Landslide recognition using SVM, Random Forest, and Maximum Likelihood classifiers on high-resolution satellite images: A case study of Itaóca, southeastern Brazil

Helen Cristina Dias[1,2]* , Lucas Henrique Sandre[3] , Diego Alejandro Satizábal Alarcón[4] ,
Carlos Henrique Grohmann[1,2] , José Alberto Quintanilha[1]

Abstract

Landslide identification is important for understanding their conditioning factors, and for constructing susceptibility, risk, and vulnerability maps. In remote sensing this can be accomplished manually or through classifiers. This study compares three image classifiers (Maximum Likelihood, Random Forest, and Support Vector Machines (SVM)) used in identifying landslides in Itaóca (São Paulo, Brazil). Two datasets were used: a RapidEye-5 (5 m) image and a Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) (12.5 m). Seven pixel-based classifications were produced, two for each classifier and a binary class that identified only landslides and non-landslides. One classification contained five spectral bands (5B), while the other contained six bands (6B) and included the slope derived from the DEM. The results were validated using Kappa index and F1 score. The SVM 6B classification achieved the best results among the validation indices used herein. It identified a landslide area of 399,325 m². The results contribute to landslide mapping in tropical environments using pixel-based classifiers. However, although the SVM classification was successful, only landslides with larger areas were captured by the algorithms, confirming the importance of conducting further analyses using images with finer spatial resolution.

KEYWORDS: mass movement; pixel classification; supervised classification; RapidEye Multispectral image; digital elevation model.

INTRODUCTION

Shallow landslides are very common in mountainous regions with steep slopes and in densely occupied cities. They are responsible for moving material from different origins and with different volumes and have long-range power, which can cause severe social and economic damage in a region. Shallow landslides can be defined as the movement of a cohesive mass in a sloping area with variable speed and planar rupture surface, which can be controlled by discontinuities or more fragile layers (Hungr *et al.* 2014). They occur frequently in Brazil because its climate and physical features are often reported to trigger landslides in some regions (IBGE, 2019). The often-disordered occupation of some cities tends to increase the occurrence of landslides and cause damage to the residents (IBGE, 2019). Landslides of large magnitude have been triggered in several cities and states in Brazil, such as Caraguatatuba, São Paulo (1967), Cubatão, São Paulo (1985), Angra dos Reis, Rio de Janeiro (2010), Nova Friburgo, Rio de Janeiro (2011), and Antonina, Paraná (2011) (Vieira and Gramani 2015).

Certain morphological, pedological, and geological conditions increase the susceptibility of regions to shallow landslides (Dias *et al.* 2017, Hussain *et al.* 2019, Cerri *et al.* 2020). Morphological constraints affect slope conditions that influence its stability (Sidle *et al.* 1985). One of the most important constraints is slope angle. Shallow landslides are directly affected by gravity, and a higher occurrence and susceptibility to these processes has been observed in several locations with slope angles of more than 25° (Pachauri and Pant 1992, Fernandes *et al.* 2001, Zhou *et al.* 2002, Nery and Vieira 2015, Dias *et al.* 2017).

Identifying susceptibility, vulnerability, and risk of landslide processes is based on different datasets, including inventories (van Westen *et al.* 2008). Susceptibility mapping for shallow landslides is a very common approach for mass movements studies (Dou *et al.* 2020, Merghadi *et al.* 2020, Wang *et al.* 2020, Dias *et al.* 2021). Landslide detection and mapping are the first steps to assess the susceptibility, vulnerability, and risk of this type of natural hazard. For example, susceptibility assessment methods are developed based on the premise that future landslides will occur under the same conditions that triggered landslides in the past (Aleotti and Chowdhury 1999, Guzzetti *et al.* 2012). Therefore, the use of inventories forms the basis for any predictive analysis of landslides. The landslide inventory is a crucial factor since mapping accuracy directly

[1]Institute of Energy and Environment, Universidade de São Paulo – São Paulo (SP), Brazil. E-mails: helen.dias@usp.br, guano@usp.br, jaquinta@usp.br

[2]Spatial Analysis and Modelling Laboratory, Institute of Energy and Environment, Universidade de São Paulo – São Paulo (SP), Brazil.

[3]Institute of Geosciences, Universidade Estadual de Campinas – Campinas (SP), Brazil. E-mail: lucashsandre@hotmail.com

[4]Institute of Geosciences, Universidade de São Paulo – São Paulo (SP), Brazil. E-mail: diego.satizabal@usp.br

*Corresponding author.

depends on its results, and inventories with errors can modify the susceptibility analyses to create distortions in the mapping (Steger *et al.* 2016).

According to Guzzetti *et al.* (2012), landslide inventories can be constructed with different objectives, from a geomorphological or a historical (*e.g.*, event; seasonal; multi-temporal) perspective. Mapping shallow landslides in the landscape is crucial for the prevention and improved understanding of these processes' dynamics. Scars on the landscape can be identified manually, semi-automatically, or automatically (Hölbling *et al.* 2017). Conventionally, geomorphological field mapping and visual interpretations of aerial photographs have also been used for landslide identification (Guzzetti *et al.* 2012). Recently, high or very high-resolution satellite images have been frequently used for landslide mapping (Hölbling *et al.* 2017).

Although many studies on inventories have been conducted, the criteria for mapping are yet incipient and variable (Guzzetti *et al.* 2012). Different methods have been used to analyze multispectral images aiming at landslide recognition, including the application of indices, such as pixel-based classifications (Marcelino *et al.* 2009), object-based classifications (Stumpf and Kerle 2011, Dou *et al.* 2015, Comert *et al.* 2019), and convolution networks (Lei *et al.* 2019, Wang *et al.* 2021). The objective of this study was to compare the performances of three image classifiers (Maximum Likelihood, Random Forest, and Support Vector Machines) in identifying shallow landslides generated during a 2014 high precipitation event in the city of Itaóca (São Paulo, southeastern Brazil).

STUDY AREA

The city of Itaóca is located in the southern part of São Paulo State, in Brazil, and has a total area of 183 km² (Fig. 1). The city borders four municipalities: Apiaí (SP), Iporanga (SP), Ribeira (SP), and Adrianópolis (PR) and is part of the Ribeira do Iguape River Basin.

The relief in this region is strongly undulating and mountainous, with steep slopes, high drainage density, and deep, narrow valleys. Local geology is characterized by Itaóca granite (80% of the area), metasedimentary rocks of the Serra da Boa Vista Hill and Betari Hill formations, and to a lesser extent by Quaternary contact metamorphic rocks (hornfels) and alluvial sediments (Brollo *et al.* 2015). The municipality has a humid subtropical climate (Cfa) with hot summers (Köppen 1936). Temperatures are higher than 22°C during the summer and the area receives more than 30 mm of rain in the driest month.

