Remote Sensing Applications: Society and Environment 36 (2024) 101314

Contents lists available at ScienceDirect

Remote Sensing Applications: Society and Environment

journal homepage: www.elsevier.com/locate/rsase

Land use and land cover changes without invalid transitions: A case study in a landslide-affected area

Renata Pacheco Quevedo [a,b,c,*], Daniel Andrade Maciel [a], Mariane Souza Reis [a], Camilo Daleles Rennó [a], Luciano Vieira Dutra [a], Clódis de Oliveira Andrades-Filho [d,e], Andrés Velástegui-Montoya [f,g], Tingyu Zhang [h,i], Thales Sehn Körting [a], Liana Oighenstein Anderson [j]

[a] Earth Observation and Geoinformatics Division, National Institute for Space Research (INPE), São José Dos Campos, 12227010, São Paulo, Brazil
[b] Department of Structural and Geotechnical Engineering, Polytechnic School of the University of São Paulo, Avenida Prof. Luciano Gualberto, Travessa do Politécnico 380, São Paulo, 05508-010, São Paulo, Brazil
[c] Georisk and Environment Group, Geology Department, University of Liège, Liège, Belgium
[d] Department of Geodesy, Federal University of Rio Grande Do Sul (UFRGS), Porto Alegre, 91509-900, Rio Grande do Sul, Brazil
[e] Postgraduate Program in Remote Sensing, Federal University of Rio Grande Do Sul (UFRGS), Porto Alegre, 91509-900, Rio Grande do Sul, Brazil
[f] Centro de Investigación y Proyectos Aplicados a las Ciencias de la Tierra (CIPAT), ESPOL Polytechnic University, P.O. Box 09-01-5863, Guayaquil, Ecuador
[g] Faculty of Engineering in Earth Sciences FICT, ESPOL Polytechnic University, P.O. Box 09-01-5863, Guayaquil, Ecuador
[h] Key Laboratory of Degraded and Unused Land Consolidation Engineering, the Ministry of Natural Resources, Xi'an, Shaanxi, China
[i] Institute of Land Engineering and Technology, Shaanxi Provincial Land Engineering Construction Group Co., Ltd, Xi'an, Shaanxi, China
[j] National Center for Monitoring and Early Warning of Natural Disaster, São José dos Campos, 12247-016, Brazil

ARTICLE INFO

Keywords:
Mass movement
Land cover transition
Time series analysis
Remote sensing
Landsat
Classification

ABSTRACT

Land use and land cover (LULC) analysis provides valuable information to understand environmental changes and their effects on landslide occurrence. However, LULC time series can be affected by errors in classifications that lead to invalid transitions and, therefore, to misinterpretations. One solution is to include temporal approaches that reduce the effects of invalid transitions. Here, we aimed to evaluate how such methods can improve the LULC analysis for a landslide-affected area. For that, we integrated the Random Forest (RF) class likelihoods with the temporal approach provided by the Compound Maximum a Posteriori (CMAP) algorithm, named here as RF-CMAP. Results from RF-CMAP were compared to those obtained from the traditional RF in a post-classification comparison approach. Although both methods presented high performance, with overall accuracy (OA) values greater than 0.87, RF-CMAP reached higher OA than RF for all the analysed years and corrected 99.92 km^2 (12% of the total area) of invalid transitions presented by the traditional RF. Furthermore, RF-CMAP was capable of correctly classifying more areas than RF in landslides (e.g., 66% and 21% for RF-CMAP and RF in 2000, respectively). Finally, this study contributes to exploring the integration between RF and CMAP algorithms to avoid invalid transitions and to assess how the existence of LULC invalid transitions can impact subsequent analyses.

* Corresponding author. Earth Observation and Geoinformatics Division, National Institute for Space Research (INPE), São José Dos Campos, 12227010, São Paulo, Brazil.
E-mail address: renatapquevedo@gmail.com (R.P. Quevedo).

https://doi.org/10.1016/j.rsase.2024.101314
Received 24 January 2024; Received in revised form 11 July 2024; Accepted 29 July 2024
Available online 3 August 2024

1. Introduction

Land use and land cover (LULC), climate changes, and disasters related to natural hazards interact in complex ways (Jia et al., 2022). It is estimated that almost a third of the global land area was significantly affected by LULC changes (LUCC) (Winkler et al., 2021). LUCC present cumulative effects (Leemans and Zuidema, 1995), which can impact not only the global climate (Lambin et al., 2001) but also surface runoff (Garg et al., 2019) and ecosystem services (Sharma et al., 2019). As such, LUCC act as a causal and consequential factor of climate changes (Dale, 1997) and landslide disasters (Quevedo et al., 2023b).

Changes in LULC may modify the relationship between triggering factors and landslide occurrence, affecting the evolution of slopes (in)stability (Wasowski, 1998). In addition, LUCC can have effects on slope equilibrium conditions (Tarantino et al., 2007). The changes alone cannot predict landslides but can increase the probability of re-activation or first landslide activation (Tarantino et al., 2007). In the Flemish Ardennes, Belgium, an analysis carried out by Van den Eeckhaut et al. (2010) using LUCC and other conditioning factors, pointed out that human activities can increase slope instability and, even, initiate landslides. Other recent studies have been pointing that some kinds of LUCC can improve landslide susceptibility, such as the conversion of forests to grass and arable lands (Chen et al., 2019), the abandonment of cultivated lands (Persichillo et al., 2017), and human engineering activities (Liu et al., 2021; Vuillez et al., 2018; Zeng et al., 2023).

For example, in the *Rolante* River Basin (RRB), Brazil, a torrential rainfall occurred on 5 January 2017 triggering more than 300 landslides (Quevedo et al., 2019, 2024) with a consequent river dam and flash flood (SEMA, 2017). These landslides occurred shortly after some LUCC occur, such as forestry harvesting, which could act as a possible catalyst for landslides. Mapping these changes is the first step to understanding their relationship with landslide events (Uehara et al., 2020). However, in many landslide studies, the LULC maps used for susceptibility analysis are often sourced from pre-existing data, such as available maps from various institutions, which may not necessarily adhere to the same scale or methodology (Persichillo et al., 2017). Thus, it is necessary to obtain a consistent LULC mapping to avoid subsequent errors related to non-existent LULC or LUCC classes (Quevedo et al., 2023a).

The development of remote sensing techniques provided tools to analyse large areas and apply multi-temporal approaches through classification methods. This kind of approach can be used for LUCC assessment by evaluating the succession of LULC classes in a pixel over three or more different times of interest, which is called land cover trajectory (Mertens and Lambin, 2000). One widely used method to detect land cover trajectories consists of classifying different satellite images and comparing the classifications, pixel by pixel, known as post-classification comparison (Lu et al., 2004; Zhu, 2017).

The post-classification comparison presents a complete matrix of change information; however, the land cover trajectories' quality depends on the accuracy of each individual classification (Lu et al., 2004; Tewkesbury et al., 2015). Classifying individual LULC maps for posterior comparison may result in some invalid transitions, i.e., land cover changes that are not likely to occur in the analysed area in a given period (Liu et al., 2008; Reis, 2022). The presence of invalid transitions can conduct to a misinterpretation, e.g., relating landslide occurrence with deforestation when it is actually some erroneous LUCC classification (Quevedo et al., 2023a). Therefore, it is required post-processing steps for identification and correction (Traoré et al., 2013), which generally includes expert-based-transition rules (Hermosilla et al., 2018) and manual correction (Olofsson et al., 2013), being an expensive and lengthy process.

To overcome this issue, some temporal approaches can be used, such as the Compound Maximum *a Posteriori* (CMAP) (Reis et al., 2020) that was created to classify the whole LULC trajectory at once using the multi-temporal information available in time series. The advantages of using CMAP for LUCC analysis rely on incorporating land cover dynamics into the classification process, without needing a post-classification step. The CMAP algorithm general formulation requires the estimation of the class conditional feature distribution (or likelihood), which in Reis et al. (2020) was considered to follow a Gaussian distribution. This classification method requires statistical pre-assumptions on the data.

However, LULC mapping of landslide-affected areas are usually influenced by the terrain characteristics, being necessary to include morphometric factors (e.g., slope, aspect, terrain curvatures). For that, the Random Forest (RF) algorithm offers the advantages of being a non-parametric approach, capable of dealing with nonlinear data, computationally fast and has been showing improved accuracy in LULC classifications (Piao et al., 2021). In addition, Random Forest allow to obtain uncertainty metrics (i.e., probabilities) of a pixel to be assign to a specific class, which is useful to be combined with CMAP, with the advantage of being non-parametric. Aiming to integrate the temporal approach of CMAP with the non-parametric advantages of the RF, here, we substituted the class likelihood at each point of time with features derived from the RF algorithm, being herein referred to as RF-CMAP. This kind of classification can be more trustful for subsequent analyses, for instance, landslide susceptibility assessment.

In this sense, this study aims to evaluate how the integrated use of RF-CMAP can improve the LULC trajectory classification in landslide occurrence areas. To achieve this objective, we focused on the analysis of invalid transitions/trajectories obtained by the post-classification comparison of relatively high-accuracy LULC products from the RF in a real landslide scenario versus RF-CMAP results. Therefore, the contributions of this study are: i) to propose and evaluate the RF adaptation (RF-CMAP) necessary to include relevant features for the classification of landslide-affected areas (e.g., morphometric factors) and temporal approach; ii) to evaluate how RF classification can be improved with CMAP approach to deal with invalid LUCC in time-series analysis; iii) to estimate the occurrence of invalid transition in the landslide-affected areas considering high-accuracy (>90%) LULC products; iv) to analyse invalid transitions in the LULC maps generated for the RRB and their possible impacts in landslide analysis; v) to estimate the occurrence of invalid transition in the landslide-affected areas. The improved accuracy of the RF-CMAP can help into evaluation of long-term time-series, improving classification accuracy over time and helping into landslide mapping and susceptibility evaluation.

2. Study area

The RRB is a sub-basin of the *Sinos* River catchment, located in the *Rio Grande do Sul* state, Brazil (Fig. 1). The area covers 828 km^2 and comprehends the municipalities of *Riozinho*, *Rolante*, and *São Francisco de Paula*. The study area is composed of volcanic rocks from the Serra Geral Formation and sandstones from the *Botucatu* Formation (Uehara et al., 2022). Between the northern and southern portions, there is the Meridional Plateau Escarpment, composed of V-shaped valleys, with altitudes varying between 16 m and 988 m and slopes reaching more than 50°.

The climate is classified as a humid subtropical ocean climate without a dry season, with temperate summer (Cfb) in the northern RRB, the highest areas, and with hot summer (Cfa) in the southern RRB, the lowest areas, according to the *Köppen*'s climate classification (Alvares et al., 2013). Considering a more detailed climate classification (Rossato, 2020), the RRB would be classified as subtropical very humid, with cold winter and temperate summer, with well-distributed rainfall throughout the year, and annual averages between 1700 mm and 2000 mm.

The predominant biome is the Atlantic Forest, which includes Semideciduous Seasonal Forests and Mixed Ombrophilous Forests (forests with *Araucaria angustifolia (Bertol.) Kuntze*) (IBGE - Instituto Brasileiro de Geografia e Estatística, 2019). Moreover, there is a Federal Protected Area in the RRB, which covers mainly Mixed Ombrophilous Forest, but also wetlands, native grassland, and forest plantation (ICMBio - Instituto Chico Mendes de Conservação da Biodiversidade, 2020). This forest plantation corresponds to ecologically managed tree monocultures, composed mainly of *Pinus* sp. and *Eucalyptus* sp. (Fonseca et al., 2009).

The RRB has rural characteristics with activities related to pasture, forestry, agriculture, and native grasslands in the northern portion (Luerce, 2015). The pasture and agriculture are concentrated in the northern portion of the RRB, in the geomorphological unit of *Campos Gerais* Plateau, and in the southern portion, where the relief is more dissected. In the transition area, the LULC is mainly composed of forest and silviculture due to the high slope degrees, which is a limitation for many activities. Silviculture is the LULC class that has been showing the greatest growth in the study area (Henderson et al., 2016; Lang, 2013).

Lastly, due to its environmental characteristics, the RRB has areas with high landslide susceptibility. In fact, in January 2017, an extreme rainfall event resulted in more than 300 landslides (Quevedo et al., 2019, 2020, 2024). It was estimated that the precipitation reached values between 90 mm and 272 mm in 4 h (SEMA, 2017). The sediments displaced from the slopes generated a dam on the river with subsequent flash floods, which reached *Rolante* City (SEMA, 2017). These landslides were recorded in the steepest areas (RIFFEL et al., 2021), with a greater presence of forested areas and silviculture. Furthermore, many of the affected areas recorded forestry harvesting on hilltops shortly before the extreme precipitation event and subsequent landslides.

The RRB is a representative region Atlantic Forest in *Rio Grande do Sul* and can act as a one important local for evaluation of the proposed model. Recently, May 2024, the *Rio Grande do Sul* state suffered the biggest disaster in its history, with floods and landslides. The RRB is located in the affected area, mainly by landslides, where more than 6000 landslides were mapped until 24 June 2024 (Andrades-Filho et al., 2024; Possantti et al., 2024).

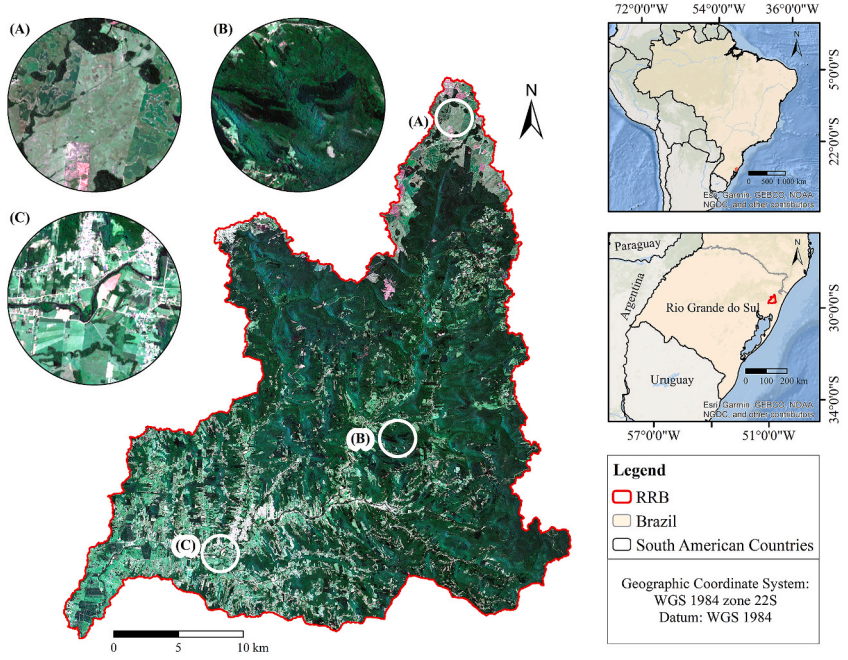

Fig. 1. Study area and its different Land Use and Land Cover characteristics: A) Northern portion: *Campos Gerais* Plateau, a flat area with natural grassland, pasture, and agriculture; B) Central portion: Meridional Plateau Escarpment, a mountainous relief with remnant Atlantic Forest natural forest and silviculture; C) Southern portion: near the *Rolante* river mouth, an area with a dissected relief, concentrating urban areas, agriculture, and pasture.

