

Article

Maximum Fraction Images Derived from Year-Based Project for On-Board Autonomy-Vegetation (PROBA-V) Data for the Rapid Assessment of Land Use and Land Cover Areas in Mato Grosso State, Brazil

Henrique Luis Godinho Cassol ^{1,*}, Egidio Arai ¹, Edson Eyji Sano ², Andeise Cerqueira Dutra ¹ , Tânia Beatriz Hoffmann ¹ and Yosio Edemir Shimabukuro ¹

¹ National Institute for Space Research (INPE), Av. dos Astronautas, 1758, CEP: 12.227-010 São José dos Campos, SP, Brazil; egidio.arai@inpe.br (E.A.); andeise.dutra@inpe.br (A.C.D.); tania.hoffmann@inpe.br (T.B.H.); yosio.shimabukuro@inpe.br (Y.E.S.)

² Embrapa Cerrados, BR-020, km 18, CEP: 73301-970 Planaltina, DF, Brazil; edson.sano@embrapa.br

* Correspondence: henrique.cassol@inpe.br

Received: 16 March 2020; Accepted: 28 April 2020; Published: 2 May 2020



Abstract: This paper presents a new approach for rapidly assessing the extent of land use and land cover (LULC) areas in Mato Grosso state, Brazil. The novel idea is the use of an annual time series of fraction images derived from the linear spectral mixing model (LSMM) instead of original bands. The LSMM was applied to the Project for On-Board Autonomy-Vegetation (PROBA-V) 100-m data composites from 2015 (~73 scenes/year, cloud-free images, in theory), generating vegetation, soil, and shade fraction images. These fraction images highlight the LULC components inside the pixels. The other new idea is to reduce these time series to only six single bands representing the maximum and standard deviation values of these fraction images in an annual composite, reducing the volume of data to classify the main LULC classes. The whole image classification process was conducted in the Google Earth Engine platform using the pixel-based random forest algorithm. A set of 622 samples of each LULC class was collected by visual inspection of PROBA-V and Landsat-8 Operational Land Imager (OLI) images and divided into training and validation datasets. The performance of the method was evaluated by the overall accuracy and confusion matrix. The overall accuracy was 92.4%, with the lowest misclassification found for cropland and forestland (<9% error). The same validation data set showed 88% agreement with the LULC map made available by the Landsat-based MapBiomass project. This proposed method has the potential to be used operationally to accurately map the main LULC areas and to rapidly use the PROBA-V dataset at regional or national levels.

Keywords: spectral unmixing; machine learning; fraction images; cloud computing

1. Introduction

The Mato Grosso state presents the second largest area of deforestation and forest degradation in the Brazilian Legal Amazonia (BLA), a Brazilian political division that encompasses the states of Acre, Amapá, Amazonas, Mato Grosso, Pará, Rondônia, Roraima, Tocantins, and part of Maranhão [1,2]. Mato Grosso is the main grain and cattle beef producer in the BLA [3,4]. Its northern part is covered by the Amazonia biome, while the southern part is covered by the Cerrado and Pantanal biomes. This state has a key role in understanding the Brazilian land use and land cover (LULC) dynamics, since several processes of change in LULC that are occurring in Mato Grosso are being repeated in other Brazilian states covered by the Amazonia and Cerrado biomes. These are the cases of the states of Pará,

Rondônia, and Amazonas, as well as the newest agricultural frontier in Brazil, the MATOPIBA region (states of Maranhão, Tocantins, Piauí, and Bahia) [5]. The negative outcomes of this large expansion in food production are habitat fragmentation, biodiversity reduction, and high pressure on protected areas (conservation units and indigenous lands) [6,7]. Therefore, a quick assessment of LULC changes in all the above-mentioned regions is important for forecasting yield, planning logistics (transportation and storage), preventing fire, and combating and controlling illegal deforestation [6,8,9]. LULC monitoring is also crucial to estimate CO₂ emissions from deforestation and degradation losses [10,11], as well as to prevent soil erosion, the introduction of invasive species, and resource losses such as timber and water supply for local inhabitants [12,13].

Various moderate spatial resolution (~30 m) remote sensing data have been widely applied to map and monitor LULC changes at the global scale. These are the cases of China's GlobeLand30 [14] and the University of Maryland's Global Land Analysis and Discovery (GLAD) [15]. Recently, a pool of Brazilian universities and nongovernment and private institutions launched the MapBiomas project [16] to produce year-based LULC maps of Brazil using cloud-computing resources and historical series of Landsat data available in the Google Earth Engine platform. These initiatives use remote sensing data with moderate spatial and temporal resolutions, that is, Landsat-like satellites, typically with 16-day repeat passes and 30-m spatial resolution. Especially in tropical regions, there is well-known persistent cloud coverage. For example, in Brazil, only 20% of its territory presents more than 12 cloud-free Landsat observations in a year (within a total of 23 scenes) [17]. The limited number of Landsat scenes is crucial not only in the Amazonia biome but also in the Cerrado biome because of its high climatic seasonality and the presence of a six-month wet season (from October to March) [18], in which most of the grain production is concentrated.

Thus, at a 16-day temporal resolution, Landsat images may not be sufficient to fully explore the variations in vegetation phenology and the temporal patterns of agricultural lands due to cloud cover constraints [19]. In this sense, satellite sensors with relatively coarse spatial resolution (~100–300 m) and high temporal resolution may be a singular alternative to monitor LULC changes in regions such as Mato Grosso state in Brazil. These sensors include the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor onboard the Terra and Aqua platforms, with spatial resolutions of 250, 500, and 1000 m and 1–2 days of repeat pass [20]; the Medium Resolution Imaging Spectrometer (MERIS) sensor onboard the Envisat platform with a spatial resolution of 300 m and three-day repeat pass [21]; and the Project for On-Board Autonomy-Vegetation (PROBA-V), with 100-m or 1-km spatial resolutions and daily observations [22]. Recently, the combination of sensors, such as Landsat and Sentinel-2 [23,24], has been a good alternative to increase the availability of temporal images, allowing an improved characterization of land surfaces. However, so far, application of this sensor combination has been limited in the detection of burned areas [24] or in the monitoring of agricultural crop cycles [23].

Most of the LULC mapping from large areas has relied on supervised or unsupervised classification algorithms applied to the entire pixel (resolution cell). The use of the entire pixel to discriminate LULC classes in coarse spatial-resolution images (and even in moderate-resolution images) is, to some extent, limited because of the frequent mixture of different spectral classes in a single pixel [25,26]. Spectral mixture analysis [27], often modeled using a linear spectral mixture model (LSMM), is one of the most common techniques for considering the nature of the obtained mixed spectral data [26]. LSMM converts the digital numbers or reflectance of an image to a few endmembers or fraction images containing proportions of different terrain components, usually vegetation, soil, and shade, at a subpixel level [27]. LSMM has been demonstrated to be a useful technique for interpreting coarse and moderate spatial-resolution remote sensing data by converting spectral information into physical parameters related to the abundance of materials on the ground surface [28,29]. The LSMM technique has been used for detecting seasonal changes in three LULC classes (open tropical forest, Cerrado grasslands, and soybean plantation) [30] to differentiate the major Cerrado physiognomies [31] and vegetation seasonality [32] and to monitor deforestation and degradation by fire and selective logging [33].

Annual or monthly composite images based on the maximum vegetation fraction values ensure that cloudy and off-nadir pixels have low chances to be selected in the final composites, as stated by [30] and allow better characterization of LULC classes with strong seasonality (e.g., tropical savannas, croplands, and pasturelands), as shown by [34]. Since different LULC classes in Mato Grosso state present distinct annual behavior in terms of the proportion of vegetation, soil, and shade, in theory, the maximum fraction images (and their corresponding standard deviations) have the potential to better discriminate the most representative LULC classes from this region (forestlands, shrublands, grasslands, rain-fed croplands, and pasturelands). In this context, this study aimed to present a new method to map the major LULC classes in Mato Grosso state based on the maximum value proportional fractions of vegetation, soil, and shade derived from the 2015 PROBA-V 100-m images.

2. Materials and Methods

2.1. Study Area

The study area corresponded to Mato Grosso state, located in midwestern Brazil in the BLA region. This area covers approximately 900,000 km², is characterized by a heterogeneous landscape (highlands, lowlands, wetlands, depressions, and plateaus), and encompasses three Brazilian biomes: Amazonia (tropical rainforest, 56.7% of the state), Cerrado (tropical savanna, 37.4%), and Pantanal (wetland, 5.9%) (Figure 1) [35]. Dense and open ombrophilous forests are found in the northern part of the state (Amazonia); savannas with diverse proportions of grasses, shrubs, and trees are found in the plateaus and lowlands in the central and southern parts of the state (Cerrado); and wetlands are found in the southwestern part of the state (Pantanal) [35]. These heterogeneous natural and anthropogenic landscapes make the LULC classification quite challenging. In addition, Mato Grosso is the largest producer and exporter of soybeans, corn, and cotton in Brazil [36]. There are also extensive areas of pasture, reforestation by *Eucalyptus*, and sparsely distributed forest logging [3,5,37,38].

