MDPI

Article

A Semi-Automated Workflow for LULC Mapping via Sentinel-2 Data Cubes and Spectral Indices

Michel E. D. Chaves [1,*], Anderson R. Soares [2], Guilherme A. V. Mataveli [1], Alber H. Sánchez [1] and Ieda D. Sanches [1]

[1]   Earth Observation and Geoinformatics Division, National Institute for Space Research (INPE),
     São José dos Campos 12227-010, São Paulo, Brazil
[2]   Digital Business Department, Cognizant Technology Solutions, São Paulo 04705-000, São Paulo, Brazil
*   Correspondence: michel.chaves@inpe.br

Abstract: Land use and land cover (LULC) mapping initiatives are essential to support decision making related to the implementation of different policies. There is a need for timely and accurate LULC maps. However, building them is challenging. LULC changes affect natural areas and local biodiversity. When they cause landscape fragmentation, the mapping and monitoring of changes are affected. Due to this situation, improving the efforts for LULC mapping and monitoring in fragmented biomes and ecosystems is crucial, and the adequate separability of classes is a key factor in this process. We believe that combining multidimensional Earth observation (EO) data cubes and spectral vegetation indices (VIs) derived from the red edge, near-infrared, and shortwave infrared bands provided by the Sentinel-2/MultiSpectral Instrument (S2/MSI) mission reduces uncertainties in area estimation, leading toward more automated mappings. Here, we present a low-cost semi-automated classification scheme created to identify croplands, pasturelands, natural grasslands, and shrublands from EO data cubes and the Surface Reflectance to Vegetation Indexes (*sr2vgi*) tool to automate spectral index calculation, with both produced in the scope of the Brazil Data Cube (BDC) project. We used this combination of data and tools to improve LULC mapping in the Brazilian Cerrado biome during the 2018–2019 crop season. The overall accuracy (OA) of our results is 88%, indicating the potential of the proposed approach to provide timely and accurate LULC mapping from the detection of different vegetation patterns in time series.

Keywords: LULC mapping; machine learning; python; semi-automated mapping; time series

1. Introduction

Brazil is a relevant global producer and exporter of commodities such as soybeans, cotton, coffee, and maize. Within Brazil, the Cerrado biome is a prominent producer region at the center of the country's recent agricultural boom [1]. This biome has experienced cropland expansion in recent decades and, in some cases, over natural vegetation, inducing illegal deforestation. The MATOPIBA, a region in the Cerrado biome which includes parts of the states of Maranhão, Tocantins, Piauí, and Bahia, is at the forefront of agricultural expansion, accounting for 25% of the soybean produced in the biome [2]. The expansion of monocultures has been affecting natural vegetation areas and threatening local biodiversity. Given this, accurate land use and land cover (LULC) information of this agricultural frontier is needed to support decision making regarding agriculture dynamics, climate change, deforestation, and food security [2].

Information of this magnitude can be obtained via maps. Worldwide, highly accurate LULC classification has strategic value for reducing uncertainties and supporting the implementation of policies [3–5]. This context reinforces the need for detailed mappings [6]. Free and open access data, analysis-ready data (ARD), and high-performance computing created a new era for LULC analysis [7]. However, accurate LULC classification remains a

challenge influenced by many factors [8]. Major factors are related to the use of time series for LULC classification at the regional or global level [9], algorithms and input data [8], requirements, user-defined parameters, and computational costs [10], Earth observation (EO) platforms and datasets, the spatial-spectral-temporal characteristics of satellite data, and approaches to change detection [11], web-based workflows [12], accurate spatiotemporal event detection [13], and estimation of the general and species-specific phenological stages [14]. The state of the art for extracting land surface information from remote sensing-based techniques indicates that LULC classifications are migrating from exclusively human-based to semi-automated approaches [14,15]. In this scenario, the use of multidimensional ARD and spectral vegetation indices (VIs) from medium-resolution EO images for detecting subtle differences in vegetation types and improving LULC classification can be automated. However, sample collection, which configures the semi-automated nature of many LULC classification approaches, remains a task that is difficult to automate.

In Brazil, improving the efforts for LULC mapping and monitoring in this region is crucial, and the adequate separability of classes is a key factor in this process. Some initiatives engaged in LULC classification in the Brazilian Cerrado and other biomes face indiscernible patterns in crop phenology classes as a limiting factor, which shows the value of accurate and precise separability of classes when detecting subtle differences in vegetation. The TerraClass [16], for example, is developed by the Brazilian National Institute for Space Research (INPE) and the Brazilian Agriculture Research Corporation (Embrapa). It complements the Brazilian Amazon Deforestation Monitoring Program (PRODES) by adding information about the previous LULC spatial distribution and regional statistics in deforested areas in the Brazilian Legal Amazon and Cerrado biomes [16]. The project proposes a mapping project for the whole country. For this, it has been trying to adopt automated mapping processes. Therefore, there is a need for timely and accurate crop mapping initiatives, which require automation in learning systems and applications.

The launch of the Sentinel-2A (2015) and Sentinel-2B (2017) satellites by the European Space Agency (ESA) provided new possibilities for this purpose [17]. Sentinel-2 carries the MultiSpectral Instrument (MSI), a sensor able to record radiance in 13 spectral bands with spatial resolutions varying between 10 m and 60 m, from the visible to shortwave infrared (SWIR) portions of the electromagnetic spectrum [18]. Sentinel's MSI (S2/MSI) sensor has three bands in the red edge region which are useful for vegetation discrimination and LULC mapping [15]. This is due to its sensitivity to chlorophyll and subtle variations among different crops and phenological states [19]. This characteristic is useful for deriving VIs and metrics to evaluate vegetation [20]. Considering open science and data-sharing policies, S2/MSI has strategic value as a cost-effective analytical instrument of global cooperation for elaborating precise and timely map pings [7]. However, managing and taking advantage of the large collection available is challenging [21]. Likewise, cloud or shadow interference and geometric and atmospheric noise hinder analysis [22]. An alternative to overcome this limitation is modeling images as data cubes and uniform spatiotemporal tessellation of EO data with common temporal and spatial reference systems determined for a specific region over a defined time interval, which allows efficient storage and access in ARD arrangements [23]. This configuration enables the production of timely and accurate LULC maps. The Brazil Data Cube (BDC) project [24], also developed by the INPE, creates multidimensional data cubes from medium-resolution satellite imagery from the satellite missions Landsat, CBERS, and Sentinel for Brazil.

Worldwide, different initiatives have been using ARD, which greatly simplifies large-area analyses to improve the level of detail of LULC classifications. For environmental applications, in Australia, Landsat and S2/MSI data cubes are being used to generate burned area and severity maps, providing information about change characterization [25]. In Europe, they provide information about protected areas [26]. In Canada, detailed information on the land cover dynamics and forest developments following disturbance events support science-based policies, forest inventories, and forest management programs [27]. In South America, they allow monitoring of deforestation at sub-annual scales [28]. For

croplands, a precise and accurate Landsat 30 m-derived cropland product was developed for Australia and China [25]. In the Cerrado, phenological shifts and interannual cropping practice changes were identified using time series [29]. The potential of this is advantageous for the implementation of public policies, especially those related to food security, conservation of natural landscapes, and agricultural prices.

Within MATOPIBA, the *Extremo Oeste Baiano* is a mesoregion of the Bahia state that can be more efficiently analyzed by using data cubes [30]. Therefore, the goal of this study was to show how to use S2/MSI-derived data cubes and VIs to detect different vegetation types in the *Extremo Oeste Baiano* mesoregion. Our hypothesis is that we can identify different vegetation classes and improve LULC mapping by exploring the electromagnetic spectrum through a semi-automated approach that uses data cubes and pattern recognition to detect subtle differences in vegetation. This permits the automation of a workflow to optimize LULC mapping approaches. In the case of this study, our approach made possible the automation of many steps of LULC classification (ARD access, VIs calculation, and accuracy assessment), except for sample collection, which configures the semi-automated nature of this LULC classification approach. We believe our approach eases the identification of LULC or, more specifically, vegetation. The success of our approach is mainly due to two things: the use of machine learning and the automation of the processing of large amounts of satellite imagery, which is performed during data cube building [24]. Before the use of data cubes, satellite images were processed individually to achieve a correct interpretation of the spatial and spectral characteristics of each image, which was both time-consuming and computer-intensive. This constrained analysis to specific regions in space and time. However, we still rely on using training samples taken by specialists in remote sensing, and for this reason, our approach is not fully automated. To perform the analysis, we used an S2/MSI data cube developed by the BDC project, which was accessed via the SpatioTemporal Asset Catalog [31], and five VIs derived from combinations including the red edge, near-infrared (NIR), and SWIR spectral bands, which were correlated with complementary constituents of the plants.

