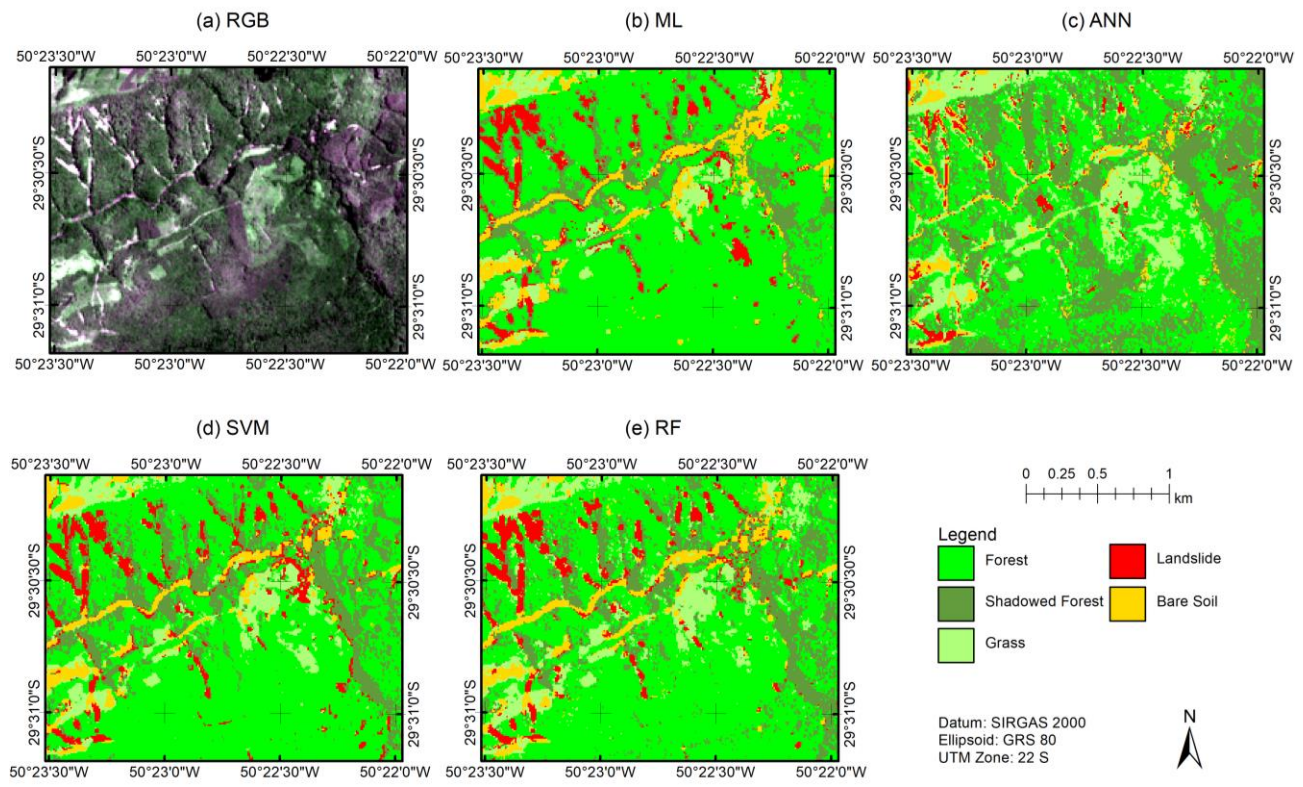
Source: The authors (2020).

Figure 5 – Classification products with B03 + NDVI variance Features for Area 1.



