Pattern Recognition Letters 148 (2021) 54–60

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Pattern Recognition and Remote Sensing techniques applied to Land Use and Land Cover mapping in the Brazilian Savannah

Leila M.G. Fonseca[a,*], Thales S. Körting[a], Hugo do N. Bendini[a], Cesare D. Girolamo-Neto[a], Alana K. Neves[a], Anderson R. Soares[a], Evandro C. Taquary[a], Raian V. Maretto[b]

[a] *Earth Observation and Geoinformatics Division - National Institute for Space Research (INPE), Avenida dos Astronautas, 1758, São Jos dos Campos, 12227-010, Brazil*
[b] *Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Hengelosestraat 99 7514AE, Enschede, The Netherlands*

ARTICLE INFO

Article history:
Received 5 March 2020
Revised 29 March 2021
Accepted 22 April 2021
Available online 8 May 2021

Keywords:
Earth Observation
Agriculture mapping
Cerrado
Time series
Deep Learning
Pastureland
Deforestation
Remote sensing

ABSTRACT

The Brazilian Savannah, or Cerrado, has gained vital importance in the discussions about sustainable land development after the conversion of half of its natural vegetation. For the last two decades, most of the agricultural expansion in Brazil has occurred in this biome. This is related to technological improvements in agriculture as well as to environmental compliance policies that have effectively reduced soybean expansion in the Brazilian Amazon biome. Therefore, remotely sensed imagery, pattern recognition and image processing techniques have been employed to analyze and monitor the land dynamics over Cerrado. In this work, we present a brief review on Land Use and Land Cover mapping (LULC) in the Cerrado biome from an application perspective: natural vegetation, pastureland, agriculture, and deforestation. In this review we selected some studies whose results could contribute to the development of more detailed and accurate LULC maps for the Cerrado biome.

1. Introduction

The Brazilian Savannah, known as Cerrado, is the second largest Brazilian biome and encompasses an area of approximately 2 million km^2 in the Brazilian territory (Fig. 1). The Cerrado is a relevant area for the biodiversity conservation and sustainable development in Brazil. Taking into account its full extent, the Cerrado biome is responsible for storing around 5.9 billion tons of carbon in the vegetation and 23.8 billion tons in the soil [72].

In just about 20 years, due to the expansion of the agricultural frontier in this region, Brazil has become one of the largest agricultural producers and exporters. The challenges to preserve the Cerrado biome and promote sustainable development are innumerable and involve actions such as land use planning, expansion of agricultural production through the increase of productivity and the integration of sustainable systems, and mainly the reduction of pressure due to the expansion of the border by the conversion of new natural areas.

Nevertheless, Cerrado lost over 50% of its natural vegetation mainly due to agricultural expansion [32]. Moreover, there is a projection that 31-34% of the remaining biome is likely to be cleared by 2050 [69]. Therefore, the Land Use an Land Cover (LULC) mapping and the monitoring of vegetation remnants are important to understand the process of land occupation and thus elaborate and implement public policies for Cerrado preservation and sustainable development.

In this context, remote sensing technologies have been crucial to ensure the effectiveness of environmental monitoring and now, new challenges have risen which consider new patterns of land conversion. TerraClass, project developed by the National Institute for Space Research (INPE), identifies the land use for those regions in which the natural vegetation was suppressed. Initially, it focused on the Legal Amazon deforestation [2] and later on Cerrado [31]. MapBiomas project [67] has carried out the annual fully automatic mapping for all Brazilian biomes using Landsat imagery. Both projects have been contributing to public policies in Brazil, however they still do not deal with more detailed LULC patterns.

This scenario underlines the need for new methods based on pattern recognition techniques in order to accurately map the land use and natural vegetation cover to ensure the Cerrado sustainable development and preservation. Due to the complexity of the Cerrado biome, increasingly advanced methods have been applied to study and monitor the changes in its vegetation. They evolved together with remote sensing (RS) and geoinformatics technologies and include traditional supervised and non-supervised classi-

* Corresponding author.
E-mail address: leila.fonseca@inpe.br (L.M.G. Fonseca).

https://doi.org/10.1016/j.patrec.2021.04.028

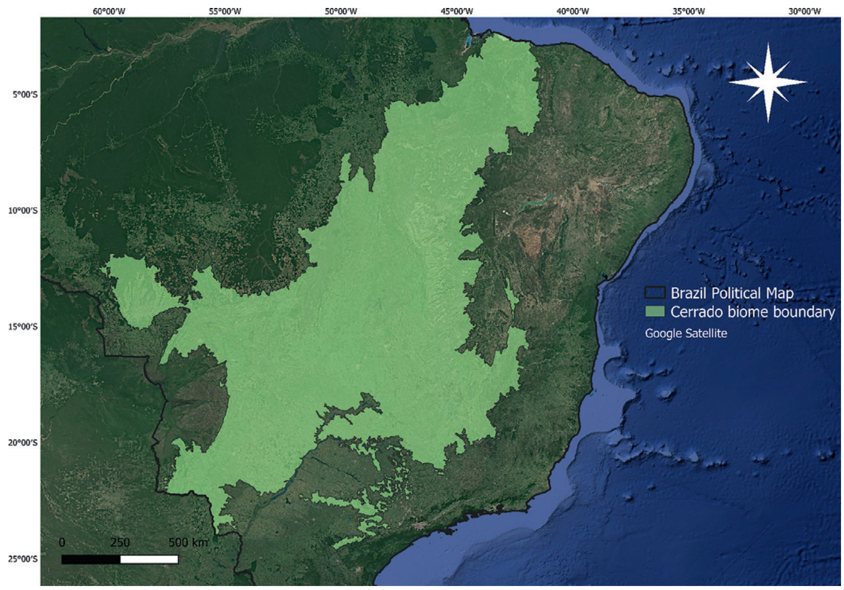

Fig. 1. Cerrado biome in Brazil.