2. Materials and Methods

2.1. Study Area

A case study was conducted in an area located in the leading crop producer and prominent exporter zone of the MATOPIBA: the *Extremo Oeste Baiano* mesoregion in the state of Bahia. This study area corresponds to an S2/MSI tile of a BDC grid (Figure 1).

This mesoregion has a heterogeneous landscape resulting from a modern and active pioneer frontier shaped by the abrupt expansion of distinct summer crops (corn, cotton, and mainly soybean) at the expense of native vegetation [1] and varying natural conditions [32]. Coffee, essentially cultivated in central pivots, is the main perennial crop produced in the region, albeit in a smaller area than summer agricultural crops [33]. Aside from that, the landscape covers other crops and a natural mosaic of vegetation that can be divided into forestlands (ciliary, dry forests, gallery, and Cerradão), shrublands and savanna (Cerrado sensu stricto, palm, park savanna, and *vereda*), grasslands (*campo rupestre*, *campo limpo*, and *campo sujo*, notably), and pasturelands [34]. The climate type in the site is Aw (tropical savanna marked by dry winters) [35], with an average annual temperature of 24 °C and average annual precipitation of 1145 mm [36]. The main local rainfed crops are soybean, corn, and cotton. In areas with irrigated systems, farmers rotate soybean with cotton and, occasionally, corn during the second growing seasons [36]. In the west, the study area borders the Goiás and Tocantins states, where the landscape combines inner ecosystems and subtypes of the abovementioned pasturelands, shrublands and savanna, and grasslands [37]. This scenario characterizes this region as suitable to analyze the relevance of S2/MSI bands for detecting subtle changes in phenologies. Then, four broad LULC classes were considered in this study: Croplands, Shrublands, Natural Grasslands, and Pasturelands.
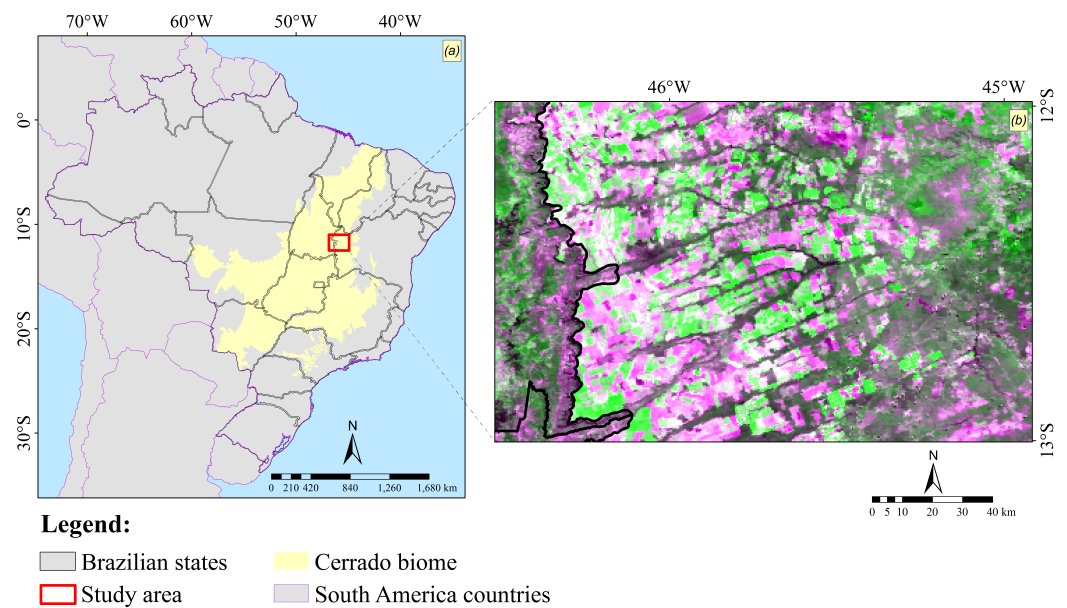
Figure 1. Location of the study area from the perspective of (**a**) South America, Brazil, Cerrado biome, and Bahia state and (**b**) landscape composition, with croplands, pasturelands, natural grasslands, and shrublands.

2.2. Method

Our methodological procedures (Figure 2) included satellite data accessing, sample filtering, temporal analysis, VIs calculation, mapping, and accuracy evaluation.
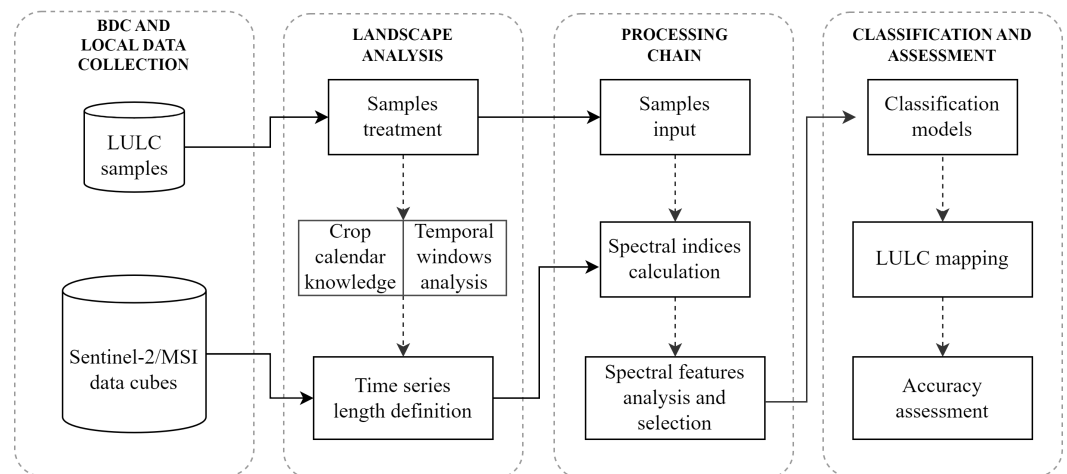

Figure 2. Data workflow to generate the LULC classification from S2/MSI data cubes and LULC samples.

2.3. Satellite Data and Classification Approach

The data cube used (S2/MSI) is composed of VIs with a spatial resolution of 10 m projected and clipped to a previously defined reference grid [24]. It covers part of the 23LLF, 23LMF, 23LLG, and 23LMG tiles of the Military Grid Reference System (MGRS) used by the ESA, and it was formed by monthly composites corresponding to the 2018–2019 crop season (from 1 September 2018 to 1 August 2019). We excluded the period between 2 August 2019 and 31 August 2019 to reduce dimensionality, as it represents a transition to the next crop season or agricultural year in the region. During this period, the harvest process was already concluded, and the sanitary emptiness (*Vazio Sanitário*, in Portuguese) is in effect [38]. The selected VIs correlated with the vegetation characteristics, being measures of the chlorophyll content and vegetation health [39]. They were selected based on tests with

more than 80 VIs made available in the Surface Reflectance to Vegetation Indexes (*sr2vgi)* Python package [40]. Each test analyzed their ability to identify land surface characteristics and assess the heterogeneous and continuous nature of the landscape composition and cover type-specific LULC change processes. VIs formulated with combinations including red edge bands can reduce saturation compared with those derived from red ones [19]. Previous research has successfully used most of these VIs to map croplands [17,41] and natural vegetation [37,42] in semi-automated approaches.