In 2014, between January 12th and 13th, a large magnitude event occurred in the area, with high rainfall rates of 105 mm/h (Gramani and Martins 2015). In the city, intense rain resulted in a sudden flood, generalized landslides, debris flows, and floods with large movements of material (Fig. 2). The event resulted in 25 deaths, 3 missing people, and left 203 homeless, in addition to the damage caused to local housing and infrastructure (Brollo *et al.* 2015).

IMAGE CLASSIFICATION

The image classification process consists of assigning each pixel of an image to one among a group of classes (Jensen 2007, Meneses and Almeida 2012, Lillesand *et al.* 2015). The classification methods can be grouped according to different criteria, such as supervised or unsupervised classification, or classification by pixel and region. Pixel classification uses only the spectral information from each pixel to identify homogeneous areas based on distance measurements or the probability of a pixel belonging to a specific class (Meneses and Almeida 2012, Lillesand *et al.* 2015). In classification by region, the decision process uses a group of pixels as a classification unit. Supervised classification requires prior knowledge of the target classes to classify the image in pre-defined classes of interest. Unsupervised classification requires little or no participation of the user in the classification process, and is generally used when information is lacking on the number and nature of the target classes present at a site, or when an exploratory classification of the image is necessary to readily identify target classes in the image (Meneses and Almeida 2012). In this study, three supervised and pixel classifiers were used to identify shallow landslides in the image, namely: Maximum Likelihood, Random Forest, and Support Vector Machines (SVM).

Maximum likelihood

Maximum likelihood is a classification tool that uses the weighted distances between the mean values of the pixel values of the classes, which are defined using statistical parameters (Meneses and Almeida 2012). This method determines the probability of a pixel belonging to a certain class pre-defined by the user (Ahmad and Quegan 2012, Meneses and Almeida 2012). This classifier is widely used to identify features in satellite images due to its accessibility (Lu and Weng 2007) and previously obtained favorable results in the classification of landslides (Marcelino *et al.* 2009, Pawluszek *et al.* 2018).

Random forest

Random Forest classifier is a combination of decision trees where each tree provides a classification label, with most of the chosen classes being considered a decision. It consists of a collection of tree-structured classifiers that build rules through binary partitioning in regions that are homogeneous in relation to the variable class (Cutler *et al.* 2007). After generating a large number of trees, the most popular class is voted for. This method is widely applied to identify landslides (Stumpf and Kerle 2011, Wang *et al.* 2021).

Support vector machines

SVM is a classifier that features learning algorithms and analyzes the data used for classification and regression (Boser *et al.* 1992). SVM uses sample data for training to create a reliable model that accurately classifies the samples in each class. In the case of a two-class pattern recognition problem in which the classes are linearly separated, SVM selects the result from an infinite number of linear decision limits which minimizes the generalization error (Pal and Mather 2005).
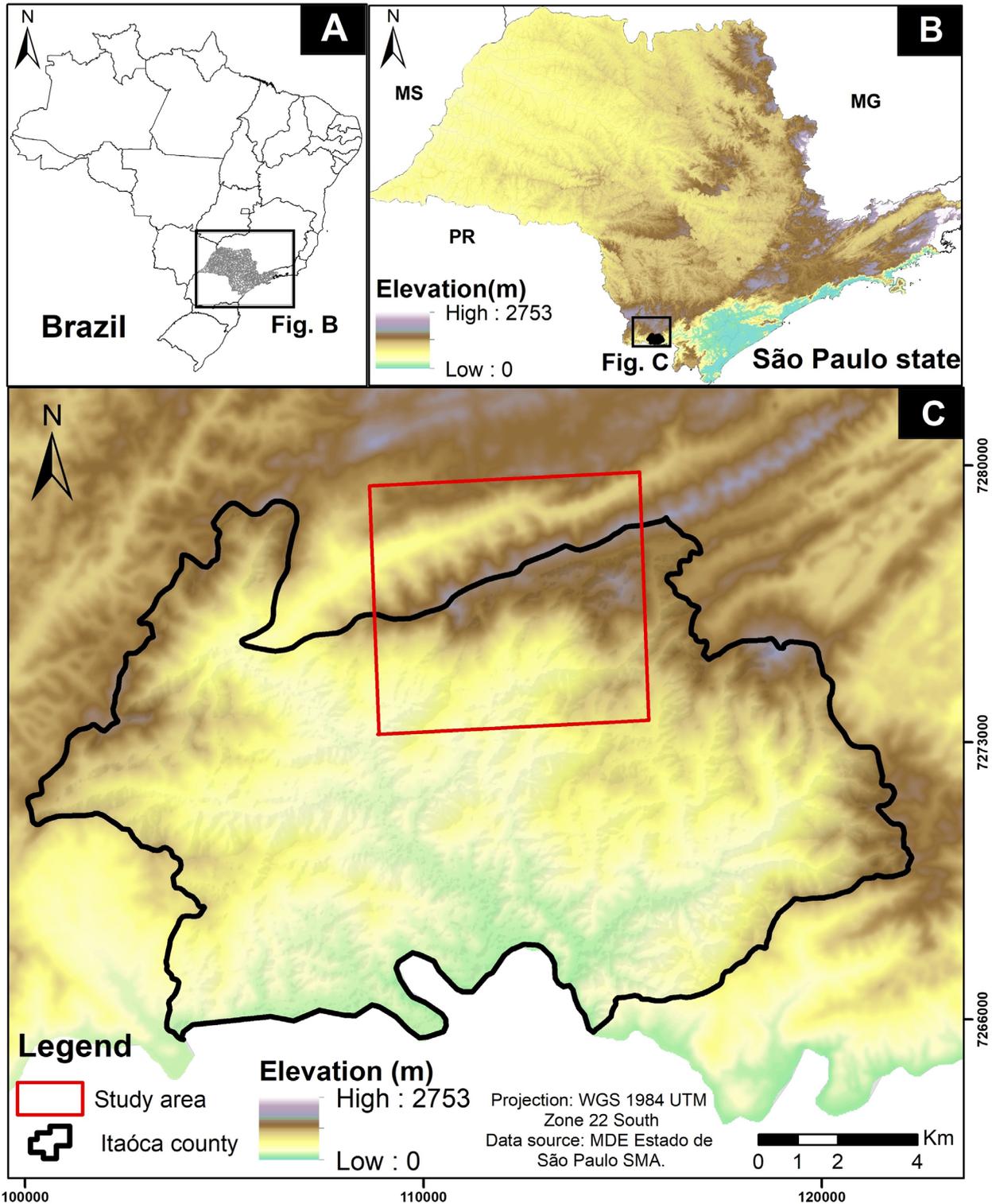
Figure 1. Location of the city of Itaóca in relation to Brazil and the state of São Paulo.

SVM is currently considered to be a stable and accurate classifier, with superior performance compared to traditional pattern recognition techniques (Chunhui *et al.* 2018), and has been widely applied to identify landslides (Pawluszek *et al.* 2018, Wang *et al.* 2021).

METHODS

Identification and evaluation of landslides were performed in five steps:

- Database acquisition;
- Definition of criteria for landslide identification;
- Definition of classes mapped by the classifiers;
- Application of the three classifiers (Maximum Likelihood, Random Forest, and SVM);
- Validation;
- Analysis and comparison of the results.