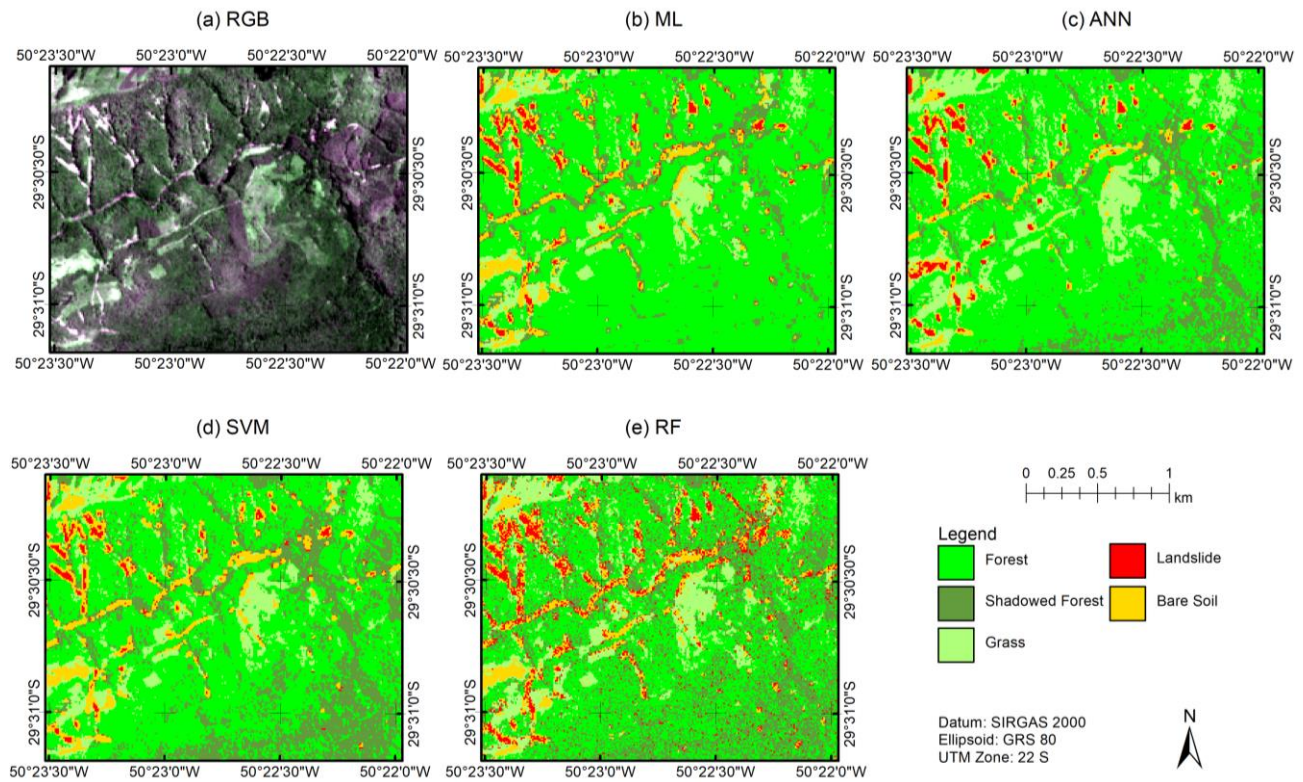
Source: The authors (2020).

Figure 6 – Classification products with all Features for Area 2.



Source: The authors (2020).

Figure 7 – Classification products with B03 + NDVI variance Features for Area 2.



Source: The authors (2020).

For Area 2 (Figure 6), the classifiers also had, generally, a similar result, except for the ANN, which classified less Landslide areas and a higher proportion of Shadowed Forest. Moreover, the ANN presented the noisiest output and could not properly delineate the Bare Soil trail crossing the scene horizontally (clearly

represented in the others). Analyzing Figures 5 and 7, where only two features were used, one can notice an increase in the proportion of Bare Soil classified in comparison with Figures 4 and 6 (without feature selection). This could be explained by the absence of a geomorphometric attribute, which could help to discriminate both classes by associating Bare Soil characteristics on high slope values as Landslide. Besides, the Shadowed Forest class is significantly underrepresented in Area 1, specially by the ANN, where almost any pixel was identified on that class. RF, on the contrary, classified a considerable amount of Shadowed Forest, however it showed the highest sensibility to salt-and-pepper effect for both areas. The difficulty to detect Shadowed Forest well might occur due to the absence of the B02 (green), since this type of vegetation might not present variations on the NDVI compared to Grass or Forest, however could have lower reflectance values for this band (i.e., dark green).

It is important to notice that, even with a significant reduction on the number of features used, the classifiers were still able to detect the majority of landslide scars, specially ML and SVM. This emphasizes the potential of this features for landslide detection. Besides, data spatialization in maps can provide relevant information regarding the classifier's performance, which cannot be identified only through the OA and kappa index values. This can be illustrated by observing Figure 5, where the SVM and the RF presented noticeably different results, however both showed OA values around 77%.

5 CONCLUSION

Most of the evaluated classifiers could effectively detect the Landslides scars, which is the main target of this research. Special attention should be given to SVM, which achieved among the best performances for both Area 1 and 2. Even though for the classification with all features, the RF algorithm outperformed the SVM in Area 2, SVM presented a more stable behavior of output values, being less affected to the changes due to the feature selection procedures.