3. Methods

The methodology carried out in this study comprises four main steps (Fig. 2): (i) data acquisition; (ii) used algorithms for LULC classification; (iii) model evaluation metrics; iv) analysis of invalid transitions in landslides.

3.1. Data acquisition

3.1.1. Land use and land cover classes

The RRB comprehends seven LULC classes: i) silviculture; ii) forest; iii) water; iv) bare soil; v) construction; vi) agriculture; and vii) grassland. Silviculture represents forest plantation, composed of *Pinus* sp. and *Eucalyptus* sp. Forest comprises natural forest formation, both primary and secondary. Water comprehends rivers, streams, and little agriculture reservoirs. Bare soil represents areas of exposed soil, whether from deforestation, forestry harvesting, or crop rotation. Construction mainly covers the cities of *Riozinho*, *Rolante*, and *São Francisco de Paula*, in addition to sheds and silos. Agriculture is related to crops, with an emphasis on wheat, soybeans, and vegetables. Grassland comprises natural meadows and small pastures. Finally, we included a shadow class representing areas with relief shadows due to the mountainous characteristics.

3.1.2. Image and variable selection

Since the landslides occurred on 5 January 2017 (Quevedo et al., 2019, 2020), we considered three previous years for LULC trajectory analysis: i) 2000; ii) 2008; iii) 2016. The year 2016 was selected for being the year before the landslide occurrence, and 2000 and 2008 were selected to identify one or two previous silviculture cycles in the RRB. Capturing forestry cycles in the RRB is important because silviculture is one of the frequent LULC classes on hilltops (Fig. 3). According to the chosen period, we selected Landsat-family imagery Collection 2 Level 2, atmospherically corrected surface reflectance (Wulder et al., 2019), made available by the United States Geological Survey and downloaded from the Microsoft Planetary Computer platform (Microsoft Open Source et al., 2022).

For the year 2000, we used Landsat 7 satellite Enhanced Thematic Mapper Plus (ETM+) sensor, and made a composition based on the median of four images: i) 22 December 1999, ii) 24 February 2000, and iii) two scenes of 20 March 2000. This composition aimed to minimise the presence of clouds and relief shadow effects caused by the mountainous relief pattern. The year 2008 was analysed using an image from the Landsat 5 satellite, Thematic Mapper (TM) sensor, sensed on 6 February 2008. For 2016, we selected two images from the Landsat 8 satellite, Operational Land Imager (OLI) sensor. We selected the image from 20 January 2016 that covers approximately 90% of the RRB. For the area not covered by this image, we used an image from 15 March 2016. The image selection prioritized summer images, because, due to the rugged terrain, the lighting in winter changes drastically, increasing the number of shaded areas.

A total of 18 variables were used in the classification, including Landsat surface reflectance (LSR) bands, spectral indices, and morphometric factors (Uehara et al., 2022). The LSR bands included: blue, green, red, near-infrared (NIR), short-wave infrared 1 (SWIR1), and short-wave infrared 2 (SWIR2). In addition, we selected nine spectral indices (Phan et al., 2020): Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Soil Adjusted Vegetation Index (SAVI), Modified Soil-Adjusted Vegetation Index-2 (MSAVI2), Normalized Difference Built-Up Index (NDBI), Normalized Difference Water Index (NDWI), Modified Normalized Difference Water Index (mNDWI), Normalized Difference Moisture Index (NDMI), and Simple Ratio (SR). Although the indices might co-vary (e.g., NDVI and EVI), they are expected to produce slightly different results due to their distinct formulation. For instance, EVI is less susceptible to vegetation saturation when increasing leaf area index, although EVI is known to be susceptible to bidirectional effects (Petri and Galvão, 2019). On the other hand, while mNDWI often outperforms NDWI in detecting water bodies (Xu, 2006), NDWI can still yield reasonable results in certain water types, as well being less sensitive to the low signal-to-noise ratio of TM data. Finally, considering the varied relief along the RRB, we also included three morphometric factors derived from the NASADEM (NASA JPL, 2020): elevation, slope, and Terrain Ruggedness Index (TRI). The variables are detailed in Table 1.

3.1.3. Sampling process

An active learning process (Su et al., 2020) (Fig. 4) was used to select a representative sample set, improving the classification performance by iteratively adding or removing new samples in the input set. In other words, our approach involved selecting an initial

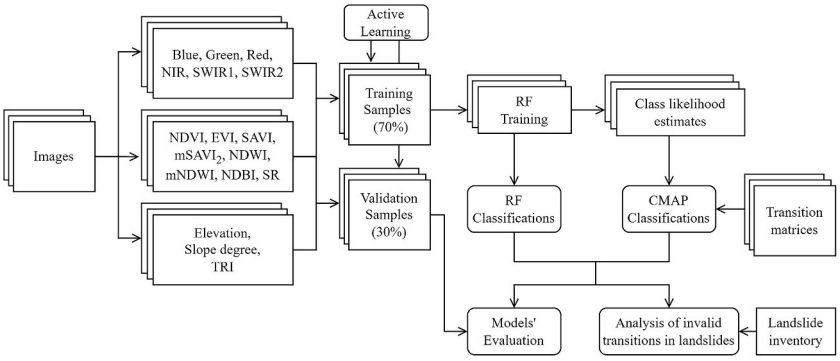

Fig. 2. Methodological flowchart.

Fig. 3. Land cover changes in hilltop areas. A) presence of silviculture on a hilltop in 2013. B) forestry harvesting and the presence of landslides in 2017.

Table 1
Spectral and morphometric indices used as variables for LULC classifications.

Index Name	Acronym	Equation	Meaning	Reference
Normalized Difference Vegetation Index	NDVI	(NIR – red)/(NIR + red)	It quantifies vegetation's greenness, using normalization to reduce errors.	Rouse Jr. et al. (1973)
Enhanced Vegetation Index	EVI	2.5*((NIR – red)/(NIR + 6 * red - 7.5 * blue + 1))	It minimizes the atmospheric influences existent in NDVI	Kaufman and Tanre (1992)
Soil-Adjusted Vegetation Index	SAVI	(NIR – red)/(NIR + red + 0.5) * (1.5)	It reduces the soil influence on canopy spectra.	Huete (1988)
Modified Soil-Adjusted Vegetation Index-2	MSAVI2	(2 * NIR + 1 – SQRT((2 * NIR + 1)^2 - 8 * (NIR – red)))/2	It increases vegetation sensitivity by minimizing the soil background influences.	Qi et al. (1994)
Normalize Difference Build-up Index	NDBI	(SWIR1 – NIR)/(SWIR1 + NIR)	It was developed to automatically map build-up areas.	Zha et al. (2003)
Normalized Difference Water Index	NDWI	(green – NIR)/(green + NIR)	It delineates open water, emphasizing the difference between water bodies and other objects.	McFeeters (1996)
Modified Normalized Water Index	MNDWI	(green – SWIR1)/(green + SWIR1)	It reduces the built-up land noise present in the NDWI when extracting bodies of water.	Xu (2006)
Normalized Difference Moisture Index	NDMI	(NIR – SWIR1)/(NIR + SWIR1)	Also called NDWI, it emphasizes the water content in the vegetation canopies.	Gao (1996)
Simple Ratio	SR	NIR/red	It was developed to measure the leaf-area index, which emphasizes vegetative vigour.	Jordan (1969)
Elevation	–	–	It is represented by the Digital Elevation Model, in which each pixel value corresponds to the altitude	Miller and Laflamme (1958)
Slope	–	Slope = arctan(Rise/Run)	It represents the slope angle or the angle of inclination, which considers the elevation (rise) and distance (run) differences between two points.	Miller and Laflamme (1958)
Terrain Ruggedness Index	TRI	$\left[\frac{1}{N^2-1}\sum_{k=1}^{N^2}(z_o - z_k)^2\right]^{1/2}$	It quantitatively estimates the terrain heterogeneity.	Riley et al. (1999)

NIR: near-infrared; SWIR1: short-wave infrared 1; N is the moving window size (we used $N = 3$); z_o is the elevation of the central cell and z refers to the elevation of each neighbour cell (Habib, 2021).

sample set, running an RF model, and assessing both model performance and classification uncertainty. After that, new samples were selected in high entropy areas and added to the input set to improve model performance and decrease classification uncertainty. If the entropy of this region increases, we return to the previous sample set. The model was considered reliable when changes in the sample set showed no significant improvements in accuracy or visual classification, stopping the active learning process.

After applying the active learning process for each year, we obtained sample datasets composed of 1809 samples for 2000, 1874 for 2008, and 2003 for 2016 (Fig. 5). Samples were collected to provide a spatial representativeness of the study area, as well as to select only pure pixels (i.e., avoiding spectral mixture effects). In addition, according to Ebrahimy et al. (2021), Random Forest is robust against medium sample sizes, such as we collected in this study. Our reference data were based on expert visual interpretation, who knows the study area and has experience in field surveys in the region. Each sample set was randomly divided into training (70%) and validation (30%) sample sets, maintaining this proportion for each LULC class (Table 2).

In addition, we created a larger validation dataset to guarantee a more suitable model performance evaluation, uniting the three validation samples and extracting LULC classes referring to each of the analysed years. This dataset was, therefore, composed of 1706 validation samples, which contained the three-year information. Thus, we were able to evaluate two transition periods: i) 2000 to 2008, and ii) 2008 to 2016.

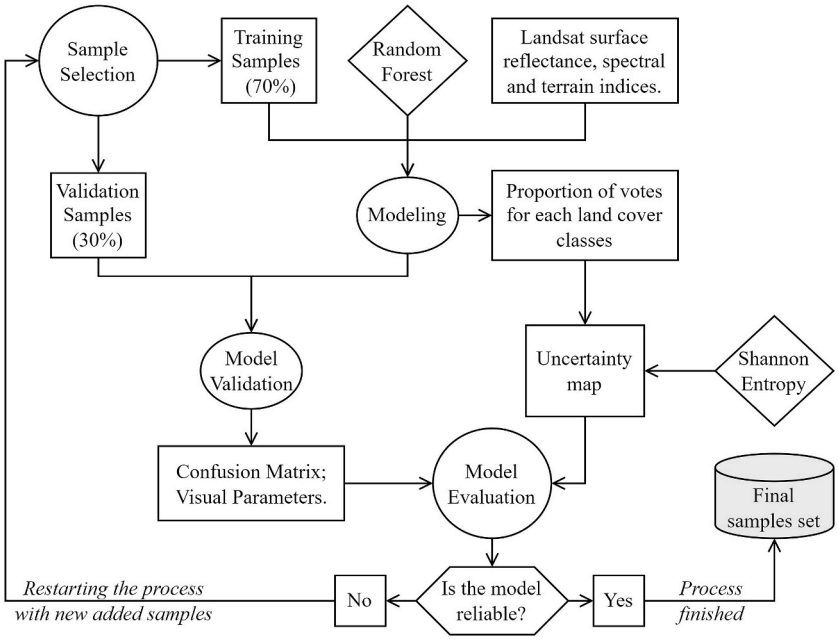

Fig. 4. Sampling process using Active Learning technique.

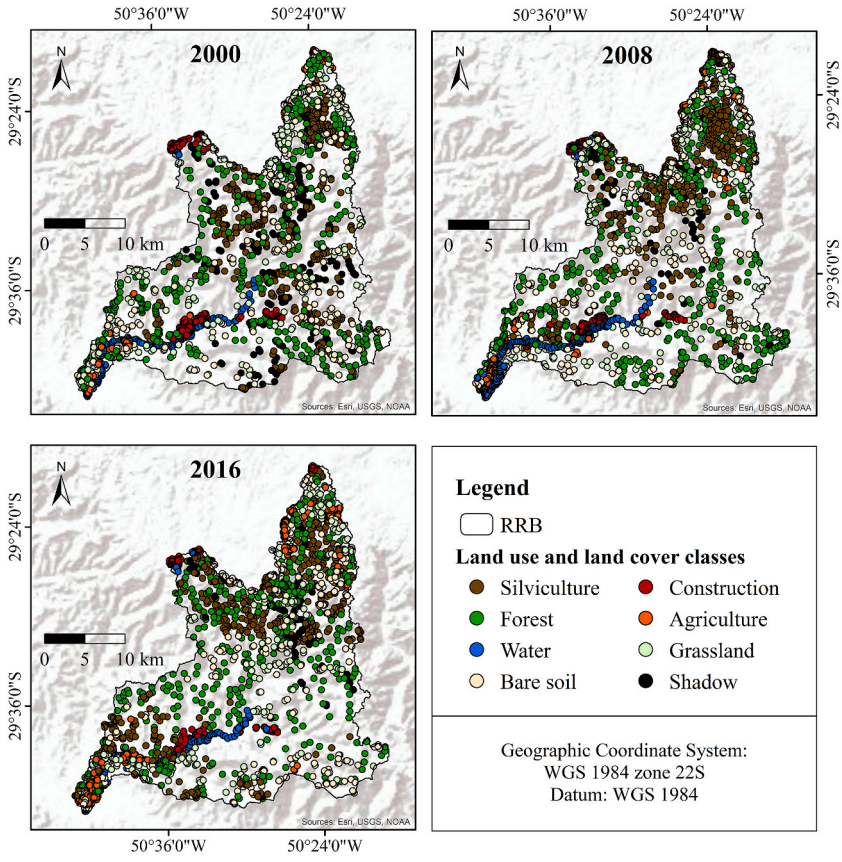

Fig. 5. LULC sample set for each year.

Table 2
Training and validation samples used for each year.