To explore this potential, we used five VIs (Table 1) formulated from a combination among the red edge, NIR, and SWIR spectral bands. These spectral bands are correlated with the chlorophyll and biomass conditions and permit more accurate detection of the temporal behavior of vegetation [20]. The VIs used to detect subtle differences in their inter-relation were the Normalized Multi-Band Drought Index (NMDI) [43], Normalized Difference Vegetation Index Red Edge (NDVIre) [44], Red Edge Vegetation Index (RERVI) [45], Core Red Edge Triangular Vegetation Index (RTVIcore) [46], and Vegetation Index 700 (VI700) [47]. The NMDI assesses the effect of drought on vegetation [43] and estimates the vegetation moisture and soil moisture content [48]. NDVIre is a red edge-based VI that has the potential to identify and enhance the characterization of crop [49] and bare land [50] patterns. The RERVI correlates with the nitrogen status [51] and canopy chlorophyll content [52]. The RTVIcore has linear relationships with the chlorophyll content and leaf biomass, with reduced saturation in high-biomass areas [53]. The VI700 assesses the vegetation status and differentiates crops from other vegetation types [54].

Table 1. VIs that comprised the data cube. S2/MSI spectral bands: b3 (green), b4 (red), b5 (red edge 1), b8 (NIR), b8a (NIR narrow), b11 (SWIR 1), and b12 (SWIR 2).

VI	Formula with Sentinel-2 Bands
NMDI	$(b8a - (b11 - b12))/(b8a + (b11 - b12))$
NDVIre	$b8 - b5/b8 + b5$
RERVI	$b8/b5$
RTVIcore	$100 * (b8a - b5) - 10 * (b8a - b3)$
VI700	$(b5 - b4)/(b5 + b4)$

These VIs were automatically calculated by our tool, *sr2vgi*, in the processing chain step of the classification scheme represented in Figure 2. The automatic calculation of complementary VIs allows the classifier algorithm to detect subtle variations, and it is useful for improving classifications [41]. To train the classifier, we used a dataset of 360 samples, including the following LULC classes: (1) Croplands (220 samples), (2) Natural Grasslands (9 samples), (3) Pasturelands (90 samples), and (4) Shrublands (41 samples). These samples were previously taken by experts on remote sensing via visual inspection of S2/MSI images, time series analysis from S2/MSI data cubes, and the Temporal Vegetation Analysis System (SATVeg), a free web-based tool designed to provide instantaneous access to MODIS VIs time series in South America [55]. To the best of our knowledge, there are not statistics for all LULC classes assessed. Therefore, we did not follow a methodology that considered the area proportion of each class or stratified random sampling. While the area of Croplands is prominent in the region, that of Natural Grasslands is small. Consequently, the total samples for each one was uneven. Collecting samples for landscape analysis, represented in Figure 2, is the only non-automated step of our LULC classification scheme and is what defines its supervised (semi-automated) nature.

We used the random forest (RF) classification algorithm [56], available in the *scikit-learn* Python package [57]. RF is an ensemble classifier widely used for LULC classification using remote sensing data [9]. It constructs a set of decision trees (DTs) to make a prediction using a randomly selected subset of training samples and variables (Nvar). When the forest grows to a user-defined number of trees (Ntree), the RF creates trees with a high variance but low bias. The classification results from the average class assignment probabilities calculated across all trees. The RF evaluates unlabeled data inputs against all DT created

in the ensemble, and each tree votes to define the class membership. The one with the most votes is selected [58]. Here, we used 500 trees to perform the classification. We also used stratified proportional sampling. The train and test dataset ratio was 70/30; that is, we had 252 training and 108 testing samples. The configuration of these parameters was implemented in the classification and assessment step of the scheme represented in Figure 2.

3. Results

3.1. VIs Temporal Patterns of the Assessed LULC Classes

Among all the VIs used, NDVIre presented the most representative behavior in the time series analysis extracted from the samples used for each assessed LULC class. This time series analysis is a relevant step for the RF algorithm in the supervised step to generate the semi-automated LULC classification. The time series in this VI indicates the differences and similarities between the patterns of each class (Figure 3), including subtle changes that permit their differentiation. Croplands and Pasturelands presented distinct temporal patterns, while Natural Grasslands and Shrublands presented similarities, which indicates the source of possible confusion for the RF algorithm in the classification process.
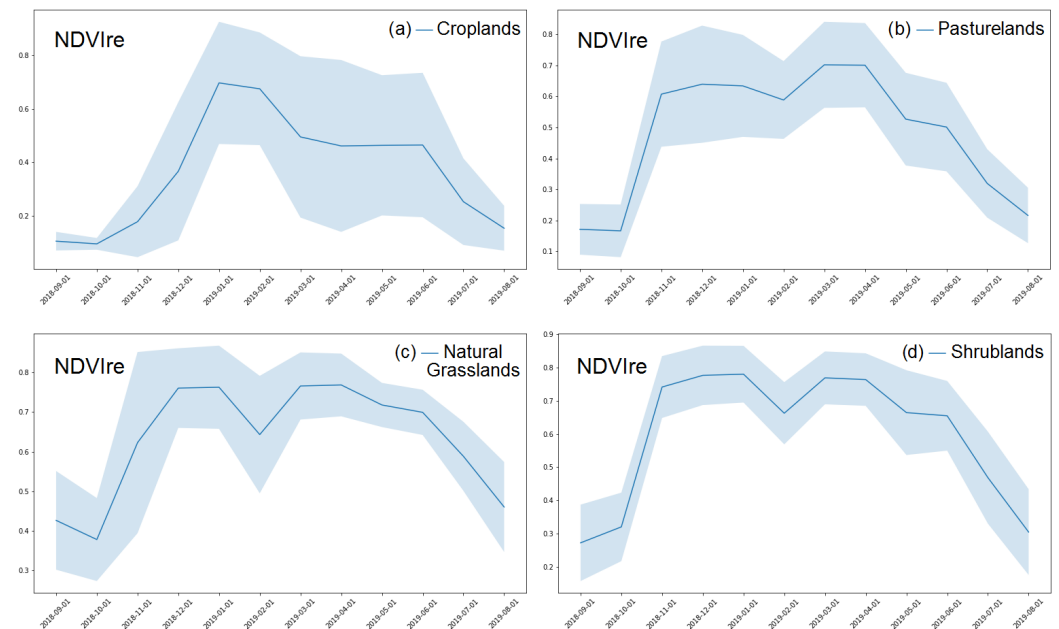

Figure 3. Time series derived from samples of the four assessed LULC classes: (**a**) Croplands, (**b**) Pasturelands, (**c**) Natural Grasslands, and (**d**) Shrublands in NDVIre, showing the median profile with the blue line and all values in a blue shadow.

3.2. Croplands and Pasturelands

Despite presenting differences in NDVIre, the time series of Croplands and Pasturelands showed similarities in some VIs. The time series analysis of the NMDI, for example, illustrated this. Even so, it was possible to note subtle differences between both LULC classes (Figure 4), such as differences in amplitude, the peak of maximum vegetative vigor, the duration of this peak, and the natural decrease as a consequence of the senescence of the vegetation and the harvest, in the case of Croplands. This LULC class presented a peak (between January and February) with higher values in comparison with Pasturelands. These differences may have helped the RF algorithm detect differences that permit their separation.
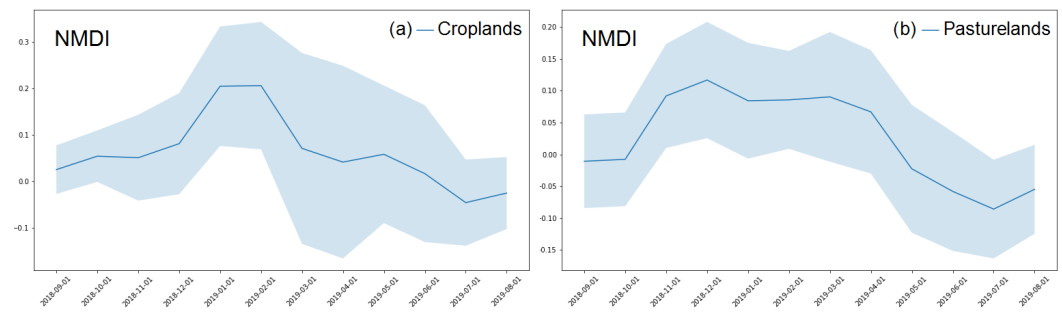
Figure 4. Time series derived from the samples of the (**a**) Croplands and (**b**) Pasturelands LULC classes in NMDI, showing the median profile in the blue line and all values in a blue shadow.

