AUTOMATED LANDSLIDE DETECTION USING ARTIFICIAL NEURAL NETWORKS AND SENTINEL-2 IMAGERY: A CASE STUDY IN MARIANA, BRAZIL

Mateus Oliveira Xavier[1] and César Falcão Barella[2]
[1]Department of Architecture and Urban Planning, Federal University of Ouro Preto, MG-Brazil
mateus.xavier@ufop.edu.br
[2]Department of Environmental Engineering, Federal University of Ouro Preto, MG-Brazil
cesarbarella@ufop.edu.br

ABSTRACT

Landslide inventories are crucial for risk management and for developing and validating susceptibility and hazard models. Traditional methods, like visual interpretation of aerial photos and satellite imagery, field mapping, topographic surface interpretation, and literature reviews, are often time-consuming, costly, and subjective. Automated methods are becoming more important for saving time and resources. In this study, we used an Artificial Neural Network (ANN) classification algorithm to detect landslide scars caused by rain events in Mariana, Minas Gerais, Brazil. Input parameters included the slope from local topography and products from change detection techniques using Sentinel-2 images taken before and after the landslides. These images provided data on visible and near-infrared spectral bands and four vegetation indices. Sixteen ANN models were trained using various dataset configurations in binary point formats representing landslide and non-landslide data. The optimal dataset included 25 points on five pre-selected landslides and 50 non-landslide points from diverse land uses. This model achieved 89.41% precision, 90.48% sensitivity, and an F1-Score of 0.8994. These results were derived by comparing 84 landslides in a restricted portion of the study area and juxtaposing individual landslide detection using the automated model with a manually created inventory based on high spatial resolution images from a Remotely Piloted Aircraft System (RPAS) and fieldwork. The automated inventory matched well with the manual inventory, demonstrating the effectiveness of the ANN in quickly, safely, and cost-effectively detecting landslides using free satellite images and open-source machine learning tools.

KEYWORDS: *Landslide Inventory; Artificial Neural Networks; Change Detection; Sentinel-2 Images.*

I. INTRODUCTION

Landslides are geodynamic processes involving the deformation and movement of soil, rock, debris, or a combination of these materials, along one or more surfaces, moving downward and outward from a slope [1–3]. These events occur when the resisting forces of the materials are surpassed by gravitational forces or a combination of gravitational forces and other agents [2,4], which may include triggers such as earthquakes, volcanic eruptions, or intense rainfall. Landslides pose significant damage risks in mountainous regions [5,6] and are considered the second most destructive natural hazard globally, following earthquakes [7].

The starting point for defining strategies aimed at risk management and minimizing natural landslide disasters involves the preparation of inventory maps [8,9]. This process is also fundamental for developing and validating reliable predictive models of susceptibility, hazard, and risk, as it provides key data such as the spatial distribution, geometry, and extent of landslides [10,11]. The development of inventories initially seeks to identify and locate the morphological features left in past records [12,13], allowing for the later addition of more detailed information when available [14,15].

Various techniques are employed to create landslide inventories, including field surveys, historical record analysis, photo interpretation of analog or digital aerial images, and satellite image interpretation, which may be used individually or in combination [16,17]. With advancements in remote sensing and

digital image processing, automatic and semi-automatic methods for detecting landslide scars have become increasingly important in inventory development [18–21]. Among these techniques, machine learning (ML) models have been widely used in landslide-related studies owing to their high robustness and predictive capacity [22–24]. The most commonly used algorithms in this context include support vector machines, artificial neural networks (ANN), random forests, and decision trees, in addition to deep learning techniques and ensemble models [25–28].

However, these approaches are still underexplored in the academic literature, particularly in the context of landslide inventorying, and often present limitations, such as issues related to the access and cost of orbital images, methodological complexity, and the heterogeneity of the investigated terrains. Therefore, the primary aim of this study is to develop a straightforward, free, effective, and reproducible methodology for the automatic identification of landslide scars, focusing on the production of landslide inventories to bolster predictive statistical models (susceptibility, hazard, and risk). In this study, we employed certain spectral indices in conjunction with an image change detection technique to generate independent input variables. We further examined the impact of the number of landslides on the input dataset and the optimal ratio between landslide scars and the remaining study area. The methodology was devised to compile landslide inventories swiftly and dependably in regions with significant variability in spectral signatures owing to land-use changes.

This article is structured into five sections, starting with this introduction. The next section details the study area, followed by the description of the material and methods in Section 3. Section 4 is dedicated to the presentation and discussion of the main results. Finally, Section 5 offers the conclusions of the study.

II. STUDY AREA

The study area encompasses 52 km² and is in the southeast region of Brazil, specifically in the municipality of Mariana, state of Minas Gerais. The terrain is mountainous, with elevations ranging from 486 m to 938 m and an average slope of 21°, which can reach up to 69° locally. The climate is humid, with an average annual temperature ranging between 19.5°C and 21.8°C and precipitation concentrated in the summer months. Local land uses and occupations are diverse, with natural vegetation of many sizes, including bare soil, rivers, roads, urban areas, pasturelands, and agricultural fields. Most of the landslides identified in the study area occurred during a rainfall event in February 2020. According to records from the Brazilian Center for Monitoring and Alerting of Natural Disasters [29] and the Brazilian Water Agency [30], approximately 161 mm of rain accumulated in the study area between February 7 and 13, 2020, with 111 mm recorded on the previous day. This day had the highest number of landslide-related occurrences in the region, as reported by the Municipal Civil Defense.

III. MATERIAL AND METHODS

The objective of this study was to develop a straightforward, rapid, efficient, and cost-free method for creating landslide inventories using parameters derived from local topography (slope) and change detection in images obtained from Sentinel-2 products (visible and near-infrared spectral bands and the NDVI, EVI, GRVI, and GNDVI vegetation indices). For this purpose, we trained ANN models to classify the study area using exclusively open-source tools, such as Orange Data Mining (ODM) version 3.34 [31] for ML and QGIS version 3.12.3 [32] for geoprocessing.

3.1. Input Datasets

We selected products derived from Sentinel-2 mission images from the European Space Agency. This choice was motivated by the fact that these images were free, orthorectified, and had a spatial resolution suitable for identifying most landslides in the study area. Additionally, Sentinel-2 images are extensively utilized in academic research for a wide range of applications, including agricultural monitoring, vegetation analysis, and land use classification [33–35]. The mission consists of a constellation of two satellites in polar orbit with a 180° phase difference between them, temporal resolution of five days, and multispectral optical sensors capable of sampling 13 spectral bands [36,37].

In this study, we focused on the blue (B2), green (B3), red (B4), and near-infrared (B8) bands, which have a spatial resolution of 10 m.

Based on the landslide event that occurred in February 2020, we sought images with minimal cloud interference, which is a challenging task, especially in mountainous areas with high humidity following extreme events, such as earthquakes or heavy storms [38]. Obtaining image pairs for change detection with almost no cloud cover is challenging [39]. In this context, we selected two Sentinel-2 products with cloud coverage below 4.7%: one image from November 5, 2019, prior to the landslide occurrence, and another from April 28, 2020, after the event. The images were classified as Level-2A, bottom-of-atmosphere, and were chosen for atmospheric correction to ensure a more accurate representation of reflectance on the land surface.

We extracted four vegetation indices from the acquired images: NDVI [40], GNDVI [41], GRVI [42], and EVI [43]. Landslide scars were identified by analyzing changes in these vegetation indices over time. High values of these indices indicated healthy vegetation, while low values suggested stress or absence of vegetation, a common feature of recent landslides [44]. It is important to note that other studies have similarly employed vegetation indices for automated landslide mapping [45,46].

The change-detection technique [38,47,48] was used to obtain the majority of the input features for the developed models. Thus, we subtracted the pixel-by-pixel values of the products from November 5, 2019 (prior to the event) and from their corresponding products on April 28, 2020 (post-event). This technique was applied to the four vegetation indices (NDVI, EVI, GNDVI, and GRVI) and the four Sentinel-2 bands (B2, B3, B4 and B8). In addition to the features extracted from the change detection technique, we incorporated a 10 m spatial resolution slope raster, derived from pre-failure topography into the models. This inclusion is warranted as slope is recognized as a critical factor for landslide occurrence [49,50] and serves as a refinement criterion for model classification. Therefore, slope is an essential parameter for separating areas that exhibit land-use alternations in regions of low inclination, such as agricultural areas that seasonally alternate between vegetation and bare soil.