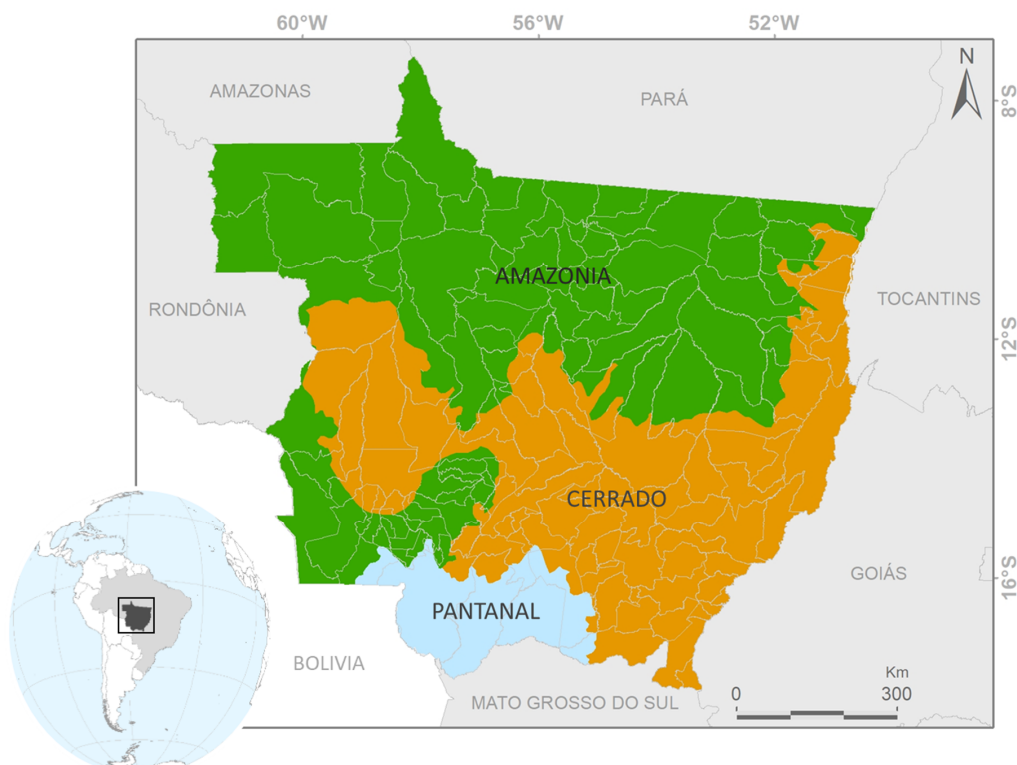


Figure 1. Location of the study area (Mato Grosso state) in Brazil (inset figure) and its coverage by three Brazilian biomes: Amazonia (rainfall forest, 57% of the state's area), Cerrado (tropical savanna, 37%), and Pantanal (wetland, 6%). The blank lines are the boundaries of municipalities. Source: [35].

2.2. Datasets

This study was based on the PROBA-V Collection 1 (C1) top of canopy daily synthesis product obtained in the entire year of 2015. PROBA-V C1 images are produced every five days with a spatial resolution of 100 m at nadir and with radiometric and geometric corrections [39]. PROBA-V is the payload of a satellite mission to map land cover and vegetation and was designed to ensure continuity of the Satellite Pour l'Observation de la Terre Vegetation (SPOT-VEGETATION) mission [22,40]. The sensor operates with four bands in the visible and infrared spectra (blue, 0.435–0.490 μm ; red, 0.610–0.700 μm ; near infrared (NIR), 0.760–0.930 μm ; and shortwave infrared (SWIR), 1.520–1.670 μm). The PROBA-V mission also provides the status map (SM) information (pixel radiometric quality data) and the normalized difference data between the NIR and red bands, called the normalized difference vegetation index (NDVI) (not used in this study).

As ancillary data, we selected two sets of Landsat-8 Operational Land Imager (OLI) satellite data obtained in 2015 (bands 4, 5, and 6 in the green, red, and NIR spectral regions, respectively; spatial resolution of 30 m; repeat pass of 16 days). One set was from the crop-growing season (July–December), and the other set was from the harvesting season (January–July). The MODIS-based time series of NDVI signatures available in the Satellite Vegetation (SATVeg) web tool, developed by the Brazilian Agricultural Research Corporation (Embrapa)¹, were also used in this study. The LULC map of 2015, produced by the MapBiomass project, was the main vector-based data considered in this study.

2.3. Data Processing

The main steps of the methodological approach applied in this study and conducted in the Google Earth Engine platform [41] are shown in Figure 2. First, we imported the PROBA-V 100-m scenes from 2015 in the Google Earth Engine platform and clipped to the Mato Grosso state boundary [42]. The following bands were considered: Blue, red, NIR, SWIR, and SM. The annual mosaic of cloud-free images from 2015 was generated by a script for masking pixels whose values stated as cloud or cirrus in the SM band were excluded [40]; thus, only pixels with good radiometric quality, that is, classified in the SM band as 0 (clear), were selected.

Next, we selected homogeneous areas of bare soils, large croplands, and clear water bodies in the PROBA-V annual mosaic to obtain the pure endmembers of soil, vegetation, and shade, respectively. Then, the soil, vegetation, and shade fraction images were obtained by selecting endmembers by visual inspection of the median values of the blue, red, NIR, and SWIR bands. The representative reflectance values of each endmember in the blue, red, NIR, and SWIR spectral regions for Mato Grosso state are shown in Figure 3.

The LSMM assumes that pixel values are a linear combination of reflectance values from a certain number of elements (endmembers). The fraction image is obtained for each endmember representing the proportion of the corresponding endmember within a pixel by minimizing the sum of squares of errors in the original data [27]. The proportion of each element is constrained to the range from 0% (absence) to 100% (maximum presence) at each cell resolution. The LSMM can be represented by Equation (1):

$$R_i = \sum_{j=1}^m f_j r_{i,j} + \varepsilon_i \quad (i = 1, \dots, p) \quad (1)$$

where:

R_i = the spectral reflectance in the i th spectral band;

$r_{i,j}$ = the spectral reflectance of the j th component in the i th spectral band (endmember);

f_j = the proportion of the j th component within the pixel;

¹ Available in www.satveg.cnptia.embrapa.br. Accessed on 30 June 2019.

$i = 1, p$ (number of spectral bands used);
 $j = 1, m$ (number of considered components); and
 ε_i = the residual for the i th spectral band.

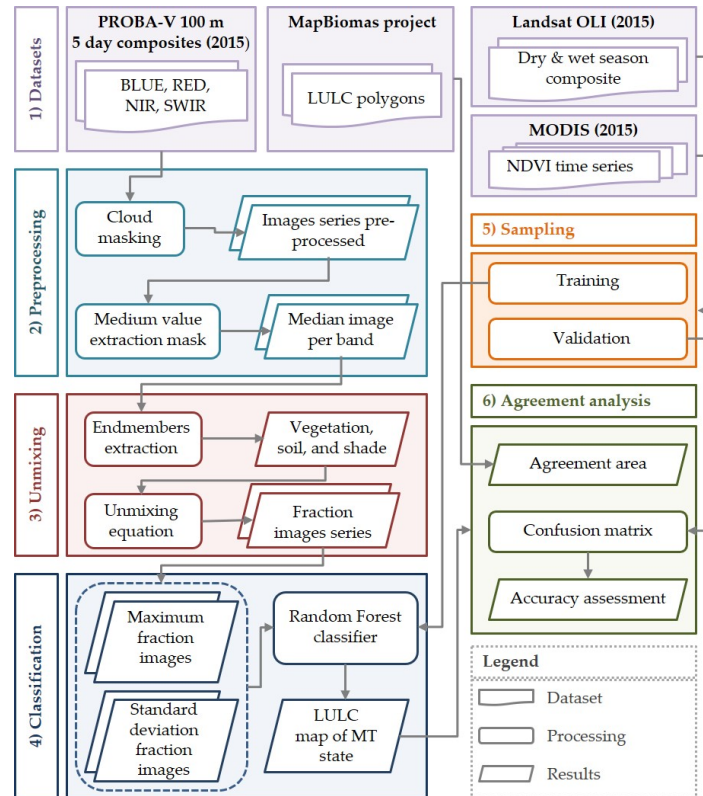


Figure 2. Processing steps of the land use and land cover classification of Mato Grosso state. NIR, near infrared; SWIR, shortwave infrared; LULC, land use and land cover; OLI, Operational Land Imager; MODIS, Moderate Resolution Imaging Spectroradiometer; NDVI, normalized difference vegetation index; and MT, Mato Grosso.

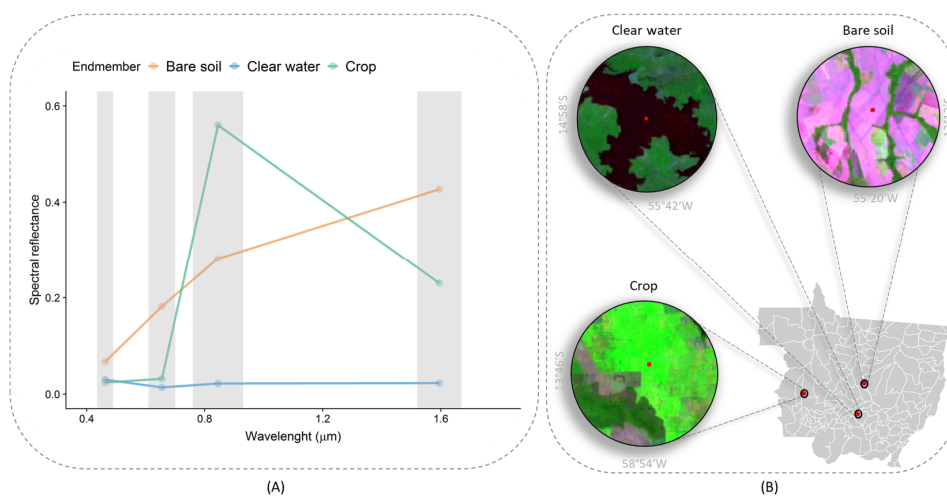


Figure 3. Endmembers collected from median reflectance values of Project for On-Board Autonomy–Vegetation (PROBA-V) annual composites (A) and their corresponding samples of “Crop”, “Bare soil”, and “Clear water” in the municipalities of Sapezal, Sorriso, and Chapada dos Guimarães (Mato Grosso state), respectively (B).

The set of PROBA-V images obtained in 2015 was reduced to three maximum and three standard deviation fraction images of vegetation, soil, and shade. This approach ensured that the highest fraction value was selected on the composites within the year [30]. Therefore, the information of one year of PROBA-V images (in theory, 73 cloud-free images) was condensed in six bands. The novel idea is that the highest vegetation fraction value refers to the most likely class, and the standard deviation of its fractions represents the seasonality of that class throughout the year. The standard deviation of the fraction images is calculated as the deviation from the average value of each fraction image in the annual mosaic composites.

2.4. LULC Description

The LULC classes considered in this study were annual croplands, forestlands, water bodies, humid savanna, pasturelands, and dry savanna. These classes were discriminated based on the variations in the proportions of the soil, vegetation, and shade endmembers, according to the seasonality, instead of analyzing their spectral behaviors, which invariably change with the sensor and spatial resolution [43]. This means that each class has a typical annual behavior according to its increase or decrease in vegetation, soil, and shade proportions throughout the year [30,31]. For instance, in Mato Grosso, croplands are mostly represented by annual, rain-fed plantations of soybean, maize, and cotton, while pasturelands are mostly represented by planted pastures with *Brachiaria* species. Therefore, the terrain conditions in croplands over the year are significantly more variable than those from planted pastures, ranging from green vegetation cover to bare soils or straws, which explains the higher seasonality of soil and vegetation fraction images throughout the year [30,31,44]. In other words, croplands have high vegetation cover, low soil cover, and low shade fraction during the peak of the growing season; however, during the harvesting and preseeding phases, croplands have low vegetation, high soil, and low shade fractions (Figure 4).