fication methods, such as pixel-based classification algorithms, like Maximum Likelihood Classification (MLC) and Spectral Angle Mapper (SAM), applied to one single scene [66], object-based methods [9], and moreover, integrating multi-temporal classification based on Satellite Image Time Series (SITS) [52]. We have also observed a rapid advance in open RS technologies such as the launch of Sentinel [71], CBERS (Chinese Brazilian Earth Resources Satellite) [22], and Amazonia-1 satellite[1], the increase of CubeSat compliant nano-satellites constellations, developments of the infrastructure for remotely sensed data storage and dissemination, such as the Analysis Ready Data (ARD) [54], and also the evolution of machine learning techniques based on satellite imagery [36].

In this context, it is important to continuously review and reflect upon recent developments on LULC mapping based on RS, and also look ahead to new advances and opportunities to improve the LULC maps. Therefore, we present a brief review on Land Use and Land Cover mapping (LULC) in the Cerrado biome from the application perspective: natural vegetation, pastureland, agriculture, and deforestation. In this review, we selected some studies whose results could contribute to the development of more detailed and accurate LULC maps for the Cerrado biome. This survey focus on methods based on optical data although we understand that radar systems can contribute to this matter. This review is organized as follows. In Sections 2–5, we briefly review some studies for mapping natural vegetation cover, pastureland, agriculture, and deforestation, respectively. In Section 6, we discuss the main key points that we would consider to develop new methods capable of producing high accuracy and detailed LULC maps for the Cerrado biome and also present a comparative summary of some studies we have analyzed in this review.

2. Natural vegetation

Large-scale mapping of the Cerrado vegetation using RS technologies is still a challenge due to the high spatial variability and spectral similarity among its vegetation types (phytophysiognomies). [55] proposed a classification legend for Cerrado vegetation with 25 phytophysiognomies classified according to horizontal and vertical vegetation patterns observed in situ (Fig. 2). The three major classes are grassland, savannah and forest. However, this system follows a hierarchical structure, e.g., Grassland is divided into open grasslands, shrub grasslands or rupestrian grasslands.

Several authors have integrated RS technologies and machine learning in order to map Cerrado vegetation. The use of spectral reflectance information alone was not suitable to improve discrimination among various types of vegetation in the Cerrado and so other methods have been adopted [20,24]. Ferreira et al. [20] evaluated how the linear spectral un-mixing method contributed to the classification of 5 phytophysiognomies using a Landsat-7 image. The authors performed an automatic classification using the Mahalanobis distance. The misclassification errors and the recall value of 75% for shrub grasslands and open savannah were highly associated with transition areas between these two classes and the alternation between shrubs and small trees.

In order to tackle this problem, [26] proposed a method based on texture features derived from the Gray Level Co-occurrence Matrix (GLCM). Texture features like Entropy and Contrast were able to capture the smooth increase on arboreal elements percentage from classes of shrub grasslands, shrub savannah, typical savannah and wooded savannah. The classification accuracy of eleven phytophysiognomies, using WorldView-2 images, spectral unmixing components, and the Random Forest (RF) algorithm [12] was of 67%. When the authors added texture features, the accuracy increased to 75% and the misclassification for all classes decreased.

Schwieder et al. [63] used Tasseled Cap components of greenness, brightness and wetness [34] derived from Landsat-5 and Landsat-7 time series. Seven phytophysiognomies were classified using the Support Vector Machines (SVM) algorithm [50] and the overall accuracy was of 63%. Transition areas between two vegetation types were still a source of misclassification similar to what [20] achieved.

In addition to the hierarchy, another important aspect in the Cerrado vegetation classification is the context in which each physiognomy occurs, since the [55] legend is based on in situ observations. For example, the existence of some phytophysiognomy pattern relies on the proximity of water bodies and the presence of certain vegetation species or rocky outcrops. Thus, a per pixel classification is not able to correctly identify the context patterns. In order to compare the hierarchical and non-hierarchical approaches and consider the context in the classification, [44] employed a RF

[1] http://www.inpe.br/amazonia1

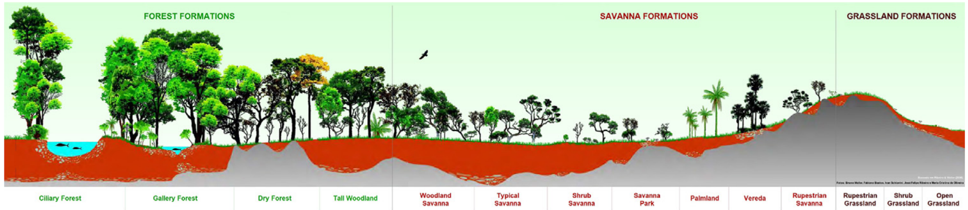

Fig. 2. Physiognomies according to the Ribeiro and Walter classification system, represented in a biomass gradient (growing from right to left).

algorithm and GEOBIA (Geographic Object-based Image Analysis) techniques to classify 7 phytophysiognomies. A Super pixel segmentation method, called SLIC (Simple Linear Iterative Clustering) [1], was used in order to isolate spectrally homogeneous regions, and thus create meaningful objects of similar sizes. The hierarchical and non-hierarchical approaches overall accuracies were of 68.95% and 66.40%, respectively. When analyzing the phytophysiognomies individually, six of seven classes showed higher accuracies in the hierarchical approach.