3.3. Shrublands and Natural Grasslands

Considering the spectro-temporal profiles automatically extracted from all samples used for each LULC class, the time series of Natural Grasslands and Shrublands showed more similarities. The time series of the VI700 showed the potential of VIs to distinguish these two LULC classes (Figure 5), which can help the RF algorithm to detect differences that permit their separation among profiles of vegetation. From using this VI, it was possible to detect differences in amplitude, peak, and duration of the peak, as well as the natural decrease as a consequence of the senescence of the vegetation. Both LULC classes presented two peaks: between November and January and between March and April, interrupted by abrupt minimum values that could be associated with outliers. These could have been caused by noises produced by clouds during the rainy period. In this case, outliers represent the discontinuity of VI signals in the time series, an effect derived from the covering of the land by clouds. In general, the Shrublands presented higher values in comparison with Natural Grasslands.
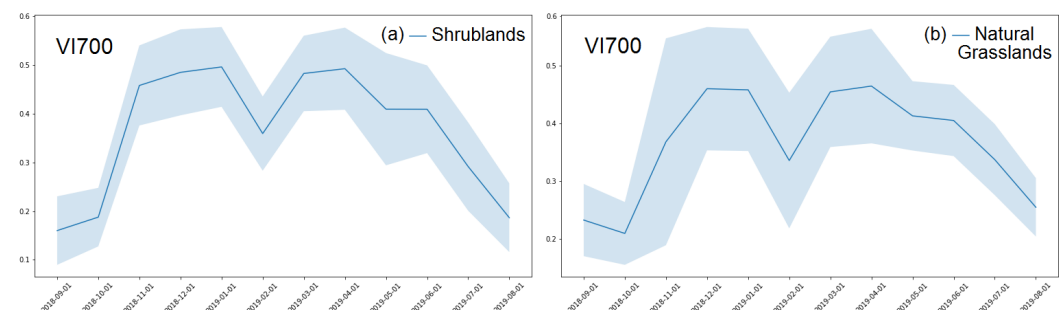

Figure 5. Time series derived from the samples of the (**a**) Shrublands and (**b**) Natural Grasslands LULC classes in the VI700, showing the median profile with a blue line and all values in a blue shadow.

3.4. LULC Classification

The semi-automated LULC classification resulting from our data cube-derived approach emphasizes the heterogeneity of the vegetational component of the mesoregion's landscape (Figure 6). Overall, the time series patterns of the VIs provided sufficient information for LULC classification. The target classes were distinguished, despite the occurrence of granular effects and confusion involving Croplands and Pasturelands (western portion) as well as Shrublands and Natural Grasslands (eastern portion).

The overall accuracy (OA) of the LULC classification was 88%. Most of the errors were derived from confusion between Croplands and Pasturelands (especially in the western portion of the study area) as well as between Natural Grasslands and Shrublands (especially in the eastern portion of the study area). A plausible explanation, aside from the similarities in time series, is the uneven amount of samples used to train the RF algorithm in the semi-automated LULC classification approach. To represent the accuracy of individual

classes, we also evaluated the producer (PA) and user (UA) accuracies of each class. Figure 7 presents the error matrix derived from the LULC classification.

Figure 6. Land use and land cover classification of the assessed study area in the 2018–2019 crop season.
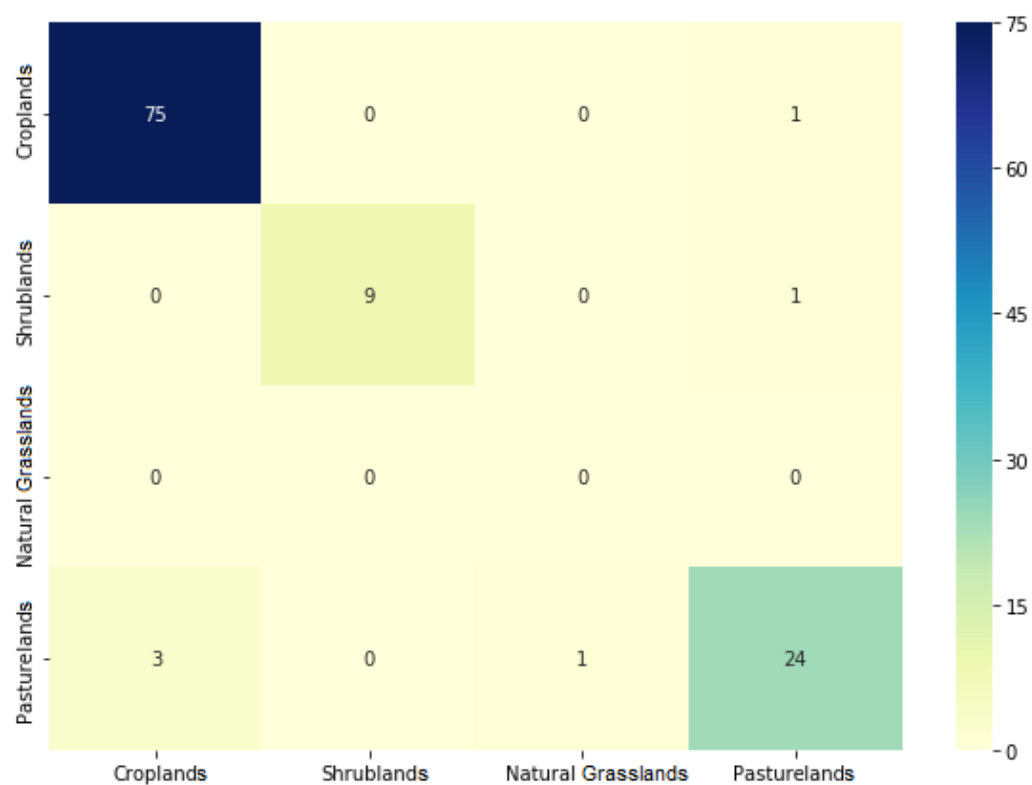

Figure 7. Error matrix derived from comparing the classification and the 30% share of the data used for validation.

The UA and PA values were as follows: Croplands (UA = 97% and PA = 90%), Shrublands (UA = 90% and PA = 82%), Natural Grasslands (UA = 25% and PA = 100%), and Pasturelands (UA = 78% and PA = 86%). To assess and determine the contribution of each VI for discriminating the four assessed LULC classes, we calculated the Gini index (Table 2). The analysis of the VIs in each month revealed that, in general, the RERVI, VI700, and NMDI features were the three most significant ones for the RF algorithm. Months from the beginning of the interval of observation (from September to November) presented special relevance to the classification algorithm.