	2000		2008		2016		Total validation sample set
	Training	Validation	Training	Validation	Training	Validation	
Silviculture	268	115	246	105	315	135	355
Forest	262	112	269	115	277	119	346
Water	96	41	102	44	111	48	133
Bare soil	188	80	178	77	220	95	252
Construction	113	48	115	50	105	45	143
Agriculture	69	30	160	69	147	63	162
Grassland	130	55	178	76	158	67	198
Shadow	141	61	63	27	69	29	117
Total	1267	542	1311	563	1402	601	1706

3.2. Algorithms for LULC classification

3.2.1. Random Forest algorithm

The RF algorithm is an ensemble of Decision Tree models, which may be used for classification or regression (Breiman, 2001). Some advantages of RF consist of being a non-parametric supervised classifier, and potentially more suitable to non-Gaussian variables, allowing the use of categorical attributes, and ranking variables according to their importance for the modelling process (Rodriguez-Galiano et al., 2012). This algorithm uses bagging for randomly selecting training sample subsets to grow each tree, avoiding overfitting, and decreasing the correlation among Decision Tree. These features have made the RF algorithm popular in remote sensing image classification (Belgiu and Drăgu, 2016).

The bagging method allows each tree to grow with a different sample set, which is randomly drawn (with replacement). This reduces the estimator variance, providing more stability in the RF outputs (Breiman, 1996). The RF outputs are based on the most frequently voted class for classification and the average result for the regression method (Breiman, 2001; Lee et al., 2020). Using the classification approach, it is possible to obtain the proportion of votes for each class from the generated trees for each pixel (Fig. 6).

A set of samples is left out of the bagging, being called Out-of-Bag (OOB), which is used for model performance evaluation. OOB concentrates approximately one-third of the original training set and can be used for estimating the classification error rate. Unlike cross-validation, OOB produces unbiased estimates (Breiman, 2001) and is also used to estimate variable importance through the Mean Decrease in Accuracy (MDA). The MDA considers the difference between the OOB error using the original training subset and a sample set with randomly permuted variable values (Boonprong et al., 2018; Calle and Urrea, 2011). The greater the changes in accuracy, the more important the variable that has been permuted. The variable is considered insignificant if there are no changes in the OOB error (Georganos et al., 2018).

In this study, we ran the RF algorithm using the *randomForest* package (version 4.7–11) (Liaw and Wiener, 2002) in R (version 4.0.2) (R Core Team, 2014). For running the RF, the user needs to define the number of trees to be generated and the number of randomly selected variables used during the splitting process for each node. The default number of variables is equivalent to the square root of the total used variables. However, it is possible to tune RF parameters using the *Caret* package (version 6.0–94) (Kuhn, 2012) in R to check with how many trees the model stabilizes and how many variables per split the OOB is reduced. Therefore, after tests using the *Caret* package, we defined 1500 trees (*ntree* = 1500) and eight variables out of 18 variables for the splits (*mtry* = 8).

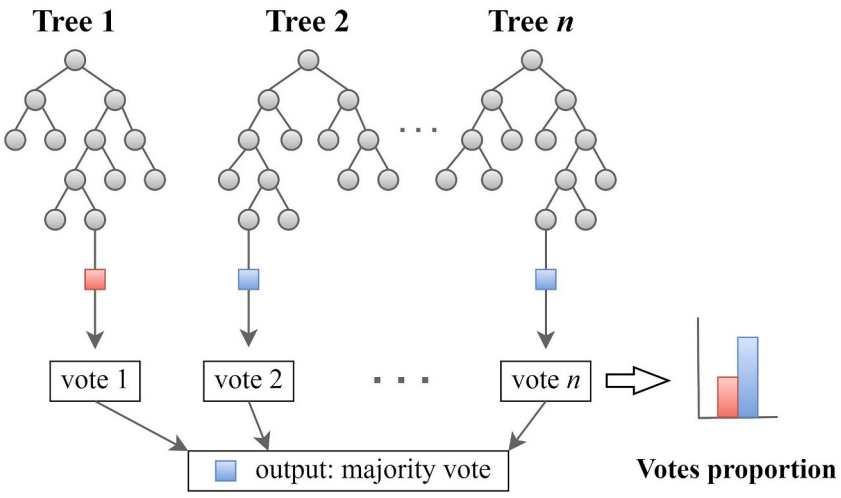

Fig. 6. RF classification: each tree votes in one class and the most voted class represents the RF output. Adapted from Belgiu and Drăgu (2016).

3.2.2. Compound Maximum a Posteriori

The CMAP algorithm was developed by Reis et al. (2020) as a supervised classifier used to estimate the trajectory $s = \left\{w_1^{k_1}, \ldots, w_t^{k_t}, \ldots, w_T^{k_T}\right\}$ in a singular geographic position and T times of interest. Here, s is the trajectory that contains the actual class $w_t^{k_t}$ at each time $t \in T$ and k_t is the indicator of the class in the set W_t that holds all K_t possible LULC classes for each time t. An organized temporal sequence of T observations for a given geographic position (e.g., correspondent pixels from an image time series) is represented by $\vec{X} = \left\{\vec{x}_1, \ldots, \vec{x}_t, \ldots, \vec{x}_T\right\}$. CMAP estimates the trajectory using Equation (01).

$$\hat{s} = arg_s \max \left(P\left(\vec{X}\middle|s\right) \times P(s), s \in \boldsymbol{\Omega}\right) \tag{1}$$

in which Ω is the set that holds all possible trajectories, derived from the Cartesian product of all sets of classes for each time, and $P(s)$ is the *a priori* probability of the trajectory s.

Reis et al. (2020) proposed different ways to calculate $P(s)$. Here, we considered the so-called "discriminative" method, which is derived from the basic assumption of a Markov chain that the *a priori* probability of a class occurring at a given time t depends only on the class at $t-1$, or, for cases in which there is stronger evidence of the classes in more recent times, on the class at $t+1$. For this last case, we have that:

$$P(s) = P\left(w_1^{k_1} \cap \ldots \cap w_t^{k_t} \cap \ldots \cap w_T^{k_T}\right) \tag{2}$$

and

$$P(s) = P\left(w_T^{k_T}\right) \times P\left(w_{T-1}^{k_{T-1}}\middle|w_T^{k_T}\right) \times \ldots \times P\left(w_1^{k_1}\middle|w_2^{k_2}\right) \tag{3}$$

In this case, $P\left(w_t^{k_t}\middle|w_{t+1}^{k_{t+1}}\right)$ is known and can be represented by matrices that tabulate the probability of each transition for times t and $t+1$. In these matrices:

$$\sum_{k_t=1}^{K_t} P\left(w_t^{k_t}\middle|w_{t+1}^{k_{t+1}}\right) = 1 \tag{4}$$

One interesting characteristic of this approach is that the user can set different *a priori* probabilities for each transition/trajectory. As such, it is possible to avoid very rare or invalid trajectories (i.e., those that could never happen in the field) by assigning them very low values of $P(s)$. In our case study, any trajectory that contains at least one transition from Construction to Water, for example, will receive a $P\left(\vec{X}\middle|s\right) \times P(s) = 0$, which is not expected to reach the maximum value among other possible trajectories, and, therefore, not classified.

Based on fieldwork analysis and expert experiences in the study area, we considered that this period is not enough to grow a forest in areas with bare soil or silviculture. On the other hand, areas with agriculture or grassland might become secondary forests due to the study area's climatic conditions. Moreover, both the constructed areas, as well as water bodies, are unlikely to be substituted for another LULC. Therefore, the LUCC were defined as valid or invalid (Supplementary Material 1) and then, the probability (1) was divided by the number of valid transition (Supplementary Material 2). In this sense, our transition matrix (Table 3) accounts for two subsequent analysed times, i.e., a transition period of eight years, showing the respective LUCC probabilities.

3.2.3. Integrating Random Forest within Compound Maximum a Posteriori

In their original study, Reis et al. (2020) calculated $P\left(\vec{x}_t\middle|w_t^{k_t}\right)$ considering the Probability Density Function of a Gaussian distribution. As an alternative to using the Gaussian hypothesis, we propose to calculate $P\left(\vec{x}_t\middle|w_t^{k_t}\right)$ based on the votes of an RF model. In this study, no statistical (parametric) model is considered for $\vec{x}_t$.

Table 3
Backward transition matrix: probabilities for each transition.

t	t+1							
	Sc	F	W	BS	C	A	G	Sd
Silviculture (Sc)	0.17	0.00	0.00	0.17	0.00	0.00	0.17	0.13
Forest (F)	0.17	0.25	0.00	0.17	0.17	0.20	0.17	0.13
Water (W)	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.13
Bare Soil (BS)	0.17	0.00	0.00	0.17	0.17	0.20	0.17	0.13
Construction (C)	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.13
Agriculture (A)	0.17	0.25	0.00	0.17	0.17	0.20	0.17	0.13
Grassland (G)	0.17	0.25	0.00	0.17	0.17	0.20	0.17	0.13
Shadow (Sd)	0.17	0.25	0.50	0.17	0.17	0.20	0.17	0.13

As explained in Section 3.2.1, RF models are based on assigning the most voted class to an object. Under relaxed assumptions, this can be interpreted as the estimation of $P\left(w_t^{k_t}\middle|\vec{x}_t\right)$: in any Decision Tree-based model, class probabilities are tied to a set of pixels (the set determined by the feature sub-space spanned by the Decision Tree leaf), not to any pixel individually, which, originally, is the necessary element to calculate CMAP. In this study, we try to use the CMAP structure assuming $\vec{x}_t$ not as pixel, but, implicitly, as an indicator of the feature sub-space to which $\vec{x}_t$ belongs to. Following the Bayes theorem, we have that:

$$P\left(\vec{x}_t\middle|w_t^{k_t}\right)=\frac{P\left(w_t^{k_t}\middle|\vec{x}_t\right)\times P\left(\vec{x}_t\right)}{P\left(w_t^{k_t}\right)} \tag{5}$$

where $P\left(\vec{x}_t\right)$ is constant and does not interfere with the maximum calculation. For cases in which we can suppose that $P\left(w_t^{k_t}\right)$ is also equiprobable, we may be able to estimate $arg_w \max\left(P\left(\vec{x}_t\middle|w_t^{k_t}\right)\right)$ as the $arg_w \max\left(P\left(w_t^{k_t}\middle|\vec{x}_t\right)\right)$, which is enough for the maximum calculation necessary in CMAP. This approach will be referred to as RF-CMAP in this study.

3.3. Model evaluation

The classifications were evaluated using the confusion matrix and derived metrics, i.e., the overall accuracy (OA), the user's accuracy (UA), and the producer's accuracy (PA) (Congalton, 1991), in addition to F1 Score, calculated as follows:

$$\mathrm{OA}=\sum_{i=1}^{q}p_{ii}, \tag{6}$$

$$PA_i=\frac{p_{jj}}{p_{+j}}, \tag{7}$$

$$UA_i=\frac{p_{ii}}{p_{i+}} \tag{8}$$

$$F1\ Score_i=\frac{2*PA_i*UA_i}{PA_i+UA_i} \tag{9}$$

where q represents the number of LULC classes, p_{jj} is the proportion of pixels correctly classified as class j, p_{+j} represents the total of reference pixels of class j, p_{ii} is the proportion of pixels correctly classified as class i, and p_{i+} represents the total of pixels classified as class i (Olofsson et al., 2014).

Some advantages of using a confusion matrix rely on its quantitative nature and that it provides classification accuracy and misclassification between classes (Foody, 2002). The OA represents the proportion of correctly classified pixels. The PA denotes the classifier sensitivity, being associated with omission error. The UA represents the reliability of each classified class and is associated with commission error. Furthermore, we analysed the area of each LULC class, the disagreement in pixel classification, and the presence of invalid transitions between RF and RF-CMAP.

In addition to the evaluation by year, we compared model's performance considering both areas with and without changes between 2000 and 2008, and 2008 and 2016. The areas with changes were also analysed according to the RF' errors in classification. In this part, we aimed to evaluated how many of the samples wrongly classified by the RF were corrected by the RF-CMAP. Also, we verified how many errors presented invalid transitions and which of them were well corrected by RF-CMAP.

Furthermore, an additional validation was carried out focused on the areas where the RF and RF-CMAP classifications were discordant. In this sense, another sample set was acquired considering the discordant pixels. For that, we first calculated the discordant area, i.e., areas differently classified for each model in each year. Then, we randomly selected 1% of the samples for each discordant class (e.g., discordant class 1: RF classified the pixel as class w_t^1 and RF-CMAP as class w_t^2) and included these samples in the second validation. Finally, we evaluated each sample by visually interpreting Landsat images and assigning them to the appropriate LULC class.

3.4. Estimation of invalid transition in landslides

Since the presence of invalid transitions can lead to misunderstandings regarding the relationship between LUCC and landslide occurrence (Quevedo et al., 2023a), we focused on identifying invalid transitions within landslide areas. For that, we utilised the classification results from RF and the landslide inventory provided by Quevedo et al. (2024).

Our approach involved counting how many landslides covered at least one pixel with an invalid transition. Then, we calculated the percentage of landslide polygons with invalid transitions during two periods: i) between 2000 and 2008, and ii) between 2008 and 2016. Additionally, we identified the most common types of invalid transitions found within the landslides and how this can impact susceptibility analysis.

4. Results

4.1. LULC classification

In all the six classifications (RF and RF-CMAP, 2000, 2008, and 2016), the predominant classes were forest, covering more than 430 km^2 (52% of the RRB), and grassland with almost 20% of the total area (>150 km^2) (Supplementary Material 3). Two classes presented different trends between both models: silviculture and water (Supplementary Material 4). Silviculture decreased between 2000 and 2008 and increased in 2016 in the RF classifications; however, the RF-CMAP showed that silviculture tended to increase over the years. This different trend is related to approximately 30 km^2 classified as silviculture by RF in 2000 and corrected by RF-CMAP. In a similar pattern, water class was overestimated for 2008 using RF (+4.4 km^2), which made it looks like there was a rising trend in 2008 and a decrease in 2016.

We analysed how much RF-CMAP increased (or decreased) the classified area of a given LULC class each year in comparison with RF (Fig. 7). In 2000, RF-CMAP classified silviculture with an area 46% (32 km^2) smaller compared to RF; water and construction were also expressive in percentual (39% and 38% smaller), but in the total area represented a decrease in approximately 3 km^2; bare soil was 13% (10 km^2) smaller using RF-CMAP in 2000. On the other hand, there was an increase in the forest (33 km^2) and shadow (5 km^2). Other differences in classification relied on areas of water and construction in 2008, and construction and shadow in 2016.