To ensure that no features gained undue significance during the processing phase, data normalization was conducted. This step was critical in preventing scale differences between the input variables from affecting the learning capabilities of the models.

To deepen our understanding of the operations and choices made by the models, we assessed the relevance of the input parameters using the Gini Index (GI) [51], a widely used metric in the field of machine learning [13,52]. The GI values assigned to each parameter were derived from global assessments that considered the averages of all the developed models. This approach provides a comprehensive and well-founded insight into the influence of attributes on the decision-making process of models, thus solidifying our understanding of the process.

3.2. Artificial Neural Network

An Artificial Neural Network (ANN) is a nonparametric supervised learning classification method introduced by McCulloch and Pitts [53]. In this study, we utilized a multilayer perceptron ANN algorithm with backpropagation. This powerful algorithm is effective for landslide analysis [54] and can learn nonlinear models through a feedforward network consisting of input and output layers, along with one or more hidden layers containing a sufficient number of neurons [13,55]. The configurations of the ANN algorithm used in this study were optimized to make the best use of available resources, ensuring that the models achieved high operational performance.

To reduce the risk of overfitting and maintain simple models without increasing computational costs, a single hidden layer of neurons was adopted, with the number determined by equation $H \leq 2n + 1$, as proposed by Hecht-Nielsen [56], where H = number of neurons in hidden layer, and n = number of input parameters. Given the incorporation of nine input parameters into the models, the number of neurons in the hidden layer was set to its maximum value, totalling 19.

A rectified linear activation function (ReLu) was used as the activation function [57,58]. This function is one of the most widely adopted activation functions in ANN training due to its simplicity and ease of computation [59,60]. For the optimization process, the Adaptive Moment Estimation (Adam) optimizer,

introduced by Kingma and Ba [61], was used to properly calibrate the weights associated with the neurons, aiming for a systematic minimization of the prediction error. Compared to other optimization algorithms, Adam offers superior computational efficiency, memory efficiency, and reduced need for tuning [62]. Furthermore, Adam is frequently used in ML studies focused on landslide analysis, demonstrating better performance compared to other optimizers [63,64]. Finally, an early stopping technique was implemented as a regularization strategy to prevent overfitting of the models to the training dataset [65,66]. After testing, stopping criteria were established with a threshold of 500 epochs and a learning rate of $\alpha = 0.5$, which proved to be suitable for the effective performance of the models.

The ANN models were trained for the automated identification of landslides in the study area using various configurations of input-dependent variables. Therefore, the number of records with and without landslides was varied to assess the performance and precision of the developed models. Records without landslides were derived from random points distributed over polygons delineated in areas identified in the post-event image. Six distinct zones were identified: high vegetation, low vegetation, water bodies, urban areas, burned areas, and seasonally bare soil.

The records pertaining to landslides originated from five occurrences stemming from a pluviometric event recorded in February 2020. The landslides were identified through fieldwork and photointerpretation of images obtained by the RPAS. Five occurrences were manually outlined, and random points were placed within them. A total of 16 ANN models were trained and categorized into four groups based on the number of landslide points used: 25, 50, 75, and 100. The amount of non-landslide data in each model was determined using a multiplier factor in relation to the number of landslides, which varied from 1 to 4. In this context, the proportions of the dependent input variables (landslides:non-landslides) were 1:1, 1:2, 1:3, and 1:4. Thus, the number of landslide and non-landslide points used in the training of the 16 ANN models was as follows: 25:25, 25:50, 25:75, 25:100, 50:50, 50:100, 50:150, 50:200, 75:75, 75:150, 75:225, 75:300, 100:100, 100:200, 100:300, and 100:400.

3.3. Performance Evaluation

A performance evaluation of the models was conducted by comparing the results achieved with a manually crafted landslide inventory meticulously compiled through fieldwork and photointerpretation of images obtained via RPAS. This investigation focused on a limited section of the study area, indicated in red in Figure 1, where 84 landslides were identified, including one emphasized in the same figure.

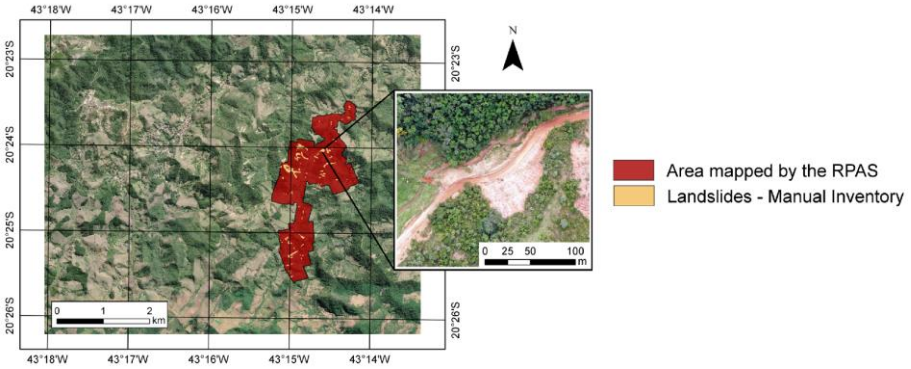


Figure 1. Study area highlighting the region mapped by RPAS for model performance validation

Following the manual inventory of landslides, two validation approaches were performed. The first approach aimed to evaluate the accuracy of the automatically inventoried landslide geometry compared to the manually mapped features in the official inventory obtained through fieldwork and photointerpretation. To this end, statistical analyses were conducted to compare the 84 identified and mapped landslides on a pixel-by-pixel basis with the inventory produced by various ANN models. The quality of the models was assessed using four traditional metrics derived from the confusion matrix, namely accuracy, precision, sensitivity, and specificity.

Additionally, we assessed the agreement of the results of each model using the Kappa coefficient (k), a statistical index introduced by Cohen [67], which measures the agreement between two independent

classifications. For a clearer interpretation of the results, we adopted the evaluation criteria suggested by Landis and Koch [68], which have been used in cartographic studies [69–72]. Kappa values less than 0 indicate "poor" agreement; between 0 and 0.2, the agreement is considered "slight"; between 0.2 and 0.4, it is deemed "fair"; between 0.4 and 0.6, the agreement is "moderate"; between 0.6 and 0.8, it becomes "substantial"; and values between 0.8 and 1 indicate "almost perfect" agreement.

The second validation approach was conducted to verify only the ability of the models to identify landslides in the form of point features without pixel-by-pixel comparison. The goal was to assess the capability of the models to recognize the presence of landslides in the landscape, regardless of whether there was a perfect spatial match between the geometries of the automated and manual inventories. This approach is justified because most landslide inventories used in susceptibility analyses consist of point-based geometries [73]. Therefore, true positives, false positives, and false negatives were computed as discrete point values. The statistical metrics of precision, sensitivity, and their harmonic mean, known as the F1-score, were utilized at this stage. Due to the inability to calculate true negatives, which involve area units while the rest of the confusion matrix consists of discrete point values, accuracy and specificity were not measurable in this phase.

IV. RESULTS AND DISCUSSION

4.1. Assessment of Reflectance Differences Between the Adopted Sentinel-2 Images

Understanding the complexities and specific characteristics of the study area is crucial for understanding the applicability of the parameters used in ML models. Figure 2 presents the spectral alterations between the adopted Sentinel-2 images in our study concerning wavelengths across the bands of the visible spectrum (blue, green, and red) and near-infrared spectrum. The reflectance differences displayed correspond to known targets in the study area and demonstrate their changes in areas of high vegetation, low vegetation, burned areas, seasonally bare soil, water bodies, urban areas, and landslides.