Database

A RapidEye Analytic Ortho Tile multispectral image collected by the RapidEye-5 satellite on August 22nd, 2014 was used. The image has a spatial resolution of 5 m, radiometric

Source: Marcelo F. Gramani.
Figure 2. 2014 mass movement event. Shallow landslides triggered in Itaóca, São Paulo.

resolution of 16 bits, and contains five spectral bands (blue, red, green, red edge, and near infrared) (Planet 2020). Before the image was made available to the user, it was pre-processed to obtain a radiometrically and geometrically corrected image (Planet 2020). Morphological information (slope angle) was derived from a digital elevation model (DEM) from the SRTM mission (Farr *et al.* 2007), processed to remove outlier data, re-sampled to a 12.5 m spatial resolution, and distributed by the Alaska Satellite Facility (ASF DAAC 2020).

Criteria for the identification of landslide scars

The identification of scars from shallow landslides was based on two visual criteria: the absence of vegetation and the shape of the slope features. These two criteria were used due to the characteristics of shallow landslide processes, as shallow landslides are the gravitational movement of a mass of rock, debris, or earth down a slope (Hungr *et al.* 2014) that has an elliptical or conical signature. These visual criteria were also used by Lopes *et al.* (2007) and Carou *et al.* (2017). Training samples were collected manually and randomly. An example of landslide scar used as sample is presented on Fig. 3.
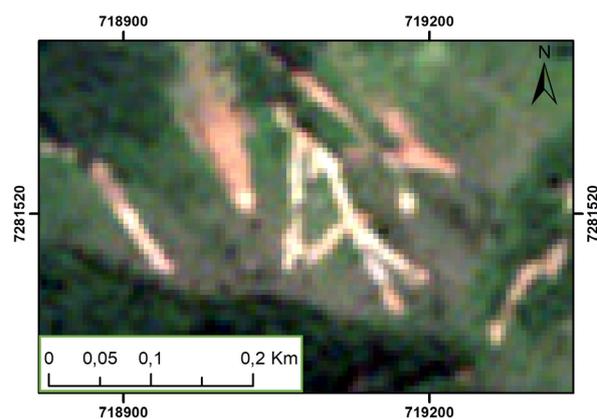

Figure 3. Shallow landslides (orange/red) used as training samples. Satellite image: RapidEye (5m) – 08/22/2014.

Definitions of the classes to be mapped

Five classes were selected for mapping by the classifiers: landslides, hydrography, vegetation, pastures, and clouds (Tab. 1). These classes were selected based on visual observations of the land cover pattern in the study area. A binary classification was also performed to evaluate the impact of using only two classes (landslides and non-landslides).

Table 1. Classes used for classification and their respective characteristics.

Class	Characteristics
Landslides	Features with no vegetation and an elliptical or conical shape.
Hydrography	Any feature with a continuous drainage pattern.
Vegetation	Areas of dense vegetation of medium to large size.
Pastures	Open fields, often used for livestock.
Clouds	Low altitude clouds.

Classifiers

Three pixel-based supervised multispectral image classifiers were applied in QGIS 3.14 (QGIS 2020): Maximum Likelihood, using the SCP plugin (Luca 2016), and Random Forest and SVM using the Dzetsaka plugin (Karasiak 2016). Six classifications were made, two for each classifier: one that contained five bands (5B) and another that contained six bands (6B), including slope information.

Furthermore, a binary classification was performed using maximum likelihood (6B) where only two classes were identified, landslides and non-landslides. Maximum likelihood was selected to the binary analysis as it presented good results in the multiclass classifications.

Validation

Validation was carried out using 502 random points. First, the classifications were validated using the Kappa index (K), a method used to evaluate classification results, with values ranging from 0 (null agreement) to 1 (perfect agreement) (Meneses and Almeida 2012, Lillesand et al. 2015).

In addition to the K value, user accuracy was determined, which considers inclusion or commission errors (Hord and Brooner 1976, Rosenfield et al. 1982, Rosenfield and Fitzpatrick-Lins 1986, Congalton 1991). User accuracy (also called sensitivity) indicates the probability that a classified class (map/image) actually represents the same category. Producer accuracy (or accuracy) was also evaluated, which considers exclusion or omission errors. Accuracy indicates the probability that a given class has been correctly classified according to the training data (Rosenfield et al. 1982, Story and Congalton 1986). The relationship between producer accuracy (accuracy) and user accuracy (sensitivity) was calculated using the F1 score, according to Equations 1, 2, and 3:

$$Accuracy = TP / (TP + FP) \qquad (1)$$

$$Sensivity = TP / (TP + FN) \qquad (2)$$

$$F1 Score = 2 * \frac{Accuracy * Sensivity}{(Accuracy + Sensivity)} \qquad (3)$$

where:
TP: the pixels that were correctly identified by the classifier (true positive);
FP: the pixels that were classified incorrectly (false positive);
FN: the pixels that belong to the class and were not classified correctly (false negative).

RESULTS AND DISCUSSION

Shallow landslides were identified in two datasets using the three aforementioned supervised pixel-based classifiers, one being the multispectral image (RapidEye) and the other including slope as a sixth band (RapidEye + SRTM slope). Morphological characteristics are important and should be considered for landslides recognition. In this paper, 6B classifications considered slope information, which is one of the most important predisposing factors to landslides (Sidle et al. 1985). Slope provides information about the angle of the slope, some angles can be more prone to landslides than others and directly influence their occurrence. Additionally, a binary classification was performed to analyze the differences in the identification of only two classes, landslides, and non-landslides. Fig. 4 presents a visual comparison of the classifiers with the RapidEye (5B), RapidEye + SRTM slope (6B), and binary classification datasets only for those areas of the image where the landslides are visible.

The classifications defined different areas for each class (Tab. 2). The Maximum Likelihood, SVM, and Random Forest classifiers applied to the 5B image identified areas for the landslide class of 679,100 m², 327,475 m², and 423,975 m², respectively. For the 6B image, these classifiers identified areas of 1,151,150 m², 443,375 m², and 399,325 m², respectively. For the binary classification (6B), an area of 549,075 m² was identified for landslides. The large differences in the values reflected the randomness of the classifiers used for the identification process.

The Kappa index (K) was then determined to validate the results (Tab. 3). The best classifications were found to be those of the Random Forest 5B (75.6%) and SVM 5B (75.5%), followed by the Maximum Likelihood 5B (74.2%). The binary classification had the lowest K value (3.75%). The first three classifications produced good results in the study area. The overall accuracies were 90.2% for the Random Forest 5B and SVM 5B classifications, 88.2% for the Maximum Likelihood 5B classification, and 60.8% for the binary classification. When observing the user accuracy for the landslide class (Tab. 3), the highest value (83.3%) was obtained by the SVM classification of the 6-band image and the lowest value (37.5%) was obtained for Maximum Likelihood from the 6-band image. The other classifications varied between 60 and 66.7% for both images. Moreover, the landslide class had a producer accuracy of 30.8% for both the Random Forest (5B and 6B) and SVM 5B classifications. The highest accuracy was 46.2% for the Maximum Likelihood 5B and 6B classifications.