The main differences were observed in i) areas in which RF generated invalid transitions and ii) in the reduction of salt and pepper effects when using RF-CMAP (Fig. 8). There was an overestimation of water in 2000 when using RF. In 2008, there were differences in water and a salt and pepper effect in the bare soil area near the river, which was improved using RF-CMAP. Finally, RF-CMAP maintained a more homogeneous construction area when compared to RF in 2016.

4.2. Model performance

In general, the classification results presented greater values for RF-CMAP algorithm than for RF, presenting OA values higher than 0.87 for all analysed years (Table 4). Similar OA values were expected for models of the same year due to the high accuracy of RF models. Although CMAP, by design, tends to improve only the classification of classes constrained by LULC dynamics, this improvement reflected in improved OA values for the entire study area in each year. For instance, for almost all the LULC classes, the RF-CMAP presented higher PA than RF, i.e., RF-CMAP classified more correctly the study area according to the reference samples (Supplementary Material 5). Complementary, F1 Score showed highest values in RF-CMAP classification, highlighting 2016, when all classes reached greater value than RF (Table 4).

The variable importance values (Fig. 9) were normalised according to the greatest value in each year. Elevation, red band, and SWIR2 appeared amongst the five most important variables for the three analysed years. Elevation appeared as the first most important for both 2008 and 2016, and the third in 2000. The red band was the second most important for 2008 and 2016, and the fourth for 2000. SWIR2 was the most important in 2000, the fourth most important variable in 2008, and the third in 2016. TRI was important for models of 2008 and 2016. Other variables appeared in only one year, such as blue band (2000), mNDWI (2000), slope (2016), and SWIR1. On the other hand, MSAVI2, NDVI, NIR, SAVI, and SR were the least important variables.

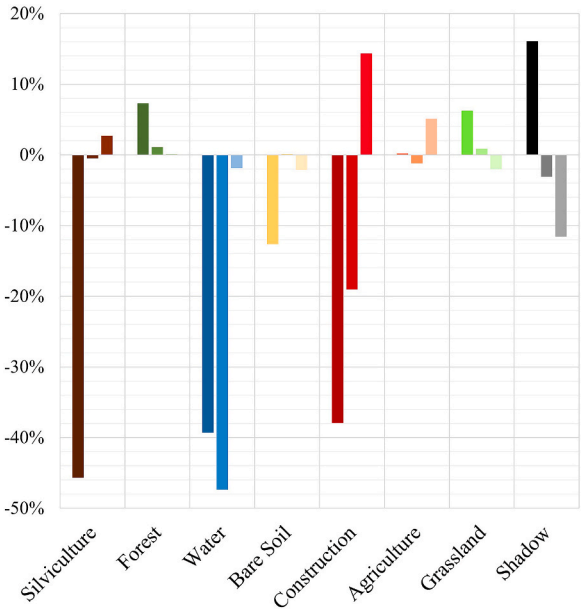

Fig. 7. Difference between RF and RF-CMAP classifications. Positive values indicated that RF-CMAP classified more pixels as a given class than RF. In the same way, negative values mean that RF-CMAP classified a smaller area as a determined class compared to RF. Read from left to right: 2000 (darker tones), 2008 (intermediate tones), and 2016 (lighter tones).

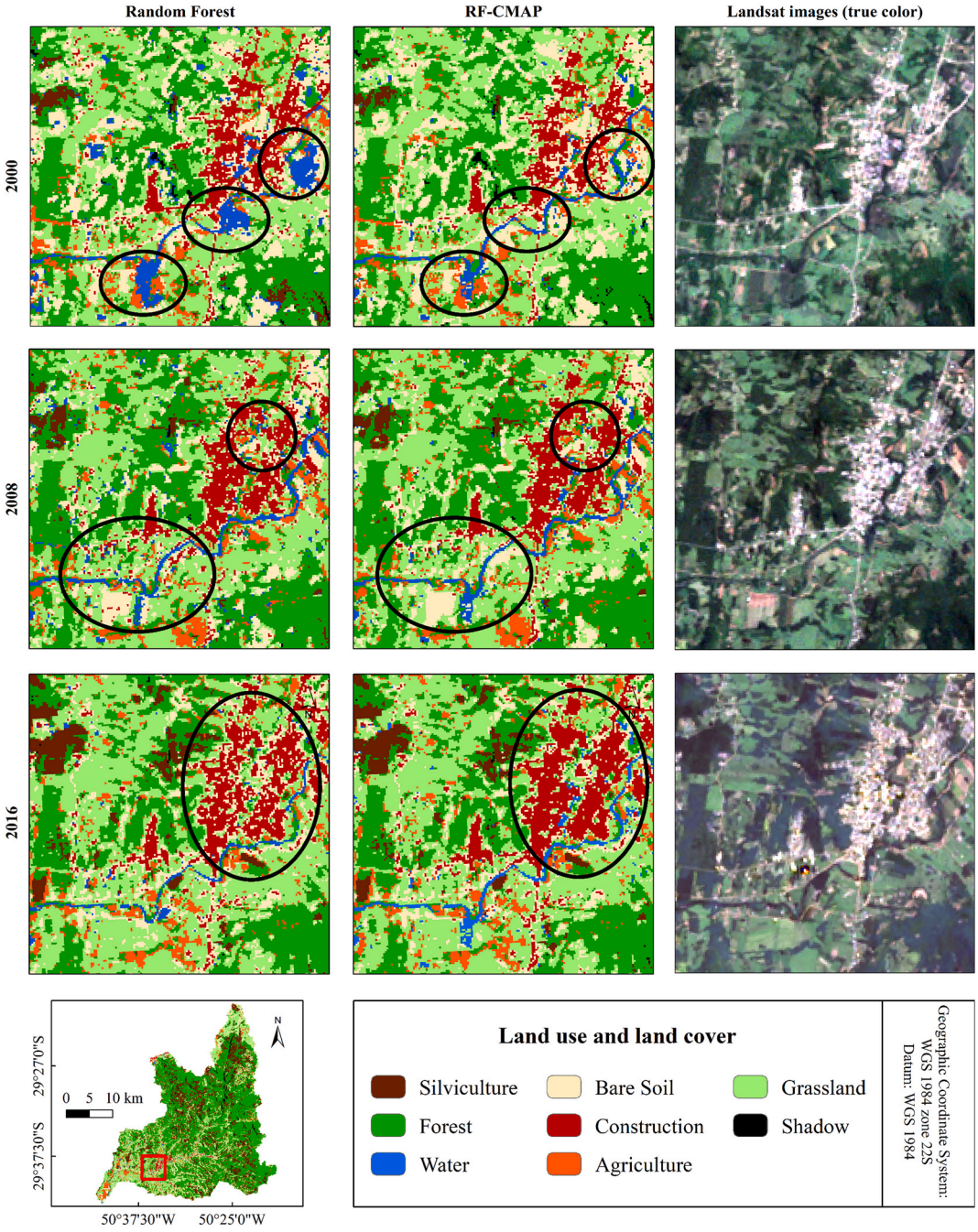

Fig. 8. RF and RF-CMAP classification subsets in an area near Rolante City. The left and central columns show RF and RF-CMAP results, respectively. The right column refers to the Landsat image used for classification. Each row represents an analysed year. Black circles indicate areas with main differences between the models' classification.

4.3. Accuracy assessment of areas with and without transitions

Among all the 1706 validation samples, 63.36% presented no changes in all the analysed period, i.e., in the trajectory 2000-2008-2016. In this stable areas, RF-CMAP performed better than RF in most of the LULC classes (Table 5). RF classified more correctly the grassland and shadow classes, although the differences with RF-CMAP classification relied only in 2 and 3 samples, respectively.

The temporal analysis considered also the transition between two years. Therefore, we analysed two transition periods: i) 2000 to 2008, and ii) 2008 to 2016. In a similar way, most classes were more correctly classified by RF-CMAP, with exception only for grassland and shadow classes (Table 6). Note that the F1 Score for the RF was higher than RF-CMAP only for shadow class in the first transition (2000–2008). In addition, the improvement in classification is represented also by the slightly highest overall accuracy

Table 4
Accuracy metrics for LULC classes in each analysed year and classification model.

	RF									RF-CMAP								
	2000			2008			2016			2000			2008			2016		
	PA	UA	F1	PA	UA	F1	PA	UA	F1	PA	UA	F1	PA	UA	F1	PA	UA	F1
Sc	0.956	0.889	0.921	0.881	0.956	0.917	0.890	0.976	0.931	0.956	0.960	0.958	0.937	0.958	0.947	0.952	0.970	0.961
F	0.947	0.906	0.926	0.950	0.840	0.892	0.942	0.880	0.910	0.975	0.908	0.940	0.989	0.888	0.936	0.958	0.956	0.957
W	0.839	0.956	0.894	0.915	0.944	0.929	0.892	0.928	0.910	0.939	0.984	0.961	0.946	0.976	0.961	0.946	0.976	0.961
BS	0.759	0.835	0.795	0.751	0.876	0.809	0.789	0.856	0.821	0.793	0.855	0.823	0.765	0.912	0.832	0.793	0.925	0.854
C	0.822	0.841	0.831	0.893	0.906	0.899	0.879	0.912	0.895	0.793	0.964	0.870	0.921	0.949	0.935	0.986	0.909	0.946
A	0.762	0.674	0.715	0.828	0.884	0.855	0.843	0.833	0.838	0.774	0.707	0.739	0.841	0.910	0.874	0.849	0.859	0.854
G	0.842	0.854	0.848	0.902	0.786	0.840	0.867	0.768	0.815	0.850	0.843	0.846	0.914	0.793	0.849	0.886	0.778	0.828
Sd	0.992	0.975	0.983	0.864	1.000	0.927	0.943	0.892	0.917	0.992	0.936	0.963	0.818	1.000	0.900	0.933	0.916	0.924
OA	0.874			0.880			0.883			0.894			0.906			0.916		

PA: Producer's accuracy; UA: user's accuracy; F1: F1 score; Sc: silviculture, F: forest, W: water, BS: bare soil; C: construction, A: agriculture, G: grassland, Sd: shadow; OA: overall accuracy.

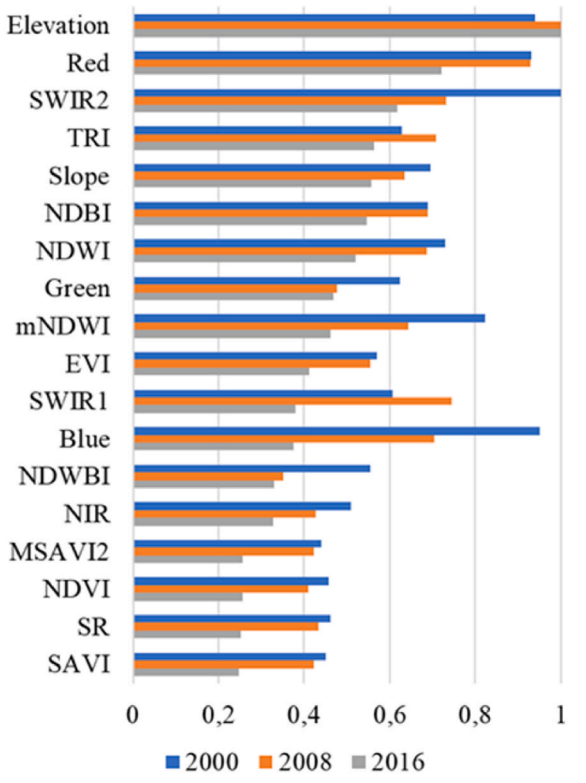

Fig. 9. Variable importance ranking given by RF modelling for each year. The highest importance value was used for standardizing the values each year.

Table 5
Corrected classified LULC classes that presented no changes in the entire analysed trajectory (2000-2008-2016).

	Reference samples	RF		RF-CMAP	
Silviculture	145	120	82.76%	137	94.48%
Forest	333	301	90.39%	324	97.30%
Water	130	95	73.08%	122	93.85%
Bare Soil	83	48	57.83%	58	69.88%
Construction	135	89	65.93%	107	79.26%
Agriculture	51	26	50.98%	28	54.90%
Grassland	105	83	79.05%	81	77.14%
Shadow	99	86	86.87%	82	82.83%
Overall Accuracy		0.784		0.869	

values reached by RF-CMAP.

In the study area, we found 24 possible LUCC between 2000 and 2008, which were analysed by transition period. Among the 24 possible LUCC, RF-CMAP classified 10 classes more correctly than RF, 13 class presented the same number of correct classified sample in both models, and in just in one transition class RF performed better than RF-CMAP (Table 7). This models' performance can be verified in the F1 score, which presented highest values for RF-CMAP in 12 transition classes, same values for both models in seven LUCC classes, and higher values for RF than RF-CMAP in five classes.

The second analysed period, 2008 to 2016, presented 21 possible LUCC in the study area. In this scenario, RF-CMAP outperformed the RF in 13 transition classes, more than half of the LUCC classes, according to the F1 Score values (Table 8). RF presented higher F1 Score values than RF-CMAP in six classes, and the same value in two classes. Also, RF-CMAP presented a slightly higher OA when compared with RF (0.718 and 0.669, respectively).

4.4. Validation of discordant areas between the classifications

The discordant areas between the classifications were analysed considering each year separately (Supplementary Material 6). The greatest discordant area occurred in 2000, when the classifications from RF and RF-CMAP discorded in 54.43 km^2 (6.6% of the total area) (Table 9). The area with differences decreased to 41.14 km^2 (5%) in 2008 and 30.95 km^2 (3.7%) in 2016. Subsequently, we

Table 6
Accuracy metrics for LULC classes that presented no changes in each analysed transition periods (2000–2008 and 2008–2016).