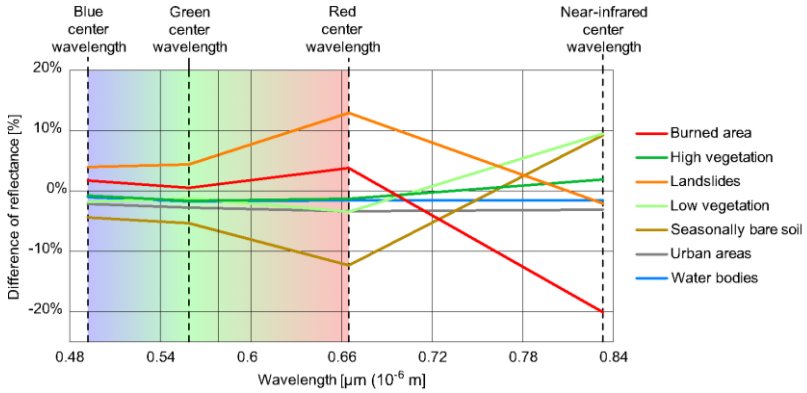


Figure 2. Spectral signature differences between Sentinel-2 images from April 28, 2020, and November 5, 2019

The reflectivity differences between urban areas and water bodies were relatively moderate in the four analyzed bands, indicating that these targets did not undergo significant changes between the image acquisition dates. High vegetation did not undergo significant changes in any of the bands; this was because the vegetation in the study area is primarily composed of evergreen trees with dense canopies, which are characteristic of humid Atlantic Forest regions [74,75], and did not result in significant changes in the reflectance characteristics of this vegetation.

In contrast to high vegetation, which remained unchanged between the acquisition dates of the two analyzed images, vegetative growth was observed in both low vegetation and areas of seasonally bare soil, resulting in increased photosynthetic activity and chlorophyll content. These characteristics are particularly detectable in the red and near-infrared spectral bands [76,77]. For low vegetation, this growth is primarily indicated by an increase in reflectance in the near-infrared spectrum. For seasonally bare soil, in addition to the increased reflectance in this spectral range, vegetative growth caused a notable decrease in the reflectance across all bands of the visible spectrum, especially in the red band.

In the burned area, there was a slight increase in reflectance in the red band and a significant decrease in reflectance in the near-infrared region. These changes were attributed to a reduction in vegetation owing to fires that occurred between the analyzed periods. This is because charred vegetation has low ability to reflect light in the near-infrared range [78,79].

The landslides exhibited significant increases in reflectance differences within the visible spectrum bands, owing to the displacement of the vegetative cover that previously occupied their areas before the development of geodynamic processes. However, for the near-infrared band, there was no large variation in reflectance as the soil exposed to the landslides showed similar values to the vegetation that previously occupied these areas. The reflectance difference signatures of the landslides and burned areas exhibited similar behaviors. Nevertheless, a distinction between these two regions is possible because of the higher reflectance values of the landslides.

One notable aspect is the inverse relationship between the signature differences between areas affected by landslides and those with seasonally bare soil, particularly in the visible spectrum bands. It was observed that, while vegetative cover for seasonally bare soil increased after the rainy season, vegetative cover disappeared in landslide areas. Thus, in the process of change detection, the reflectance values found in these areas for the visible spectrum were similar but with opposite signs. This means that if landslides occurred during the same period when seasonally bare soil was exposed, the task of identifying and differentiating these two soil covers would become challenging, as both would be easily confused in manual or automated classification processes.

4.2. Assessment of Change Detection Products

Figure 3 shows the normalized results obtained through the application of the image differencing technique across the blue, green, red, and near-infrared spectral bands and for the vegetation indices EVI, GNDVI, NDVI, and GRVI. Additionally, the GI values for the same spectral bands and vegetation indices are displayed in the same figure.

As indicated in Section 4.1, urban areas, water bodies, and high vegetation did not exhibit significant reflectance changes in any of the spectral bands. This was reflected in all image-differencing products in which these regions displayed pixel values closer to zero.

For the visible spectrum bands, particularly red, the greatest discrepancies were observed in the landslides and the smallest in the seasonal bare soil, where the landslides exhibited higher pixel values and the seasonal bare soil displayed lower values. However, these regions are not easily distinguishable when analysing difference images in the near-infrared region. In this spectral range, the reflectance differences between landslides and high vegetation were not as pronounced as expected, resulting in difficulty in distinguishing these regions. The same occurred in areas with low vegetation and seasonally bare soil, which also became confounded due to similarities in pixel values. Conversely, the burned area was prominent in the near-infrared difference image, showing the lowest reflectance difference values for this band of the spectrum.

Upon applying the change detection technique to the vegetation indices EVI, GNDVI, NDVI, and GRVI, substantial changes, such as increases or decreases in vegetation cover, were highlighted because these indices are capable of quantifying the health and quantity of vegetation. Consequently, in the different images related to the vegetation indices, the landslide and burned areas were distinguished by reduced pixel values, as these regions experienced vegetation loss. Conversely, areas with seasonally uncovered soil showed higher pixel values due to vegetative growth in these zones.

Furthermore, it is important to highlight that the GI results indicate that the visible spectrum bands exerted the greatest influences on the developed ML models. Immediately thereafter, the slope (GI=0.229), which is not represented in Figure 3, was followed by the vegetation indices, and lastly, the near-infrared band. This underscores the significance of the visible spectral features and slopes in the decisions made by the models.

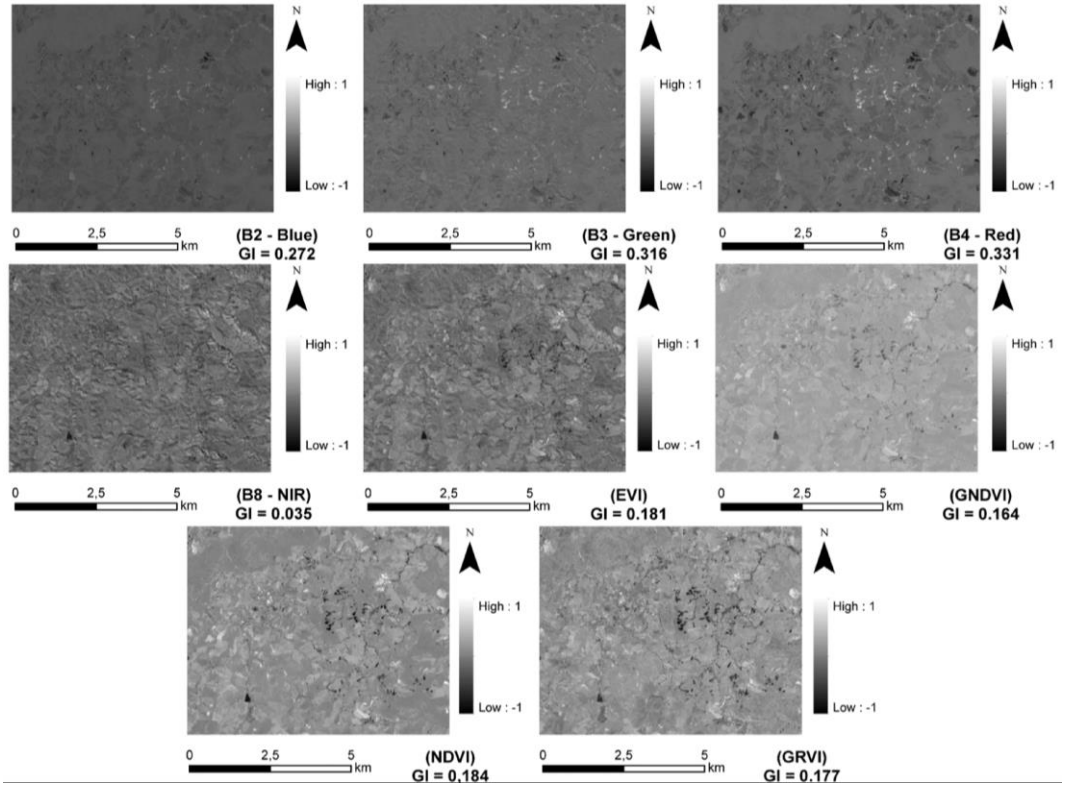


Figure 3. Normalized results from image differencing for the blue, green, red, and near-infrared bands, and for the vegetation indices EVI, GNDVI, NDVI, and GRVI.