Although these approaches have shown useful in the individual case studies, they were usually restricted to comparably small extents and did not account for variations in environmental characteristics that exist across the entire Cerrado extent. The development of robust mapping approaches that account for these limitations is part of a project supported by the Forest Investment Program (FIP) aiming to promote sustainable land use and forest management. Bendini et al. [8] analyzed the potential of Landsat analysis-ready-data (ARD) in combination with different environmental data (terrain data, soil data and vertical distance to the nearest drainage) to classify Forest, Savannah and Grassland with an overall accuracy of 86.04%. On the other hand, the classification of 12 vegetation types achieved an overall accuracy of 76.63%.

In recent years, Deep Learning (DL) methods have thrived in the RS field [36]. For mapping Cerrado vegetation, [48] used DL-based method to discriminate the three major types of vegetation (Forest, Savannah and Grassland). The authors classified patches of Landsat images that were entirely designated as one of the three classes. Semantic segmentation, i.e., the assignment of a separate class per pixel, was not performed and a considerable mixture of classes in a single patch could be observed.

The semantic segmentation was performed by Neves et al. [45] to identify Forest, Savannah and Grassland in a Cerrado conservation unit. They used an adapted U-net architecture and achieved an overall accuracy of 87% [56]. Neves [47] adapted the U-net architecture to hierarchically classify the vegetation physiognomies. Superpixel centroids were randomly selected and used as sample patches center to generate the training dataset. RGB bands combined with 2-band Enhanced Vegetation Index (EVI2) were used as input data. For the first and second hierarchical levels the overall accuracy were 92.8% and 86%, respectively.

3. Pastureland

In recent years, over 500,000 km2 of the Cerrado natural vegetation have been converted into cultivated pastures [51]. Sano et al. [59] and Scaramuzza et al. [61] carried out pasture mapping in Cerrado. Almeida et al. [2] mapped four different pasture cover types in the Amazon region: herbaceous pasture, shrubby pasture, pasture with bare soil and regeneration with pasture. They applied methodologies based on image segmentation and visual interpretation, which not only takes a considerable amount of time but also relies on the interpreters experience.

In order to improve the mapping quality as well as to reduce the time spent on visual interpretation, time series techniques have been employed to map pastures in Brazil. Müller et al. [42] used time series of Landsat-5 and Landsat-7 images in order to map pastures, agriculture, savannah, forest, water, and non-vegetated lands in the state of Mato Grosso, Brazil. A Normalized Difference Vegetation Index (NDVI) time series and the Random Forest (RF) algorithm were applied to perform a pixel wise classification. Using several statistical metrics from time series, the authors obtained an overall accuracy of approximately 93%. Parente et al. [51] adopted a similar methodology based on a pixel-wise classification to map pasture areas in the whole Brazilian territory. They used NDVI time series and also RF algorithm in a cloud computing platform (Google Earth Engine) to annually map Brazilian pastures, from 1985 to 2017, with a classification accuracy of about 87%.

Neves et al. [43] applied GEOBIA to classify pasture, forest and agriculture in the state of Mato Grosso. Using time series of Enhanced Vegetation Index 2 (EVI2) from MODIS images, polar time series metrics [35], and RF classifier, the authors compared a pixel-wise classification with GEOBIA approach. Considering the pasturelands, GEOBIA techniques outperformed the pixel wise classification with an accuracy of 92% and 75%, respectively. Costa et al. [15] also used several time series metrics obtained from MODIS images to classify vegetation, pastures and other land uses in the state of Minas Gerais, and the overall accuracy was 85%.

Despite all efforts, automatic methods for discriminating different types of pasture still pose a challenge. To classify pasture lands, [46] used GEOBIA and Landsat 8 surface reflectance image time series of three path/rows from 2013 to 2015. The methodology was hierarchically performed in two steps: in the first step, all pasture lands were detected and differentiated from other targets; in the second step, pasture was differentiated between Herbaceous Pasture and Shrubby Pasture. The following features were used in the classification: vegetation indices, fractions of the Linear Spectral Mixture Model, components of the Tasseled Cap Transformation, other spectral attributes and texture features. In the first step, the overall accuracy was up to 87.76%, while, in the second step, it was up to 73.15%.

Girolamo-Neto et al. [25] verified that image texture when used with other features extracted from Landsat-8 and Sentinel-2 can improve the discrimination between herbaceous pasture and shrubby pasture in the Brazilian Amazon and Cerrado biomes, respectively. Besides, the use of texture information allowed identifying the first signs of pasture degradation, such as the presence of trees and bare soil patches on shrubby pastures. On the other hand, [17] evaluated spectral unmixing components [64] to identify degraded pastures. High values of soil and vegetation components were obtained for degraded pasture and cultivated pastures, respectively, demonstrating their importance to differentiate both types of pasture. Rufin et al. [58] evaluated the Tasseled Cap transformation [34] to discriminate herbaceous pasture from shrubby pasture in the Amazon, which was able to identify the transition between both pastures. Even though various automatic methodologies have been proposed to identify different types of pasture, none of the previous studies have shown better results than those obtained by Almeida et al. [2].