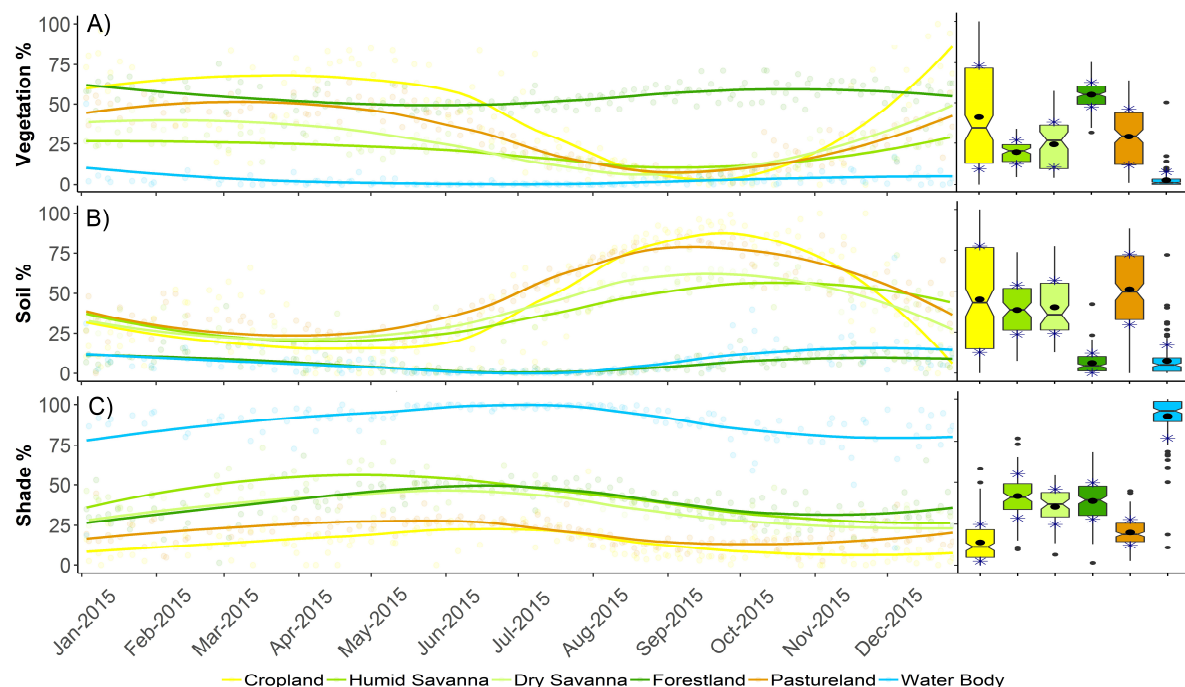


Figure 4. Annual time series of the linear spectral mixture model (LSMM) fraction images of vegetation (A), soil (B), and shade (C) of six representative land use and land cover classes of Mato Grosso state derived from the PROBA-V 100-m datasets. Figures on the right side are the boxplots of annual fraction images of croplands, humid savanna, dry savanna, forestland, pastureland, and water bodies. Filled dots and asterisks represent the mean and the standard deviation, respectively.

Forestlands include the Amazon tropical rainforest located in the northern part of Mato Grosso state, Cerrado woodlands, riparian forests, and forest regenerations with varying ages. They are characterized by deciduous and semideciduous forest with trees between 15 and 20 m height and no leaf losses greater than 50% at the Cerrado biome and by the open forest with trees between 25 and 30 m height in the Amazon forest, with canopy cover greater than 80% at the Amazon biome [45,46]. In addition, forestlands have a medium vegetation fraction, low soil proportion, medium shade fraction, and low standard deviation for all fractions throughout the year because of the weak seasonality, as shown in the boxplots of Figure 4. Water bodies include major rivers and perennial lakes from the wetlands of the Pantanal biome and present the highest shade proportion throughout the year (boxplot of Figure 4C) and almost null proportions of vegetation and soil (Figure 4A,B, respectively).

Humid savanna corresponds to the natural grass-dominated strata mostly found in the wetlands of the Pantanal biome, highly susceptible to flooding during the wet season (October to March), and occurs mainly in poorly drained soils [31]. Because of the sparse, up to 3-m-high trees covering less than 20% of the surface, these areas can be identified by a slightly higher fraction of shade compared with that of other vegetated classes, as shown in the boxplots of Figure 4C [45]. Dry savanna corresponds to the mixture of grass-, shrub-, and tree-dominated layers in varying proportions that occur mainly in the Cerrado biome [6,45]. Savannas and pasturelands usually present medium soil, vegetation, and shade fractions and moderate seasonality, which results in higher standard deviations than that of forestlands. Their location in Figure 4 is between the cropland and forestland classes. Additionally, a lower soil fraction on managed pastures is expected in relation to natural grasslands, as well as a lower shade fraction in relation to the humid savannas. Dry savannas are characterized by a higher proportion of shrubs and trees in relation to pasturelands and humid savannas, which may be characterized by an intermediate proportion of the soil and vegetation fraction [45].

2.5. Sampling Design and Classification

To ensure the representativeness of sampling, a set of 622 samples (348 in the Amazonia biome, 234 in the Cerrado, and 40 in the Pantanal) was randomly distributed over the study area by visual inspection (Figure 5).

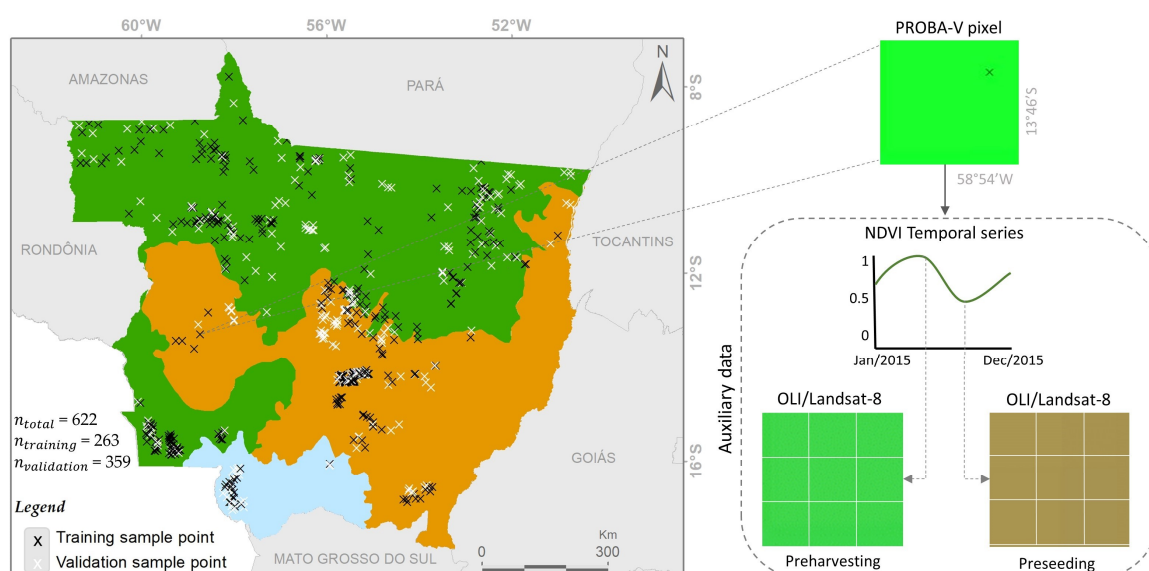


Figure 5. Location of 622 training and validation samples in the study area. In the upper right, the RGB (Red-Green-Blue) color composite of median PROBA-V data: Red, NIR, and blue of an agricultural sample, with respective conceptual behavior of the ancillary data from MODIS NDVI and the two seasons of the Landsat-8 OLI (RGB:654) in the bottom right.

In total, we selected 151 (84 for validation sampling) samples of croplands; 128 (56) forestland samples; 100 (36, 37, and 36) pastureland, humid, and dry savanna samples; and 43 (14) water body samples. The number of samples for water bodies was low because they occupy a relatively small area in Mato Grosso state. We collected samples only from perennial large rivers, streams, and lakes. For the dry savanna, the small samples were due to confusion with pastures, especially because that class comprises several other classes with a gradient of shrubs and grasslands. Training samples were chosen empirically by observing the sample distribution, accuracy assessment, and classification performance.

On the other hand, the number of validation samples was calculated based on the statistical approach proposed by [47], which consists of generating an adequate probability dataset among stratified random sampling. The sample size for validation was based on the extension area of each stratum and their expected accuracy to ensure sufficient size to produce sufficiently precise estimates. As there was no reference map with the same LULC classes for the study area, a previous classification map was developed using two-thirds of total training samples and one-third for testing (Figure 5) to obtain the expected overall accuracies by each LULC. Thus, the set parameters were 2% of the standard error of expected overall accuracy; 90% of the expected user's accuracy of the high-confidence accuracies for forestlands, croplands, and water bodies; and 70% of the expected user's accuracy to the low-confidence accuracies for humid savanna, pastureland, and dry savanna. Since some classes are identified more easily than others, this measure will influence the overall size of the samples.

Finally, the mosaics were processed by the random forest classifier in the Google Earth Engine platform. Random forest is a supervised classifier where each tree handles a random selection from the training data by creating a set of a multiple decision trees using bootstrap [48]. The random replacement of trees and samples makes the classifier robust for dealing with a low number of training samples [49]. At each bootstrap replacement, a new tree node is built by gaining information from the comparison of the multiple attributes (spectral bands). This procedure is repeated for the other nodes of the tree, which grow without pruning. Then, k number of trees is generated, forming the (random) forest [48]. The parameter k was set to 30, which is the maximum number of trees to be created.

The accuracy was evaluated by the confusion matrix and by the comparison of results obtained by the MapBiomass project (Table 1). The same validation dataset was used to assess the agreement with MapBiomass project classes according to [50,51]. As MapBiomass raster-based LULC products are available at 30-m spatial resolution, the data were degraded to 100-m spatial resolution by using the most likely class in the pixel (mode), enabling intercomparison between datasets (https://code.earthengine.google.com/?accept_repo=users/mapbiomas/user-toolkit). The legend consistency between this study and the MapBiomass project is shown in Table 1.