4.3. Performance of ANN Models for Automatic Detection of Landslides

Using the established input parameters, 16 ANN models were developed for automatic landslide inventory. All models achieved accuracy results above 97%, with an average value of 97.41% and a standard deviation of 0.05%. Although this metric is widely adopted in many studies, it is highly dependent on the ratio of positives to negatives [80]. The results for the specificity, sensitivity, precision, and kappa agreement index are shown in Figure 4. The plotted lines represent the models created with sets of 25, 50, 75, and 100 points located within the five landslides identified in the field. The number of non-landslide points was determined using a multiplication factor relative to the number of points in the landslides. The multiplication factors, which range from 1 to 4, are indicated along the x-axis in all the charts presented. The y-axis represents the values of the metrics evaluated for each model configuration.

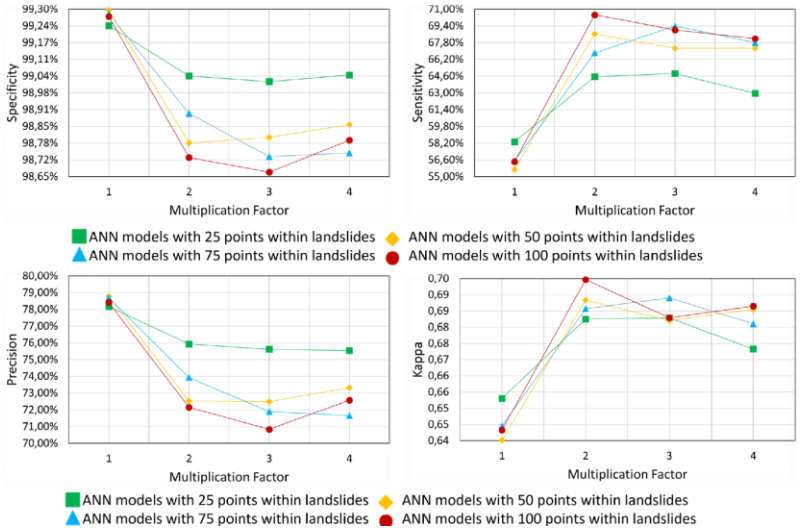


Figure 4. Performance metrics for the different ANN models (First validation approach).

The high specificity levels observed across all models indicated that the number of false positives was substantially lower than the true negative rate, resulting in specificity values above 98.67%. This demonstrates that the models performed very well in accurately identifying areas without landslides, minimizing the error rate associated with incorrectly assigning landslide features to unaffected regions. Similar to the accuracy result, this outcome should be interpreted with caution, as it tends to be influenced by the non-landslide area, which is proportionally larger. Therefore, it is recommended to evaluate the models performance using precision and sensitivity metrics, which are not influenced by the disproportion between landslide and non-landslide points.

Upon analyzing the precision, it was found that these same false positives constituted a significant portion of the positive classifications, reflected in precision values ranging between 70.84% and 78.78%. This is due to the relatively smaller area occupied by landslides compared to the total non-landslide area, resulting in false positives having a more pronounced impact on the precision of the models than on its specificity. The sensitivity varied between 55.68% and 70.45%, indicating that while all models were capable of correctly identifying most areas with landslides, there was a degree of omission in the results, indicating that some of the areas affected by landslides were not detected by the models.

However, even the most precise landslide inventory maps typically do not encompass all movements that have occurred in a region [16]. Thus, the results are highly dependent on the method used and the specialist knowledge of the responsible technician [17,19,81]. Additionally, given that statistical landslide susceptibility models aim to identify patterns in the predisposing factors of landslide areas to predict future occurrences, minor omissions will not necessarily result in deficient modeling. Therefore, while the inventory may not capture all landslide features, it should minimize, whenever possible, the identification of false positives. In this context, it is preferable for the inventory to prioritize greater specificity over sensitivity.

In addition to the aforementioned performance metrics, the calculation of the kappa agreement index for all models showed that they had a substantial level of agreement, with values fluctuating between 0.64 and 0.70. Therefore, it can be stated that all models produce automatically generated inventories that are spatially similar to the manually compiled inventory.

Visually, both false positives and false negatives were concentrated at the edges of the identified areas. This phenomenon is visually demonstrated in Figure 5, which provides a comparison between landslides from the manually constructed inventory and those identified by the model developed with 25 landslide points and 50 non-landslide points (multiplication factor of 2). The visual analysis in Figure 5 shows that both false positives and false negatives are particularly associated with the peripheries of the manually mapped landslides.

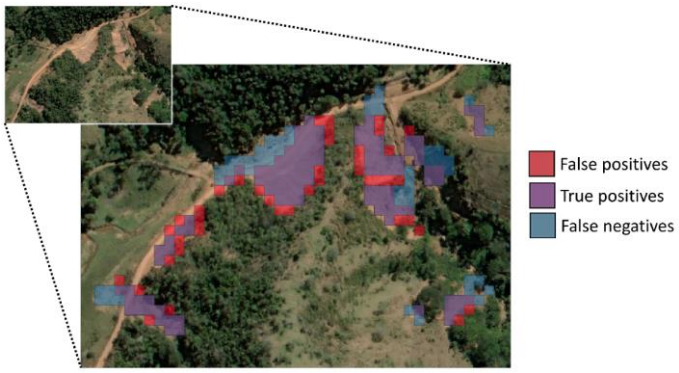


Figure 5. Comparison of landslides identified in the manually constructed inventory and by the model with a ratio of 25 landslide points to 50 non-landslide points

Given that the purpose of the proposed inventory is to integrate it into susceptibility models and considering that many such models are constructed using one point per landslide [73,82], the imperfect identification of landslide scar edges does not preclude the development of robust susceptibility models as long as the reference point is carefully positioned in locations that have been genuinely affected. In this context, the precision and sensitivity values were recalculated based on the individual identification

of each landslide (Figure 6), rather than a pixel-by-pixel comparison between manually and automatically produced inventories. Any intersection between landslides identified manually and by the model was considered a true positive entry in the confusion matrix, regardless of the perfect overlay between the two.

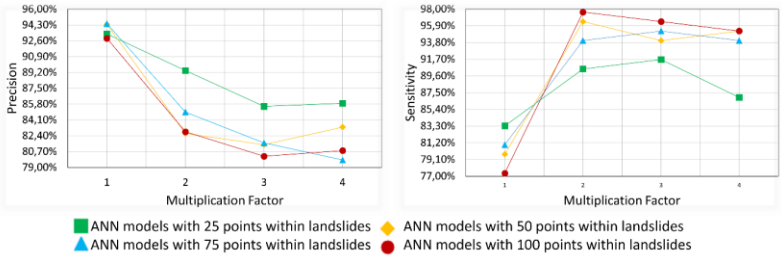
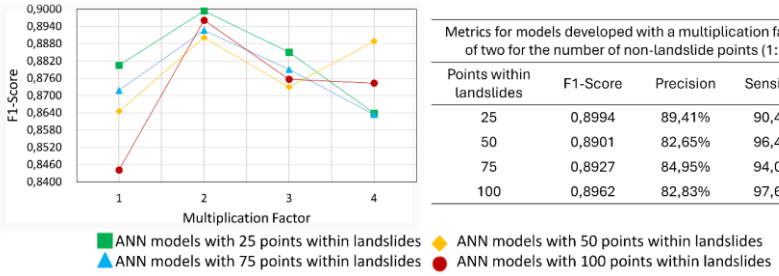


Figure 6. Precision and sensitivity recalculated based on the individual identification of each landslide

The precision of the models varied from 79.80% to 94.44%. Models that adopted balanced input data ratios between landslide and non-landslide areas exhibited the highest precision scores, exceeding 92.86%. However, it is important to note that these same models (1:1) recorded the lowest sensitivity values, ranging from 77.38% to 83.33%. Such results of high precision and relatively low sensitivity, as also found in the works in [24,28], suggest that these models minimize the misclassification of features identified as landslides by the algorithm, although they are not capable of identifying all existing landslides within the study area. The analysis of the graphs also indicated that increasing the number of non-landslide points (1:2, 1:3, and 1:4) used in the model input correlated with an increase in false positives and a reduction in false negatives. This pattern translates into a trade-off in which models with lower precision exhibit higher sensitivity. This implies that as the models become more sensitive to landslide detection, they tend to be less precise due to the increased occurrence of false positives.