4. Agriculture

The paradox between environmental conservation and economic development is a challenge in Brazil, where there is a complex and dynamic agricultural scenario. This reinforces the need for effective methods for detailed agricultural mapping [7]. RS data is an important solution for agricultural mapping, once it captures the seasonal vegetation behavior. Lower spatial resolution sensors (250m - 1 km) with up to daily revisit times have been widely used for agricultural mapping in the Cerrado biome [4,10,13,19,53,57]. Although low spatial resolution images have been sufficient for mapping large-scale agricultural practices, such as double and single cropping systems, they do not allow the detection of small fields due to the spectral mixture of different targets, which includes crops with significant importance for family farming.

Most LULC mapping initiatives [2,67], particularly agriculture, have been applied at low level of thematic detail as a broad class for annual agriculture, which does not take into account cropping practices and rotation systems patterns. Advances in computational processing performance and development of consistent methods to integrate data of different sensors have improved the LULC mapping approaches [5,6,42,58,63]. However, high temporal resolution is necessary due to the high cloud cover as well as the variability in the agricultural calendar that produces different crop seasons within the same year and different cultivation practices (e.g., no tillage, center pivot irrigation, crop-livestock integration and the use of early varieties). In this case, the integration of Sentinel, Landsat, CBERS imagery is a potential solution to increment the revisit frequency. However, great processing efforts to be made to perform spectral, geometric and radiometric imagery normalization, and also to correct the Landsat-7 scan Line Corrector (SLC)-off data [6,40].

Besides that, phenological metrics (phenometrics) extracted from image time series [33] are valuable features to improve agricultural mapping methods. Phenometrics have been explored for this purpose, using MODIS time series [10] and recently also with Landsat-like images [5,6,42].

In this context, [7] employed for the first time land surface phenological metrics derived from dense satellite image time series to classify agricultural land in the Cerrado biome. They used all available Landsat images between April 2013 and April 2017, applying a weighted ensemble of Radial Basis Function (RBF) convolution filters as a kernel smoother to fill data gaps such as cloud cover and Scan Line Corrector (SLC)-off data. Through this approach, they created a dense Enhanced Vegetation Index (EVI) data cube with an 8-day temporal resolution and derived phenometrics for a RF classification using different thematic levels in a hierarchical approach. This work showed that phenometrics derived from dense Landsat image time series were able to describe complex agricultural phenological patterns in the Brazilian Cerrado, and to map the distribution of crops over this area with an accuracy greater than 90%. With these results, the authors opened a perspective of integrating features from multiple crop seasons showing that phenometrics combined with hierarchical classification approaches allows mapping agriculture with a higher thematic detail. Furthermore, in the agricultural domain, hierarchical approach enables the evaluation of cropping practices and crop types regardless of the use of a priori cropland mask. This methodology was already used for subsiding a public policy concerning to water use assessment in the Cerrado biome by the Brazilian Water Agency [3].

5. Deforestation

Producing accurate deforestation maps is a crucial step for providing information to enable public policies for preventing and monitoring deforestation. INPE has been estimating deforestation rates in the Brazilian Amazon biome since 1988, through the PRODES project, on an annual basis [29]. PRODES have been used together with the Near Real-time Deforestation Detection System (DETER), which has played an important role in the reduction of deforestation rates in the Brazilian Amazon biome in the early 2000s [11]. Considered as the main references on large-scale accurate mapping of deforestation in tropical forests [14], PRODES and DETER methodologies have been adapted to be applied to the Cerrado biome, generating a very consistent temporal series of natural vegetation suppression also for this biome, since 2018 [30].

Deforestation data provided by PRODES and DETER constitute a powerful information to support the development of automatic methods to map deforested areas. Several initiatives have been launched to automate this process, such as the Global Land Analysis and Discovery (GLAD), developed by Global Forest Watch [28], the Deforestation Alert System (SAD), developed by Imazon [68], and the SIAD, developed by LAPIG [27]. To assess state-of-art pattern recognition methods, [49] evaluated three deep learning techniques for automatic deforestation detection in the Brazilian Amazon and Cerrado biomes. The authors used two Landsat 8 images acquired at different dates. The strategies based on Deep Learning achieved the best performance in comparison with other methods and achieved an overall accuracy up to 95% for Cerrado. Following a similar change detection approach, [38] proposed two variations of the U-Net to map deforestation in the Brazilian Amazon biome, using PRODES maps as training data and Landsat 8 imagery. Their methodology was applied to an area in the Southestern of the state of Pará, achieving approximately 95% of overall accuracy. [37] extended this method for large-scale areas, applying it for the entire state of Pará, which comprises approximately 1.25 km^2, with an overall accuracy of approximately 94%. Furthermore, the author adapted the method for applying to a region comprising approximately 130000 km^2 in the east of the Cerrado biome, in the borders among the states of Bahia, Minas Gerais, Goiás and Tocantins, where an overall accuracy of approximately 92% was achieved.

On the other hand, [70] developed a new method based on DL strategy using high spatiotemporal resolution imagery, provided by the PlanetScope constellation of CubeSat compliant nanosatellites. The author implemented a hybrid approach combining two types of artificial neural networks: a Recurrent Neural Network (LSTM) and a Convolutional Neural Network (U-Net). The author produced a deforestation map for an area of 18000 km^2 of the Cerrado and a period span from August 2017 to August 2018. The results achieved an overall accuracy of 99.92%.