Table 1. Legend consistency between this study and the MapBiomass[†] land use and land cover mapping project.

This Study	MapBiomass Project
Croplands	Annual and perennial crop and semiperennial crop
Forestland	Forest formation and forest plantation
Water bodies	River, lake and ocean
Humid savanna	Mangrove, wetland, and other nonforest natural formation
Pastureland	Grassland formation, pasture, and mosaic of agriculture and pasture
Dry savanna	Savanna formation

[†] Collection version 3.1 (<https://mapbiomas.org/en>).

3. Results

3.1. Maximum and Standard Deviation Fractions of Vegetation, Soil, and Shade

The annual time-series composites of fraction images were reduced by three maximum fractions and by three standard deviation bands. Figure 6 shows the frequency histogram of each of the LULC classes in these six bands. In terms of vegetation fraction, forestland, cropland, and pastureland

classes presented relatively high responses due to the high absorption of photosynthetically active radiation by leaves and canopies throughout the year. However, these LULC classes differ by their seasonality, which is more evident in croplands and pasturelands than in forests, as noted by the high standard deviation of the vegetation fraction (Figure 6D). The forestland defined here is characterized by semideciduous forest and open or dense tropical forest with no pronounced winter, so the proportion of vegetation is quite constant throughout the year, with a medium proportion of shade in the dry season, as 20–50% of the leaves fall from May to September.

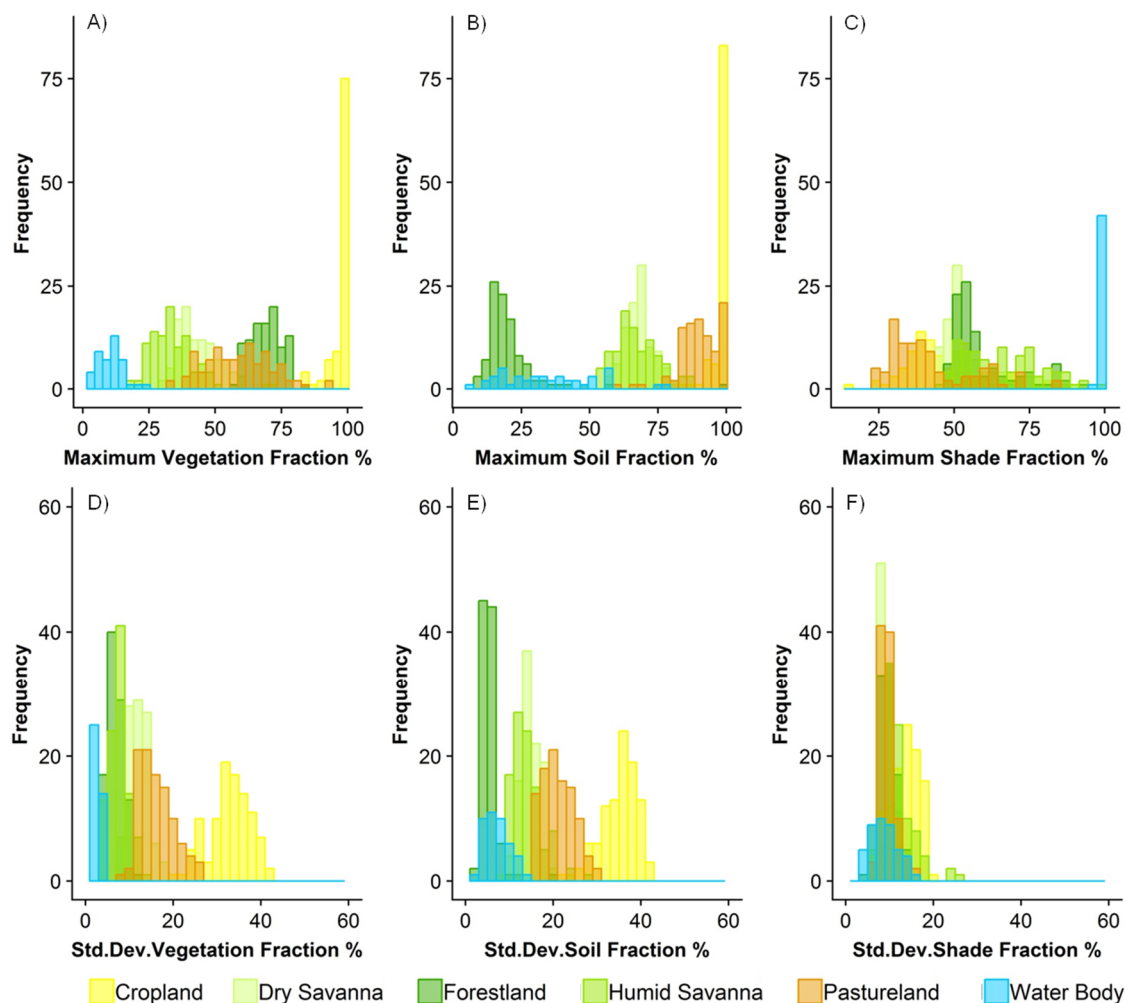


Figure 6. Maximum vegetation (A), maximum soil (B), and maximum shade (C) fractions of the one-year composite mosaic of the land use and land cover classes of Mato Grosso state and their corresponding standard deviations of the vegetation (D), soil (E), and shade (F) fractions.

The use of the maximum values of vegetation and soil fractions within a year allows a separation between cropland and forestland or between forestland and pastureland. The soil fraction presented the opposite behavior of maximum vegetation fraction in vegetated classes; the proportion of soil increases during the harvesting and seeding period when the soil becomes exposed, and the vegetation proportion decreases. At the end of the crop cycle (January–April), the maximum vegetation proportion occurs in the cropland class and vice versa in the soil fraction (Figure 6A,B).

Pastureland and savannas presented behavior similar to that of the maximum and standard deviation of the vegetation and soil fractions, however, with lower levels than those observed from the cropland (Figure 6A,B,D,E). Pastureland and savanna canopies are more heterogeneous than those from cropland, presenting higher levels of mixed vegetation, soil, and shade fractions during the year. Therefore, pastureland can be discriminated from forestland and cropland by the

intermediate values of the standard deviation of the vegetation and soil fractions (Figure 6C,D). Similarly, the maximum fraction of the shade image was helpful to differentiate pastureland from humid savanna and dry savanna because of the stronger seasonality of the pastureland. Pastureland presented higher seasonality of the soil fraction, as depicted in the standard deviation of the soil fraction in Figure 6E, than that from the humid savanna and dry savanna and lower maximum shade fractions. The humid savanna is differentiated by the relatively high amount of shade in the maximum annual composite. As the water bodies absorb most of the incident radiation in the visible and infrared spectra, they are clearly separated by their maximum shade fraction due to the low reflectance of this target; consequently, the water bodies also presented low proportions of vegetation and soil fractions.

Figure 7 shows the RGB (Red-Green-Blue) color composite of the maximum fractions of soil, vegetation, and shade fractions in Mato Grosso state. Water bodies (rivers and lakes) and forestland appear blueish and greenish in this false-color composite, respectively. Cropland has a similar proportion of maximum fractions of soil and vegetation throughout the year, so it appears yellowish.

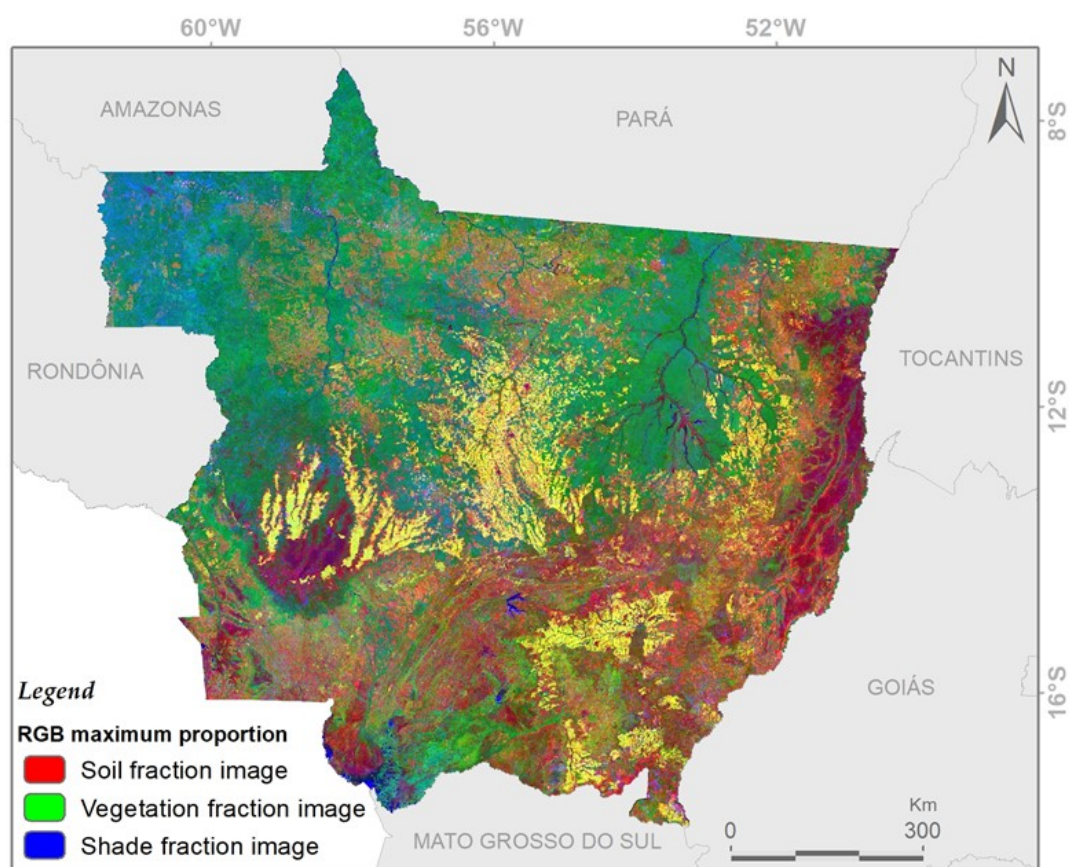


Figure 7. RGB color composite of the maximum soil, vegetation, and shade proportions derived from fraction images of Mato Grosso state. Cropland is highlighted in yellow, and forestland is highlighted in green.