There are no established standards for acceptable precision and sensitivity values in the automation of landslide inventories using machine learning. However, regardless of the spatial comprehensiveness of the landslide inventories, the inventories used to construct statistical models of susceptibility should allow for the recognition of patterns in landslides that are used as dependent input variables [49]. Thus, inventories must be representative of the phenomenon investigated and cover a significant number of events that have occurred in the region. Consequently, although a high precision rate is desirable in an automated landslide inventory, it should be balanced with sensitivity, underscoring the importance of an adequate and representative inventory for the construction of susceptibility models.

To assess the trade-off between precision and sensitivity and select a balanced model, the F1-score metric can be used. The F1-score provides a harmonic measure that balances the ability to correctly identify landslides (avoiding false positives) and detect the highest number of events in the study area (avoiding false negatives). The four models developed with a multiplication factor of two for the number of non-landslide points (1:2) showed the highest values for the metric, as detailed in Figure 7. Although the models achieved similar F1-scores, the standout was constructed with 25 points located within 5 landslides, with precision and sensitivity values of 89.41% and 90.48%, respectively. This demonstrates that the model achieved a remarkable balance between the ability to make precise classifications and effectively identify point landslides in the study area.



Metrics for models developed with a multiplication factor of two for the number of non-landslide points (1:2)

Points within landslides	F1-Score	Precision	Sensitivity
25	0,8994	89,41%	90,48%
50	0,8901	82,65%	96,43%
75	0,8927	84,95%	94,05%
100	0,8962	82,83%	97,62%

Figure 7. Balance between precision and sensitivity assessed with the F1 score metric

4.4. Comparative Visualization of Landslide Inventories Generated by the Top Four ANN Models Constructed

The inventories produced by the top four ANN models (25:50, 50:100, 75:150 e 100:200) are shown in Figure 8. Although the model constructed with 25 points located within 5 landslides achieved the highest F1 score, visually, the four inventories displayed considerable similarities in the locations and dimensions of the detected landslides.

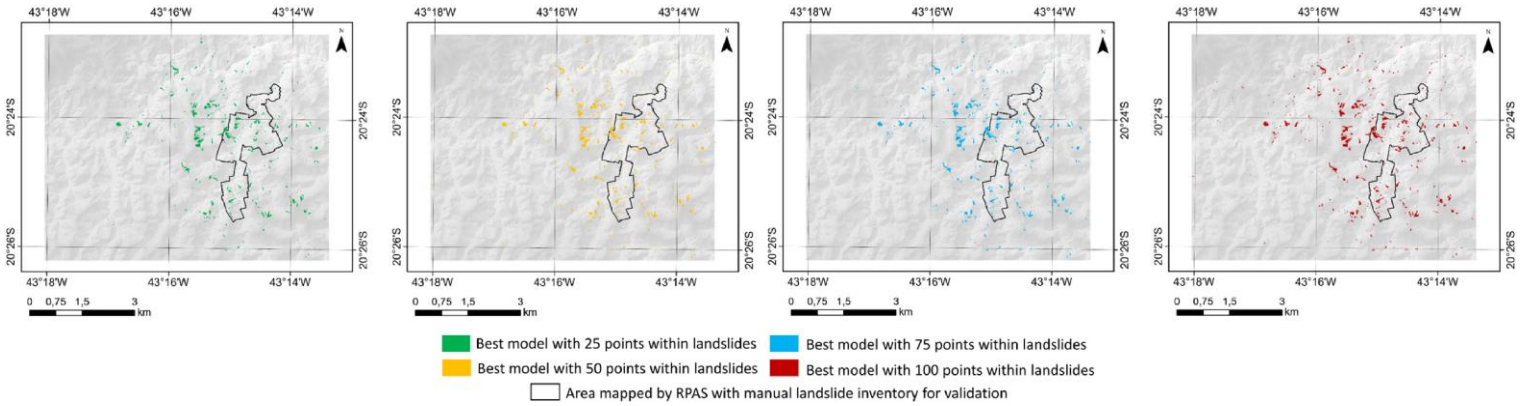


Figure 8. Visual comparison of landslide inventories generated by the four ANN models

Figure 9 highlights a specific region within the study area to emphasize the overlap resulting from the top four constructed models. The shaded area indicates congruence among the four inventories produced, signaling a common territory identified by all models. Furthermore, Figure 9 illustrates that the noticeable discrepancies in the results were primarily situated at the edges of the marked landslides. It is evident that the models, when identifying landslides, exhibit remarkable consistency in their central regions, but manifest more significant variations at their peripheries or outlines.

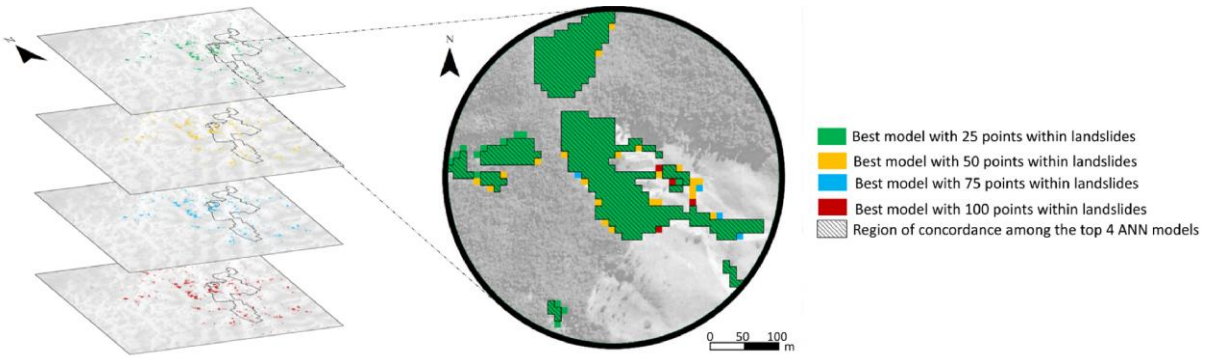


Figure 9. Overlap and variations in landslide detection by the top four models

Thus, the spatial agreement of the predictions also reinforces the competence of this approach in efficiently and consistently mapping landslides. Furthermore, if the decision is to develop a point-based landslide inventory, as implemented in most statistical susceptibility models, automated landslide detection becomes even less susceptible to edge errors. Therefore, the presented automation process is valuable and enables the rapid inventory of landslides. It can be applied to the production of regional and national landslide inventories, especially in areas characterized by rapid vegetation recovery or frequently affected by intense precipitation. This approach is increasingly relevant as climate change predicts more intense rainfall in certain areas due to global warming [83].

Due to the 10-meter resolution of Sentinel-2 images, the models are limited in detecting small landslide scars (< 100m²). This limitation can be overcome by using higher-resolution input raster parameters. Moreover, it was observed in the four main models performed (25:50, 50:100, 75:150 e 100:200) that the majority of false positives are linked to 100 m² landslide scars, corresponding to a single pixel (10m x 10m). Thus, applying a filter to remove scars with an area of 100 m² could enhance the precision of

the models. However, filtering would also remove true positives, consequently reducing the sensitivity of the models. This approach may prove beneficial and aligns with findings from other authors, suggesting that statistical susceptibility models can be developed without complete inventories, provided a representative dataset of landslides is used [84]. In summary, the inventory may not identify all existing landslides in the landscape, but it should prevent the mapping of incorrect features. It is important to emphasize that the application of area filters is specific to both the study region and the mapped landslides, with the responsible technician for the inventory making the decision.

V. CONCLUSIONS

This study evaluated the applicability of a machine learning tool for the automated identification of landslides resulting from rainfall events using free and open-source satellite imagery and an ANN algorithm. The results indicate that the employed technique demonstrates competence in automatically identifying landslides, showing an effectiveness comparable to that of manual inventorying performed through the visual interpretation of aerial images and field mapping. The main conclusions are as follows: (1) The difference image technique used in the sampled input parameters proved to be simple and efficient for the automatic identification of landslide features in the landscape. The adopted approach highlighted landslide areas in contrast to other forms of land use and cover and can be easily replicated in other regions. (2) The visible spectrum bands, followed by slope angle and vegetation indices, were the most influential parameters in the automatic landslide scar detection models. Near-infrared change detection proved less influential. (3) Sixteen ANN models were trained with variations in the input dataset proportion. The models that adopted a dataset based on a 1:2 ratio of landslides to non-landslides (25:50, 50:100, 75:150 e 100:200) produced the best results. (4) Despite the similarity observed in validation results of the models, a significant portion of errors was associated with delineating the boundaries of landslide features. However, these errors can be mitigated through the adoption of point features rather than polygonal ones. This strategy commonly enhances validation metrics and is widely utilized in statistically based landslide susceptibility mapping. (5) Although the top four developed models exhibited similar statistical and visual results, the model constructed with the least amount of input data (25 points within 5 landslides and 50 points outside of landslides) achieved the highest precision (89.41%) and the best balance between precision and sensitivity (90.48%), with an F1 score of 0.8994. This indicates that for landslide identification in the study area, an extensive database is not required, but rather a representative dataset is sufficient.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to the European Space Agency for providing free Sentinel-2 data and the Federal University of Ouro Preto (UFOP) and the Coordination for the Improvement of Higher Education Personnel (CAPES) of the Federal Government of Brazil for their support.