However, despite the promising results, these methods need further improvements to compose a trustworthy fully automated approach. Methods that consider short-term temporal contexts, such as the ones proposed by [38] are still susceptible to errors due to the occurrence of clouds in one of the considered timestamps, as well as to false positives because of the vegetation seasonal behavior in the Cerrado biome. Likewise, making use of dense time-series along with spatial convolutions constitutes a big challenge when it comes to mapping wider areas as it demands a great effort to gather training samples for both spatial and temporal models. Therefore, the main challenge is to come up with a feasible way to get the most of both spatial and temporal information.

6. Final remarks and future opportunities

Performing LULC mapping in the Cerrado biome is a challenge due to its complex vegetation cover, differently from the Amazon biome. Different strategies for LULC mapping in the Cerrado have been proposed and three relevant contributions can be pointed out: 1) Time series analysis and phenological metrics arise as valu-

Table 1
Summary of contributions and limitations of selected studies.

	Author	Input Data	Techniques	Contributions	Limitations
Natural Vegetation	[47]	WorldView-2 (2 m)	Deep Learning and GEOBIA	Hierarchical semantic segmentation; Differentiate 9 vegetation types	Small study area due to lack of reference data
	[20]	Landsat (30 m)	Mahalanobis distance	Differentiate 5 vegetation types	Confusion between Shrub Cerrado and Wooded Cerrado
	[8]	Landsat (30 m) and Environmental data	Random Forest	Differentiation of 12 vegetation types and combination of environmental data in the data cube	Limitations on mapping classes related to savanna gradients and also classes driven by spatial context
	[26]	Landsat-8 (30 m) and WorldView-2 (2 m)	Random Forest, GEOBIA and Texture Analysis	Use of GLCM entropy; Differentiate 10 vegetation types	Require the use of squared segments which mix many classes into a single object
Pastureland	[51]	Landsat (30 m) time series	Random Forest	Large-scale mapping for pastureland	Low level of semantic detail; Salt-pepper noise due to cloud cover; Artifacts in the scene transitions
	[46]	Landsat-8 (30 m) time series	Random Forest and GEOBIA	Hierarchical approach; Time series and GEOBIA improved discrimination between Shrubby Pasture and Herbaceous Pasture	Segmentation considered only dry season which may not represent the entire time series
	[25]	Sentinel-2 (20 m) time series	Random Forest, GEOBIA and Texture Analysis	Entropy and Contrast (GLCM) captured the difference between shrub and herbaceous pastures	Concentration of shrub-arboreal elements on small portions of herbaceous pasture caused classification errors
Agriculture	[7]	Landsat (30 m) dense time series	Random Forest	Use of phenometrics extracted from dense time series of medium spatial resolution satellites ; Hierarchical strategy; Higher level of detail	Time series still present noise after filtering, in areas with a higher incidence of clouds; Limitations on the second season detection
	[4]	MODIS (250 m) time series	Maximum Likelihood	Use of time series metrics; High accuracies for crop masking	Low accuracy for mapping crop types (confusion with non- commercial crops); Low spatial resolution images
	[35]	MODIS (250 m) time series	Decision Trees	Use of polar metrics from time series	Low spatial resolution; Lack of semantic detail
Deforestation	[40]	Landsat-8 OLI (30 m)	Multitemporal U-Net variation (Deep Learning)	Use of semantic segmentation adapted to multitemporal images	Cloud cover is a critical issue
	[70]	PlanetScope (3 m)	Hybrid Archi- tecture (Deep Learning)	Combine spatial context and time series analysis	Require big amount of training data
	[49]	Landsat-8 OLI(30 m)	Deep Learning and SVM	Compared two Deep Learning strategies to SVM	Require big amount of training data; small study area
	[28]	Landsat-8 OLI (30m)	Decision trees over time series and spectral features	Fully operational	Limitations for mapping vegetation of low and medium height

able assets to perform LULC mapping, particularly for agriculture; 2) Hierarchical classification approaches are an efficient strategy to deal with classification problems when involving a high number of classes, making it possible to achieve higher accuracies in comparison with those of the classical approach; 3) Contextual information is a key factor to accurately classify the natural vegetation and deforestation patterns in the Cerrado biome. Table 1 shows a comparative summary of some studies discussed in our review, pointing out their main contributions and limitations.

Dense image data cubes composed of long-term SITS can improve extraction of LULC data and provide a better understanding of landscape dynamics [73]. In this context, numerous projects have been developed to produce Earth Observation Data Cubes such as the Australian Geoscience Data Cube (AGDC) [18] and the Framework for Operational Radiometric Correction for Environmental monitoring (FORCE) [23]. INPE researchers are currently developing the Brazilian Data Cube (BDC) from medium spatial resolution images, including Landsat, CBERS and Sentinel satellites imagery [21].

Schultz et al. [62] demonstrated that it is possible to improve the performance of LULC mapping through the combination of spatial and temporal heterogeneity and correlation. Thus, it is a suitable solution to incorporate contextual information and dense satellite image time series to perform LULC mapping. For such a task, we believe that methods that incorporate spatial context such as multi-temporal segmentation, GEOBIA and semantic segmentation must be explored. Recently, [16] proposed a multi-temporal image segmentation method, combining the well-known region growing approach and the Dynamic Time Warping to produce homogeneous regions in space and time. Since the context obtained by spatiotemporal segmentation includes more than a single time series, we can avoid the strong influence of noise by suppressing common artifacts in SITS classification. Recently, [65] developed the Stmetrics package Python3[2], which provides an environment to extract state-of-the-art time-series features.

As presented in our previous discussion, new perspectives came up with the development of DL approaches. To make these methods more accessible for RS users, [39] developed the DeepGeo[3] framework which is distributed as a free and open source package. The system consists on a package for Python programming language, which provides a high level and user friendly API (Application Programming Interface). Besides, it provides tools to easily scale up analysis to process large amount of data while remaining flexible and easily extensible.