3.2. LULC Map

Figure 8 shows the LULC map of Mato Grosso state derived from the random forest classification of the maximum and standard deviation fractions of soil, vegetation, and shade, in roughly 90 MP (megapixels). The areas occupied by each LULC class in the Amazonia, Cerrado, and Pantanal biomes are shown in Table 2.

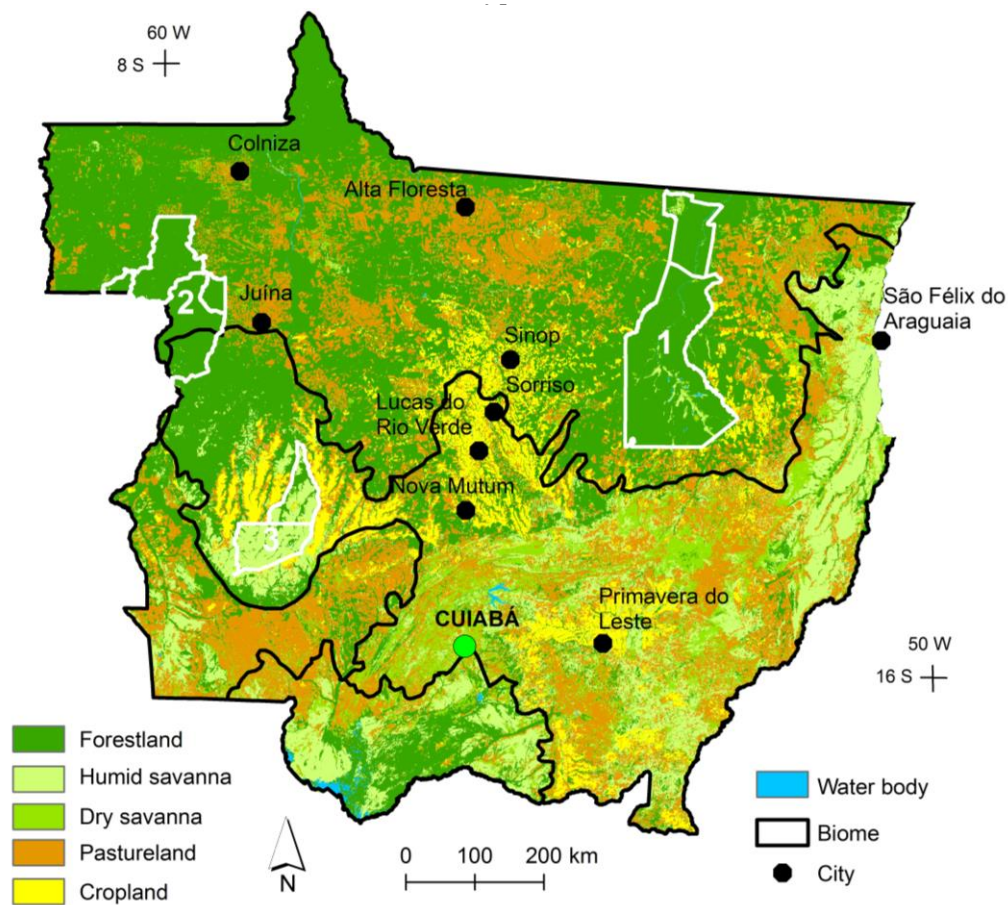


Figure 8. Land use and land cover map of Mato Grosso state derived from the maximum and standard deviation fraction images of 2015. Identification of indigenous lands: 1 = Parque do Xingu; 2 = Parque do Aripuanã; and 3 = Paresi. Cuiabá is the capital of Mato Grosso state.

Table 2. Area (km²) occupied by the representative LULC classes of Mato Grosso state, estimated according to the PROBA-V 100-m classification by the random forest algorithm.

LULC class	Amazonia	Cerrado	Pantanal	Subtotal
Croplands	14,178	25,111	175	39,464
Forestland	310,353	89,329	22,272	421,954
Water bodies	955	773	1701	3429
Humid savanna	17,406	93,790	20,679	131,875
Pastureland	124,133	92,772	11,304	228,208
Dry savanna	18,584	50,735	8016	77,336
Subtotal	485,609	352,510	64,147	902,266

The dominant LULC class in Mato Grosso is forestland, covering an area of approximately 422,000 km² (45% of the state), mostly in the northern and southern parts of the state, followed by pastureland (~228,000 km², 24% of the state), basically found throughout the state. The spatial distribution of forestland and pastureland agrees quite well with the MODIS-based LULC map of Mato Grosso from 2017 [52]. Forestland and pastureland are also the dominant LULC classes in the part of the state covered by the Amazonia biome: 64% and 26%, respectively. Forestland, humid savanna, and pastureland are the three dominant LULC classes in areas occupied by the Cerrado biome: 25%, 27%, and 26%, respectively. In the Pantanal biome, we found 35% forestland and 32% humid savanna, the two most dominant LULC types in this biome.

A large portion of forestland in Mato Grosso is protected by indigenous lands, especially by the Parque do Xingu and the Parque do Aripuanã indigenous lands in the northeastern and northwestern regions of Mato Grosso, as well as the Paresi indigenous land in the western part of the state (Cerrado biome). Surprisingly, there is a lack of permanently protected conservation units in Mato Grosso, which is commonly found in other states in the Amazonia biome. The cropland occurs mostly in the central part of the state, adjacent to the cities of Nova Mutum, Lucas do Rio Verde, Sorriso, and Sinop, in the Cerrado/Amazonia ecotone region, and along the BR-163 highway (Cuiabá-Santarém highway, not shown in Figure 8), where most of the cropland and forest losses due to clear cutting and selective logging are found [53,54]. Primavera do Leste is the major grain-producing region in the state, located in the Cerrado biome. Juína, Colniza, and Alta Floresta, mostly occupied by pasturelands, are important in terms of Amazon biodiversity conservation. These are priority regions of the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA) to combat and control illegal deforestation in the BLA since they are located near the Amazonas state, where we find large fragments of pristine rainfall forest. In the region of São Felix do Araguaia, we find the largest areas of humid savannas in Mato Grosso, located over the alluvial sediments of the Araguaia River basin [55,56]. In Brazil, the Araguaia River Basin and the Pantanal are the two largest wetlands influenced by tropical, seasonal floods.

3.3. Accuracy

The LULC map of the Mato Grosso state showed an overall accuracy of 92.4% for 2015, with user accuracy and producer accuracy above 78% (Table 3). The highest confusion was among dry savanna, humid savanna, and pastureland (~86–95% user accuracy), as such classes present a mixture of woods, shrubs, and grasslands [31].

Table 3. Confusion matrix among the major land use and land cover classes from Mato Grosso state using the validation dataset. The values in parentheses indicate the number of training and validation samples, respectively. UA, user accuracy; PA, producer accuracy. Overall accuracy = sum of diagonal/total number of samples = 243/263 = 92.4%.

Reference/ Classified	Cropland (67/84)	Forest-Land (72/56)	Water Body (29/14)	Humid Savanna (64/36)	Pasture-Land (63/37)	Dry Savanna (64/36)	Total (263/359)	UA
Cropland	76	0	0	0	8	0	84	90
Forestland	0	55	0	0	1	0	56	98
Water Body	0	0	14	0	0	0	14	100
Humid Savanna	0	0	0	31	0	5	36	86
Pastureland	0	0	0	0	35	2	37	95
Dry Savanna	0	0	0	3	1	32	36	89
Total	76	55	14	34	45	39	263	
PA	100	100	100	91	78	82		92

Water bodies presented 100% confidence, although we acknowledge that the spatial resolution of 100 m limits the classification of small rivers and lakes. For instance, only a few parts of the Xingu River in the Parque do Xingu indigenous land, corresponding to its largest riverbeds, could be mapped in this study (Figure 8). Croplands, forestlands, and water bodies obtained high user accuracy and producer accuracy >90%. Pasturelands, on the other hand, showed high user accuracy (95%) but low producer accuracy (78%).

4. Discussion

4.1. Classification Performance and Uncertainties

The proposed approach demonstrated that the highest value of the fraction images on the annual composite allowed the discrimination of LULC classes based on their highest peaks during the year, while standard deviation emphasized the seasonality of these classes, improving the separability of LULC classes with similar spectral responses. However, some misclassifications remained, for example, the confusion between pasturelands and croplands that has already been reported by several authors [5,6,17]. The managed pastures promote vegetation growth similar to that of the croplands. Such homogeneous vegetation covering confuses the classifier, hence reducing the accuracy of this class [5,6]. For other natural classes from the Cerrado biome, such as the dry and humid savannas reported here, the misinterpretation may be even worse. According to the Cerrado vegetation classification system proposed by [45], there are at least seven different phytophysiognomies in the Cerrado biome, which differ in terms of vegetation cover, the amount of biomass, and proportion of trees, shrubs, and grasses. In addition, the changes in LULC classes during the observed year can produce error in the classification, confusing the classifier when assigning the last change as a correct class. For instance, the conversion of forest in pasture could classify this class as forest due to the highest vegetation fraction in that year.

Compared with the validation sample design applied in the MapBiomias project, we noticed that cropland, forestland, water bodies, and pastureland presented similar user and producer accuracies (92–100%) (Table 4). However, dry and humid savannas presented high levels of mismatches (53–64%), probably due to the legend definitions. The overall accuracy classification performance for Mato Grosso state was similar to that reported by the MapBiomias project, which was 90% in 2015 [16]. The overall accuracy in terms of biome was 95% for the Amazonia, 83% for the Cerrado, and 81% for the Pantanal, demonstrating again the greater difficulty in mapping savanna formations (https://mapbiomas.org/en/estatistica-de-acuracia?cama_set_language=en).

Table 4. Confusion matrix among the major land use and land cover classes with the MapBiomias project.