	2000–2008									2008–2016								
		RF				RF-CMAP					RF				RF-CMAP			
	RS	CC	PA	UA	F1	CC	PA	UA	F1	RS	CC	PA	UA	F1	CC	PA	UA	F1
Sc	177	161	0,910	0,970	0.939	168	0,949	0,955	0.952	263	206	0,783	0,990	0.874	242	0,920	0,968	0.943
F	358	330	0,922	0,962	0.942	352	0,983	0,946	0.964	341	317	0,930	0,961	0.945	332	0,974	0,971	0.972
W	130	105	0,808	1000	0.894	122	0,938	0,984	0.960	130	107	0,823	0,991	0.899	123	0,946	0,976	0.961
BS	124	74	0,597	0,871	0.708	88	0,710	0,880	0.786	105	73	0,695	0,924	0.793	75	0,714	0,926	0.806
C	135	102	0,756	0,962	0.847	107	0,793	0,964	0.870	140	110	0,786	0,973	0.870	129	0,921	0,949	0.935
A	62	41	0,661	0,774	0.713	43	0,694	0,768	0.729	109	74	0,679	0,925	0.783	77	0,706	0,928	0.802
G	181	158	0,873	0,854	0.863	157	0,867	0,863	0.865	127	107	0,843	0,748	0.793	106	0,835	0,752	0.791
Sd	108	93	0,861	1000	0.925	89	0,824	1000	0.904	101	88	0,871	1000	0.931	83	0,822	0,954	0.883
OA		0.835				0.883					0.822				0.887			

RS: Reference samples; CC: correct classified; PA: producer's accuracy; UA: user's accuracy; F1: F1 score; Sc: silviculture, F: forest, W: water, BS: bare soil; C: construction, A: agriculture, G: grassland, Sd: shadow; OA: overall accuracy.

Table 7
Accuracy metrics for LULC classes that presented changes between 2000 and 2008.

2000	2008	RS	RF				RF-CMAP			
			CC	PA	UA	F1	CC	PA	UA	F1
Silviculture	Bare Soil	10	7	0,700	1000	0.824	7	0,700	1000	0.824
	Grassland	15	10	0,667	**0,714**	0.690	**12**	**0,800**	0,667	**0.727**
Forest	Silviculture	39	33	0,846	0,589	0.694	33	0,846	**0,647**	**0.733**
	Bare Soil	20	12	0,600	0,800	0.686	**13**	**0,650**	**0,813**	**0.722**
	Construction	1	1	1000	0,333	0.500	1	1000	**0,500**	**0.667**
	Agriculture	2	1	0,500	0,125	0.200	1	0,500	**0,250**	**0.333**
	Grassland	14	10	0,714	0,455	0.556	10	0,714	0,455	0.556
	Shadow	2	**1**	**0,500**	**0,500**	**0.500**	0	0,000	0,000	0.000
Bare Soil	Silviculture	47	23	0,489	1000	0.657	**35**	**0,745**	1000	**0.854**
	Agriculture	47	20	0,426	**0,833**	0.564	**22**	**0,468**	0,815	**0.595**
	Grassland	43	30	0,698	0,508	0.588	**32**	**0,744**	0,508	**0.604**
Agriculture	Silviculture	5	5	1000	0,556	0.715	5	1000	0,556	0.715
	Forest	2	1	0,500	**0,167**	**0.250**	1	0,500	0,125	0.200
	Bare Soil	11	3	0,273	0,500	0.353	**6**	**0,545**	**0,667**	**0.600**
	Construction	1	0	0,000	0,000	0.000	**1**	**1000**	**1000**	**1.000**
	Grassland	3	2	0,667	0,222	0.333	2	0,667	0,222	0.333
Grassland	Silviculture	49	21	0,429	0,700	0.532	**29**	**0,592**	**0,967**	**0.734**
	Forest	11	4	0,364	**0,500**	0.421	**8**	**0,727**	0,421	**0.533**
	Bare Soil	51	40	0,784	**0,816**	**0.800**	40	0,784	0,800	0.792
	Construction	3	0	0,000	0,000	0.000	0	0,000	0,000	0.000
	Agriculture	46	38	0,826	0,655	0.731	38	0,826	0,655	0.731
Shadow	Silviculture	2	2	1000	**0,222**	**0.363**	2	1000	0,182	0.308
	Forest	6	4	0,667	**0,267**	**0.381**	**5**	**0,833**	0,238	0.370
	Bare Soil	1	0	0,000	0,000	0.000	0	0,000	0,000	0.000
Overall Accuracy				0,622				0,703		

RS: Reference samples; CC: correct classified; PA: producer's accuracy; UA: user's accuracy; F1: F1 score; OA: overall accuracy.

Table 8
Accuracy metrics for LULC classes that presented changes between 2008 and 2016.

2008	2016	RS	RF				RF-CMAP			
			CC	PA	UA	F1	CC	PA	UA	F1
Silviculture	Bare Soil	34	22	0,647	**1000**	0.786	**26**	**0,765**	0,963	**0.853**
	Grassland	20	12	0,600	0,630	0.615	**16**	**0,800**	**0,696**	**0.744**
	Shadow	2	2	1000	**0,18**	**0.305**	2	1000	0,167	0.286
Forest	Silviculture	16	11	0,688	0,220	0.333	**13**	**0,813**	**0,382**	**0.520**
	Bare Soil	10	7	0,700	0,636	0.666	7	0,700	**0,778**	**0.737**
	Agriculture	1	1	1000	**0,077**	**0.143**	1	1000	0,071	0.133
	Grassland	7	6	0,857	0,462	0.600	6	0,857	0,462	0.600
	Shadow	2	1	0,500	**0,143**	**0.222**	1	0,500	0,125	0.200
Bare Soil	Silviculture	39	19	0,487	**0,950**	0.644	**20**	**0,513**	0,870	**0.645**
	Agriculture	32	22	0,688	**0,917**	0.786	**25**	**0,781**	0,893	**0.833**
	Grassland	41	25	0,610	0,694	0.649	**27**	**0,659**	**0,711**	**0.684**
Agriculture	Silviculture	5	4	0,800	0,333	0.470	4	0,800	**0,400**	**0.533**
	Forest	2	2	1000	0,500	0.667	2	1000	**1000**	**1.000**
	Bare Soil	26	22	0,846	**0,917**	**0.880**	22	0,846	0,880	0.863
	Grassland	15	7	0,467	**0,318**	**0.378**	7	0,467	0,292	0.359
Grassland	Silviculture	48	33	0,688	**0,733**	0.710	**34**	**0,708**	0,723	**0.715**
	Forest	7	5	0,714	**0,500**	0.588	**6**	**0,857**	0,462	**0.600**
	Bare Soil	43	31	0,721	**0,705**	**0.713**	31	0,721	0,689	0.705
	Construction	1	0	0,000	0,000	0.000	0	0,000	0,000	0.000
	Agriculture	30	28	0,933	0,609	0.737	28	0,933	**0,622**	**0.746**
Shadow	Forest	9	1	0,111	0,500	0.182	**2**	**0,222**	0,667	**0.333**
Overall Accuracy			0,669					0,718		

RS: Reference samples; CC: correct classified; PA: producer's accuracy; UA: user's accuracy; F1: F1 score; OA: overall accuracy.

randomly selected validation subsets to analyse this area and identified the most accurate model (Table 9).

In general, RF-CMAP classified more correctly than RF in the discordant areas (Fig. 10). In 2000, RF-CMAP correctly classified 66% of these areas, followed by 21% from RF. In approximately 14% of the discordant area, both models presented error, i.e., incorrectly classified the area. In addition, RF-CMAP also presented the highest percentage of correctly classified areas in 2008 (46%) and 2016

Table 9
Discordant areas and used samples for validation.

	Area (km^2)	Samples
2000	54.43	600
2008	41.14	517
2016	30.95	424

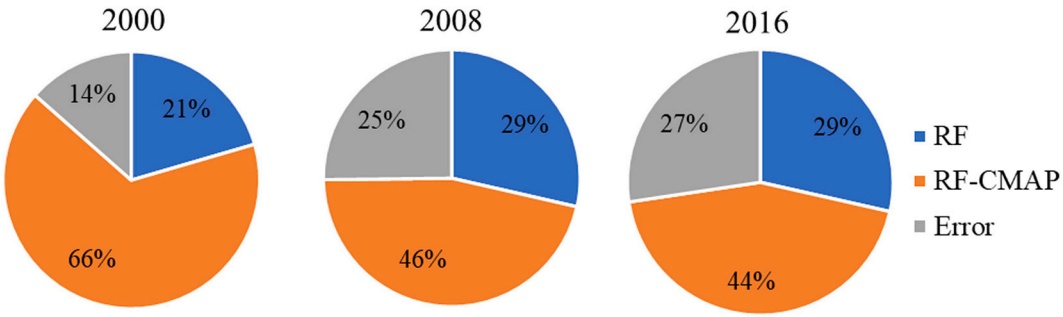

Fig. 10. Validation of discordant areas between RF and RF-CMAP classifications. The percentage represents the corrected classification according to each model. Error represents areas wrongly classified by both models.

(44%).

The percentage of correct classification in the discordant areas per LULC class (Fig. 11) is discussed separately as follows. In 2000, the RF-CMAP performed better than RF mainly for shadow (89%) and forest (88%), in addition to grassland (56%) and water (55%). On the other hand, RF stood out in construction (93%), silviculture (77%), and bare soil (64%). The areas wrongly classified by both models remained similar in all classes, with a maximum value of 29% in the grassland, followed by 20% for agriculture and 18% for water.

In 2008, RF-CMAP correctly classified more areas of forest (66%), silviculture (60%), construction (54%), and grassland (39%) than RF. The RF correctly classified more shadow (63%) and water (55%). The areas wrongly classified by both models were greater than the correct classification in agriculture (47%) and bare soil (36%). Finally, in 2016, RF-CMAP well-classified construction (70%), forest (54%), and silviculture (48%); RF had more correct classification in shadow (67%), water (48%), and grassland (45%). Both models misclassified bare soil (76%) and agriculture (43%).

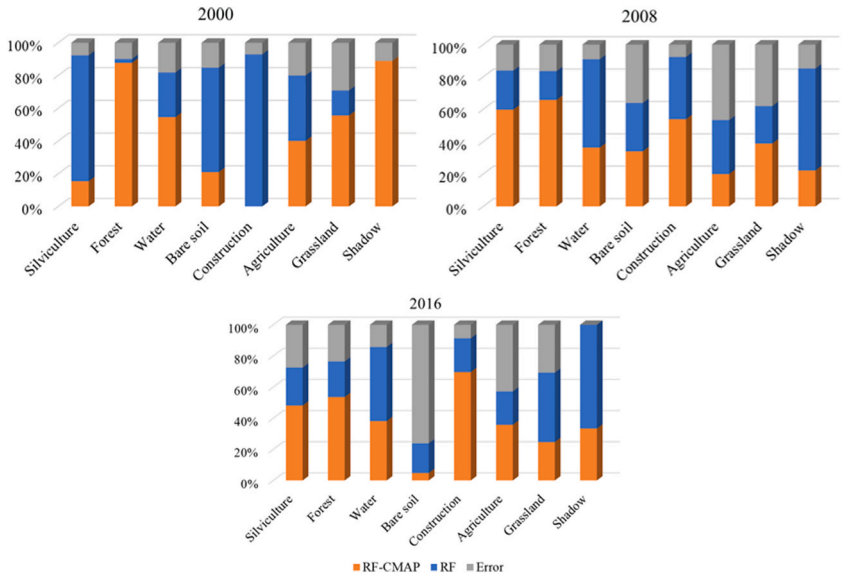

Fig. 11. Validation of discordant areas between RF and RF-CMAP classifications by LULC classes. Each colour represents the corrected classification according to the models. Error is the areas wrongly classified by both models.

4.5. Land cover transition

As previously discussed, individually stacking the RF-generated classifications can lead to the generation of invalid trajectories. In this study, invalid trajectories covered a total area of 99.92 km^2 (12%), which corresponded to 71.57 km^2 (8.72%) between the years 2000 and 2008, and 28.35 km^2 (3.46%) between 2008 and 2016 (Fig. 12). RF-CMAP presented only valid LULC transitions (i.e., considering each period of eight years), although some invalid trajectories (between 2000 and 2016, three sequences in analysis) were identified due to the inclusion of the shadow class, which corresponded to 9.64 km^2 (1.16%). Since this covers a small area, we consider that this limitation does not affect our analysis. However, this may be a problem for studies in other areas/periods. The treatment of non-LULC classes that can disrupt the calculation of $P(s)$ by the adopted methodology should be further investigated using alternative methods, such as the "simplified" one previously proposed in Reis et al. (2020).

The invalid transitions between 2000 and 2008 in RF classifications were mainly related to increases in forest class in 2008. There was indeed an increase in forest area in this period. However, considering eight years, just areas previously occupied by agriculture, grassland, and shadow were capable of such transformation in the RRB. Then, areas with silviculture, water, bare soil, and constructions were not able to become a forest in the subsequent eight years, being an invalid transition.

In the first period, RF classification showed that 38.45 km^2 of silviculture and 17.28 km^2 of bare soil in 2000 changed and became forest in 2008. Similarly, 11.77 km^2 of silviculture and 4.07 km^2 of bare soil in 2008 became forest in 2016. Forestry crops are usually cut when the trees reach a large size. Thus, for these areas to become a forest, the planted trees must be removed in their entirety, leaving the soil exposed. Although the RRB has a humid climate and increases in forest areas were recorded, it is unlikely that there will be large forest growth in areas of bare soil over eight years.

Other invalid transitions in both analysed periods were related to an increase in the water surface. We considered that only shadow could be transformed into water, since none of the analysed years presented floods. Also, the RRB consists of a main river with few tributaries. As such, the agricultural ponds in the region often comprise few pixels and there was no significant production of new ponds.

In addition, RF-CMAP was capable of correct some errors presented by RF classification. When considering the stable classes, i.e., areas with no LUCC in the analysed period, RF-CMAP corrected more than 50% of the RF classification errors related to forest and water class between 2000 and 2008, and regarding to silviculture, forest, water, and construction between 2008 and 2016 (Fig. 13).

Finally, RF-CMAP corrected near to 78% and 81% of the RF errors that contained invalid transition in the periods 2000–2008 and 2008–2016, respectively. The greatest improvement occurs in areas with forest, water, and bare soil in the first period, presenting correction in more than 80% of the samples. Between 2008 and 2016, almost all the classes presented more than 75% of correction by RF-CMAP, the exception relied in bare soil (33% - two corrections of six invalid transitions).

4.6. LULC in landslide-affected areas

Upon examining the landslide-affected area in 2017, we conducted an analysis using RF classification (Fig. 14) and discovered the existence of invalid transitions. When observed within the landslide polygons, these invalid transitions can lead to misinterpretations by establishing a connection between LULC/LUCC and landslides. In total, we identified 335 landslide polygons, out of which a significant number, at least 118 (35%), showed invalid transitions between 2000 and 2008. Similarly, from 2008 to 2016, we found that 54 (16%) of the landslides exhibited invalid transitions (Fig. 15).