REFERENCES

[1]. Varnes, D.J. (1978) *Slope Movement Types and Processes. In Landslides: Analysis and Control*, Schuster, R.L., Krizker, R.J., Eds.; Transp. Res. Board, Nat. Ac. Sc.: Washington, DC, USA, pp.11–33.

[2]. Terzaghi, K. (1950) *Mechanism of Landslides. In Application of geology to engineering practice*; Paige, S., Ed.; Geological Society of America: New York, NY, USA, Vol. Berkey volume, pp.83–123.

[3]. Hungr, O.; Leroueil, S.; Picarelli, L. (2014) "The Varnes Classification of Landslide Types, an Update", Landslides, 11, pp.167–194, doi:10.1007/s10346-013-0436-y.

[4]. Turner, A.K. (2018) "Social and Environmental Impacts of Landslides", Innovative Infrastructure Solutions, 3, 25p., doi:10.1007/s41062-018-0175-y.

[5]. Laimer, H.J. (2017) "Anthropogenically Induced Landslides – A Challenge for Railway Infrastructure in Mountainous Regions". Eng Geol, 222, pp.92–101, doi:10.1016/j.enggeo.2017.03.015.

[6]. Shano, L.; Raghuvanshi, T.K.; Meten, M. (2020) "Landslide Susceptibility Evaluation and Hazard Zonation Techniques-a Review", Geoenvironmental Disasters, 7, 19p., doi:10.1186/s40677-020-00152-0.

[7]. Qi, T.; Zhao, Y.; Meng, X.; Shi, W.; Qing, F.; Chen, G.; Zhang, Y.; Yue, D.; Guo, F. (2021) "Distribution Modeling and Factor Correlation Analysis of Landslides in the Large Fault Zone of the Western Qinling Mountains: A Machine Learning Algorithm", Remote Sens, 13, 24p., doi:10.3390/rs13244990.

[8]. Calligaris, C.; Poretti, G.; Tariq, S.; Melis, M.T. (2013) "First Steps towards a Landslide Inventory Map of the Central Karakoram National Park", Eur J Remote Sens, 46, pp.272–287, doi:10.5721/EuJRS20134615.

[9]. Kamp, U.; Growley, B.J.; Khattak, G.A.; Owen, L.A. (2008) "GIS-Based Landslide Susceptibility Mapping for the 2005 Kashmir Earthquake Region", Geomorphology, 101, pp.631–642.

[10]. Anshori, R.M.; Samodra, G.; Mardiatno, D.; Sartohadi, J. (2022) "Volunteered Geographic Information Mobile Application for Participatory Landslide Inventory Mapping", Comput Geosci, 161, 12p.

[11]. van Westen, C.J.; Castellanos, E.; Kuriakose, S.L (2022) "Spatial Data for Landslide Susceptibility, Hazard, and Vulnerability Assessment: An Overview", Eng Geol 2008, 102, pp.112–131.

[12]. Ramos-Bernal, R.N.; Vázquez-Jiménez, R.; Romero-Calcerrada, R.; Arrogante-Funes, P.; Novillo, C.J. (2018) "Evaluation of Unsupervised Change Detection Methods Applied to Landslide Inventory Mapping Using ASTER Imagery", Remote Sens, 10, 24p., doi:10.3390/rs10121987.

[13]. Uehara, T.D.T.; Passos Corrêa, S.P.L.; 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, pp.665–680, doi:10.14393/rbcv72n4-54037.

[14]. Hervás, J. (2013) *Landslide Inventory*. In Encyclopedia of Natural Hazards; Bobrowsky, P.T., Ed.; Encyclopedia of Earth Sciences Series; Springer Netherlands: Dordrecht, pp.610–611 ISBN 978-90-481-8699-0.

[15]. Fell, R.; Corominas, J.; Bonnard, C.; Cascini, L.; Leroi, E.; Savage, W.Z. (2008) "Guidelines for Landslide Susceptibility, Hazard and Risk Zoning for Land Use Planning", Eng Geol, 102, pp.85–98, doi:10.1016/j.enggeo.2008.03.022.

[16]. Galli, M.; Ardizzone, F.; Cardinali, M.; Guzzetti, F.; Reichenbach, P. (2006) "Comparing Landslide Inventory Maps", Geomorphology, 94, pp.268–289, doi:10.1016/j.geomorph.2006.09.023.

[17]. Guzzetti, F.; Cardinali, M.; Reichenbach, P.; Carrara, A. (2000) "Comparing Landslide Maps: A Case Study in the Upper Tiber River Basin, Central Italy", Environ Manage, 25, pp.247–263.

[18]. Dias, H.C.; Hölbling, D.; Grohmann, C.H. (2021) "Landslide Susceptibility Mapping in Brazil: A Review". Geosciences, 11, 15p.

[19]. 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 Sci Rev, 112, pp.42–66, doi:10.1016/j.earscirev.2012.02.001.

[20]. Wang, H.; Zhang, L.; Yin, K.; Luo, H.; Li, J. (2021) "Landslide Identification Using Machine Learning", Geoscience Frontiers, 12, pp.351–364, doi:10.1016/j.gsf.2020.02.012.

[21]. Yu, B.; Chen, F.; Xu, C.; Wang, L.; Wang, N. (2021) "Matrix Segnet: A Practical Deep Learning Framework for Landslide Mapping from Images of Different Areas with Different Spatial Resolutions", Remote Sens, 13p., doi:10.3390/rs13163158.

[22]. Ermini, L.; Catani, F.; Casagli, N. (2005) "Artificial Neural Networks Applied to Landslide Susceptibility Assessment", Geomorphology, 66, pp.327–343, doi:10.1016/j.geomorph.2004.09.025.

[23]. Su, Z.; Chow, J.K.; Tan, P.S.; Wu, J.; Ho, Y.K.; Wang, Y.H. (2021) "Deep Convolutional Neural Network–Based Pixel-Wise Landslide Inventory Mapping", Landslides, 18, pp.1421–1443.

[24]. Prakash, N.; Manconi, A.; Loew, S. (2021) "A New Strategy to Map Landslides with a Generalized Convolutional Neural Network", Sci Rep, 11p., doi:10.1038/s41598-021-89015-8.

[25]. Goetz, J.N.; Brenning, A.; Petschko, H.; Leopold, P. (2015) "Evaluating Machine Learning and Statistical Prediction Techniques for Landslide Susceptibility Modeling", Comput Geosci, 81, 11p., doi:10.1016/j.cageo.2015.04.007.

[26]. Meena, S.R.; Ghorbanzadeh, O.; van Westen, C.J.; Nachappa, T.G.; Blaschke, T.; Singh, R.P.; Sarkar, R. (2021) "Rapid Mapping of Landslides in the Western Ghats (India) Triggered by 2018 Extreme Monsoon Rainfall Using a Deep Learning Approach", Landslides, 18, pp.1937–1950, doi:10.1007/s10346-020-01602-4.

[27]. Nhu, V.H.; Mohammadi, A.; Shahabi, H.; Ahmad, B. Bin; Al-Ansari, N.; Shirzadi, A.; Geertsema, M.; Kress, V.R.; Karimzadeh, S.; Kamran, K.V. (2020) "Landslide Detection and Susceptibility Modeling on Cameron Highlands (Malaysia): A Comparison between Random Forest, Logistic Regression and Logistic Model Tree Algorithms", Forests, 11, 28p., doi:10.3390/F11080830.