Reference/ Classified	Cropland	Forestland	Water Body	Humid Savanna	Pastureland	Dry Savanna	Total	UA
Cropland	83	0	0	1	0	0	84	99
Forestland	0	56	0	0	0	0	56	100
Water Body	0	0	14	0	0	0	14	100
Humid Savanna	0	0	0	19	2	15	36	53
Pastureland	0	0	0	1	36	0	37	97
Dry Savanna	0	0	0	12	1	23	36	64
Total	83	56	14	33	39	38	263	
PA	100	100	100	58	92	61		88

The total area classification performance agreed with that of the previous study [16] for most classes (Table 5). One of the main differences between those mappings was that the extent of cropland estimates was 32%. Here, it is important to note that the MapBiomias maps do not differentiate between annual and perennial crops. Perennial crops are classified as forestlands in our classification. Our approach did not discriminate among perennial and semiperennial crops and annual crops because of the low seasonality of the perennial and semiperennial crops. If the semiperennial crops (sugarcane, for example) were excluded from the area comparison based on the MapBiomias project, the difference would be 10% higher. Furthermore, both initiatives have the limitation of not classifying more than one annual harvest cycle (typically, double cropping), a common practice in Mato Grosso state [34,57]. Other studies showed larger cropland estimates than those in our method: 98,488 km² [58] and 98,723 km² [59], in which double-cropped areas were counted twice [59]. To overcome this limitation, one possibility is the use of a crop calendar (wet-dry seasons), as proposed by [60], including the short

and long crop cycles in the estimates, or using a metric that clearly identifies two cycles in the same time series [19,52,57].

Table 5. Comparison of the areas mapped by this study and by the MapBiomass project.

Classes	This study (km ²)	MapBiomass (km ²)	Difference
Cropland	39,464	52,219	−32%
Forestland	421,954	432,923	−3%
Water Body	3429	17,711	−416%
Dry Savanna	131,875	97,944	26%
Pastureland	228,208	231,189	−1%
Humid Savanna	77,336	68,962	11%
Total	902,266	900,947	

The forestland showed a 3% difference between the two studies. In the MapBiomass project, the tree-dominated savannas, tree-dominated wetlands, and the Amazon rainforest were grouped in the forestland class. The area estimation of pastureland differed by only 1% in relation to the MapBiomass project, even though pasturelands are difficult to map because of the similar spectral response with the savanna grasslands and other land covers [5]. Farmers often use natural fields for cattle ranching, making it difficult to differentiate them from native savanna grassland [61]. In that manner, pastures were confused with dry savanna (Table 1). The commission error was better than the omission error, leading to a most likely overestimation of the area of pastures. Pasture conversion represented an average of 79% of new cropland area in Brazil and 20% of natural vegetation conversion during the 2000–2014 period, which represents approximately 8% of pasture conversion per year [59].

The greatest difference in area estimation was found in the water body class, which was 4.1 times lower than the MapBiomass project estimation. The main reason for this difference was probably related to the lower spatial resolution of the PROBA-V (100 m) in comparison to that of the Landsat collection (30 m) classification of the MapBiomass. Another possible reason is the underestimation in this study of the Pantanal biome because of the use of the maximum shade fraction in the annual composite, different from that of the other initiatives [e.g., 16], which perform classification using wet season datasets when there is a lack of data. Another Brazilian initiative, called TerraClass, computed 4747 km² of water bodies in Mato Grosso state in 2014 using Landsat mosaic composites and visual interpretation [62]. In this sense, the proposed approach minimizes the influence of the seasonality of the native vegetation of the savanna and wetlands by not determining single dry or wet periods, which, in turn, vary throughout Mato Grosso state [63] and may affect the area estimation in LULC mappings.

4.2. Novel Approach

Although the results have shown inconsistencies in the classification of certain classes, uncertainties may be associated with the heterogeneity of the landscape within the 100 m × 100 m cell of the sensor used, making it difficult to map small environmental changes that may not be detected due to the spatial resolution. On the other hand, the better temporal resolution increased the probability of finding pixels without the influence of clouds, making it possible to extract values at least once a year, which might not occur in sensors with lower temporal resolution.

Thus, from the point of view of quickly assessing the main LULC classes with the low computational cost of cloud computing and human resources, the proposed approach has the advantage of presenting a higher availability of cloud-free pixels. For example, Landsat-like sensors can have a maximum of 23 cloud-free pixels in a year and up to 60 in the Sentinel-2A and Sentinel-2B satellites (only Sentinel-2A scenes are available in Brazil, so far), while PROBA-V can obtain up to 73 synthetic 5-day composites [64]. In fact, these are hypothetical values that depend on the frequency of cloud cover, especially during the rainy season (October–April), which reduces data availability. Figure 9 shows the frequency of the cloud-free and cloud shadow-free pixels over Mato Grosso state in 2015 and

their comparison with the moderate spatial resolution. The average number of clear observations was 29 pixels without cloud and cloud shadows over the year (Figure 9B). In addition, there was a clear gradient of the spatial distribution of data availability from east to west (Figure 9A). In the east, a drier region with a predominance of Cerrado physiognomies, we found clearer observations than those in the west, which is characterized by dense Amazon forest (Figure 9A). Data availability is a limiting factor in reconstructing the phenological behavior of the main LULC classes in a time series classification [57]. Nevertheless, PROBA-V imagery has an excellent trade-off between data availability and spatial resolution, with a minimum of 20 “clear” pixels per year (Figure 9A).

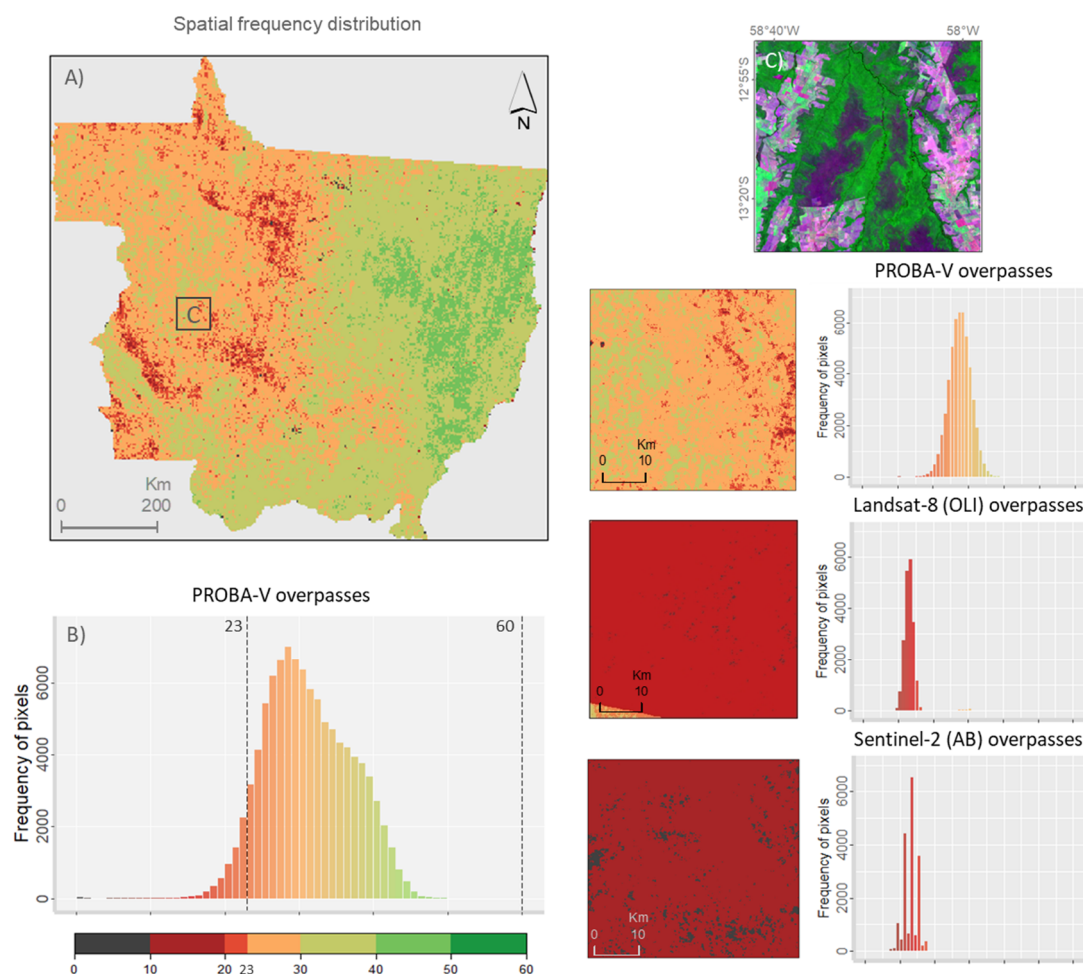


Figure 9. (A) Spatial distribution of cloud-free and cloud shadow-free pixels in an annual PROBA-V 100-m composite. (B) Histogram of the frequencies of cloud-free and cloud shadow-free pixels in Mato Grosso state. Vertical, dashed lines represent a hypothetical maximum number of pixels in the Landsat-8 OLI and Sentinel-2 collection. (C) Enlargement of a cropland-dominated region located in the Campo Novo do Parecis municipality. In the upper right, the red-NIR-SWIR color composite of PROBA-V. In the center right, the spatial distribution of data availability of a year of PROBA-V images and Sentinel-2 (bottom right).

4.3. Perspective

The method presented encouraging results for Mato Grosso state, but it may face some difficulty in other regions or in discriminating more detailed LULC classes. For example, the Cerrado biome is characterized by landscape patterns varying from open grassland to dense canopies; therefore, obtaining accurate LULC classification in this biome may be slightly challenging [65]. In this study, we grouped Cerrado physiognomies into dry savanna and humid savanna to produce LULC classes

with good performance. However, if there is a necessity to discriminate among the seven types of physiognomies (two grassland classes, three shrubland classes, and two forestland classes) according to the Cerrado vegetation classification system proposed by [45] or to precisely discriminate between natural grasslands and planted pasturelands, the accuracy of the method may decrease substantially because of the similarities in the spectral responses.

The use of spectral mixture models has been reported to estimate woody and herbaceous cover using the photosynthetic vegetation (PV) fraction in addition to the nonphotosynthetic vegetation (NPV) fraction [66,67], which can include materials such as dry organic matter, litter, wood, and dry logs. According to [67], using only the vegetation fraction can overestimate seasonal vegetation, and the use of NPV can be useful to analyze the dynamics, function, and structure of vegetation. The use of four components (PV, NPV, shade, and soil) to generate land cover classifications in forested areas with high biomass, such as the Amazon rainforest, has been considered [68]. Recent research has focused on collecting photosynthetic and nonphotosynthetic endmembers of vegetation to enable the discrimination of dry and wet vegetation in environments with strong seasonality during the year [69]. However, few studies have investigated the applicability of these four components to quantify and monitor changes in seasonal vegetation and with low levels of biomass, as is the case for the Cerrado biome.