On the other hand, the reduction in approximately 80% of the input features did not affect significantly the overall accuracy. When analyzing the result from the classification with only two features, the most parsimonious classification algorithms were the ML for Area 1 (OA of 84.00%), and ANN and SVM for Area 2 (OA of 81.66%). This demonstrates that the models used are robust and capable of performing the classification with few variables, as long as they are adequate.

Both RF GINI and Weka feature selection methods were able to rank relevant features, providing, in general, satisfactory accuracy results for the classifications with only five features (less than half of the original dataset). Moreover, the comparison of both methods and study areas allowed the selection of the two (B03 and NDVI variance) most frequent features amongst the first five ranked in all cases. These two features showed a great potential for landslide detection, providing more than 77% of OA in 75% of the cases analyzed.

It is important to emphasize that finding the most appropriate classifier is problem-dependent, meaning that a classifier can be best recommended for a specific study area and time. Therefore, it is crucial to test the best option for their specific case.

Moreover, it is important to point out that the classes used must be chosen thoroughly, as the supervised classification quality depends directly on that. Likewise, this study proves that good reference data (i.e. training and test samples) are crucial for achieving better outcomes.

For future studies, it is recommended to add segmentation processes before the classification in addition to testing other state-of-the-art pattern recognition approaches. Nevertheless, even though semi-automatic classification methods have proven to display satisfactory results, it does not exclude completely the importance of manual processing and interpreter interference. Semi-automatic algorithms still show some problems, which can be better managed through auxiliary data such as field work and visual interpretation corrections, in order to produce better classification results.

Acknowledgment

This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*

- Brasil (CAPES) - Finance Code 001 and the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq). The authors also thank the *Fundação de Amparo à Pesquisa do Estado de São Paulo* (FAPESP, grant no 2017/24086-2), and the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq, grant no 303360/2019-4).

Author's Contribution

Tatiana Dias Tardelli Uehara, Sabrina Paes Leme Passos Corrêa, Renata Pacheco Quevedo, Thales Sehn Körting, Luciano Vieira Dutra and Camilo Daleles Rennó.

Conceptualization; T.D.T.U, S.P.L.C and R.P.Q. Data curation; T.D.T.U, S.P.L.C and R.P.Q. Formal analysis; T.D.T.U, S.P.L.C and L.V.D. Investigation; T.D.T.U, S.P.L.C and R.P.Q. Methodology; T.D.T.U, S.P.L.C, R.P.Q., T.S.K., L.V.D. and C.D.R. Project administration; T.D.T.U. Supervision; T.S.K., L.V.D. and C.D.R. Validation; T.D.T.U. and S.P.L.C. Visualization; T.D.T.U, S.P.L.C and R.P.Q. Writing – original draft; T.D.T.U, S.P.L.C and R.P.Q. Writing – review & editing; T.S.K., L.V.D. and C.D.R.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- ABE, S. Support vector machines for pattern classification. **Springer**, 2005.
- ALDEHIM, G.; WANG, W. Determining appropriate approaches for using data in feature selection. **International Journal of Machine Learning and Cybernetics**, v. 8, n. 3, p. 915–928, 2017. Springer Berlin Heidelberg.
- ALER, R.; VALLS, J. M.; BOSTRÖM, H. Study of Hellinger Distance as a splitting metric for Random Forests in balanced and imbalanced classification datasets. **Expert Systems with Applications**, v. 149, p. 113264, 2020. Elsevier Ltd. Disponível em: <<https://doi.org/10.1016/j.eswa.2020.113264>>. .
- AGRESTI, A. **An introduction to categorical data analysis**. John Wiley & Sons, 2018.
- AURORA, M.; DAS GUPTA, A.; GUPTA, R. An artificial neural network approach for landslide hazard zonation in the Bhagirathi (Ganga) Valley, Himalayas. **International Journal of Remote Sensing**, v. 25, n. 3, p. 559–572, 2004. Taylor & Francis.
- BHANDARI, A.; KUMAR, A.; SINGH, G. Feature extraction using Normalized Difference Vegetation Index (NDVI): A case study of Jabalpur city. **Procedia technology**, v. 6, p. 612–621, 2012. Elsevier.
- BRADLEY, J. V. **Distribution-free statistical tests**. Prentice-Hall, Englewood Cliffs, New Jersey, 388 p, 1968.
- BREIMAN, L. Random forests. **Machine learning**, v. 45, n. 1, p. 5–32, 2001. Springer.
- BREIMAN, L. **Classification and Regression Treest**. Routledge, 2017.
- CHEN, W.; POURGHASEMI, H.; KORNEJADY, A.; GEODERMA, N. Z.-; 2017, U. Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. **Elsevier**, 2017. Available at: <<https://www.sciencedirect.com/science/article/pii/S0016706117303890>>.
- CHEN, F.; YU, B.; LI, B. A practical trial of landslide detection from single-temporal Landsat8 images using contour-based proposals and random forest: a case study of national Nepal. **Landslides**, v. 15, n. 3, p. 453–464, 2018. Springer.
- CHEN, W.; LI, X.; HE, H.; WANG, L. A review of fine-scale land use and land cover classification in open-pit mining areas by remote sensing techniques. **Remote Sensing**, v. 10, n. 1, 2018.
- DAVIES, T. Landslide hazards, risks and disasters: Introduction, In: SHRODER, J. F.; DAVIES, T.