[28]. Amankwah, S.O.Y.; Wang, G.; Gnyawali, K.; Hagan, D.F.T.; Sarfo, I.; Zhen, D.; Nooni, I.K.; Ullah, W.; Zheng, D. (2022) "Landslide Detection from Bitemporal Satellite Imagery Using Attention-Based Deep Neural Networks", Landslides, doi:10.1007/s10346-022-01915-6.

[29]. CEMADEN (2024) "Mapa Interativo Da Rede Observacional Para Monitoramento de Risco de Desastres Naturais Do CEMADEN" [Online] Available: http://www.cemaden.gov.br/mapainterativo (accessed on 26 March 2024).

[30]. ANA (2023) "Portal HidroWeb"[Online]Available: https://www.snirh.gov.br/hidroweb/serieshistoricas (accessed on 3 July 2023).

[31]. Demsar, J.; Curk, T.; Erjave, A.; Gorup, C.; Hocevar, T.; Milutinovic, M.; Mozina, M.; Polajnar, M.; Toplak, M.; Staric, A. (2013) "Orange: Data Mining Toolbox in Python", Journal of Machine Learning Research, 14, pp.2349–2353.

[32]. QGIS Development Team (2020), QGIS 3.12.3 - Geographic Information System.

[33]. Phiri, D.; Simwanda, M.; Salekin, S.; Nyirenda, V.R.; Murayama, Y.; Ranagalage, M. (2020) "Sentinel-2 Data for Land Cover/Use Mapping: A Review", Remote Sens, doi:10.3390/rs12142291.

[34]. Qu, F.; Qiu, H.; Sun, H.; Tang, M. (2021) "Post-Failure Landslide Change Detection and Analysis Using Optical Satellite Sentinel-2 Images", Landslides, 18, pp.447–455, doi:10.1007/s10346-020-01498-0.

[35]. Segarra, J.; Buchaillot, M.L.; Araus, J.L.; Kefauver, S.C. (2020) "Remote Sensing for Precision Agriculture: Sentinel-2 Improved Features and Applications", Agronomy.

[36]. Immitzer, M.; Vuolo, F.; Atzberger, C. (2016) "First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe", Remote Sens, doi:10.3390/rs8030166.

[37]. Vuolo, F.; Neuwirth, M.; Immitzer, M.; Atzberger, C.; Ng, W.T. (2018) "How Much Does Multi-Temporal Sentinel-2 Data Improve Crop Type Classification?", International Journal of Applied Earth Observation and Geoinformation, 72, pp.122–130, doi:10.1016/j.jag.2018.06.007.

[38]. Yang, X.; Chen, L. (2010) "Using Multi-Temporal Remote Sensor Imagery to Detect Earthquake-Triggered Landslides", International Journal of Applied Earth Observation and Geoinformation, 12, pp.487–495, doi:10.1016/j.jag.2010.05.006.

[39]. Ju, J.; Roy, D.P. (2008) "The Availability of Cloud-Free Landsat ETM+ Data over the Conterminous United States and Globally", Remote Sens Environ, 112, pp;1196–1211, doi:10.1016/j.rse.2007.08.011.

[40]. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. (1973) "Monitoring Vegetation Systems in the Great Plains with ERTS", In Proceedings of the Third Earth Resources Technology Satellite—1 Symposium; Greenbelt, MD, USA, pp.309–317.

[41]. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N.; Blaustein, J. (1995) "Use of a Green Channel in Remote Sensing of Global Vegetation from EOS-MODIS", Remote Sens Environ, 58, pp.289–298.

[42]. Tucker, C.J. (1979) "Red and Photographic Infrared Linear Combinations for Monitoring Vegetation", Remote Sens Environ, 8, pp.127–150, doi:10.1016/0034-4257(79)90013-0.

[43]. Huete, A.; Justice, C.; Van Leeuwen, W. (1999) "MODIS Vegetation Index (MOD13). Algorithm Theoretical Basis Document."

[44]. Uehara, T.D.T.; Soares, A.R.; Quevedo, R.P.; Korting, T.S.; Fonseca, L.M.G.; Adami, M. (2020) "Land Cover Classification of an Area Susceptible to Landslides Using Random Forest and NDVI Time Series Data", In Proceedings of the International Geoscience and Remote Sensing Symposium; Institute of Electrical and Electronics Engineers Inc., September 26; pp.1345–1348.

[45]. Deijns, A.A.J.; Bevington, A.R.; van Zadelhoff, F.; de Jong, S.M.; Geertsema, M.; McDougall, S. (2020), "Semi-Automated Detection of Landslide Timing Using Harmonic Modelling of Satellite Imagery, Buckinghorse River, Canada", International Journal of Applied Earth Observation and Geoinformation, doi:10.1016/j.jag.2019.101943.

[46]. Yang, W.; Wang, Y.; Sun, S.; Wang, Y.; Ma, C. (2019) "Using Sentinel-2 Time Series to Detect Slope Movement before the Jinsha River Landslide", Landslides, 16, pp.1313–1324.

[47]. Mas, J.F. (1999) "Monitoring Land-Cover Changes: A Comparison of Change Detection Techniques", Int J Remote Sens, 20, pp.139–152, doi:10.1080/014311699213659.

[48]. Rosendo, J. dos S. (2005) "Índices de Vegetação e Monitoramento Do Uso Do Solo e Cobertura Vegetal Na Bacia Do Rio Araguari - MG - Utilizando Dados Do Sensor MODIS". Master's Thesis, Universidade Federal de Uberlândia: Uberlândia, MG, Brazil.

[49]. Corominas, J.; van Westen, C.; Frattini, P.; Cascini, L.; Malet, J.P.; Fotopoulou, S.; Catani, F.; Van Den Eeckhaut, M.; Mavrouli, O.; Agliardi, F. (2014) "Recommendations for the Quantitative Analysis of Landslide Risk", Bulletin of Engineering Geology and the Environment, 73, pp.209–263, doi:10.1007/s10064-013-0538-8.

[50]. Pham, B.T.; Prakash, I.; Tien Bui, D. (2018) "Spatial Prediction of Landslides Using a Hybrid Machine Learning Approach Based on Random Subspace and Classification and Regression Trees", Geomorphology, 303, pp.256–270, doi:10.1016/j.geomorph.2017.12.008.

[51]. Quinlan, J.R. (1993) *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers Inc.: San Mateo, CA, USA, 1993; ISBN 1-55860-238-0.

[52]. Ye, C. ming; Wei, R. long; Ge, Y. gang; Li, Y.; Junior, J.M.; Li, J. (2022) "GIS-Based Spatial Prediction of Landslide Using Road Factors and Random Forest for Sichuan-Tibet Highway", J Mt Sci, 19, pp.461–476, doi:10.1007/s11629-021-6848-6.

[53]. Mcculloch, W.S.; Pitts, W. (1943) "A Logical Calculus of the Ideas Immanent in Nervous Activity", Bulletin of Mathematical Biophysics, 5, pp.115–133, doi:10.1007/BF02478259.

[54]. Hong, H.; Xu, C.; Revhaug, I.; Bui, D.T. (2015) "Spatial Prediction of Landslide Hazard at the Yihuang Area (China): A Comparative Study on the Predictive Ability of Backpropagation Multi-Layer Perceptron Neural Networks and Radial Basic Function Neural Networks In Cartography-maps connecting the world". *Lecture Notes in Geoinformation and Cartography*, Sluter, C.R., Cruz, C.B.M., Menezes, P.M.L. de, Eds.; Springer, Cham; pp.175–188 ISBN 9783319177373.

[55]. Gardner, M.W.; Dorling, S.R. (1998) "Artificial Neural Networks (the Multilayer Perceptron) - a Review of Applications in the Atmospheric Sciences", Atmos Environ, 32, pp.2627–2636, doi:10.1016/S1352-2310(97)00447-0.

[56]. Hecht-Nielsen, R. (2018) "Kolmogorov's Mapping Neural Network Existence Theorem". In Proceedings of the IEEE First Annual International Conference on Neural Networks; IEEE press.: New York, NY; pp. pp.11–13.

[57]. Nair, V.; Hinton, G.E. (2010) "Rectified Linear Units Improve Restricted Boltzmann Machines", In Proceedings of the 27th ICML; Haifa, Israel; pp.807–814.