The classifications of the 5B image obtained K values higher than those of the 6B image. Although the same was true for overall accuracy, there was a significant increase in user accuracy (83.3%) in the SVM 6B, compared to the SVM 5B (66.7%). According to the F1 score, the classification with the highest overall accuracy was the SVM 6B at 52.6%, followed by the Maximum Likelihood 5B at 52.2%. The binary classification
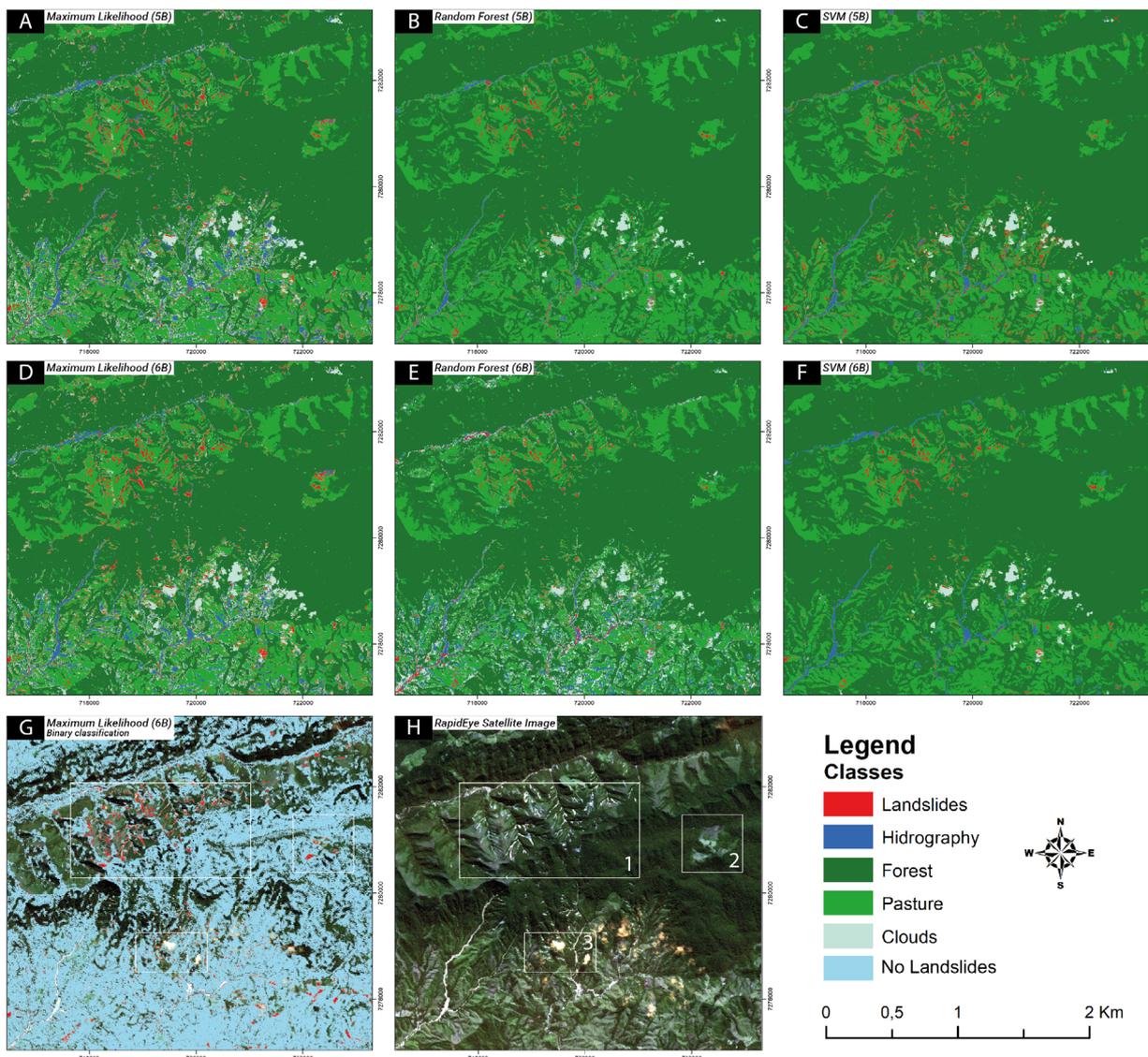
Figure 4. Classifications generated by the three classifiers. (A) Maximum Likelihood classification of the 5-band image (5B - RapidEye); (B) Maximum Likelihood classification of the 6-band image (6B - RapidEye + SRTM slope); (C) Random Forest classification of the 5-band image; (D) Random Forest classification of the 6-band image; (E) SVM classification of the 5-band image; (F) SVM classification of the 5-band image; (G) Maximum Likelihood classifier (Binary classification)- 6B; (H) RapidEye satellite image.

Table 2. Areas (m²) identified for each class in each multi-class classification.

Classes	Maximum Likelihood (5B)	Random Forest (5B)	SVM (5B)	Maximum Likelihood (6B)	Random Forest (6B)	SVM (6B)
Landslides	679,100	423,975	327,475	1,151,150	443,375	399,325
Hidrography	1,057,450	376,375	405,000	1,122,375	1,038,500	505,250
Vegetation	29,835,850	33,463,500	33,337,775	2,544,150	2,578,100	391,875
Pastures	10,069,600	8,879,600	9,043,075	9,759,975	8,651,725	9,721,750
Clouds	1,816,100	314,650	344,775	28,880,450	30,746,400	32,439,900
TOTAL	**43,458,100**	**43,458,100**	**43,458,100**	**43,458,100**	**43,458,100**	**43,458,100**

exhibited low *K* values (3.8%) and overall accuracy (60.8%). These results could be attributed to the classification of a large number of non-landslides areas, such as clouds and rivers, as landslides. In the binary classification, classes with similar spectral response could not be distinguished, which may have led to further errors.

A detailed analysis of three sectors with different characteristics was also conducted. The aim of this analysis was to demonstrate the difficulty in identifying classes using the selected classifiers. According to the sectors defined in Fig. 4H, Sector 1 is the area with the highest concentration of landslides (Fig. 5). Although the Maximum Likelihood 6B classifier defined the landslide class in considerably wider areas than the other classifiers, several areas corresponding to other classes were classified as false positives (FP). The SVM 6B classifier identified areas corresponding to landslides more

Table 3. Accuracies of the three classifications in the landslide class images.