Several research groups have been implemented their own systems for analyzing RS data using friendly development frameworks. In this sense, [35] developed the Geographic Data Mining Analyst (GeoDMA)[4], a framework that integrates image segmentation, feature extraction and selection, landscape and multitemporal features and data mining, allowing pattern recognition tasks and multi-temporal analysis. But the RS community still demands for frameworks to easily perform LULC analysis, considering that most of remote sensing specialists do not have skills on computing. Therefore, most of them have utilized different tools to run all processes necessary for mapping LULC. The integration of all these tools in a framework in order to provide a friendly environment for processing and analyzing a large amount of data has been developed by INPE researchers [21].

Finally, it is important to mention that Synthetic Aperture Radar (SAR) images can be used together with optical images to improve the Cerrado vegetation mapping [41,60]. Cerrado vegetation exhibits a clear biomass gradient throughout different phytophysiognomies and SAR systems are able to detect different degrees of surface roughness. Therefore, we are also investigating SAR and LIDAR GEDI data (Global Ecosystem Dynamics Investigation) as biomass proxies for Cerrado.

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.

Acknowledgments

This work is supported by the Amazon Fund in collaboration with the Brazilian Development Bank (BNDES) (No. 17.2.0536.1), the Brazilian National Research Council (CNPq) (grants #140372/2017-2, #303360/2019-4), the Coordination for the Improvement of Higher Education Personnel (CAPES) (Finance Code 001), and also the project Development of systems to prevent forest fires and monitor vegetation cover in the Brazilian Cerrado (Forest Investment Program (FIP), World Bank (#P143185), Ministry of Science, Technology and Innovation).

References

[1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk, SLIC superpixels compared to state-of-the-art superpixel methods, IEEE Trans. Pattern Anal. Mach.Intell. 34 (11) (2012) 2274–2282.
[2] Almeida, et al., 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 (3) (2016) 291–302.
[3] A.N.d.e.S.B.B. ANA, Atlas Irrigação: 2ª Edição Uso da Água na Agricultura, Technical Report CDU 631.67(084.4), Agência Nacional de Águas e Saneamento Básico (Brasil), Brasilia, DF, 2021.
[4] Arvor, et al., Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil, Int. J. Remote Sens. 32 (22) (2011) 7847–7871.
[5] Bendini, et al., Using landsat 8 image time series for crop mapping in a region of Cerrado, Brazil., Embrapa Meio Ambiente-Artigo em anais de congresso (ALICE), International Archives of the Photogrammetry Remote Sensing and Spatial, 2016.
[6] Bendini, et al., Assessment of a multi-sensor approach for noise removal on landsat-8 OLI time series using CBERS-4 MUX data to improve crop classification based on phenological features, Revista Brasileira de Cartografia 69 (5) (2017).
[7] Bendini, et al., Detailed agricultural land classification in the Brazilian cerrado based on phenological information from dense satellite image time series, IJAEOG 82 (2019) 101872.
[8] H.N. Bendini, L.M.G. Fonseca, M. Schwieder, P. Rufin, T.S. Korting, A. Koumrouyan, P. Hostert, Combining environmental and landsat analysis ready data for vegetation mapping: a case study in the Brazilian savanna biome, ISPRS - Int. Arch.Photogramm. Remote Sens. Spatial Inf. Sci. XLIII-B3-2020 (2020) 953–960, doi:10.5194/isprs-archives-XLIII-B3-2020-953-2020.
[9] T. Blaschke, Object based image analysis for remote sensing, ISPRS 65 (1) (2010) 2–16.
[10] E.F. Borges, E.E. Sano, Séries temporais de evi do modis para o mapeamento de uso e cobertura vegetal do oeste da bahia, Boletim de Ciências Geodésicas 20 (3) (2014) 526–547.
[11] Boucher, et al., Brazil's success in reducing deforestation, Trop. Conserv. Sci. 6 (3) (2013) 426–445, doi:10.1177/194008291300600308.