Four different invalid transitions were identified within the landslides, the conversion of: i) bare soil to forest; ii) water to forest; iii) silviculture to forest; and iv) forest to water. It is noteworthy that these transition dynamics were not observed in the RRB when considering a relatively short period of eight years (i.e., invalid transitions). Consequently, the landslide analysis could be affected if we assume, for example, that 35% of the landslides were associated with reforestation, even though there was no evidence of forest growth over eight years.

Furthermore, these inconsistencies in LULC transition maps could result in misinterpretation according to the landslide mechanical areas (initiation, transport, and deposition) (Supplementary Material 7). For example, 90% of the pixels defined as a transition from silviculture to forest between 2000 and 2008 located in landslide areas were found in the middle or lower part of the slope, corresponding to transport (65%) and deposition (25%) zones. This could lead us to interpret that an overweight due to the forest growing or the remotion of silviculture generated an instability in the slope. Similarly, the misclassified pixels found in the landslide rupture area (initiation zone) between 2008 and 2016 corresponded to forest regrowth related to the conversion of bare soil and silviculture. Therefore, it is noteworthy that the subsequent analysis, such as landslide susceptibility, will be directly associated with LULC mapping consistency.

5. Discussion

5.1. RF improvements by adding the CMAP temporal approach

Several approaches exist to classify LULC, aiming to ensure temporal representativeness and realistic transitions over time (Boori and Voženílek, 2014). However, traditional classifiers can result in invalid transitions, leading us to misinterpret the real LULC dynamics in the analysed area. In this sense, Reis et al. (2020) developed the Compound Maximum *a Posteriori* (CMAP) algorithm as a trajectory classifier, designed to incorporate multi-temporal datasets to classify the whole LULC trajectory at once. The CMAP was applied to Amazon (Reis, 2022) and Atlantic (Maciel et al., 2021) forest environments.

Nevertheless, some drawbacks of CMAP consisted of expect unimodal classes, considering only Gaussian data (Reis, 2022; Reis et al., 2020). In that sense, some studies are adapting the CMAP to consider non-Gaussian data, such as Dutra et al. (2023), who

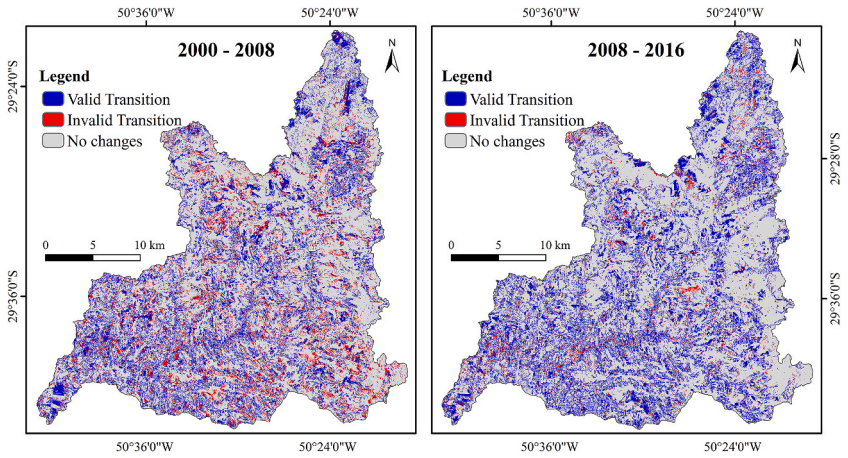

Fig. 12. Validity of land cover transitions for each eight-year analysed period, considering only RF classifications.

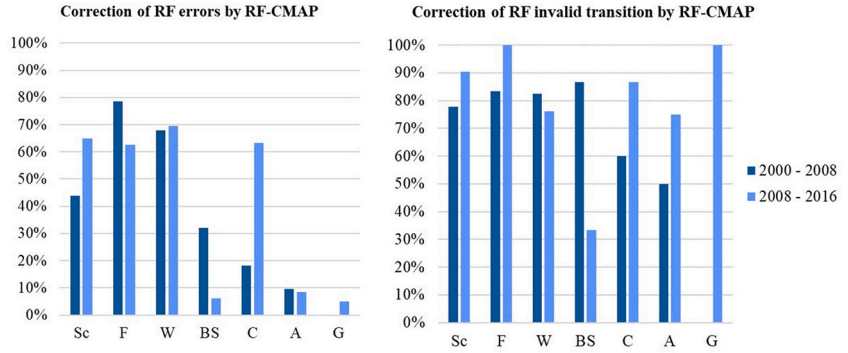

Fig. 13. The amount of validation samples containing errors in RF classification, which was corrected by the RF-CMAP, regarding to areas without LULC transition.

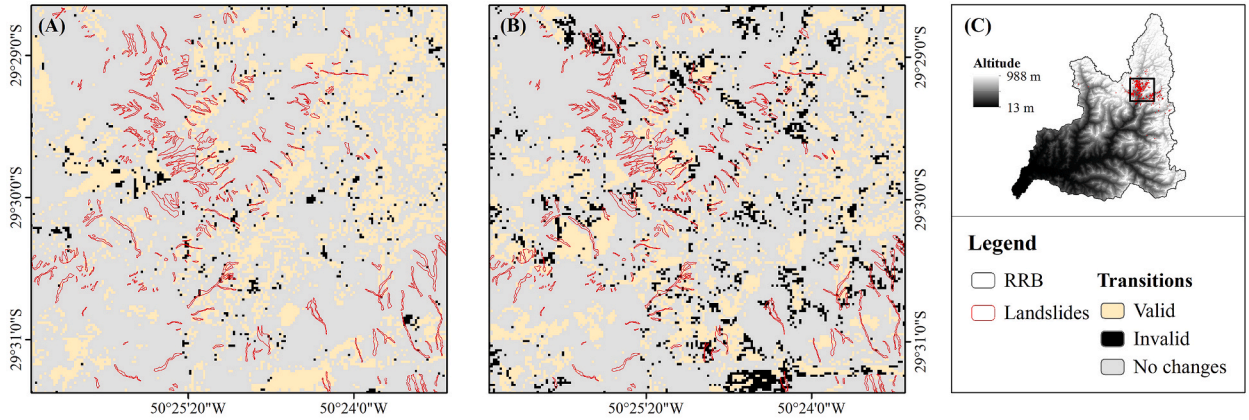

Fig. 14. Presence of invalid transitions in RF outputs between 2000 and 2008 (A) and between 2008 and 2016 (B) in landslide-affected areas (C).

developed a non-parametric version of CMAP, using Decision Trees. The authors found that this version is efficient for inconsistencies-free classifications. Following this approach, we substituted the class likelihood at each point of time with features derived from the Random Forest (RF) algorithm (Breiman, 2001), which does not assume any specific distribution for the data, calling this RF-CMAP. The RF algorithm offers the advantages of being capable of dealing with nonlinear data, which can be necessary for mountainous areas (e.g. morphometric factors).

Generally, the combination of spectral bands and indices provides accurate LULC modeling (Chaves et al., 2020), which can be improved by including morphometric parameters in mountainous areas (Wang et al., 2020). The variable importance analysis derived

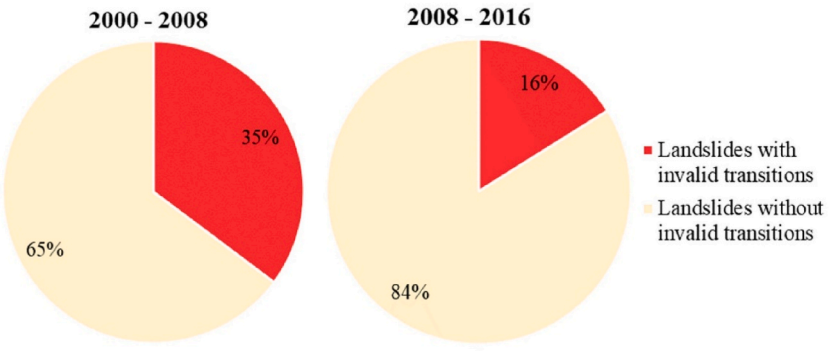

Fig. 15. Percentage of landslide polygons with invalid transitions in RF results.

from the RF showed the importance of including morphometric factors in our analysis. Elevation was one of the most important variables, as well as the slope and Terrain Ruggedness Index (TRI) (Fig. 9). This can be justified by the different LULC along the RRB, e.g., grassland and agriculture are mostly located in the higher and lower altitudes, while forest and silviculture are predominantly in intermediate altitudes, on the plateau escarpment. Uehara et al. (2022) applied time-series metrics for LULC mapping in the RRB, focusing on landslide identification, and found slope and altitude among the ten most important predictors.

The use of the RF-CMAP is also useful for mountainous areas, as it could avoid common problems in classifications for this kind of region, such as cloud shadow confusion with water bodies due to its low reflectance. In addition, in very dynamic environments it could help to post-process the classification of the retrieved time-series to ensure that only possible transitions are contained in the results. Moreover, the RF-CMAP improves the spatiotemporal robustness of the time-series analysis, allowing a better understanding of land use changes over time and avoiding errors in these temporal analyses. This is important for a careful analysis of land changes, the relationships between LUCC and biogeochemical parameters (e.g., water quality), and policy making.

Although the OA values for both algorithms were similar, RF-CMAP tended to overpermed the RF and avoided invalid transitions, which presented a valuable improvement in the final map (Fig. 8). The analysis of discordant areas between the classifications showed that, in general, RF-CMAP was more capable of correctly classifying these areas, e.g., the forest class was more correctly classified by RF-CMAP in all analysed years (Fig. 11). The evaluation of discordant areas acts as a second model validation and allows us to verify which model performed better.

Furthermore, LULC misclassifications can cause errors in any derived analysis. The estimation of deforestation area (Tavares et al., 2022), urban growth (Ishtiaque et al., 2017), floods (Thiam et al., 2022), and landslides (Sangeeta and Singh, 2023), among other subjects, are examples of subsequent studies that can be affected by LULC classification errors. Therefore, guaranteeing the correctness of LULC classification by adding a temporal approach, such as the CMAP, can improve subsequent studies and analysis.

5.2. Impacts of invalid transitions in landslide analysis

Human activities have significantly influenced landslide occurrence (Glade, 2003), with recent studies evaluating these influences through LULC and LUCC (Quevedo et al., 2023a, 2023b). LULC and LUCC directly impact soil mechanical behaviour and moisture. Vegetation, for example, plays a critical role in safeguarding soil against erosion and enhancing slope stability through mechanical anchoring and soil suction by roots (Löbmann et al., 2020; Masi et al., 2021; Parra et al., 2021). Conversely, deforestation, construction, and the abandonment of croplands compromise slope stability and improve landslide susceptibility (Chen et al., 2019; Persichillo et al., 2017).

Then, understanding the impacts of LULC and LUCC on landslide occurrences is important for disaster mitigation (Chen and Huang, 2013). However, it is equally important to ensure that the analysis of LULC/LUCC influence is based on suitable and accurate maps. Because of that, we aimed to guarantee a reliable LULC mapping, without invalid transitions to avoid misinterpretation about this relationship.

The RRB landslide inventory (Quevedo et al., 2024), composed of 335 polygons, showed invalid transitions in 35% of the landslides between 2000 and 2008, according to the RF classification. In other words, at least 118 landslides could be associated with a non-realistic LULC transition. When considering the interval between 2008 and 2016, at least 16% (54) of the landslides contained invalid transitions. Since many studies are seeking relationships between landslides and LULC/LUCC (Chen et al., 2019; Gariano et al., 2018; Persichillo et al., 2017; Reichenbach et al., 2014), it is essential to guarantee a reliable LULC mapping, with only valid transitions and trajectories.

In addition, it is important to understand the entire slope failure process, i.e., consider landslide mechanical areas: initiation, transport, and deposition. This importance relies on the different nature of the process that occurs in each part of the hillslope and needs to be considered in landslide susceptibility modelling (Sun et al., 2022). For example, we found that most of the invalid transition associated with forest regrowth in forestry areas is in the transport (65%) or deposition (25%) areas between 2000 and 2008. The rupture areas were marked by the presence of forest regrowth associated with invalid transitions between 2008 and 2016.

Finally, the presence of invalid transitions and/or trajectories could lead to misinterpretation about the land cover dynamics in an analysed area. Such analysis was presented by Quevedo et al. (2023a), in which overlaying MAPBIOMAS (Souza et al., 2020) LULC

maps provided invalid transitions in Petrópolis municipality. Furthermore, studies related to LULC future scenarios, with subsequent susceptibility future scenarios and/or recommendations given to local managers can be spoiled (Gariano et al., 2018; Rabby et al., 2022).

5.3. Limitations and uncertainties

Some limitations were found in this study, such as using one different training sample set for each year, including collecting samples and applying the active learning method for each data set, which is time and resource-consuming. Although it tends to be more accurate, calibrating and using only one training sample set for all the analysed years can speed up and decrease the costs of the LULC classification, as demonstrated by Reis et al. (2019).

Moreover, the shadow class can generate invalid trajectories in the analysis for more than two consecutive years. Here, we treated shadow as a LULC class, therefore allowing any transitions including this class. This scenario was not originally accounted for in CMAP (Reis et al., 2020) and may result in the classification of trajectories that could be considered invalid when analysing the three years jointly (e.g. Water in 2000, Shadow in 2008, and Forest in 2016). Although this effect was small in our study, it can be more expressive for analysis using longer time series. One solution could be to use different approaches to calculate the *a priori* probabilities of the trajectories, as proposed by Reis et al. (2020), and further exemplified by the classification of 37 annual observations of Landsat data containing similar non-LULC class Clouds/Cloud Shadows, in Reis (2022).

Finally, it is necessary to ensure a high-quality classification of the first image, as it is used as the main parameter of valid/invalid trajectories. Alternatives include, despite the manual classification of the first map, a bootstrap approach that could ensure an interactive process between all the dates and the reduction of uncertainty of the first date. Despite these unsolved uncertainties, the proposed model presents an easy-to-implement approach to deal with invalid transitions in LULC using the common RF approach. The proposed model could be used in RF-based remote sensing LULC products (e.g., MAPBIOMAS (Souza et al., 2020) and TERRACLAS (Almeida et al., 2016) in Brazil) to improve classification accuracy over time.