- (Ed.). **Landslides hazards, risks and disasters**. Elsevier Inc., p. 1-16, 2015. DOI. 10.1016/B978-0-12-396452-6.00001-X.
- FEIZIZADEH, B.; ROODPOSHTI, M. S.; BLASCHKE, T.; ARYAL, J. Comparing GIS-based support vector machine kernel functions for landslide susceptibility mapping. **Arabian Journal of Geosciences**, v. 10, n. 5, p. 122, 2017.
- FOODY, G. Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy. **Photogrammetric Engineering Remote Sensing**, v. 70, n. 5, p. 627-633, 2004. DOI. 0099-1112/04/7005-0627/\$3.00/0.
- FOX, G. A.; SABBAGH, G.; SEARCY, S.; YANG, C. An automated soil line identification routine for remotely sensed images. **Soil Science Society of America Journal**, v. 68, n. 4, p. 1326-1331, 2004. Soil Science Society.
- GAMEIRO, S.; QUEVEDO, R. P.; OLIVEIRA, G. G.; RUIZ, L. F. C.; GUASSELLI, L. A. Análise e correlação de atributos morfométricos e sua influência nos movimentos de massa ocorridos na Bacia do Rio Rolante, RS. In: XIX SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 19, 2019, Santos. **Anais...** Santos: INPE, 2019. p. 2880-2883.
- GERENTE, J.; PLETSCH, M. A. J. S.; SOTHE, C.; FRANCISCO, C. N. Classificação Semiautomática De Cicatrizes De Movimentos De Massa Por Meio De Técnicas De Detecção De Mudanças. **Revista Brasileira de Geomorfologia**, v. 18, n. 4, 2017a.
- GERENTE, J.; SÖTHER, C.; NEGRÃO, P.; KÖRTING, T. S. Mass movements' scars classification using data mining techniques. In: XVIII SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 18, 2017. **Anais...** Santos: INPE, 2017b. p. 3553-3560.
- GHORBANZADEH, O.; BLASCHKE, T.; GHOLAMNIA, K.; et al. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. **Remote Sensing**, v. 11, n. 2, p. 196, 2019. Multidisciplinary Digital Publishing Institute.
- GUZZETTI, F.; MONDINI, A. C.; CARDINALI, M.; FIORUCCI, F.; SANTANGELO, M.; CHANG, K. Landslide inventory maps: New tools for an old problem. **Earth-Science Reviews**, v. 112, p. 42-66, Apr 2012. DOI. 10.1016/j.earscirev.2012.02.001
- HALL, M. A. **Correlation-based feature selection for machine learning**. 1999. 198 f. Tese (Doutorado em Filosofia) - University of Waikato. Hamilton, NewZealand, 1999.
- HAQUE, U.; SILVA, P. F.; DEVOLI, G.; PILZ, J.; ZHAO, B.; KHALOUA, A.; WILOPO, W.; ANDERSEN, P.; LU, P.; LEE, J.; YAMAMOTO, T.; KEELLINGS, D.; WU, J.; GLASS, G. E. The human cost of global warming: deadly landslides and their triggers (1994-2014). **Science of the Total Environment**, v. 682, p. 673-684, Mar 2019. DOI. 10.1016/j.scitotenv.2019.03.415
- HASTIE, T.; TIBDHIRANI, R.; FRIEDMAN, J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2. ed. [S.l.]: **Springer**, 2009. ISSN 00349437
- HUANG, Y.; ZHAO, L. Review on landslide susceptibility mapping using support vector machines. **CATENA**, v. 165, p. 520-529, 2018.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 19157:2013**: Geographic information - Data quality. [S.l.], 2013. 146 p. Available in: <<https://www.iso.org/obp/ui/#iso:std:iso:19157:ed-1:v1:en>>.
- JOYCE, K. E.; BELLISS, S. E.; SAMSONOV, S. V.; MCNEILL, S. J.; GLASSEY, P. J. A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters. **Progress in Physical Geography**, v. 33, n. 2, p. 183-207, 2009. Sage Publications Sage UK: London, England.
- KALANTAR, B.; UEDA, N.; AL-NAJJAR, H. A. H.; et al. Landslide susceptibility mapping at Dodangeh watershed, Iran using LR and ANN models in GIS. v. 10790, p.107901D, 2018. **International Society for Optics and Photonics**. v. 10790. p. 107901D