[12] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[13] Brown, et al., Classifying multiyear agricultural land use data from Mato Grosso using time-series MODIS vegetation index data, RSE 130 (2013) 39–50.
[14] O. Carvalho, et al., Slowing Amazon deforestation through public policy and interventions in beef and soy supply chains, Science 344 (6188) (2014) 1118–1123, doi:10.1126/science.1248525.
[15] Costa, et al., Classifying grasslands and cultivated pastures in the Brazilian Cerrado using support vector machines, multilayer perceptrons and autoencoders, in: International Workshop on Machine Learning and Data Mining in Pattern Recognition, Springer, 2015, pp. 187–198.
[16] Costa, et al., Segmentation of optical remote sensing images for detecting homogeneous regions in space and time, Embrapa Solos-Artigo em periódico indexado (ALICE), 2019.
[17] Davidson, et al., Objective indicators of pasture degradation from spectral mixture analysis of landsat imagery, J. Geophys. Res. 113 (G1) (2008).
[18] Dhu, et al., Digital earth Australia – unlocking new value from earth observation data, Big Earth Data 1 (1-2) (2017) 64–74, doi:10.1080/20964471.2017.1402490.
[19] Esquerdo, et al., Use of NDVI/AVHRR time-series profiles for soybean crop monitoring in Brazil, Int. J. Remote Sens. 32 (13) (2011) 3711–3727.
[20] Ferreira, et al., Spectral linear mixture modelling approaches for land cover mapping of tropical savanna areas in Brazil, Int. J. Remote Sens. 28 (2) (2007) 413–429.
[21] Ferreira, et al., Earth observation data cubes for Brazil: requirements, methodology and products, Remote Sens. 12 (24) (2020), doi:10.3390/rs12244033.
[22] Fonseca, et al., Earth observation applications in Brazil with focus on the CBERS program, IEEE Geosci. Remote Sens. Mag. 2 (2) (2014) 53–55, doi:10.1109/mgrs.2014.2320924.
[23] D. Frantz, FORCE—Landsat + sentinel-2 analysis ready data and beyond, Remote Sens. 11 (9) (2019) 1124, doi:10.3390/rs11091124.
[24] Girolamo-Neto, et al., Assessment of texture features for Brazilian savanna classification: a case study in Brasília National Park, Braz. J. Cartogr 69 (2017) 891–901.
[25] Girolamo-Neto, et al., Object based image analysis and texture features for pasture classification in brazilian savannah, XXIV ISPRS Conference (2020).
[26] C.D. Girolamo-Neto, Identification of Brazilian savannah physiognomies on Brasilia Nationa Park using random forest on high and medium spatial resolution images, National Institute for Space Research, 2018 Ph.D. thesis.
[27] M. Gomes, L. Guimarães, F. Junior, M. Aguiar, M. Ferreira, N. Ferreira, Sistema automático de detecção de desmatamentos para o estado de goiás a partir de imagens Índices de vegetação MODIS, in: Anais XII Simpósio Brasileiro de Sensoriamento Remoto, 2005, pp. 3025–3027.
[28] M.C. Hansen, et al., High-resolution global maps of 21st-century forest cover change, Science 342 (6160) (2013) 850–853, doi:10.1126/science.1244693.
[29] INPE, Amazon Program - Monitoring the Brazilian Amazon by Satellite: The PRODES, DETER, DEGRAD and TerraClass Systems, National Institute for Space Research, 2019.
[30] INPE, PRODES Annual Increment of deforested areas in the Brazilian Cerrado, National Institute for Space Research, 2019.
[31] INPE, Projeto Terraclass Cerrado-Mapeamento do Uso e Cobertura Vegetal do Cerrado, 2019c, (http://www.dpi.INPE.br/tccerrado/c).
[32] INPE, National Institute for Space Research - Annual Deforestation in Brazilian Savannah, 2020, (http://www.obt.inpe.br/cerrado). An optional note.
[33] P. Jönsson, L. Eklundh, Timesata program for analyzing time-series of satellite sensor data, Comput. Geosci. 30 (8) (2004) 833–845.
[34] R.J. Kauth, G. Thomas, The tasselled cap–a graphic description of the spectral-temporal development of agricultural crops as seen by LANDSAT, in: LARS Symposia, 1976, p. 159.
[35] T. Körting, L. Fonseca, G. Câmara, GeoDMA - geographic data mining analyst, Comput. Geosci. 57 (2013) 133–145, doi:10.1016/j.cageo.2013.02.007.