Monthly boxplots of NDVIre (Figure 8), a VI representative of the others, showed the variability of the samples from each LULC class. Croplands and Pasturelands were the LULC classes with the higher variability, concentrating the higher presence of outliers.
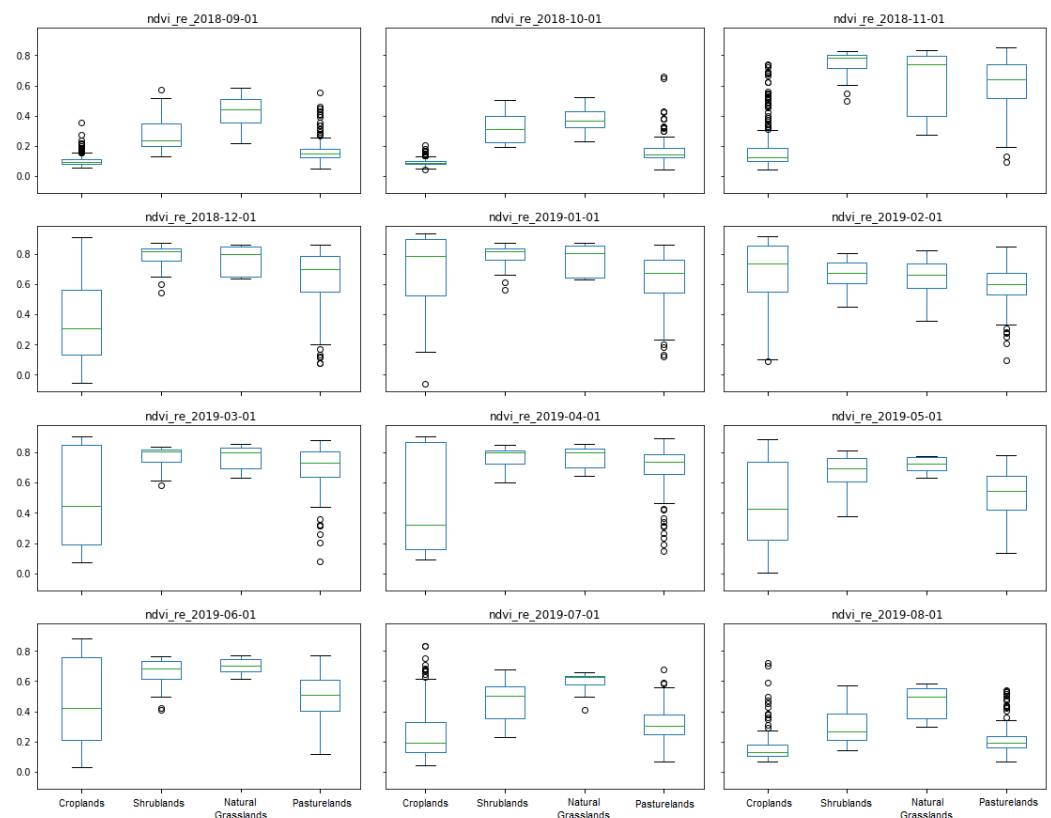

Figure 8. Boxplot of NDVIre analysis from September 2018 to August 2019. Central marks in green represent the median. Edges in blue symbolize the 25th and 75th percentiles. Upper and lower lines in black delimit the most extreme value contained in the limits determined by the sum or difference between the 75th and 25th percentiles and the difference between the 75th and 25th (or vice versa) percentiles multiplied by 1.5. Outliers (circles in black) are values outside the limits demarcated by lines.

Table 2. Significance of each monthly feature (VI) for the RF classification algorithm measured via the Gini index.

	NDVIre	NMDI	RERVI	RTVIcore	VI700
Sep	0.004	0.010	0.045	0.007	0.047
Oct	0.005	0.025	0.036	0.015	0.042
Nov	0.007	0.008	0.081	0.012	0.076
Dec	0.005	0.011	0.019	0.005	0.022
Jan	0.008	0.015	0.011	0.006	0.018
Feb	0.013	0.019	0.009	0.008	0.017
Mar	0.008	0.020	0.014	0.007	0.013
Apr	0.006	0.019	0.019	0.010	0.017
May	0.007	0.008	0.008	0.007	0.008
Jun	0.006	0.007	0.011	0.005	0.008
Jul	0.005	0.018	0.020	0.007	0.012

4. Discussion

Mapping efforts in the Cerrado biome occur on a smaller scale than in the Amazon biome [32]. Meanwhile, the rhythm of LULC changes in Cerrado is more abrupt [59]. The agricultural development in the MATOPIBA region is an example of this [60]. The growth in energy production and land use conversion activities also increased greenhouse gas (GHG) emissions in this region [61]. Given the scarceness of already-converted lands suited for agriculture, most of the agricultural expansion in MATOPIBA was over native vegetation [62]. To avoid trade embargoes, Brazil needs to develop policies for increasing agricultural expansion over already-converted lands to produce deforestation-free supplies [60,63]. The reversion of this scenario involves, in the initial stage, accurate LULC mapping to identify and monitor the regions where illegal activity has increased more significantly. Due to the expanse of the Cerrado biome, encompassing more than two million km², remote sensing is the only viable way to monitor it. Given this, the development of initiatives to automate the use of remote sensing-based ARD for the mapping process is important because it can enhance the generation of accurate near-real-time LULC information.

4.1. Croplands and Pasturelands Time Series Analysis

For the five VIs analyzed, the Croplands LULC class presented incremental growth in vegetative vigor during the months with rainfall accumulation (from September 2018 to January 2019, where the peak maximum vigor occurred). In the 2018–2019 crop season, the harvesting procedure started exactly in January 2019. The main crops in the region were soybean, cotton, and maize. Differences in cropping management and phenological cycles created less symmetrical VI dynamics when we analyzed all these crops as a wider class named Croplands (Figures 3 and 4). On the other hand, Pasturelands presented abrupt growth in vegetative vigor as a response to rainfall that occurred from September to December. The main difference in comparison with Croplands was the absence of a maximum vigor peak, which implies a more constant behavior. Both classes presented greater amplitude in values than the other analyzed LULC classes (Figures 3 and 4), which was present in the variations in Croplands (Figure 8). A key to discriminating these classes in the study area occurred after the beginning of the harvest (from February to April). During this period, Croplands still decreased, given the senescence and harvesting. Meanwhile, Pasturelands presented less variation due to the higher presence of vegetation than that in Croplands in this month, when the crop fields had fallow to refill soil fertility through non-commercial crops. These two classes had spectral similarities that complicated differentiating LULC classes in the Brazilian Cerrado [64]. The Croplands LULC class is composed of several crop types, such as soybean, maize, and coffee. This grouping generates two problems. First, the summer crops have a variant spectro-temporal behavior because of their cycles, phenology, and response to rainfall. Second, perennial crops, such as

coffee, vary less, given their phenological cycle and the irrigation (locally, coffee cultivation occurs mainly in irrigated pivots), making it hard to identify patterns in the VI time series.

4.2. Natural Grasslands and Shrublands Time Series Analysis

The Natural Grasslands and Shrublands time series had a similar pattern. This was expected, because both are LULC compositions that have a certain degree of anthropization [32]. However, any differences allow their separation. Shrublands presented a rapid response to rainfall, achieving high VI values in November 2018. Composed of *campo rupestre*, *campo limpo*, and *campo sujo*, the Natural Grasslands LULC class presented a gradual response to rainfall. Both LULC classes presented a vegetative vigor peak in January 2019 and a drop in February 2019, a month with a short drought. The Natural Grasslands LULC class decreased more at this moment, being more responsive to the absence of rainfall. This occurred due to their more shallow rooting [65]. Additionally, given its composition marked by grasses and forbs, this class is influenced by soil background effects [66,67].

Another factor that should be considered to analyze the time series of these two LULC classes is the response to fire. Most of the Cerrado gradient of vegetation is characterized by an inflammable grassy layer that can ignite in the dry season (from May to September) [68]. In addition, anthropic activities became the major source of fires to suppress natural vegetation and change the LULC composition [69]. In the *Extremo Oeste Baiano* mesoregion, fire management is not applied in Croplands. In turn, Shrublands and Natural Grasslands have characteristics that contribute to a fire spreading. The lack of an effective fire management policy in the Cerrado could lead to wildfires in the remaining native vegetation areas during the late dry seasons caused by the accumulation of dry fuel loads, especially in areas neighboring crop plantations [70]. This helps to explain the abrupt decrease in VI values between May and September.