Landslide class

| Classification	User accuracy (Sensitivity)	Producer accuracy (Accuracy)	Overall accuracy	Kappa	F1 score	TP	FN	FP
Maximum Likelihood 5B	60.0%	46.2%	88.2%	74.2%	52.2%	6	4	7
Random Forest 5B	66.7%	30.8%	90.2%	75.7%	42.1%	4	2	9
Support Vector Machines 5B	66.7%	30.8%	90.2%	75.6%	42.1%	4	2	9
Maximum Likelihood 6B	37.5%	46.2%	84.5%	67.0%	41.4%	6	10	7
Random Forest 6B	66.7%	30.8%	80.9%	58.1%	42.1%	4	2	9
Support Vector Machines 6B	83.3%	38.5%	87.5%	70.3%	52.6%	5	1	8
Maximum Likelihood binary 6B	66.7%	46.1%	60.8%	3.7%	54.5%	6	12	7
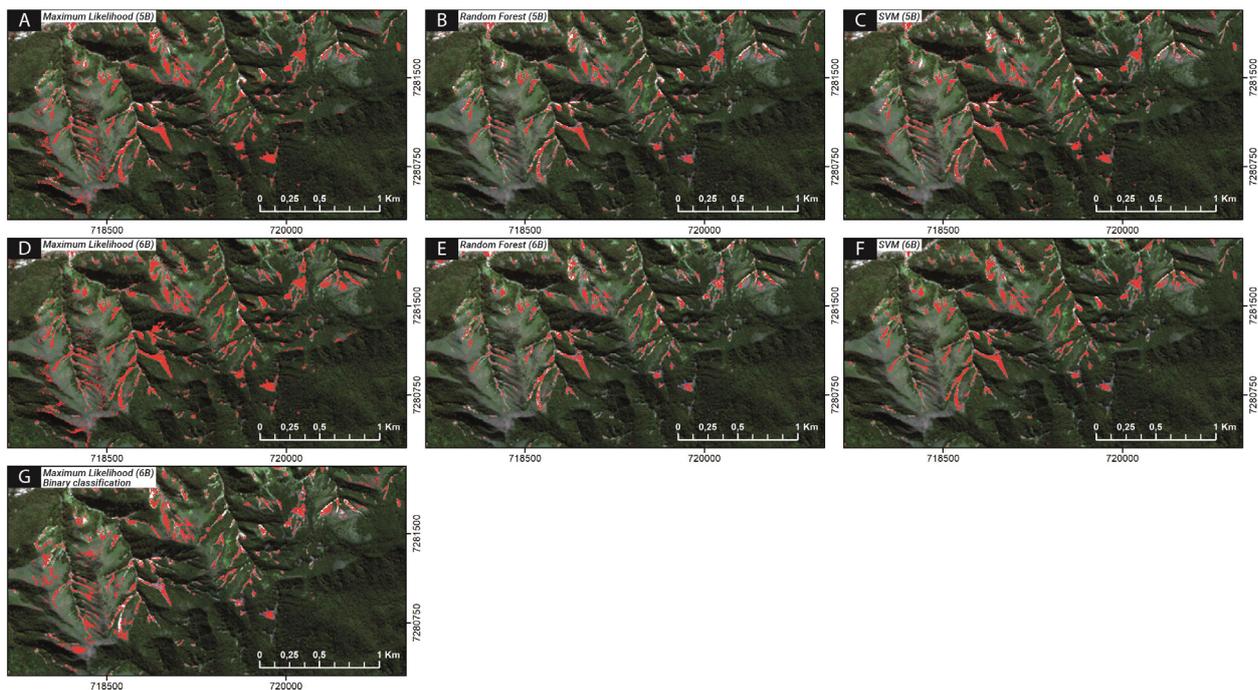
Figure 5. Classifications of shallow landslides (red) generated by the three classifiers applied to Sector 1, where SVM 6B (subfigure F) achieved the best results. (A) Maximum Likelihood classifier of the 5B image; (B) Random Forest classifier of the 5B image; (C) SVM classifier of the 5B image; (D) Maximum Likelihood classifier of the 6B image; (E) Random Forest classifier of the 6B image; (F) SVM classifier of the 6B image; (G) Maximum Likelihood classifier (Binary classification) 6B.

accurately and did not include areas of other classes to the highest possible extent (Fig. 5). Sector 2 was selected because the classifiers identified pastures as landslides in this region (Fig. 6). There were no landslide scars in this area; therefore, all of the classifiers identified FP. The most consistent result for this sector was the SVM 5B, as this was the classification that least defined FP. Furthermore, Sector 3 was selected because it included exposed soil and rocky surfaces classified as clouds, primarily using the Maximum Likelihood 5B and 6B methods, and the Random Forest 6B (Fig. 7).

Maximum Likelihood 5B and SVM 6B were found to be the best classifications for detecting shallow landslides in the study area, in agreement with previous studies in other areas (Marapareddy *et al.* 2017, Uehara *et al.* 2020). Although DEM derivatives are commonly applied to landslides recognition worldwide (Kurtz *et al.* 2014, Mezaal *et al.* 2017), in this case only SVM presented an improvement in performance with the addition of slope information.