6. Conclusions

The guarantee of a reliable LULC transition map is essential for subsequent analysis, such as landslide susceptibility modelling. In this study, we found that invalid transitions could lead us to misinterpret the relationship between LULC and landslide occurrence in the *Rolante* River Basin (RRB). In this sense, the integration of the Random Forest (RF) with Compound Maximum *a Posteriori* (CMAP) presented a high capacity for LULC classification, avoiding invalid transitions when using variables with different distributions, such as the morphometric parameters. The overall accuracy (OA) was slightly higher for RF-CMAP than RF, and the results demonstrated that RF-CMAP not only avoided invalid transitions in transition labelling, improving the classification maps over time, but also correct some areas with errors in RF classification. In addition, the RF-CMAP also performed better than RF in discordant areas and avoided invalid transitions where landslides occurred, preventing errors in subsequent analysis. As such, we conclude that RF-CMAP performed better than the traditional RF for the analysed scenario (study area, classes, image, and analysed times). Furthermore, we recommend that future studies based on multi-temporal datasets focus on algorithms that consider the LULC dynamics in the classification process, to avoid errors derived from invalid transitions.

Disclosure statement

The authors declare no conflict of interest.

CRediT authorship contribution statement

Renata Pacheco Quevedo: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Daniel Andrade Maciel:** Writing – review & editing, Validation, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Mariane Souza Reis:** Writing – original draft, Software, Formal analysis. **Camilo Daleles Rennó:** Writing – review & editing, Methodology, Formal analysis. **Luciano Vieira Dutra:** Writing – review & editing, Methodology, Formal analysis. **Clódis de Oliveira Andrades-Filho:** Writing – review & editing, Validation, Methodology, Formal analysis, Data curation. **Andrés Velástegui-Montoya:** Writing – review & editing, Funding acquisition, Formal analysis. **Tingyu Zhang:** Writing – review & editing, Formal analysis. **Thales Sehn Körting:** Writing – review & editing, Formal analysis. **Liana Oighenstein Anderson:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This research was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), Brazil - Finance Code 001 and in part by the National Council of Technological and Scientific Development (CNPq), Brazil project number 422354/2023-6 (Monitoring and Alerts of Land Cover Changes in Brazilian Biomes - Training and Semi-Automation of the BiomasBR Program), supported by the National Institute for Space Research (INPE), Brazil. This research was partly financed by the ESPOL Polytechnic University, Ecuador project "IPUS: Identification and Prediction of Urban Sprawl", with code FICT-8-2023. LOA acknowledge the National Council for Scientific and Technological Development (CNPq), process number 314473/2020-3. The authors thank the São Paulo Research Foundation (FAPESP), Brazil grant no 2023/09118-6, and the Brazilian National Council for Scientific and Technological Development (CNPq, grant no 302205/2023-3).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rsase.2024.101314.

References

Almeida, C.A. de, Coutinho, A.C., Esquerdo, J.C.D.M., Adami, M., Venturieri, A., Diniz, C.G., Dessay, N., Durieux, L., Gomes, A.R., 2016. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. Acta Amazonica 46, 291–302. https://doi.org/10.1590/1809-4392201505504.

Alvares, C.A., Stape, J.L., Sentelhas, P.C., De Moraes Gonçalves, J.L., Sparovek, G., 2013. Köppen's climate classification map for Brazil. Meteorol. Z. 22, 711–728. https://doi.org/10.1127/0941-2948/2013/0507.

Andrades-Filho, C. de O., Mexias, L.F.S., Iescheck, A.L., Viero, A.P., Giaccom, B., Cargnin, B. da R., Michelin, C.R.L., Sluter, C.R., Gonzatti, C., Cavalheiro, D., Cremon, É., Barboza, E.G., Schwarzer, G., Jacques, F. de M., Oliveira, G.G., Dorneles, J.I., Cacciatore, J.A., Schmitt, H., Novakoski, K.R., Duarte, L. da C., Guasselli, L.A., Petry, L., Mevel, L.B., Peres, L.D. da S., Reis, M. da S., da Rosa, M.L.C.C., Righi, M., Silva, M.C., Ribeiro Junior, M., Gruber, N.L.S., Herrmann, P.B., Schumacher, R.G., Quevedo, R.P., Philipp, R.P., Cárdenas, S.M.M., Soares, V.M., Peres-Núñez, W., 2024. Mapeamento das cicatrizes de movimentos de massa decorrentes do acumulado de chuva no RS entre 27/04 e 13/05 de 2024. Porto Alegre.

Belgiu, M., Drăgu, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogrammetry Remote Sens. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Boonprong, S., Cao, C., Chen, W., Bao, S., 2018. Random forest variable importance spectral indices scheme for burnt forest recovery monitoring—multilevel RF-VIMP. Rem. Sens. 10, 807. https://doi.org/10.3390/rs10060807.

Boori, M.S., Voženílek, V., 2014. Remote sensing and land use/land cover trajectories. J. Geophy. Rem. Sens. https://doi.org/10.4172/2169-0049.1000123.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24, 123–140.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1201/9780367816377-11.

Calle, M.L., Urrea, V., 2011. Letter to the editor: stability of random forest importance measures. Briefings Bioinf. 12, 86–89. https://doi.org/10.1093/bib/bbq011.

Chaves, M.E.D., Picoli, M.C.A., Sanches, I.D., 2020. Recent applications of landsat 8/OLI and sentinel-2/MSI for land use and land cover mapping: a systematic review. Rem. Sens. 12, 3062. https://doi.org/10.3390/rs12183062.

Chen, C.Y., Huang, W.L., 2013. Land use change and landslide characteristics analysis for community-based disaster mitigation. Environ. Monit. Assess. 185, 4125–4139. https://doi.org/10.1007/s10661-012-2855-y.

Chen, L., Guo, Z., Yin, K., Shrestha, D.P., Jin, S., 2019. The influence of land use and land cover change on landslide susceptibility: A case study in Zhushan Town, Xuan'en County (Hubei, China). Nat. Hazards Earth Syst. Sci. 19, 2207–2228. https://doi.org/10.5194/nhess-19-2207-2019.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens. Environ. 37, 35–46. https://doi.org/10.1016/0034-4257(91)90048-B.

Dale, V.H., 1997. The relationship between land-use change and climate change. STUDIES of CLIMATE CHANGE 753 Ecological Applications.

Dutra, L.V., Rennó, C.D., Reis, M.S., Gamba, P., 2023. A generative method for simultaneous classification of Remote Sensing time series data using an ensemble of Decision Tree classifiers. In: Anais Do XX Simpósio Brasileiro de Sensoriamento Remoto, pp. 179–182. Florianópolis.

Ebrahimy, H., Mirbagheri, B., Matkan, A.A., Azadbakht, M., 2021. Per-pixel land cover accuracy prediction: a random forest-based method with limited reference sample data. ISPRS J. Photogrammetry Remote Sens. 172, 17–27. https://doi.org/10.1016/j.isprsjprs.2020.11.024.

Fonseca, C.R., Ganade, G., Baldissera, R., Becker, C.G., Boelter, C.R., Brescovit, A.D., Campos, L.M., Fleck, T., Fonseca, V.S., Hartz, S.M., Joner, F., Käffer, M.I., Leal-Zanchet, A.M., Marcelli, M.P., Mesquita, A.S., Mondin, C.A., Paz, C.P., Petry, M.V., Piovensan, F.N., Putzke, J., Stranz, A., Vergara, M., Vieira, E.M., 2009. Towards an ecologically-sustainable forestry in the Atlantic Forest. Biol. Conserv. 142, 1209–1219. https://doi.org/10.1016/j.biocon.2009.02.017.

Foody, G.M., 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. 80, 185–201. https://doi.org/10.1016/S0034-4257(01)00295-4.

Gao, B., 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3.

Garg, V., Nikam, B.R., Thakur, P.K., Aggarwal, S.P., Gupta, P.K., Srivastav, S.K., 2019. Human-induced land use land cover change and its impact on hydrology. HydroResearch 1, 48–56. https://doi.org/10.1016/j.hydres.2019.06.001.

Gariano, S.L., Petrucci, O., Rianna, G., Santini, M., Guzzetti, F., 2018. Impacts of past and future land changes on landslides in southern Italy. Reg. Environ. Change 18, 437–449. https://doi.org/10.1007/s10113-017-1210-9.

Georganos, S., Grippa, T., Vanhuysse, S., Lennert, M., Shimoni, M., Kalogirou, S., Wolff, E., 2018. Less is more: optimizing classification performance through feature selection in a very-high-resolution remote sensing object-based urban application. GIsci. Rem. Sens. 55, 221–242. https://doi.org/10.1080/15481603.2017.1408892.

Glade, T., 2003. Landslide occurrence as a response to land use change: a review of evidence from New Zealand. Catena 51, 297–314. https://doi.org/10.1016/S0341-8162(02)00170-4.

Habib, M., 2021. Quantifying topographic ruggedness using principal component analysis. Adv. Civ. Eng. 2021, 1–20. https://doi.org/10.1155/2021/3311912.

Henderson, K.A., Reis, M., Blanco, C.C., Pillar, V.D., Printes, R.C., Bauch, C.T., Anand, M., 2016. Landowner perceptions of the value of natural forest and natural grassland in a mosaic ecosystem in southern Brazil. Sustain. Sci. 11, 321–330. https://doi.org/10.1007/s11625-015-0319-3.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., 2018. Disturbance-informed annual land cover classification maps of Canada's forested ecosystems for a 29-year landsat time series. Can. J. Rem. Sens. 44, 67–87. https://doi.org/10.1080/07038992.2018.1437719.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309. https://doi.org/10.1016/0034-4257(88)90106-X.

IBGE - Instituto Brasileiro de Geografia e Estatística, 2019. Biomas e sistema costeiro-marinho do Brasil : compatível com a escala 1:250 000.

ICMBio - Instituto Chico Mendes de Conservação da Biodiversidade, 2020. Plano de Manejo da Floresta Nacional de São Francisco de Paula.

Ishtiaque, A., Shrestha, M., Chhetri, N., 2017. Rapid urban growth in the kathmandu valley, Nepal: monitoring land use land cover dynamics of a himalayan city with landsat imageries. Environments 4, 72. https://doi.org/10.3390/environments4040072.