4.3. LULC Mapping

Our approach accurately mapped the Shrublands (UA = 90% and PA = 82%) and Pasturelands LULC classes (UA = 78% and PA = 86%). Errors were caused by the Cerrado heterogeneity due to its biodiversity and phytophysiognomies. Many of them derive from spectral similarities between Croplands, Pasturelands, and Natural Grasslands, which have similar spectral responses and subtle seasonal variations over the dry and wet seasons [29,64,71] as well as seasonality-derived issues [32]. The use of a data cube and VIs supported the detection of a temporal pattern to separate the assessed classes. However, the different sensitivity levels to the water content under a monthly interval perspective bring confusion. The errors between Croplands (UA = 97% and PA = 90%) and Pasturelands (UA = 78% and PA = 86%), for example, were already expected in the function of already-mentioned issues related to seasonal variations [64]. An important source of confusion between the four assessed classes is the uneven number of samples in the reference dataset. The crop mapping performance, for example, depends on previous information about cropping systems [41,72,73]. Assuming that some crops will have similar phenological cycles, a representative and balanced sampling strategy is vital for a good mapping performance [29,74].

Previous tests suggested that the combination between the red, red edge, NIR, and SWIR spectral bands of S2/MSI effectively detects vegetation types such that the five selected VIs were formulated with combinations among them. The RERVI and VI700 are indicators of the sensitivity of the leaf area index, biomass, and nitrogen status. The NMDI was chosen to improve the detection of sparse and superficial vegetation, as occurred in the work of Zhang et al. [75]. However, this index also presented inconsistent relationships with soil and vegetation moisture changes in areas with moderate vegetation coverage [43,75]. The RTVIcore presented the potential for improving the separability between Croplands, Natural Grasslands, and Pasturelands [17]. Our results showed that the spectral band of red edge 1 presented superior significance to red edge 2 and 3 because the VIs formulated

with this spectral band presented more importance. An explanation for this is the higher ratio between the reflectance in NIR and red edge 1 (spectral band 5 of S2/MSI), which is more significant than the ratio between NIR and other red edge channels as well as between NIR and SWIR. We expected that these other differences could be more useful to evidence the differences within the phenologies, such as to detect different crops and gradients of natural vegetation.

The use of a data cube and VIs also supported the detection of a temporal pattern to separate the classes. Despite the results, some challenges hinder the best explanation for the similarities. The first is related to the number of samples. Given this, we are working to improve the analysis by incorporating more representative samples of each broad class, detailing the cultivated crops (i.e., soybean, maize, millet, cotton, and coffee) and the natural vegetation gradient that compounds the Shrublands and Natural Grasslands LULC classes (i.e., woodlands and grasslands). We expect that this improvement could reduce the standard deviation of the phenological patterns, a factor that generates confusion between the classes. The Cerrado has unique inter-annual and seasonal variability, presenting unique challenges for LULC mapping [76]. Thus, the next step is to deepen the analysis to detect the phytophysiognomies that compound the broad classes. Following that, we aim to achieve the next level of hierarchical classification, transforming the data into information to extract accurate and precise LULC maps of the entire Cerrado biome.

5. Conclusions

This study presents a semi-automated approach that combines human knowledge and skills for the collection of LULC samples and machine learning for using multidimensional data cubes and VIs in order to detect the patterns of the LULC classes Croplands, Pasturelands, Natural Grasslands, and Shrublands in the Cerrado biome, an intensive agricultural frontier in Brazil. Considering the heterogeneous and dynamic nature of the study area, characterized by a vast gradient of vegetation types, the OA of 88% and the UA and PA values support that our strategy enables exploring the potential of ARD in the optical spectral region to enhance the dissimilarity between similar vegetation classes, being appropriate for LULC classification at this level of detail (broad level and macro-classes). It was possible to detect subtle differences in vegetation types and optimize the delineation of large individual features. The limitations found occurred because of the uneven sampling and the use of a monthly temporal aggregation to extract LULC information. The collection of new samples and the processing of different temporal aggregations (composite products) can overcome this issue.

Moreover, as the need for timely and accurate landscape information requires optimization in learning systems and applications, the results obtained indicate that the low-cost semi-automated classification scheme developed is an alternative to automating the crucial steps of LULC mapping, continuing advances in the science and engineering of automation.

References

1. Soterroni, A.; Ramos, F.; Mosnier, A.; Fargione, J.; Andrade, P.; Baumgarten, L.; Pirker, J.; Obersteiner, M.; Kraxner, F.; Câmara, G.; et al. Expanding the Soy Moratorium to Brazil's Cerrado. *Sci. Adv.* **2019**, *5*, eaav7336. [CrossRef] [PubMed]
2. Zalles, V.; Hansen, M.; Potapov, P.; Stehman, S.; Tyukavina, A.; Pickens, A.; Song, X.P.; Adusei, B.; Okpa, C.; Aguilar, R.; et al. Near doubling of Brazil's intensive row crop area since 2000. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 428–435. [CrossRef] [PubMed]
3. Becker-Reshef, I.; Justice, C.; Barker, B.; Humber, M.; Rembold, F.; Bonifacio, R.; Zappacosta, M.; Budde, M.; Magadzire, T.; Shitote, C.; et al. Strengthening agricultural decisions in countries at risk of food insecurity: The GEOGLAM Crop Monitor for Early Warning. *Remote Sens. Environ.* **2020**, *237*, 111553. [CrossRef]
4. Naikoo, M.; Rihan, M.; Ishtiaque, M. Analyses of land use land cover (LULC) change and built-up expansion in the suburb of a metropolitan city: Spatio-temporal analysis of Delhi NCR using landsat datasets. *J. Urban Manag.* **2020**, *9*, 347–359. [CrossRef]
5. Baker, E.; Cappato, A.; Todeschini, S.; Tamellini, L.; Sangalli, G.; Reali, A.; Manenti, S. Combining the Morris method and multiple error metrics to assess aquifer characteristics and recharge in the lower Ticino Basin, in Italy. *J. Hydrol.* **2022**, *614*, 128536. [CrossRef]
6. Szantoi, Z.; Geller, G.; Tsendbazar, N.E.; See, L.; Griffiths, P.; Fritz, S.; Gong, P.; Herold, M.; Mora, B.; Obregón, A. Addressing the need for improved land cover map products for policy support. *Environ. Sci. Policy* **2020**, *112*, 28–35. [CrossRef]
7. Wulder, M.; Roy, D.; Radeloff, V.; Loveland, T.; Anderson, M.; Johnson, D.; Healey, S.; Zhu, Z.; Scambos, T.; Pahlevan, N.; et al. Fifty years of Landsat science and impacts. *Remote Sens. Environ.* **2022**, *280*, 113195. [CrossRef]
8. Khatami, R.; Mountrakis, G.; Stehman, S. A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research. *Remote Sens. Environ.* **2016**, *177*, 89–100. [CrossRef]
9. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]
10. Maxwell, A.; Warner, T.; Fang, F. Implementation of machine-learning classification in remote sensing: An applied review. *Int. J. Remote Sens.* **2018**, *39*, 2784–2817. [CrossRef]
11. Pandey, P.; Koutsias, N.; Petropoulos, G.; Srivastava, P.; Ben Dor, E. Land use/land cover in view of earth observation: Data sources, input dimensions, and classifiers—A review of the state of the art. *Geocarto Int.* **2021**, *36*, 957–988. [CrossRef]
12. Sudmanns, M.; Tiede, D.; Lang, S.; Bergstedt, H.; Trost, G.; Augustin, H.; Baraldi, A.; Blaschke, T. Big Earth data: Disruptive changes in Earth observation data management and analysis? *Int. J. Digit. Earth* **2020**, *13*, 832–850. [CrossRef]
13. Yu, M.; Bambacus, M.; Cervone, G.; Clarke, K.; Duffy, D.; Huang, Q.; Li, J.; Li, W.; Li, Z.; Liu, Q.; et al. Spatiotemporal event detection: A review. *Int. J. Digit. Earth* **2020**, *13*, 1339–1365. [CrossRef]
14. Zeng, L.; Wardlow, B.; Xiang, D.; Hu, S.; Li, D. A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sens. Environ.* **2020**, *237*, 111511. [CrossRef]
15. Chaves, M.; Picoli, M.; Sanches, I. Recent applications of Landsat 8/OLI and Sentinel-2/MSI for land use and land cover mapping: A systematic review. *Remote Sens.* **2020**, *12*, 3062. [CrossRef]
16. Almeida, C.; Coutinho, A.; Esquerdo, J.; Adami, M.; Venturieri, A.; Diniz, C.; Dessay, N.; Durieux, L.; Gomes, A. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. *Acta Amaz.* **2016**, *46*, 291–302. [CrossRef]
17. Radoux, J.; Chomé, G.; Jacques, D.; Waldner, F.; Bellemans, N.; Matton, N.; Lamarche, C.; D'Andrimont, R.; Defourny, P. Sentinel-2's potential for sub-pixel landscape feature detection. *Remote Sens.* **2016**, *8*, 488. [CrossRef]
18. Defourny, P.; Bontemps, S.; Bellemans, N.; Cara, C.; Dedieu, G.; Guzzonato, E.; Hagolle, O.; Inglada, J.; Nicola, L.; Rabaute, T.; et al. Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. *Remote Sens. Environ.* **2019**, *221*, 551–568. [CrossRef]
19. Xie, Q.; Dash, J.; Huang, W.; Peng, D.; Qin, Q.; Mortimer, H.; Casa, R.; Pignatti, S.; Laneve, G.; Pascucci, S.; et al. Vegetation Indices Combining the Red and Red-Edge Spectral Information for Leaf Area Index Retrieval. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 1482–1492. [CrossRef]
20. Fernández-Manso, A.; Fernández-Manso, O.; Quintano, C. SENTINEL-2A red-edge spectral indices suitability for discriminating burn severity. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *50*, 170–175. [CrossRef]
21. Bolton, D.; Gray, J.; Melaas, E.; Moon, M.; Eklundh, L.; Friedl, M. Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. *Remote Sens. Environ.* **2020**, *240*, 111685. [CrossRef]
22. Sanchez, A.; Picoli, M.; Camara, G.; Andrade, P.; Chaves, M.; Lechler, S.; Soares, A.; Marujo, R.; Simões, R.; Ferreira, K.; et al. Comparison of Cloud cover detection algorithms on sentinel–2 images of the amazon tropical forest. *Remote Sens.* **2020**, *12*, 1284. [CrossRef]