SVM was the best classifier for application to the image comprising multispectral bands and the slope (6B), as it was the classifier with the best results determined by the various validation indices. The SVM 6B classifier identified an area of 399,325 m² that contained landslides. This classifier can be considered as the most suitable for identifying shallow landslides when spectral information is used with a DEM (Pawluszek *et al.* 2018). The SVM classifier for the 6B image had a higher number of TPs compared to FN. This method had the highest validation index improvements with the inclusion of the slope band and achieved the best results in this initial analysis. The Maximum Likelihood 5B classifier obtained validation indices similar to those of SVM 6B and identified sectors with an area of 679,100 m² that contained landslides. From the perspective of natural disaster prevention, this is an important finding that requires further consideration for city planning and prevention of natural hazards.
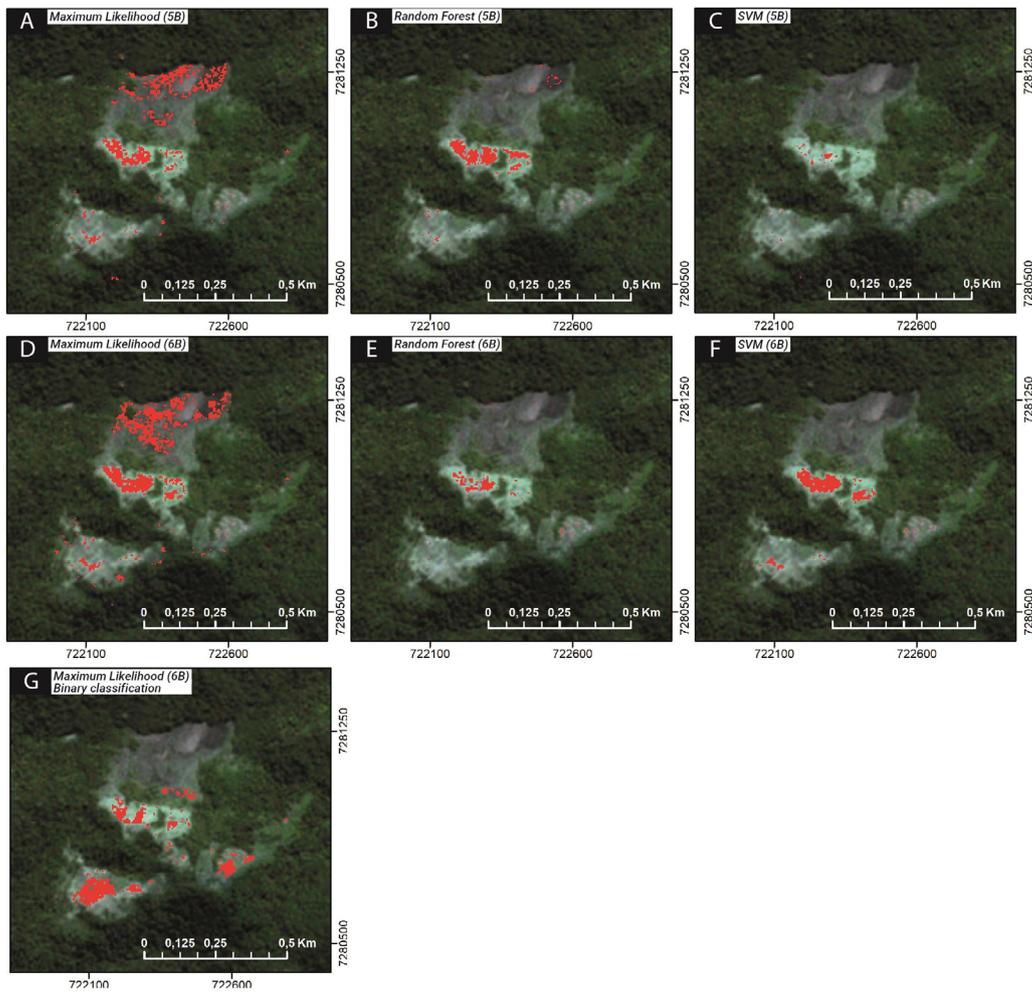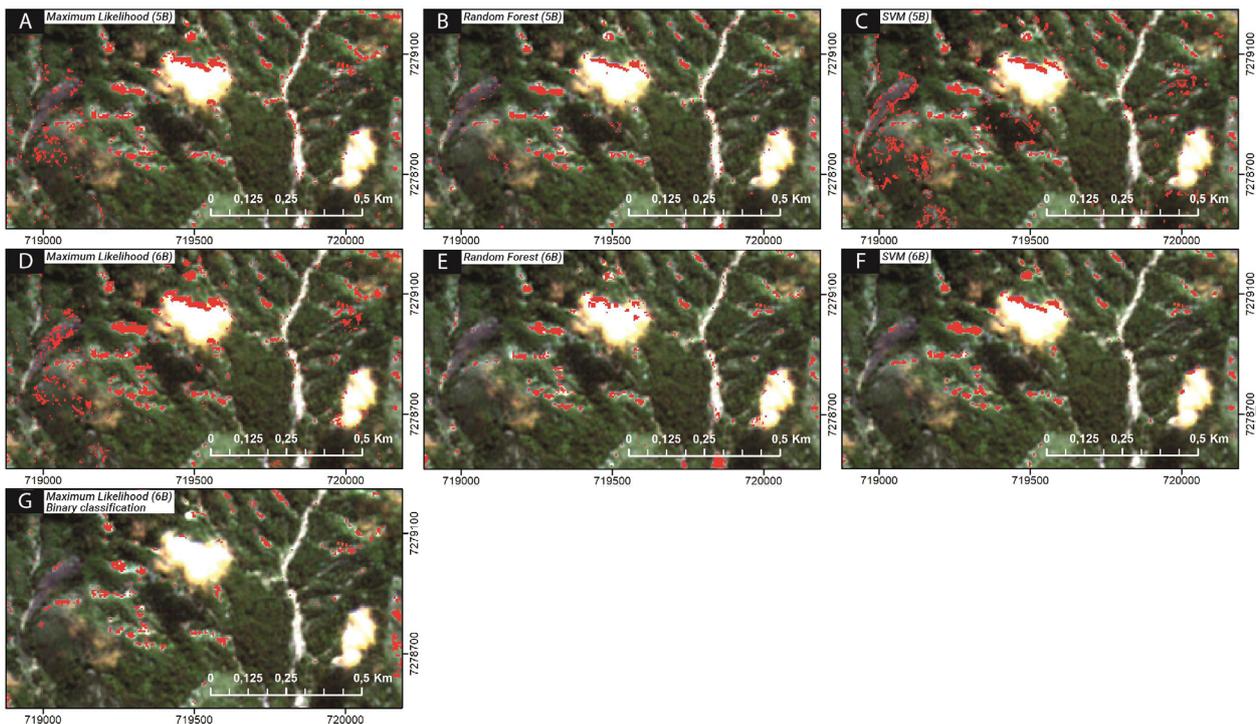
Figure 6. Classifications of shallow landslides (red) generated by the three classifiers applied to Sector 2, where SVM 5B (subfigure C) produced the lowest number of FPs. (A) Maximum Likelihood classifier of the 5B image; (B) Random Forest classifier of the 5B image; (C) SVM classifier of the 5B image; (D) Maximum Likelihood classifier of the 6B image; (E) Random Forest classifier of the 6B image; (F) SVM classifier of the 6B image; (G) Maximum Likelihood classifier (Binary classification) 6B.

Figure 7. Classifications of shallow landslides (red) generated by the three classifiers applied to Sector 3, where Maximum Likelihood 5B and 6B and Random Forest 6B had more difficulties defining the correct classes in an environment where several classes coexisted. (A) Maximum Likelihood classifier of the 5B image; (B) Random Forest classifier of the 5B image; (C) SVM classifier of the 5B image; (D) Maximum Likelihood classifier of the 6B image; (E) Random Forest classifier of the 6B image; (F) SVM classifier of the 6B image; (G) Maximum Likelihood classifier (Binary classification) 6B.

CONCLUSIONS

This study verified that the method of applying pixel-based supervised classifiers to multispectral images with high spatial resolutions was successful and effective in identifying shallow landslides. A comparison of the three classifiers enabled us to observe the behavior of each classifier in identifying landslides across the selected landscape in Itaóca, São Paulo, southeastern Brazil. Thus, we conclude that the three classifiers (Random Forest, SVM, and Maximum Likelihood) achieved good results. Maximum Likelihood 5B and SVM 6B were found to produce the best classifications for shallow landslide detection in the study area. Adding DEM (slope) information only improved the results in SVM classifier. However, this study has some limitations. Although landslides were identified in the studied landscape, only the scars with sufficiently large areas were captured by the algorithm. This confirmed the need for additional analyses on multispectral images with finer spatial resolutions, which may be conducted in future studies.