[58]. Xu, B.; Wang, N.; Chen, T.; Li, M. (2015) "Empirical Evaluation of Rectified Activations in Convolutional Network", doi:10.48550/arXiv.1505.00853.

[59]. Nola, I.T. de S. (2022) "Análise Multicritério e Aprendizado de Máquina Aplicados Na Predição Potencial Espeleológico Da Região Do Parque Nacional Serra Do, Quadrilátero Ferrífero/MG", Doctoral Thesis, Universidade Federal de Ouro Preto: Ouro Preto, MG, Brazil.

[60]. Tan, H.H.; Lim, K.H. (2019) "Vanishing Gradient Mitigation with Deep Learning Neural Network Optimization", In Proceedings of the 7th International Conference on Smart Computing & Communications; 4p..

[61]. Kingma, D.P.; Ba, J.L. (2015) "Adam: A Method for Stochastic Optimization", In Proceedings of the 3rd International Conference on Learning Representations (ICLR), San Diego, CA,15p.

[62]. Paul, S.G.; Biswas, A.A.; Saha, A.; Zulfiker, M.S.; Ritu, N.A.; Zahan, I.; Rahman, M.; Islam, M.A. (2023) "A Real-Time Application-Based Convolutional Neural Network Approach for Tomato Leaf Disease Classification", Array, doi:10.1016/j.array.2023.100313.

[63]. Nhu, V.H.; Hoang, N.D.; Nguyen, H.; Ngo, P.T.T.; Thanh Bui, T.; Hoa, P.V.; Samui, P.; Tien Bui, D. (2020) "Effectiveness Assessment of Keras Based Deep Learning with Different Robust Optimization Algorithms for Shallow Landslide Susceptibility Mapping at Tropical Area", Catena, doi:10.1016/j.catena.2020.104458.

[64]. Yi, Y.; Zhang, W.; Xu, X.; Zhang, Z.; Wu, X. (2022) "Evaluation of Neural Network Models for Landslide Susceptibility Assessment", Int J Digit Earth, 15, pp.934–953, doi:10.1080/17538947.2022.2062467.

[65]. Haykin, S. (2001) *Redes Neurais: Princípios e Prática [Translation]*, Trans. Paulo Martins Engel; Röhnelt, M., Grassi, D., Affonso, A.J., Eds.; 2nd ed.; Bookman: Porto Alegre, RS, Brazil; Vol. 1; ISBN 978-85-7307-718-6.

[66]. Prechelt, L. (1998) "Early Stopping - But When?" In *Neural Networks: Tricks of the Trade*; Orr, G.B., Müller, K.-R., Eds.; Springer: Berlin; Heidelberg; New York; Barcelona; Hong Kong; London; Milan; Paris; Singapore; Tokyo, Vol. 1524, pp.55–69 ISBN 978-3-540-65311-0.

[67]. Cohen, J. A (1960) "Coefficient of Agreement for Nominal Scales", Educ Psychol Meas, 20, pp.37–46.

[68]. Landis, J.R.; Koch, G.G. (1977) "The Measurement of Observer Agreement for Categorical Data", Biometrics, 33, pp.159–174, doi:10.2307/2529310.

[69]. Barella, C.F.; Sobreira, F.G.; Zêzere, J.L. (2019) "A Comparative Analysis of Statistical Landslide Susceptibility Mapping in the Southeast Region of Minas Gerais State, Brazil", Bulletin of Engineering Geology and the Environment, 78, pp.3205–3221, doi:10.1007/s10064-018-1341-3.

[70]. Bordoni, M.; Galanti, Y.; Bartelletti, C.; Persichillo, M.G.; Barsanti, M.; Giannecchini, R.; Avanzi, G.D.A.; Cevasco, A.; Brandolini, P.; Galve, J.P. (2020) "The Influence of the Inventory on the Determination of the Rainfall-Induced Shallow Landslides Susceptibility Using Generalized Additive Models", Catena, doi:10.1016/j.catena.2020.104630.

[71]. Zêzere, J.L.; Pereira, S.; Melo, R.; Oliveira, S.C.; Garcia, R.A.C. (2017) "Mapping Landslide Susceptibility Using Data-Driven Methods", Science of the Total Environment, 589, pp.250–267.

[72]. Guzzetti, F.; Reichenbach, P.; Ardizzone, F.; Cardinali, M.; Galli, M. (2006) "Estimating the Quality of Landslide Susceptibility Models", Geomorphology, 81, pp.166–184.

[73]. Huang, J.; Wu, X.; Ling, S.; Li, X.; Wu, Y.; Peng, L.; He, Z. (2022) "A Bibliometric and Content Analysis of Research Trends on GIS-Based Landslide Susceptibility from 2001 to 2020", Environmental Science and Pollution Research, 29, pp.86954–86993,doi:10.1007/s11356-022-23732-z.

[74]. Morellato, L.P.C.; Talora, D.C.; Takahasi, A.; Bencke, C.C.; Romera, E.C.; Zipparro, V.B. (2000) "Phenology of Atlantic Rain Forest Trees: A Comparative Study", Biotropica, 32, pp.811–823, doi:10.1111/j.1744-7429.2000.tb00620.x.

[75]. Pereira, A.B. (2009) "Mata Atlântica: Uma Abordagem Geográfica", Nucleus, 6, pp.27–52.

[76]. Badrul Hisham, N.H.; Hashim, N.; Saraf, N.M.; Talib, N. (2022) "Monitoring of Rice Growth Phases Using Multi-Temporal Sentinel-2 Satellite Image", In Proceedings of the IOP Conference Series: Earth and Environmental Science; Institute of Physics.

[77]. Graetz, R.D.; Gentle, M.R. (1982), "The Relationships between Reflectance in the Landsat Wavebands and the Composition of an Australian Semi-Arid Shrub Rangeland", Photogramm Eng Remote Sensing, 48, pp.1721–1730.

[78]. Mpakairi, K.S.; Ndaimani, H.; Kavhu, B. (2020) "Exploring the Utility of Sentinel-2 MSI Derived Spectral Indices in Mapping Burned Areas in Different Land-Cover Types", Sci Afr.

[79]. Ruscalleda-Alvarez, J.; Moro, D.; Van Dongen, R. (2021) "A Multi-Scale Assessment of Fire Scar Mapping in the Great Victoria Desert of Western Australia", Int J Wildland Fire, 30, pp.886–898.

[80]. Beguería, S. (2006) "Validation and Evaluation of Predictive Models in Hazard Assessment and Risk Management", Natural Hazards, 37, pp.315–329, doi:10.1007/s11069-005-5182-6.

[81]. Wills, C.J.; Mccrink, T.P. (2002) "Comparing Landslide Inventories: The Map Depends on the Method", Environmental & Engineering Geoscience, 8, pp.279–293, doi:10.2113/8.4.279.

[82]. Lima, P.; Steger, S.; Glade, T.; Murillo-García, F.G. (2022) "Literature Review and Bibliometric Analysis on Data-Driven Assessment of Landslide Susceptibility", J Mt Sci, 19, pp.1670–1698.

[83]. IPCC (2023) *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Core Writing Team, Lee, H., Romero, J., Eds.; Intergovernmental Panel on Climate Change: Geneva, Switzerland.

[84]. Van Den Eeckhaut, M.; Hervás, J.; Jaedicke, C.; Malet, J.P.; Montanarella, L.; Nadim, F. (2012) "Statistical Modelling of Europe-Wide Landslide Susceptibility Using Limited Landslide Inventory Data", Landslides, 9, pp.357–369, doi:10.1007/s10346-011-0299-z.

AUTHORS

Mateus Oliveira Xavier graduated in Civil Engineering from the Federal University of Ouro Preto-UFOP (2016), where he also completed his master's degree in Geotechnics in 2018. He is currently pursuing a Ph.D. at the same institution. In addition to his academic background, Mateus is currently a professor in the Department of Architecture and Urbanism at UFOP.

César Falcão Barella graduated in Environmental Engineering from the Federal University of Ouro Preto (2008), where he also completed his Master's (2011) and Ph.D. (2016) degrees in Geotechnics. He is currently a professor in the Department of Environmental Engineering at UFOP.