23. Appel, M.; Pebesma, E. On-Demand Processing of Data Cubes from Satellite Image Collections with the gdalcubes Library. *Data* **2019**, *4*, 92. [CrossRef]

24. Ferreira, K.; Queiroz, G.; Vinhas, L.; Marujo, R.; Simoes, R.; Picoli, M.; Camara, G.; Cartaxo, R.; Gomes, V.; Santos, L.; et al. Earth observation data cubes for Brazil: Requirements, methodology and products. *Remote Sens.* **2020**, *12*, 4033. [CrossRef]

25. Teluguntla, P.; Thenkabail, P.; Oliphant, A.; Xiong, J.; Gumma, M.; Congalton, R.; Yadav, K.; Huete, A. A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. *ISPRS J. Photogramm. Remote Sens.* **2018**, *144*, 325–340. [CrossRef]

26. Maso, J.; Zabala, A.; Serral, I.; Pons, X. A Portal Offering Standard Visualization and Analysis on top of an Open Data Cube for Sub-National Regions: The Catalan Data Cube Example. *Data* **2019**, *4*, 96. [CrossRef]

27. Hermosilla, T.; Wulder, M.; White, J.; Coops, N.; Hobart, G. Disturbance-Informed Annual Land Cover Classification Maps of Canada's Forested Ecosystems for a 29-Year Landsat Time Series. *Can. J. Remote Sens.* **2018**, *44*, 67–87. [CrossRef]

28. Hamunyela, E.; Verbesselt, J.; Herold, M. Using spatial context to improve early detection of deforestation from Landsat time series. *Remote Sens. Environ.* **2016**, *172*, 126–138. [CrossRef]

29. Chaves, M.; Alves, M.; Sáfadi, T.; de Oliveira, M.; Picoli, M.; Simoes, R.; Mataveli, G. Time-weighted dynamic time warping analysis for mapping interannual cropping practices changes in large-scale agro-industrial farms in Brazilian Cerrado. *Sci. Remote Sens.* **2021**, *3*, 100021. [CrossRef]

30. Chaves, M.; Soares, A.; Sanches, I.; Fronza, J. CBERS data cubes for land use and land cover mapping in the Brazilian Cerrado agricultural belt. *Int. J. Remote Sens.* **2021**, *42*, 8398–8432. [CrossRef]

31. Holmes, C.; Mohr, M.; Hanson, M.; Banting, J.; Smith, M.; Mathot, E. SpatioTemporal Asset Catalog (STAC) Specification: Making Geospatial Assets Openly Searchable and Crawlable. Available online: https://github.com/radiantearth/stac-spec (accessed on 12 February 2023).

32. Beuchle, R.; Grecchi, R.; Shimabukuro, Y.; Seliger, R.; Eva, H.; Sano, E.; Frédéric, A. Land cover changes in the Brazilian Cerrado and Caatinga biomes from 1990 to 2010 based on a systematic remote sensing sampling approach. *Appl. Geogr.* **2015**, *58*, 116–127. [CrossRef]

33. IBGE. Instituto Brasileiro de Geografia e Estatística. Sistema IBGE de Recuperação Automática (SIDRA): Produção Agrícola Municipal, Tabela 5457. 2022. Available online: https://sidra.ibge.gov.br/tabela/5457 (accessed on 23 November 2022). (In Portuguese)

34. Ribeiro, J.; Walter, B. Fitofisionomias do bioma Cerrado: Os biomas do Brasil. 1998. Available online: https://www.alice.cnptia.embrapa.br/handle/doc/554094 (accessed on 29 November 2022).

35. Lasantha, V.; Oki, T.; Tokuda, D. Data-Driven versus Köppen–Geiger Systems of Climate Classification. *Adv. Meteorol.* **2022**, *2022*, 3581299. [CrossRef]

36. Campos, R.; Pires, G.; Costa, M. Soil Carbon Sequestration in Rainfed and Irrigated Production Systems in a New Brazilian Agricultural Frontier. *Agriculture* **2020**, *10*, 156. [CrossRef]

37. Schwieder, M.; Leitão, P.; Bustamante, M.; Ferreira, L.; Rabe, A.; Hostert, P. Mapping Brazilian savanna vegetation gradients with Landsat time series. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 361–370. [CrossRef]

38. AIBA. Associação de Agricultores e Irrigantes da Bahia. Harvest Yearbook for Western Bahia: Harvest Season 2018–2019. 2022. Available online: https://aiba.org.br/wp-content/uploads/2020/08/Anuario-Ingles-Digital.pdf (accessed on 10 December 2022).

39. Frampton, W.; Dash, J.; Watmough, G.; Milton, E. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 83–92. [CrossRef]

40. Soares, A.; Chaves, M.; Fronza, J. Surface Reflectance to Vegetation Indexes (sr2vgi). 2020. Available online: https://github.com/brazil-data-cube/sr2vgi (accessed on 18 December 2022).