REFERENCES

Ahmad A., Quegan S. 2012. Analysis of maximum likelihood classification on multispectral data. *Applied Mathematical Sciences*, **6**(129):6425-6436. https://doi.org/10.1109/ICCSCE.2012.6487156

Aleotti P., Chowdhury R. 1999. Landslide hazard assessment: summary review and new perspectives. *Bulletin of Engineering Geology and the Environment*, **58**:21-44. https://doi.org/10.1007/s100640050066

ASF DAAC. 2020. *ALOS PALSAR_Radiometric_Terrain_Corrected_high_res; Includes Material JAXA/METI 2011*.

Boser B.E., Guyon I.M., Vapnik V.N. 1992. A training algorithm for optimal margin classifiers. *In*: The Workshop on Computational Learning Theory, p. 144-152.

Brollo M.J., Santoro J., Penteado D.R., Fernandes P.C.S, Ribeiro R.R. 2015. Itaóca (SP): Histórico de acidentes e desastres relacionados a perigos geológicos. *In*: Simpósio de Geologia do Sudeste, 14., 2015, Campos do Jordão. *Annals...* p. 1-5.

Carou C.B., Vieira B.C., Martins T.D., Gramani M.F. 2017. Inventário dos Escorregamentos da Bacia do Rio Gurutuba, Vale do Ribeira (SP). *Revista do Departamento de Geografia*, 172-179.

Cerri R.I., Rosolen V., Reis F.A.G.V., Pereira Filho A.J., Vemado F., Giordano L.C., Gabelini B.M. 2020. The assessment of soil chemical, physical, and structural properties as landslide predisposing factors in the Serra do Mar mountain range (Caraguatatuba, Brazil). *Bulletin of Engineering Geology and the Environment*, **79**:3307-3320. https://doi.org/10.1007/s10064-020-01791-1

Chunhui Z., Bing G., Lejun Z., Xiaoqing W. 2018. Classification of Hyperspectral Imagery based on spectral gradient, SVM and spatial random forest. *Infrared Physics & Technology*, **95**:61-69. https://doi.org/10.1016/j.infrared.2018.10.012

Comert R., Avdan U., Gorum T., Nefeslioglu H.A. 2019. Mapping of shallow landslides with object-based image analysis from unmanned aerial vehicle data. *Engineering Geology*, **260**:105264. https://doi.org/10.1016/j.enggeo.2019.105264

Congalton R. 1991. A review of assessing the accuracy of classification remotely sensed data. *Remote Sensing of Environment*, **37**(1):35-36. https://doi.org/10.1016/0034-4257(91)90048-B

Cutler D.R., Edwards Jr. T.C., Beard K.H., Cutler K., Hess K.T., Gibson J., Lawler J.J. 2007. Random Forest for classification in Ecology. *Ecology*, **88**(11):2783-2792. https://doi.org/10.1890/07-0539.1

Dias H.C., Dias V.C., Vieira B.C. 2017. Condicionantes morfológicos e geológicos dos escorregamentos rasos na bacia do Rio Santo Antônio, Caraguatatuba / SP. *Revista do Departamento de Geografia – USP*, **volume especial** (8):157-163. https://doi.org/10.11606/rdg.v0ispe.132714

Dias H.C., Gramani M.F., Grohmann C.H., Bateira C., Vieira B.C. 2021. Statistical-based shallow landslide susceptibility assessment for a tropical environment: a case study in the southeastern Brazilian coast. *Natural Hazards*, **108**:205-223. https://doi.org/10.1007/s11069-021-04676-y

Dou J., Chang K.T., Chen S., Yunus A.P., Liu J.K., Xia H., Zhu Z. 2015. Automatic Case-Based Reasoning Approach for Landslide Detection: Integration of Object-Oriented Image Analysis and a Genetic Algorithm. In: *Remote Sensing*, **7**(4):4318-4342. https://doi.org/10.3390/rs70404318

Dou J., Yunus A.P., Bui D.T., Merghadi A., Sahana M., Zhu Z., Chen C.W., Han Z., Pham B.T. 2020. Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. *Landslides*, **17**:641-658. https://doi.org/10.1007/s10346-019-01286-5

Farr T.G., Rosen P.A., Caro E., Crippen R., Duren R., Hensley S., Kobrick M., Paller M., Rodriguez E., Roth L., Seal D., Shaffer S., Shimada J., Umland J., Werner M., Oskin M., Burbank D., Alsdorf D. 2007. The Shuttle Radar Topography Mission. *Review of Geophysics*, **45**(2). http://dx.doi.org/10.1029/2005RG000183

Fernandes N.F., Guimarães R.F., Gomes R.A.T., Vieira B.C., Montgomery D.R., Greenberg H. 2001. Condicionantes geomorfológicos dos deslizamentos nas encostas: Avaliação de metodologias e aplicação de modelo de previsão de áreas susceptíveis. In: *Revista Brasileira de Geomorfologia*, **2**(1):51-71. http://dx.doi.org/10.20502/rbg.v2i1.8

Gramani M.F., Martins V.T.S. 2015. Debris flows occurrence by intense rains at Itaoca city, São Paulo, Brazil: Field observations. *In*: Aversa F., Cascini L., Picarelli L., Scavia C. (Eds.). *Landslides and engineered slopes:* experience, theory and practice. Boca Raton: CRS Press, p. 1011-1019.

Guzzetti F., Mondini A.C., Cardinali M., Fiorucci F., Santangelo M., Chang K.T. 2012. Landslide inventory maps: new tools for an old problem. *Earth-Science Reviews*, **112**(1-2):42-66. https://doi.org/10.1016/j.earscirev.2012.02.001

Hölbling D., Eisank C., Albrecht F., Vecchiotti F., Friedl B., Weinke E., Kociu A. 2017. Comparing Manual and Semi-Automated Landslide Mapping Based on Optical Satellite Images from Different Sensors. *Geosciences*, **7**(2):37. https://doi.org/10.3390/geosciences7020037

Hord R.M., Brooner W. 1976. Land use map-accuracy criteria. *Photogrammetric Engineering and Remote Sensing*, **42**(5):671-677.

Hungr O., Leroueil S., Picarelli L. 2014. The Varnes classification of landslide types, an update. *Landslides*, **11**:167-194. https://doi.org/10.1007/s10346-013-0436-y

Hussain Y., Cardenas-Soto M., Uagoda R., Martino S., Rodriguez-Rebolledo J., Hamza O., Martinez-Carvajal H. 2019. Monitoring of Sobradinho landslide (Brasília, Brazil) and a prototype vertical slope by time-lapse interferometry. *Brazilian Journal of Geology*, 49(2):1-11. https://doi.org/10.1590/2317-4889201920180085

Instituto Brasileiro de Geografia e Estatística (IBGE). 2019. *Suscetibilidade a deslizamentos do Brasil:* primeira aproximação. Rio de Janeiro: Coordenação de Recursos Naturais e Estudos Ambientais, 56 p.

Jensen J.R. 2007. *Remote sensing of environment:* an earth resource perspective. Hoboken: Prentice Hall, 619 p.