[2] https://www.github.com/brazil-data-cube/stmetrics
[3] https://github.com/rvmaretto/deepgeo
[4] http://wiki.dpi.inpe.br/doku.php?id=geodma

[36] Ma, et al., Deep learning in remote sensing applications: a meta-analysis and review, ISPRS 152 (2019) 166–177.
[37] R.V. Maretto, Automating land cover change detection: a deep learning based approach to map deforested areas, Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2020 Ph.D. thesis.
[38] R.V. Maretto, L.M.G. Fonseca, N. Jacobs, T.S. Körting, H.N. Bendini, L.L. Parente, Spatio-temporal deep learning approach to map deforestation in amazon rainforest, IEEE Geosci. Remote Sens. Lett. (2020) 1–5, doi:10.1109/LGRS.2020.2986407.
[39] R.V. Maretto, T.S. Körting, L.M.G. Fonseca, An extensible and easy-to-use toolbox for deep learning based analysis of remote sensing images, in: IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, 2019, pp. 9815–9818, doi:10.1109/IGARSS.2019.8898823.
[40] R.F.B. Marujo, L.M.G. Fonseca, T.S. Körting, H.N. Bendini, A multi-scale segmentation approach to filling gaps in landsat ETM+ SLC-off images through pixel weighting, ISPRS - Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLII-3/W11 (2020) 79–84, doi:10.5194/isprs-archives-XLII-3-W11-79-2020.
[41] S. Mendes, et al., Optical and SAR remote sensing synergism for mapping vegetation types in the endangered Cerrado/Amazon ecotone of Nova Mutum-Mato Grosso, Remote Sens. 11 (10) (2019), doi:10.3390/rs11101161.
[42] Müller, et al., Mining dense landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape, RSE 156 (2015) 490–499.
[43] Neves, et al., Combining time series features and data mining to detect land cover patterns: a case study in northern Mato Grosso state, Brazil, Revista Brasileira de Cartografia 68 (6) (2016).
[44] Neves, et al., Hierarchical classification of Brazilian savanna physiognomies using very high spatial resolution image, superpixel and GEOBIA, in: IGARSS 2019, IEEE, 2019, pp. 3716–3719.
[45] Neves, et al., Semantic segmentation of Brazilian savanna vegetation using high spatial resolution satellite data and U-net, in: XXIV ISPRS Conference, 2020.
[46] A.K. Neves, Mineração de dados de sensoriamento remoto para detecção e classificação de áreas de pastagem na Amazônia Legal, National Institute for Space Research, 2017 Master's thesis.
[47] A.K. Neves, Hierarchical mapping of Brazilian savanna (Cerrado) physiognomies based on deep learning, National Institute for Space Research (INPE), 2021 Ph.D. thesis.
[48] Nogueira, et al., Towards vegetation species discrimination by using data-driven descriptors, in: 2016 9th IAPR Workshop on Pattern Recogniton in Remote Sensing (PRRS), Ieee, 2016, pp. 1–6.
[49] M. Ortega Adarme, R. Queiroz Feitosa, P. Nigri Happ, C. Aparecido De Almeida, A. Rodrigues Gomes, Evaluation of deep learning techniques for deforestation detection in the brazilian amazon and Cerrado biomes from remote sensing imagery, Remote Sens. 12 (6) (2020), doi:10.3390/rs12060910.
[50] M. Pal, P.M. Mather, Support vector machines for classification in remote sensing, Int. J. Remote Sens. 26 (5) (2005) 1007–1011, doi:10.1080/01431160512331314083.
[51] Parente, et al., Monitoring the Brazilian pasturelands: a new mapping approach based on the landsat 8 spectral and temporal domains, IJAEOG 62 (2017) 135–143.
[52] F. Petitjean, et al., Spatio-temporal reasoning for the classification of satellite image time series, Pattern Recognit. Lett. 33 (13) (2012) 1805–1815, doi:10.1016/j.patrec.2012.06.009.
[53] Picoli, et al., Big earth observation time series analysis for monitoring brazilian agriculture, ISPRS 145 (2018) 328–339.
[54] Potapov, et al., Landsat analysis ready data for global land cover and land cover change mapping, Remote Sens. 12 (3) (2020) 426, doi:10.3390/rs12030426.
[55] Ribeiro, et al., As principais fitofisionomias do bioma cerrado, Cerrado: ecologia e flora 1 (2008) 151–212.
[56] Ronneberger, et al., U-Net: convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.
[57] Rudorff, et al., Studies on the rapid expansion of sugarcane for ethanol production in São Paulo State (Brazil) using landsat data, Remote Sens. 2 (4) (2010) 1057–1076.
[58] Rufin, et al., Land use intensity trajectories on amazonian pastures derived from landsat time series, IJAEOG 41 (2015) 1–10.
[59] Sano, et al., Mapeamento de cobertura vegetal do bioma cerrado: estratégias e resultados, Embrapa Cerrados-Documentos (INFOTECA-E), 2007.
[60] E.E. Sano, L.G. Ferreira, A.R. Huete, Synthetic aperture radar (L band) and optical vegetation indices for discriminating the Brazilian savanna physiognomies: a comparative analysis, Earth Interact. 9 (15) (2005) 1, doi:10.1175/EI117.1.
[61] Scaramuzza, et al., Land-use and land-cover mapping of the Brazilian cerrado based mainly on landsat-8 satellite images, Revista Brasileira de Cartografia 69 (6) (2017).
[62] Schultz, et al., Classificação orientada a objetos em imagens multitemporais landsat aplicada na identificação de cana-de-açúcar e soja., Revista Brasileira de Cartografia 68 (1) (2016) 131–143.
[63] Schwieder, et al., Mapping Brazilian savanna vegetation gradients with landsat time series, IJAEOG 52 (2016) 361–370.
[64] Y.E. Shimabukuro, J.A. Smith, The least-squares mixing models to generate fraction images derived from remote sensing multispectral data, IEEE Trans. Geosci. Remote Sens. 29 (1) (1991) 16–20.
[65] A.R. Soares, H.N. Bendini, D.V. Vaz, T.D.T. Uehara, A.K. Neves, T.S. Körting, L.M.G. Fonseca, Stmetrics: a python package for satellite image time-series feature extraction, IGARSS 2020, 2020.
[66] Sohn, et al., Supervised and unsupervised spectral angle classifiers, Photogramm. Eng. Remote Sens. 68 (12) (2002) 1271–1282.
[67] C.M. Souza, J. Z Shimbo, M.R. Rosa, L.L. Parente, A. A Alencar, B.F. Rudorff, H. Hasenack, M. Matsumoto, L. G Ferreira, P.W. Souza-Filho, et al., Reconstructing three decades of land use and land cover changes in Brazilian biomes with landsat archive and earth engine, Remote Sens. 12 (17) (2020) 2735.
[68] C. Souza Junior, et al., Near real-time deforestation detection for enforcement of forest reserves in Mato Grosso, in: Proceedings of Land Governance in Support of the Millennium Development Goals: Responding to New Challenges, World Bank Conference, Washington, DC, 2009.
[69] Strassburg, et al., Moment of truth for the Cerrado hotspot, Nat. Ecol. Evol. 1 (4) (2017).
[70] E. Taquary, Deep Learning Para Identificação Precisa de Desmatamentos Através do Uso de Imagens Satelitárias de Alta Resolução, Federal University of Goiás, Institute of Informatics, 2019 Master's thesis.
[71] Trémas, et al., Sentinel-2a: presentation of the CAL/VAL commissioning phase, first images, in: These proceedings, 2015, pp. 9643–9647.
[72] D. Tuchschneider, Brazil - Development of Systems to Prevent Forest Fires and Monitor Vegetation Project (English), Technical Report, World Bank Group, Washington, D.C., 2013.
[73] Xi, et al., A spatiotemporal cube model for analyzing satellite image time series: application to land-cover mapping and change detection, RSE 231 (2019) 111212.