41. Immitzer, M.; Vuolo, F.; Atzberger, C. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sens.* **2016**, *8*, 166. [CrossRef]

42. Kolecka, N.; Ginzler, C.; Pazur, R.; Price, B.; Verburg, P. Regional scale mapping of grassland mowing frequency with Sentinel-2 time series. *Remote Sens.* **2018**, *10*, 1221. [CrossRef]

43. Wang, L.; Qu, J. NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. *Geophys. Res. Lett.* **2007**, *34*, 1–5. [CrossRef]

44. Gitelson, A.; Merzlyak, M. Spectral Reflectance Changes Associated with Autumn Senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *J. Plant Physiol.* **1994**, *143*, 286–292. [CrossRef]

45. Jasper, J.; Reusch, S.; Link, A. Active sensing of the N status of wheat using optimized wavelength combination: Impact of seed rate, variety and growth stage. *Precis. Agric.* **2009**, *9*, 23–30. [CrossRef]

46. Chen, P.; Tremblay, N.; Wang, J.; Vigneault, P.; Huang, W.; Li, B. New index for crop canopy fresh biomass estimation. *Spectrosc. Spectr. Anal.* **2010**, *30*, 512–517. [CrossRef]

47. Gitelson, A.; Kaufman, Y.; Stark, R.; Rundquist, D. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* **2002**, *80*, 76–87. [CrossRef]

48. Swathandran, S.; Aslam, M. Assessing the role of SWIR band in detecting agricultural crop stress: A case study of Raichur district, Karnataka, India. *Environ. Monit. Assess.* **2019**, *191*, 1–10. [CrossRef]

49. Niazmardi, S.; Homayouni, S.; Safari, A.; McNairn, H.; Shang, J.; Beckett, K. Histogram-based spatio-temporal feature classification of vegetation indices time-series for crop mapping. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *72*, 34–41. [CrossRef]

50. Osgouei, P.; Kaya, S.; Sertel, E.; Alganci, U. Separating built-up areas from bare land in mediterranean cities using Sentinel-2A imagery. *Remote Sens.* **2019**, *11*, 345. [CrossRef]

51. Kanke, Y.; Tubana, B.; Dalen, M.; Harrell, D. Evaluation of red and red-edge reflectance-based vegetation indices for rice biomass and grain yield prediction models in paddy fields. *Precis. Agric.* **2016**, *17*, 507–530. [CrossRef]

52. Li, S.; Ding, X.; Kuang, Q.; Ata-UI-Karim, S.; Cheng, T.; Liu, X.; Tian, Y.; Zhu, Y.; Cao, W.; Cao, Q. Potential of UAV-based active sensing for monitoring rice leaf nitrogen status. *Front. Plant Sci.* **2018**, *9*, 1834. [CrossRef]

53. Kross, A.; McNairn, H.; Lapen, D.; Sunohara, M.; Champagne, C. Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *34*, 235–248. [CrossRef]

54. Sonobe, R.; Yamaya, Y.; Tani, H.; Wang, X.; Kobayashi, N.; Mochizuki, K. Crop classification from Sentinel-2-derived vegetation indices using ensemble learning. *J. Appl. Remote Sens.* **2018**, *12*, 026019. [CrossRef]

55. Esquerdo, J.C.D.M.; Antunes, J.F.G.; Coutinho, A.C.; Speranza, E.A.; Kondo, A.A.; dos Santos, J.L. SATVeg: A web-based tool for visualization of MODIS vegetation indices in South America. *Comput. Electron. Agric.* **2020**, *175*, 105516. [CrossRef]

56. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

57. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

58. Nguyen, L.; Joshi, D.; Clay, D.; Henebry, G. Characterizing land cover/land use from multiple years of Landsat and MODIS time series: A novel approach using land surface phenology modeling and random forest classifier. *Remote Sens. Environ.* **2020**, *238*, 111017. [CrossRef]

59. Strassburg, B.; Brooks, T.; Feltran-Barbieri, R.; Iribarrem, A.; Crouzeilles, R.; Loyola, R.; Latawiec, A.; Oliveira Filho, F.; Scaramuzza, C.; Scarano, F.; et al. Moment of truth for the Cerrado hotspot. *Nat. Ecol. Evol.* **2017**, *1*, 0099. [CrossRef] [PubMed]

60. Bragança, A. The economic consequences of the agricultural expansion in Matopiba. *Rev. Bras. De Econ.* **2018**, *72*, 161–185. [CrossRef]

61. Ferreira-Paiva, L.; Suela, A.; Alfaro-Espinoza, E.; Cardona-Casas, N.; Valente, D.; Neves, R. A k-means-based-approach to analyze the emissions of GHG in the municipalities of MATOPIBA region, Brazil. *IEEE Lat. Am. Trans.* **2022**, *20*, 2339–2345. [CrossRef]

62. Carneiro Filho, A.; Costa, K. *A Expansão da Soja no Cerrado: Caminhos para a Ocupação Territorial, Uso do Solo e Produção Sustentável*; INPUT do Brasil, São Paulo: São Paulo, Brazil, 2016.

63. Marengo, J.; Jimenez, J.; Espinoza, J.; Cunha, A.; Aragão, L. Increased climate pressure on the agricultural frontier in the Eastern Amazonia–Cerrado transition zone. *Sci. Rep.* **2022**, *12*, 457. [CrossRef]

64. Müller, H.; Rufin, P.; Griffiths, P.; Siqueira, A.; Hostert, P. Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape. *Remote Sens. Environ.* **2015**, *156*, 490–499. [CrossRef]

65. De Beurs, K.; Henebry, G.; Owsley, B.; Sokolik, I. Using multiple remote sensing perspectives to identify and attribute land surface dynamics in Central Asia 2001-2013. *Remote Sens. Environ.* **2015**, *170*, 48–61. [CrossRef]

66. Franzluebbers, A.; Sawchik, J.; Taboada, M. Agronomic and environmental impacts of pasture-crop rotations in temperate North and South America. *Agric. Ecosyst. Environ.* **2014**, *190*, 18–26. [CrossRef]

67. Lu, L.; Kuenzer, C.; Wang, C.; Guo, H.; Li, Q. Evaluation of three MODIS-derived vegetation index time series for dryland vegetation dynamics monitoring. *Remote Sens.* **2015**, *7*, 7597–7614. [CrossRef]

68. Pivello, V. The use of fire in the Cerrado and Amazonian rainforests of Brazil: Past and present. *Fire Ecol.* **2011**, *7*, 24–39. [CrossRef]

69. Mataveli, G.; Silva, M.; França, D.; Brunsell, N.; de Oliveira, G.; Cardozo, F.; Bertani, G.; Pereira, G. Characterization and Trends of Fine Particulate Matter (PM2.5) Fire Emissions in the Brazilian Cerrado during 2002–2017. *Remote Sens.* **2019**, *11*, 2254. [CrossRef]

70. Schmidt, I.; Eloy, L. Fire regime in the Brazilian Savanna: Recent changes, policy and management. *Flora* **2020**, *268*, 151613. [CrossRef]

71. Picoli, M.; Camara, G.; Sanches, I.; Simões, R.; Carvalho, A.; Maciel, A.; Coutinho, A.; Esquerdo, J.; Antunes, J.; Begotti, R.; et al. Big earth observation time series analysis for monitoring Brazilian agriculture. *ISPRS J. Photogramm. Remote Sens.* **2018**, *145*, 328–339. [CrossRef]

72. Petitjean, F.; Inglada, J.; Gancarski, P. Satellite Image Time Series Analysis Under Time Warping. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3081–3095. [CrossRef]

73. Maus, V.; Câmara, G.; Appel, M.; Pebesma, E. dtwSat: Time-Weighted Dynamic Time Warping for Satellite Image Time Series Analysis in R. *J. Stat. Softw.* **2019**, *88*, 1–31. [CrossRef]

74. Ndao, B.; Leroux, L.; Diouf, A.; Soti, V.; Sambou, B. A Remote Sensing Based Approach for Optimizing the Sampling Strategies in Crop Monitoring and Crop Yield Estimation Studies. In *Earth Observations and Geospatial Science in Service of Sustainable Development Goals*; Springer: Cham, Switzerland, 2019; pp. 25–36. [CrossRef]

75. Zhang, N.; Hong, Y.; Qin, Q.; Liu, L. VSDI: A visible and shortwave infrared drought index for monitoring soil and vegetation moisture based on optical remote sensing. *Int. J. Remote Sens.* **2013**, *34*, 4585–4609. [CrossRef]

76. Camara, G.; Soterroni, A.; Ramos, F.; Carvalho, A.; Andrade, P.; Souza, R.; Mosnier, A.; Mant, R.; Buurman, M.; Pena, M.; et al. Modelling Land Use Changes in Brazil: 2000–2050: A Report by the REDD-PAC Project; Available online: http://www.redd-pac. org/reports/lucc_brazil_folder.pdf (accessed on 12 December 2022).