- KAVZOGLU, T.; COLKESEN, I. A kernel functions analysis for support vector machines for land cover classification. **International Journal of Applied Earth Observation and Geoinformation**, v. 11, n. 5, p. 352–359, 2009.
- KIM, J.-H. Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. **Computational statistics & data analysis**, v. 53, n. 11, p. 3735–3745, 2009. Elsevier.
- LI, T.; FONG, S.; L., W. K. K.; et al. Fusing wearable and remote sensing data streams by fast incremental learning with swarm decision table for human activity recognition. **Information Fusion**, v. 60, p. 41–64, 2020.
- LUPIANO, V.; RAGO, V.; TERRANOVA, O. G.; IOVINE, G. Landslide inventory and main geomorphological features affecting slope stability in the Picentino river basin (Campania, southern Italy). **Journal of Maps**, v. 15, n. 2, p. 131-141, Jan 2019. DOI. 10.1080/17445647.2018.1563836.
- MANFRÉ, L.; JUN SHINOHARA, E.; BEZERRA SILVA, J.; et al. Method For Landslides Identification At The Sao Paulo State Coast, BRAZIL. 2014. **Geociencias**, v. 33, n. 1, p. 172-180.
- MARETTO, R. V.; KÖRTING, T. S.; CASTEJON, E. F.; FONSECA, L. M. G.; SANTOS, R. Spectral attributes selection based on data mining for remote sensing image classification. In: BRAZILIAN SYMPOSIUM ON GEOINFORMATICS, 16., 2015, Campos do Jordão. **Proceedings...** Campos do Jordão: INPE, 2015. p. 155–161.
- MOOSAVI, V.; TALEBI, A.; GEOMORPHOLOGY, B. S.-; 2014, U. Producing a landslide inventory map using pixel-based and object-oriented approaches optimized by Taguchi method. **Elsevier**, 2014. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169555X13004704>>.
- OLIVEIRA, G. G.; RUIZ, L. F. C.; GUASSELLI, L. A.; HAETINGER, C. 2019. Random forest and artificial neural networks in landslide susceptibility modeling: a case study of the Fão River Basin, Southern Brazil. **Natural Hazards**, v. 99, p.1049-1073, Sep 2019. DOI. 10.1007/s11069-019-03795-x.
- PAWLUSZEK, K.; BORKOWSKI, A.; TAROLLI, P. Sensitivity analysis of automatic landslide mapping: numerical experiments towards the best solution. **Landslides**, v. 15, n. 9, p. 1851–1865, 2018.
- QIN, Y.; LU, P.; LI, Z. Landslide inventory mapping from bitemporal 10 m Sentinel-2 images using change detection based Markov Random Field. In: **ISPRS TC III Mid-term Symposium “Developments, Technologies and Applications in Remote Sensing”**, v. 42, n. 3, Apr 2018. DOI. 10.5194/isprs-archives-XLII-3-1447-2018.
- QUEVEDO, R. P.; GUASSELLI, L. A.; OLIVEIRA, G. G.; RUIZ, L. F. C. Modelagem de áreas suscetíveis a movimentos de massa: avaliação comparativa de técnicas de amostragem, aprendizado de máquina e modelos digitais de elevação, **Geociências (São Paulo. Online)**, São Paulo, v. 38, n. 3, p. 781-795, 2019a. DOI. 10.5016/geociencias.v38i3.14019.
- QUEVEDO, R. P.; OLIVEIRA, G. G.; GAMEIRO, S.; RUIZ, L. F. C.; GUASSELLI, L. A. (2019b) Modelagem de áreas suscetíveis a movimentos de massa com redes neurais artificiais. In: XIX SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 19, 2019, Santos. **Anais...** Santos: INPE, 2019. p. 2910-2913.
- QUEVEDO, R. P.; OLIVEIRA, G. G.; GUASSELLI, L. A. Modelagem de áreas suscetíveis a movimentos de massa a partir de redes neurais artificiais. **Anuário do Instituto de Geociências (UFRJ. IMPRESSO)**, Rio de Janeiro, 2020. No prelo.
- RAMOS-BERNAL, R. N.; VÁZQUEZ-JIMÉNEZ, R.; ROMERO-CALCERRADA, R.; ARROGANTE-FUNES, P.; NOVILLO, C. J. Evaluation of unsupervised change detection methods applied to landslide inventory mapping using ASTER imagery. **Remote Sensing**, v. 10, p. 1-24, Dec 2018. DOI. 10.3390/rs10121987
- ROSSATO, M. S. **Os Climas do Rio Grande do Sul: variabilidade, tendências e tipologia**. 2011. 253 p. Tese (Doutorado em Geografia) – Programa de Pós-Graduação em Geografia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2011.