Currently, there are some innovative methods and image synergisms that have been tested to improve the discrimination of land cover classes. The use of L-band synthetic aperture radar (SAR) images provided a significant improvement in the discrimination of woody canopy cover from African dry savanna [70]. This radar-based approach may improve the classification of humid savannas since SAR data are sensitive to moist vegetation. The use of the tasseled cap transformation of the Landsat and Sentinel-2 time series to discriminate the Cerrado physiognomies presented an overall accuracy of 63% [43]. Previous studies also showed that Bayesian unsupervised classification and multiple platforms/sensors improved LULC discrimination [71]. Harmonic models applied to time series have been used to classify forest types [72], facilitating the overview of phenological cycles that could be used to improve interclass separation.

For future investigations, we suggest evaluating the feasibility of combining the proposed approach with other methods that rely on multiple sensors and different image processing techniques to derive accurate LULC products in complex ecosystems.

5. Conclusions

In this study, a novel method to produce LULC maps based on the classification of the three maximum fraction images and three standard deviations derived from annual composites of the 100-m spatial resolution PROBA-V dataset was presented. The innovative idea involves the analysis of the temporal (annual) behavior of the proportions of vegetation, soil, and shade of the main LULC classes to classify them according to their maximum peaks and standard deviations (seasonality), which are closely related to the biophysical and phenological behaviors of those classes.

This study presented an overall accuracy classification of 92.4% and an agreement of 88% with the LULC map produced by the MapBiomass project. The user accuracy and the producer accuracy of classifying croplands, forestlands, and water bodies were above 90%. The method is suitable for processing in cloud-computing platforms, such as the Google Earth Engine, which helps address the high dimensionality of the data; therefore, it can be easily tested for regional- and global-scale classification purposes.

One of the main limitations of the method is the low number of observations (pixels) without cloud and cloud shadows in some regions over the year to attribute the maximum and standard deviation into the correct LULC class. Data fusion or data synergism may be a proper initiative to address the lack of temporal information or the mixture complexity of the annual proportion of endmembers in the Cerrado physiognomies or even the spectral differences related to either well-managed or poorly managed pastures.

Author Contributions: Conceptualization, H.L.G.C. and Y.E.S.; methodology, H.L.G.C.; software, H.L.G.C.; validation, T.B.H. and H.L.G.C.; formal analysis, H.L.G.C., E.A., and Y.E.S.; investigation, A.C.D.; resources, Y.E.S.; data curation, E.E.S.; writing—original draft preparation, H.L.G.C. and Y.E.S.; writing—review and editing, E.E.S. and A.C.D.; visualization, A.C.D.; supervision, Y.E.S.; project administration, Y.E.S. and E.A.; funding acquisition, Y.E.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Process No. 2016/19806-3, 2018/14423-4) and by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Process No. 303299/2018-5, CNPq, Process No. 380716/2019-4).

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

1. INPE, National Institute for Space Research. Monitoramento da Floresta Amazônica Brasileira por Satélite/PRODES. 2018. Available online: <http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes> (accessed on 30 April 2019).
2. Mendes, F.S.; Baron, D.; Gerold, G.; Liesenberg, V.; Erasmi, S. Optical and SAR remote sensing synergism for mapping vegetation types in the endangered Cerrado/Amazon ecotone of Nova Mutum—Mato Grosso. *Remote Sens.* **2019**, *11*, 1161. [CrossRef]
3. Klink, C.A.; Moreira, A.G. Past and Current Human Occupation, and Land Use. In *The Cerrados of Brazil: Ecology and Natural History of a Neotropical Savanna*, 1st ed.; Oliveira, P.S., Marquis, R.J., Eds.; Columbia University Press: New York, NY, USA, 2002; pp. 69–88.
4. Gollnow, F.; Hissa, L.B.V.; Rufin, P.; Lakes, T. Property-level direct and indirect deforestation for soybean production in the Amazon region of Mato Grosso, Brazil. *Land Use Policy* **2018**, *78*, 377–385. [CrossRef]
5. Parente, L.; Mesquita, V.; Miziara, F.; Baumann, L.; Ferreira, L. Assessing the pasturelands and livestock dynamics in Brazil, from 1985 to 2017: A novel approach based on high spatial resolution imagery and Google Earth Engine cloud computing. *Remote Sens. Environ.* **2019**, *232*, 111301. [CrossRef]
6. Sano, E.E.; Rosa, R.; Brito, J.L.S.; Ferreira, L.G. Land cover mapping of the tropical savanna region in Brazil. *Environ. Monit. Assess.* **2010**, *166*, 113–124. [CrossRef]
7. Lapola, D.M.; Martinelli, L.A.; Peres, C.A.; Ometto, J.P.H.B.; Ferreira, M.E.; Nobre, C.A.; Aguiar, A.P.D.; Bustamante, M.M.C.; Cardoso, M.F.; Costa, M.H.; et al. Pervasive transition of the Brazilian land-use system. *Nat. Clim. Chang.* **2014**, *4*, 27–35. [CrossRef]
8. Kirby, K.R.; Laurance, W.F.; Albernaz, A.K.; Schroth, G.; Fearnside, P.M.; Bergen, S.; Venticinque, E.M.; Costa, C. The future of deforestation in the Brazilian Amazon. *Futures* **2006**, *38*, 432–453. [CrossRef]
9. Laurance, W.F.; Sayer, J.; Cassman, K.G. Agricultural expansion and its impacts on tropical nature. *Trends Ecol. Evol.* **2014**, *29*, 107–116. [CrossRef] [PubMed]
10. Achard, F.; Brown, S.; De Fries, R.; Grassi, G.; Herold, M.; Mollicone, D.; Pandey, D.; Souza, C., Jr. Reducing greenhouse gas emissions from deforestation and degradation in developing countries. In *A Sourcebook of Methods and Procedures for Monitoring, Measuring and Reporting*; GOFC-GOLD Report version COP14-2; GOFC-GOLD Project Office, Natural Resources Canada: Alberta, Canada, 2009; 185p.
11. Aragão, L.E.O.C.; Anderson, L.O.; Fonseca, M.G.; Rosan, T.M.; Vedovato, L.B.; Wagner, F.H.; Silva, C.V.J.; Junior, C.H.L.S.; Arai, E.; Aguiar, A.P.; et al. 21st Century drought-related fires counteract the decline of Amazon deforestation carbon emissions. *Nat. Commun.* **2018**, *9*, 536. [CrossRef]
12. Morris, R.J. Anthropogenic impacts on tropical forest biodiversity: A network structure and ecosystem functioning perspective. *Philos. Tran. R. Soc. B* **2010**, *365*, 3709–3718. [CrossRef]
13. Barlow, J.; Lennox, G.D.; Ferreira, J.; Berenguer, E.; Lees, A.C.; Nally, R.M.; Thomson, J.R.; Ferraz, S.F.D.; Louzada, J.; Oliveira, V.H.F.; et al. Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. *Nature* **2016**, *535*, 144–147. [CrossRef]
14. Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.; et al. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 7–27. [CrossRef]
15. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; et al. High-resolution global maps of 21st-century forest cover change. *Science* **2013**, *342*, 850–853. [CrossRef] [PubMed]
16. MapBiomas. MapBiomas v. 3.1. Available online: <http://mapbiomas.org/#> (accessed on 30 April 2019).

17. Parente, L.; Ferreira, L.; Faria, A.; Nogueira, S.; Araújo, F.; Teixeira, L.; Hagen, S. Monitoring the Brazilian pasturelands: A new mapping approach based on the Landsat-8 spectral and temporal domains. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *62*, 135–143. [CrossRef]
18. Sano, E.E.; Rodrigues, A.A.; Martins, E.S.; Bettiol, G.M.; Bustamante, M.M.C.; Bezerra, A.S.; Couto, A.F., Jr.; Vasconcelos, V.; Schüller, J.; Bolfe, E.L. Cerrado ecoregions: A spatial framework to assess and prioritize Brazilian savanna environmental diversity for conservation. *J. Environ. Manag.* **2019**, *232*, 818–828. [CrossRef]
19. Bendini, H.N.; Fonseca, L.M.G.; Körting, T.S.; Marujo, R.F.B.; Sanches, I.D.; Arcanjo, J.S. 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. *Proc. Brazilian Symp. Geoinform.* **2016**, 240–251.
20. Justice, C.O.; Townshend, J.R.G.; Vermote, E.F.; Masuoka, E.; Wolfe, R.E.; Saleous, N.; Roy, D.P.; Morisette, J.T. An overview of MODIS land data processing and product status. *Remote Sens. Environ.* **2002**, *83*, 3–15. [CrossRef]
21. Rast, M.; Bézy, J.; Bruzzi, S. The ESA medium resolution imaging spectroradiometer MERIS: A review of the instrument and its mission. *Int. J. Remote Sens.* **1999**, *20*, 1681–1702. [CrossRef]
22. Dierckx, W.; Sterckx, S.; Benhadj, I.; Livens, S.; Duhoux, G.; van Achteren, T.; Francois, M.; Mellab, K.; Saint, G. PROBA-V mission for global vegetation monitoring: Standard products and image quality. *Int. J. Remote Sens.* **2014**, *35*, 2589–2614. [CrossRef]
23. Bendini, H.N.; Garcia Fonseca, L.M.; Schwieder, M.; Sehn Körting, T.; Rufin, P.; Sanches, I.D.; Leitão, P.J.; Hostert, P. Detailed agricultural land classification in the Brazilian cerrado based on phenological information from dense satellite image time series. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *82*, 101872. [CrossRef]
24. Roy, D.P.; Huang, H.; Boschetti, L.; Giglio, L.; Yan, L.; Zhang, H.H.; Li, Z. Landsat-8 and Sentinel-2 burned area mapping—A combined sensor multi-temporal change detection approach. *Remote Sens. Environ.* **2019**, *231*, 111254. [CrossRef]
25. Lu, D.; Moran, E.; Batistella, M. Linear mixture model applied to Amazonian vegetation classification. *Remote Sens. Environ.* **2003**, *87*, 456–469. [CrossRef]
26. Somers, B.; Asner, G.P.; Tits, L.; Coppin, P. Endmember variability in spectral mixture analysis: A review. *Remote Sens. Environ.* **2011**, *115*, 1603–1616. [CrossRef]
27. Shimabukuro, Y.E.; Smith, J.A. The least-squares mixing models to generate fraction images derived from remote sensing multispectral data. *IEEE Trans. Geosci. Remote Sens.* **1991**, *29*, 16–20. [CrossRef]
28. Quintano, C.; Fernández-Manso, A.; Shimabukuro, Y.E.; Pereira, G. Spectral unmixing. *Int. J. Remote Sens.* **2012**, *33*, 5307–5340. [CrossRef]
29. Haertel, V.; Shimabukuro, Y.E. Spectral linear mixing model in low spatial resolution image data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2555–2562. [CrossRef]
30. Anderson, L.O.; Aragão, L.E.O.C.; Shimabukuro, Y.E.; Almeida, S.; Huete, A. Fraction images for monitoring intra-annual phenology of different vegetation physiognomies in Amazonia. *Int. J. Remote Sens.* **2011**, *32*, 387–408. [CrossRef]
31. Ferreira, M.E.; Ferreira, L.G.; Sano, E.E.; Shimabukuro, Y.E. Spectral linear mixture modelling approaches for land cover mapping of tropical savanna areas in Brazil. *Int. J. Remote Sens.* **2007**, *28*, 413–429. [CrossRef]
32. Adami, M.; Bernardes, S.; Arai, E.; Freitas, R.M.; Shimabukuro, Y.E.; Espírito-Santo, F.D.B.; Rudorff, B.F.T.; Anderson, L.O. Seasonality of vegetation types of South America depicted by moderate resolution imaging spectroradiometer (MODIS) time series. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *69*, 148–163. [CrossRef]
33. Shimabukuro, Y.E.; Arai, E.; Duarte, V.; Jorge, A.; Santos, E.G.; Gasparini, K.A.C.; Dutra, A.C. Monitoring deforestation and forest degradation using multi-temporal fraction images derived from Landsat sensor data in the Brazilian Amazon. *Int. J. Remote Sens.* **2019**, *40*, 5475–5496. [CrossRef]
34. Arvor, D.; Jonathan, M.; Meirelles, M.S.P.; Dubreuil, V.; Durieux, L. Classification of MODIS EVI time series for crop mapping in the State of Mato Grosso, Brazil. *Int. J. Remote Sens.* **2011**, *32*, 7847–7871. [CrossRef]
35. Instituto Brasileiro de Geografia e Estatística (IBGE). *Mapa de Biomas e de Vegetação*; IBGE: Rio de Janeiro, Brazil, 2004. Available online: <https://ww2.ibge.gov.br/home/presidencia/noticias/21052004biomashtml.shtm> (accessed on 29 April 2019).
36. Cohn, A.S.; Gil, J.; Berger, T.; Pellegrina, H.; Toledo, C. Patterns and processes of pasture to crop conversion in Brazil: Evidence from Mato Grosso State. *Land Use Policy* **2016**, *55*, 108–120. [CrossRef]