Lopes E.S.S., Riedel P.S., Bentz C.M., Ferreira M.V., Naleto J.L.C. 2007. Inventário de escorregamentos naturais em banco de dados geográfico – análise dos fatores condicionantes na região da Serra de Cubatão – SP. *In*: Simpósio Brasileiro de Sensoriamento Remoto, 13., 2007, Florianópolis, Brazil. *Annals...* p. 2785-2796.

Karasiak N. 2016. *Dzetsaka QGIS Classification plugin*. https://doi.org/10.5281/zenodo.2552284.

Köppen W. 1936. Das geographische System der Klimate. *In*: Köppen W., Geiger R. (eds.). *Handbuch der Klimatologie*. Berlin: Gerbrüder Bornträger, Part C, v. 1, p. C1-C44.

Kurtz C., Stumpf A., Malet J.P., Gançarski P., Puissant A., Passat N. 2014. Hierarchical extraction of landslides from multiresolution remotely sensed optical images. *ISPRS Journal Photogrammetry Remote Sensing*, **87**:122-136. https://doi.org/10.1016/j.isprsjprs.2013.11.003

Lei T., Zhang Y., Lv Z., Li S., Liu S., Nandi A.K. 2019. Landslide Inventory Mapping from Bitemporal Images Using Deep Convolutional Neural Networks. *IEEE*, **16**(6):982-986. http://dx.doi.org/10.1109/LGRS.2018.2889307

Lillesand T.M., Kiefer R.W., Chipman J.W. 2015. *Remote sensing and image interpretation*. 7ª ed. Chichester: John Wiley & Sons, 736 p.

Lu D., Weng Q. 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, **28**(5):823-870. http://dx.doi.org/10.1080/01431160600746456

Luca C. 2016. *Semi-Automatic Classification Plugin Documentation*.

Marapareddy R., Aanstoos J.V., Younan N.H. 2017. Accuracy analysis comparison of supervised classification methods for anomaly detection on levees using SAR imagery. *Electronics*, **6**(4):83. http://dx.doi.org/10.3390/electronics6040083

Marcelino E.V., Formaggio A.R., Maeda E.E. 2009. Landslide inventory using image fusion techniques in Brazil. *International Journal of Applied Earth Observation and Geoinformation*, **11**(3):181-191. https://doi.org/10.1016/j.jag.2009.01.003

Meneses P.R., Almeida T. 2012. *Introdução ao processamento de imagens de sensoriamento remoto*. Brasília: Editora UnB, 277 p.

Merghadi A., Yunus A.P., Dou J., Whiteley J., ThaiPham B., Bui D.T., Avtar R., Abderrahmane B. 2020. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Science Reviews*, **207**:103225. https://doi.org/10.1016/j.earscirev.2020.103225

Mezaal M.R., Pradhan B., Sameen M.I., Mohd Shafri H.Z., Yusoff Z.M. 2017. Optimized neural architecture for automatic landslide detection from high-resolution airborne laser scanning data. *Applied Sciences*, **7**(7):730. https://doi.org/10.3390/app7070730

Nery T.D., Vieira B.C. 2015. Susceptibility to shallow landslides in a drainage basin in the Serra do Mar, São Paulo, Brazil, predicted using the SINMAP mathematical model. In: *Bulletin of Engineering Geology and the Environment*, **74**:369-378. https://doi.org/10.1007/s10064-014-0622-8

Pachauri A.K., Pant M. 1992. Landslide hazard mapping based on geological attributes. *Engineering Geology*, **32**(1-2):81-100. https://doi.org/10.1016/0013-7952(92)90020-Y

Pal M., Mather P.M. 2005. Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, **26**(5):1007-1011. https://doi.org/10.1080/01431160512331314083

Pawluszek K., Borkowski A., Tarolli P. (2018). Sensitivity analysis of automatic landslide mapping: numerical experiments towards the best solution. *Landslides*, **15**:1851-1865. https://doi.org/10.1007/s10346-018-0986-0

Planet. 2020. *Planet Imagery product specifications*. 100 p.

QGIS. 2020. *QGIS Geographic Information System, v.3.14*. Open Source Geospatial Foundation Project. Available at: http://qgis.org. Accessed on: Oct. 2020.

Rosenfield G.H., Fitzpatrik-Lins K.A. 1986. Coefficient of agreement as measure of thematic classification accuracy. *Photogrammetric Engineering and Remote Sensing*, **52**(2):223-227.

Rosenfield G.H., Fitzptrik-Lins K.A., Ling S.H. 1982. Sampling for thematic map accuracy testing. *Photogrammetric Engineering and Remote Sensing*, **48**(1):131-137.

Sidle R.C., Pearce A.J., O'Loughlin C.L. 1985. *Hillslope stability and land use*. Washington, D.C.: John Wiley & Sons.

Steger S., Brenning A., Bell R., Glade T. 2016. The propagation of inventory-based positional errors into statistical landslide susceptibility models. *Natural Hazards and Earth System Sciences*, **16**:2729-2745. http://dx.doi.org/10.5194/nhess-2016-301

Story M., Congalton R. 1986. Accuracy assessment: a user's perspective. *Photogrammetric Engineering and Remote Sensing*, **52**(3):379-399.

Stumpf A., Kerle N. 2011. Object-oriented mapping of landslides using Random Forests. *Remote Sensing of Environment*, **115**(10):2564-2577. https://doi.org/10.1016/j.rse.2011.05.013

Uehara T.D.T, Corrêa S.P.L.P, Quevedo, R.P., Körting T.S, Dutra L.V., Rennó C.D. 2020. 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. *Revista Brasileira de Cartografia*, **72**(4):665-680. https://doi.org/10.14393/rbcv72n4-54037

Van Westen C.J., Catellanos E., Kuriakone S.L. 2008. Spatial data for landslide susceptibility, hazard and vulnerability assessment: an overview. *Engineering Geology*, **102**(3-4):112-131. http://dx.doi.org/10.1016/j.enggeo.2008.03.010

Vieira B.C., Gramani M.F. 2015. Serra do Mar: the most "tormented" relief in Brazil. *In*: Vieira B.C., Salgado A.A.R., Santos L.J.C. (Eds.). *Landscapes and landforms of Brazil:* world geomorphological landscapes. Dordrecht: Springer, 403 p. http://dx.doi.org/10.1007/978-94-017-8023-0_26

Wang H., Zhang L., Yin K., Luo H., Li J. 2021. Landslide identification using machine learning. *Geoscience Frontiers*, **12**(1):351-364. http://dx.doi.org/10.1016/j.gsf.2020.02.012

Wang W., He Z., Han Z., Li Y., Dou J., Huang J. 2020. Mapping the susceptibility to landslides based on the deep belief network: a case study in Sichuan Province, China. *Natural Hazards*, **103**:3239-3261. https://doi.org/10.1007/s11069-020-04128-z

Zhou C.H., Lee C.F., Li J., Xu Z.W. 2002. On the spatial relationship between landslides and causative factors on Lantau Island, Hong Kong. *Geomorphology*, **43**(3-4):197-207. https://doi.org/10.1016/S0169-555X(01)00130-1