Jia, S., Yang, C., Wang, M., Failler, P., 2022. Heterogeneous impact of land-use on climate change: study from a spatial perspective. Front. Environ. Sci. 10 https://doi.org/10.3389/fenvs.2022.840603.
Jordan, C.F., 1969. Derivation of leaf-area index from quality of light on the forest floor. Ecology 50, 663–666. https://doi.org/10.2307/1936256.
Kaufman, Y.J., Tanre, D., 1992. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. IEEE Trans. Geosci. Rem. Sens. 30, 261–270. https://doi.org/10.1109/36.134076.
Kuhn, M., 2012. The Caret Package.
Lambin, E.F., Turner, B.L., Geist, H.J., Agbola, S.B., Angelsen, A., Bruce, J.W., Coomes, O.T., Dirzo, R., Fischer, G., Folke, C., George, P.S., Homewood, K., Imbernon, J., Leemans, R., Li, X., Moran, E.F., Mortimore, M., Ramakrishnan, P.S., Richards, J.F., Skånes, H., Steffen, W., Stone, G.D., Svedin, U., Veldkamp, T.A., Vogel, C., Xu, J., 2001. The causes of land-use and land-cover change: moving beyond the myths. Global Environ. Change 11, 261–269. https://doi.org/10.1016/S0959-3780(01)00007-3.
Lang, M., 2013. Change of land use and land cover by silvicultural expansion on the Campos de Cima da Serra, Rio Grande do Sul. Technische Universität München, Brazil (2002/03-2008/09).
Lee, T.-H., Ullah, A., Wang, R., 2020. Bootstrap Aggregating and Random Forest, pp. 389–429. https://doi.org/10.1007/978-3-030-31150-6_13.
Leemans, R., Zuidema, G., 1995. Evaluating changes in land cover and their importance for global change. Trends Ecol. Evol. 10, 76–81. https://doi.org/10.1016/S0169-5347(00)88981-8.
Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R. News 2, 18–22.
Liu, D., Song, K., Townshend, J.R.G., Gong, P., 2008. Using local transition probability models in Markov random fields for forest change detection. Remote Sens. Environ. 112, 2222–2231. https://doi.org/10.1016/j.rse.2007.10.002.
Liu, J., Wu, Z., Zhang, H., 2021. Analysis of changes in landslide susceptibility according to land use over 38 years in lixian county, China. Sustainability 13, 1–23. https://doi.org/10.3390/su131910858.
Löbmann, M.T., Geitner, C., Wellstein, C., Zerbe, S., 2020. The influence of herbaceous vegetation on slope stability – a review. Earth Sci. Rev. 209, 103328 https://doi.org/10.1016/j.earscirev.2020.103328.
Lu, D., Mausel, P., Brondízio, E., Moran, E., 2004. Change detection techniques. Int. J. Rem. Sens. 25, 2365–2401. https://doi.org/10.1080/0143116031000139863.
Luerce, T.D., 2015. Geoturismo na bacia Hidrográfica do rio Rolante/RS: Um estudo acerca das Quedas d'água. Universidade Federal do Rio Grande do Sul. https://doi.org/10.1017/CBO9781107415324.004.
Maciel, A.L., Alves, D., Sant'Anna, S., 2021. Análise dos Processos de Desmatamento e Regeneração nas Unidades de Conservação Pertencentes no Vale do Ribeira (SP/PR) por Meio de Classificações de Imagens de Sensoriamento Remoto. Rev. Bras. Cartogr. 73, 261–277. https://doi.org/10.14393/rbcv73n1-55435.
Masi, E.B., Segoni, S., Tofani, V., 2021. Root reinforcement in slope stability models: a review. Geosciences 11, 212. https://doi.org/10.3390/geosciences11050212.
McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int. J. Rem. Sens. 17, 1425–1432. https://doi.org/10.1080/01431169608948714.
Mertens, B., Lambin, E.F., 2000. Land-cover-change trajectories in southern Cameroon. Ann. Assoc. Am. Geogr. 90, 467–494.
Miller, C.L., Laflamme, R.A., 1958. The digital terrain model: theory & application. Society's 24th Annual Meeting, Washington DC.
NASA JPL, 2020. NASADEM Merged DEM Global 1 arc second V001. https://doi.org/10.5067/MEaSUREs/NASADEM/NASADEM_HGT.001.
Olofsson, P., Foody, G.M., Stehman, S.V., Woodcock, C.E., 2013. Making better use of accuracy data in land change studies: estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sens. Environ. 129, 122–131. https://doi.org/10.1016/j.rse.2012.10.031.
Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57. https://doi.org/10.1016/j.rse.2014.02.015.
Open Source, Microsoft, McFarland, M., Emanuele, R., Morris, D., Augspurger, T., 2022. microsoft/PlanetaryComputer: October 2022. https://doi.org/10.5281/zenodo.7261897.
Parra, E., Mohr, C.H., Korup, O., 2021. Predicting patagonian landslides: roles of forest cover and wind speed. Geophys. Res. Lett. 48, 1–10. https://doi.org/10.1029/2021GL095224.
Persichillo, M.G., Bordoni, M., Meisina, C., 2017. The role of land use changes in the distribution of shallow landslides. Sci. Total Environ. 574, 924–937. https://doi.org/10.1016/j.scitotenv.2016.09.125.
Petri, C.A., Galvão, L.S., 2019. Sensitivity of seven MODIS vegetation indices to BRDF effects during the amazonian dry season. Rem. Sens. 11, 1650. https://doi.org/10.3390/rs11141650.
Phan, T.N., Kuch, V., Lehnert, L.W., 2020. Land cover classification using google earth engine and random forest classifier—the role of image composition. Rem. Sens. 12, 2411. https://doi.org/10.3390/rs12152411.
Piao, Y., Jeong, S., Park, S., Lee, D., 2021. Analysis of land use and land cover change using time-series data and random forest in North Korea. Rem. Sens. 13, 3501. https://doi.org/10.3390/rs13173501.
Possantti, I., Aguirre, A., Alberti, C., Andrades Filho, C., Azeredo, L., Balbon, J., Barbedo, R., Barcelos, M., Becker, F., Bedin, M., Bregalda, N., Cacciatore, J., Camana, M., Camargo, P., Cantor, G., Cardozo, T., Cargnin, B., Carrard, G., Castilhos, M., Cazanova, R., Chiarelli, F., Collishonn, W., Cornely, A., Cremon, É., Cunha, L., Cunha, R., Cárdenas, S., Dorneles, J., Dornelles, F., Eckhardt, R., Fan, F., Froner, M., Giaccom, B., Giasson, S., Goldenfum, J., González-Ávila, I., Gonçalves, C., Gonçalves, G., Guasselli, L., Guimarães, E., Guimarães, E., Hellmann, A., Herrmann, P., Horstmann, G., Iablonovski, G., Iescheck, A., Kipper, P., Kobyama, M., Krasner, M., Krob, L., Kuele, P., Laipelt, L., Lutz, V., Maciel, J., Magalhães, F., Mallet, J., Marques, B., Marques, G., Meirelles, F., Mexias, L., Michel, G., Michel, R., Mincarone, M., Moura, E., Müller, J., Neves, É., Nicolini, I., Nonnemacher, L., Novakoski, K., Oliveira, G., Oliveira, M., Ott, P., Paiva, R., Peres, L., Petry, L., Quevedo, R.P., Quintela, R., Ramos, M., Rauber, A., Reis, M., Ribeiro, M., Righi, M., Risso, A., Rodrigues, R., Roitman, A., Rorato, G., Royer, S., Ruhoff, A., Ruoso, E., Sampaio, M., Schabbach, L., Schiaffino, M., Schmitt, H., Schumacher, R., Schwarzer, G., Serrano, N., Sigallis, A., Silva, M., Silva, S., Sluter, C., Soares, L., Soares, V., Sousa, L., Souza, A., Tschiedel, A., Ucha, L., Umbelino, G., Utzig, E., Zambrano, F., 2024. Banco de dados das cheias na Região Hidrográfica do Lago Guaíba em Maio de 2024.
Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. Remote Sens. Environ. 48, 119–126. https://doi.org/10.1016/0034-4257(94)90134-1.
Quevedo, R.P., Guasselli, L.A., Oliveira, G.G. De, Ruiz, L.F.C., 2019. Modelagem de áreas suscetíveis a movimentos de massa: avaliação comparativa de técnicas de amostragem, aprendizado de máquina e modelos digitais de elevação. Geociências UNESP 38, 781–795. https://doi.org/10.5016/geociencias.v38i3.14019.
Quevedo, R.P., Oliveira, G.G. De, Guasselli, L.A., 2020. Mapeamento de Suscetibilidade a Movimentos de Massa a partir de Redes Neurais Artificiais. Anuário do Instituto de Geociências - UFRJ 43, 128–138. https://doi.org/10.11137/2020_2_128_138.
Quevedo, R.P., Anderson, L.O., Horta, I.T.L.G., Velástegui-Montoya, A., Veiga, R.Q., Cardozo, C.P., Sparrow, S., 2023a. The relationship between landslide occurrence and land use and land cover. In: Anais Do XX Simpósio Brasileiro de Sensoriamento Remoto, pp. 1564–1567. Florianópolis.
Quevedo, R.P., Velastegui-Montoya, A., Montalván-Burbano, N., Morante-Carballo, F., Korup, O., Daleles Rennó, C., 2023b. Land use and land cover as a conditioning factor in landslide susceptibility: a literature review. Landslides. https://doi.org/10.1007/s10346-022-02020-4.
Quevedo, R.P., Oliveira, G.G., Guasselli, L.A., 2024. Landslide Inventory - Rolante River Basin. Brazil.
R Core Team, 2014. R: A Language and Environment for Statistical Computing.
Rabby, Y.W., Li, Y., Abedin, J., Sabrina, S., 2022. Impact of land use/land cover change on landslide susceptibility in Rangamati municipality of Rangamati district, Bangladesh. ISPRS Int. J. Geo-Inf. 11 https://doi.org/10.3390/ijgi11020089.
Reichenbach, P., Busca, C., Mondini, A.C., Rossi, M., 2014. The influence of land use change on landslide susceptibility zonation: the briga catchment test site (Messina, Italy). Environ. Manag. 54, 1372–1384. https://doi.org/10.1007/s00267-014-0357-0.
Reis, M.S., 2022. Detection and analysis of forest regeneration trajectories in the lower Tapajós region. Doutorado Em Ciência Do Sistema Terrestre). Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos.

Reis, M.S., Pantaleao, E., Dutra, L.V., Sant'Anna, S.J.S., Escada, M.I.S., 2019. Effects of different methods of radiometric calibration on the use of training data for supervised classification of Landsat5/TM images from other dates. In: IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 1566–1569. https://doi.org/10.1109/IGARSS.2019.8898150.

Reis, M.S., Dutra, L.V., Escada, M.I.S., Sant'Anna, S.J.S., 2020. Avoiding invalid transitions in land cover trajectory classification with a compound maximum a posteriori approach. IEEE Access 8, 98787–98799. https://doi.org/10.1109/ACCESS.2020.2997019.

Riffel, E.S., Guasselli, L.A., Ruiz, L.F.C., Gameiro, S., 2021. Relação entre ponto de ruptura e padrão morfométrico em deslizamentos, bacia hidrográfica do Rio Rolante - RS. Revista do Departamento de Geografia 41, 1–16. https://doi.org/10.11606/eISSN.2236-2878.rdg.2021.181554.

Riley, S.J., DeGloria, S.D., Elliot, Robert, 1999. A terrain ruggedness index that quantifies topographic heterogeneity. Intermt. J. Sci. 5, 1–4.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J. Photogrammetry Remote Sens. 67, 93–104. https://doi.org/10.1016/j.isprsjprs.2011.11.002.

Rossato, M.S., 2020. Os climas do Rio Grande do Sul: uma proposta de classificação climática. ENTRE-LUGAR 11, 57–85. https://doi.org/10.30612/el.v11i22.12781.

Rouse, Jr.J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1973. Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation.

Sangeeta, Singh, S.K., 2023. Influence of anthropogenic activities on landslide susceptibility: a case study in Solan district, Himachal Pradesh, India. J. Mt. Sci. 20, 429–447. https://doi.org/10.1007/s11629-022-7593-1.

SEMA, 2017. Diagnóstico Preliminar: descritivo dos eventos ocorridos no dia 5 de janeiro de 2017 entre as regiões dos municípios de São Francisco de paula e Rolante/rs, pp. 1–26.

Sharma, R., Rimal, B., Baral, H., Nehren, U., Paudyal, K., Sharma, S., Rijal, S., Ranpal, S., Acharya, R., Alenazy, A., Kandel, P., 2019. Impact of land cover change on ecosystem services in a tropical forested landscape. Resources 8, 18. https://doi.org/10.3390/resources8010018.

Souza, C.M., Shimbo, J.Z., Rosa, M.R., Parente, L.L., Alencar, A.A., Rudorff, B.F.T., Hasenack, H., Matsumoto, M., Ferreira, L.G., Souza-Filho, P.W.M., de Oliveira, S. W., Rocha, W.F., Fonseca, A.V., Marques, C.B., Diniz, C.G., Costa, D., Monteiro, D., Rosa, E.R., Vélez-Martin, E., Weber, E.J., Lenti, F.E.B., Paternost, F.F., Pareyn, F.G.C., Siqueira, J.V., Viera, J.L., Neto, L.C.F., Saraiva, M.M., Sales, M.H., Salgado, M.P.G., Vasconcelos, R., Galano, S., Mesquita, V.V., Azevedo, T., 2020. Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine. Rem. Sens. 12 https://doi.org/10.3390/RS12172735.

Su, T., Zhang, S., Liu, T., 2020. Multi-spectral image classification based on an object-based active learning approach. Rem. Sens. 12, 504. https://doi.org/10.3390/rs12030504.

Sun, L., Liu, Q., Abdelaziz, A., Tang, X., Grasselli, G., 2022. Simulating the entire progressive failure process of rock slopes using the combined finite-discrete element method. Comput. Geotech. 141, 104557 https://doi.org/10.1016/j.compgeo.2021.104557.

Tarantino, C., Blonda, P., Pasquariello, G., 2007. Remote sensed data for automatic detection of land-use changes due to human activity in support to landslide studies. Nat. Hazards 41, 245–267. https://doi.org/10.1007/s11069-006-9041-x.

Tavares, P.A., Ferreira, J., Silva, C.V.J., Berenguer, E., Barlow, J., 2022. Exploring the role of deforestation and cropland expansion in driving a fire-transition in the Brazilian Amazon. Land 11, 2274. https://doi.org/10.3390/land11122274.

Tewkesbury, A.P., Comber, A.J., Tate, N.J., Lamb, A., Fisher, P.F., 2015. A critical synthesis of remotely sensed optical image change detection techniques. Remote Sens. Environ. 160, 1–14. https://doi.org/10.1016/j.rse.2015.01.006.

Thiam, S., Salas, E.A.L., Hounguè, N.R., Almoradie, A.D.S., Verleysdonk, S., Adounkpe, J.G., Komi, K., 2022. Modelling land use and land cover in the transboundary mono River catchment of Togo and Benin using Markov chain and stakeholder's perspectives. Sustainability 14, 4160. https://doi.org/10.3390/su14074160.

Traoré, F., Cornet, Y., Denis, A., Wellens, J., Tychon, B., 2013. Monitoring the evolution of irrigated areas with Landsat images using backward and forward change detection analysis in the Kou watershed, Burkina Faso. Geocarto Int. 28, 733–752. https://doi.org/10.1080/10106049.2012.744100.

Uehara, T.D.T., Soares, A.R., Quevedo, R.P., Körting, 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: 2020 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2020).

Uehara, T.D.T., Sehn Körting, T., Soares, A., dos, R., Quevedo, R.P., 2022. Time-series metrics applied to land use and land cover mapping with focus on landslide detection. J. Appl. Remote Sens. 16 https://doi.org/10.1117/1.JRS.16.034518.

Van Den Eeckhaut, M., Poesen, J., Vandekerckhove, L., Van Gils, M., Van Rompaey, A., 2010. Human–environment interactions in residential areas susceptible to landsliding: the Flemish Ardennes case study. Area 42, 339–358. https://doi.org/10.1111/j.1475-4762.2009.00919.x.

Vuillez, C., Tonini, M., Sudmeier-Rieux, K., Devkota, S., Derron, M.H., Jaboyedoff, M., 2018. Land use changes, landslides and roads in the Phewa Watershed, Western Nepal from 1979 to 2016. Appl. Geogr. 94, 30–40. https://doi.org/10.1016/j.apgeog.2018.03.003.

Wang, H., Liu, C., Zang, F., Yang, J., Li, N., Rong, Z., Zhao, C., 2020. Impacts of topography on the land cover classification in the Qilian Mountains, Northwest China. Can. J. Rem. Sens. 46, 344–359. https://doi.org/10.1080/07038992.2020.1801401.

Wasowski, J., 1998. Understanding rainfall-landslide relationships in man-modified environments: a case-history from Caramanico Terme, Italy. Environ. Geol. 35, 197–209. https://doi.org/10.1007/s002540050306.

Winkler, K., Fuchs, R., Rounsevell, M., Herold, M., 2021. Global land use changes are four times greater than previously estimated. Nat. Commun. 12, 2501. https://doi.org/10.1038/s41467-021-22702-2.

Wulder, M.A., Loveland, T.R., Roy, D.P., Crawford, C.J., Masek, J.G., Woodcock, C.E., Allen, R.G., Anderson, M.C., Belward, A.S., Cohen, W.B., Dwyer, J., Erb, A., Gao, F., Griffiths, P., Helder, D., Hermosilla, T., Hipple, J.D., Hostert, P., Hughes, M.J., Huntington, J., Johnson, D.M., Kennedy, R., Kilic, A., Li, Z., Lymburner, L., McCorkel, J., Pahlevan, N., Scambos, T.A., Schaaf, C., Schott, J.R., Sheng, Y., Storey, J., Vermote, E., Vogelmann, J., White, J.C., Wynne, R.H., Zhu, Z., 2019. Current status of Landsat program, science, and applications. Remote Sens. Environ. 225, 127–147. https://doi.org/10.1016/j.rse.2019.02.015.

Xu, H., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. Int. J. Rem. Sens. 27, 3025–3033. https://doi.org/10.1080/01431160600589179.

Zeng, T., Guo, Z., Wang, L., Jin, B., Wu, F., Guo, R., 2023. Tempo-spatial landslide susceptibility assessment from the perspective of human engineering activity. Rem. Sens. 15, 4111. https://doi.org/10.3390/rs15164111.

Zha, Y., Gao, J., Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. Int. J. Rem. Sens. 24, 583–594. https://doi.org/10.1080/01431160304987.

Zhu, Z., 2017. Change detection using landsat time series: a review of frequencies, preprocessing, algorithms, and applications. ISPRS J. Photogrammetry Remote Sens. 130, 370–384. https://doi.org/10.1016/j.isprsjprs.2017.06.013.