- RAYKOV, T.; MARCOULIDES, G. A. On desirability of parsimony in structural equation model selection. **Structural Equation Modeling: A Multidisciplinary Journal**, v. 6, n. 3, p. 292–300, 1999. Taylor & Francis.
- RUIZ, L. A.; FDEZ-SARRÍA, A.; RECIO, J. A. Texture feature extraction for classification of remote sensing data using wavelet decomposition: A comparative study. **International Archives of Photogrammetry and Remote Sensing**, v. 35, n. part B, p. 1109-1115, 2004.
- RUMELHART, D.; HINTON, G.E.; WILLIAMS R. J. Learning internal representation. In: **Parallel Distributed Processing: Explorations in the Microstructure of Cognition**, 1st ed., Foundations, MIT Press, Cambridge, MA, 1986.
- SECRETARIA DO AMBIENTE E DESENVOLVIMENTO SUSTENTÁVEL (SEMA). **Diagnóstico Preliminar**: Descritivo dos eventos ocorridos no dia 5 de janeiro de 2017 entre as regiões dos municípios de São Francisco de Paula e Rolante/RS. Secretaria do Ambiente e Desenvolvimento Sustentável, Porto Alegre, 2017. 26 p
- SOTHE, C. **Mapping successional forest stages and tree species in subtropical areas integrating UAV-Based photogrammetric point cloud and hyperspectral data**: comparison of Machine and Deep Learning algorithms. 2019. Tese (Doutorado em Sensoriamento Remoto) – Programa de Pós-Graduação em Sensoriamento Remoto, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2019.
- STUMPF, A.; KERLE, N. Object-oriented mapping of landslides using Random Forests. **Remote sensing of environment**, v. 115, n. 10, p. 2564–2577, 2011. Elsevier.
- TANG, J.; ALELYANI, S.; LIU, H. Feature Selection for Classification: A Review. **Data Classification: Algorithms and Applications**. p.37, 2014. Boca Raton, FL, USA: CRC Press.
- UEHARA, T. D. T.; CORREA, S. P. L. P.; QUEVEDO, R. P.; KÖRTING, T. S.; DUTRA, L. V.; RENNÓ, C. D. Classification algorithms comparison for landslide scars. In: BRAZILIAN SYMPOSIUM ON GEOINFORMATICS, 20., 2019, São José dos Campos. **Proceedings...** São José dos Campos: INPE, 2019. p. 158-169.
- VAPNIK, V. N. Statistical learning theory. Adaptive and learning systems for signal processing. **Communications and Control**, v. 2, p. 1–740, 1998.
- YU, L.; LIU, H. Efficient feature selection via analysis of relevance and redundancy. **Journal of Machine Learning Research**, v. 5, p. 1205–1224, 2004.

First author biography



Tatiana Dias Tardelli Uehara, São Paulo/SP - Brazil, 1993. Geographer by the University of São Paulo (USP), São Paulo/SP. Master student of Remote Sensing at Brazil's National Institute for Space Research (INPE), São José dos Campo/SP.



Esta obra está licenciado com uma Licença [Creative Commons Atribuição 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/) – CC BY. Esta licença permite que outros distribuam, remixem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuem o devido crédito pela criação original.