37. Koltunov, A.; Ustin, S.L.; Asner, G.P.; Fung, I. Selective logging changes forest phenology in the Brazilian Amazon: Evidence from MODIS image time series analysis. *Remote Sens. Environ.* **2009**, *113*, 2431–2440. [CrossRef]
38. Grecchi, R.C.; Beuchle, R.; Shimabukuro, Y.E.; Aragão, L.E.O.C.; Arai, E.; Simonetti, D.; Achard, F. An integrated remote sensing and GIS approach for monitoring areas affected by selective logging: A case study in northern Mato Grosso, Brazilian Amazon. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *61*, 70–80. [CrossRef]
39. Francois, M.; Santandrea, S.; Mellab, K.; Vrancken, D.; Versluys, J. The PROBA-V mission: The space segment. *Int. J. Remote Sens.* **2014**, *35*, 2548–2564. [CrossRef]
40. Wolters, E.; Dierckx, W.; Dries, J.; Swinnen, E. PROBA-V Products User Manual v3.01. VITO. 2018. Available online: http://proba-v.vgt.vito.be/sites/proba-v.vgt.vito.be/files/products_user_manual.pdf (accessed on 30 April 2019).
41. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]
42. Instituto Brasileiro de Geografia e Estatística (IBGE). *Censo Agropecuário*; IBGE: Rio de Janeiro, Brazil, 2017. Available online: <https://www.ibge.gov.br/estatisticas-novoportal/economicas/agricultura-e-pecuaria/21814-2017-censo-agropecuario.html?edicao=21858&t=sobre> (accessed on 12 November 2019).
43. Schwieder, M.; Leitão, P.J.; Bustamante, M.M.C.; Ferreira, L.G.; Rabe, A.; Hostert, P. Mapping Brazilian savanna vegetation gradients with Landsat time series. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 361–370. [CrossRef]
44. Roy, E.D.; Richards, P.D.; Martinelli, L.A.; Della Coletta, L.; Lins, S.R.M.; Vazquez, F.F.; Willig, E.; Spera, S.A.; VanWey, L.K.; Porder, S. The phosphorus cost of agricultural intensification in the tropics. *Nat. Plants* **2016**, *2*, 16043. [CrossRef]
45. Ribeiro, J.F.; Walter, B.M.T. *As Principais Fitofisionomias do Bioma Cerrado*; Sano, S.M., Almeida, S.P., Ribeiro, J.F., Eds.; Ecologia e Flora: Cerrado, Brazil, 2008; pp. 152–212.
46. Nogueira, E.M.; Nelson, B.W.; Fearnside, P.M.; França, M.B.; Oliveira, Á.C.A. Tree height in Brazil's 'arc of deforestation': Shorter trees in south and southwest Amazonia imply lower biomass. *For. Ecol. Manag.* **2008**, *255*, 2963–2972. [CrossRef]
47. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [CrossRef]
48. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *5*–32. [CrossRef]
49. Dietterich, T.G. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Mach. Learn.* **2000**, *40*, 139–157. [CrossRef]
50. Hudson, W.D.; Ramm, C.W. Correct formulation of the Kappa coefficient of agreement (in remote sensing). *Photogramm. Eng. Remote Sens.* **1987**, *53*, 421–422.
51. Congalton, R.G. Accuracy assessment and validation of remotely sensed and other spatial information. *Int. J. Wildl. Fire* **2001**, *10*, 321. [CrossRef]
52. Simoes, R.; Picoli, M.C.A.; Camara, G.; Maciel, A.; Santos, L.; Andrade, P.R.; Sánchez, A.; Ferreira, K.; Carvalho, A. Land use and cover maps for Mato Grosso State in Brazil from 2001 to 2017. *Sci. Data* **2020**, *7*, 1–10.
53. Fearnside, P.M. Brazil's Cuiabá-Santarém (BR-163) highway: The environmental cost of paving a soybean corridor through the Amazon. *Environ. Manag.* **2007**, *39*, 601–614. [CrossRef] [PubMed]
54. Pinheiro, T.F.; Escada, M.I.S.; Valeriano, D.M.; Hostert, P.; Gollnow, F.; Muller, H. Forest degradation associated with logging frontier expansion in the Amazon: The BR-163 region in southwestern Pará, Brazil. *Earth Interact.* **2016**, *17*, 1–26.
55. Latrubesse, E.M.; Stevaux, J.C. Geomorphology and environmental aspects of the Araguaia fluvial basin, Brazil. *Z. Geomorphol.* **2002**, *129*, 109–127.
56. Latrubesse, E.M.; Amsler, M.L.; Morais, R.P.; Aquino, S. The geomorphologic response of a large pristine alluvial river to tremendous deforestation in the South American tropics: The case of the Araguaia river. *Geomorphol.* **2009**, *113*, 239–252. [CrossRef]
57. Picoli, M.C.A.; Camara, G.; Sanches, I.; Simões, R.; Carvalho, A.; Maciel, A.; Coutinho, A.; Esquerdo, J.; Antunes, J.; Begotti, R.A.; et al. Big earth observation time series analysis for monitoring Brazilian agriculture. *ISPRS J. Photogramm. Remote Sens.* **2018**, *145*, 328–339. [CrossRef]

58. IBGE. *Produção Agrícola Municipal*; IBGE: Rio de Janeiro, Brazil, 2018. Available online: <https://sidra.ibge.gov.br/pesquisa/pam/tabelas> (accessed on 17 June 2019).
59. Zalles, V.; Hansen, M.C.; Potapov, P.V.; Stehman, S.V.; 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]
60. Rudorff, B.F.T.; Aguiar, D.A.; Silva, W.F.; Sugawara, L.M.; Adami, M.; Moreira, M.A. Studies on the rapid expansion of sugarcane for ethanol production in São Paulo State (Brazil) using Landsat data. *Remote Sens.* **2010**, *2*, 1057–1076. [CrossRef]
61. Scaramuzza, C.A.M.; Sano, E.E.; Adami, M.; Bolfe, E.L.; Coutinho, A.C.; Esquerdo, J.C.D.M.; Maurano, L.E.P.; Narvaes, I.d.S.; Oliveira Filho, F.J.B.D.; Rosa, R.; et al. Land-use and land-cover mapping of the Brazilian Cerrado based mainly on Landsat-8 satellite images. *Rev. Bras. Cartogr.* **2017**, *69*, 1041–1051.
62. Almeida, C.A.; Coutinho, A.C.; Esquerdo, J.C.D.M.; Adami, M.; Venturieri, A.; Diniz, C.G.; Dessay, N.; Durieux, L.; Gomes, A.R. 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]
63. Alvares, C.A.; Stape, J.L.; Sentelhas, P.C.; Moraes, G.; Leonardo, J.; Sparovek, G. Köppen's climate classification map for Brazil. *Meteorol. Z.* **2013**, *22*, 711–728. [CrossRef]
64. Sterckx, S.; Benhadj, I.; Duhoux, G.; Livens, S.; Dierckx, W.; Goor, E.; Adriaensen, S.; Heyns, W.; van Hoof, K.; Strackx, G.; et al. The PROBA-V mission: Image processing and calibration. *Int. J. Remote Sens.* **2014**, *35*, 2565–2588. [CrossRef]
65. Müller, H.; Rufin, P.; Griffiths, P.; Barros Siqueira, A.J.; 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]
66. Thorp, K.R.; French, A.N.; Rango, A. Effect of image spatial and spectral characteristics on mapping semi-arid rangeland vegetation using multiple endmember spectral mixture analysis (MESMA). *Remote Sens. Environ.* **2013**, *132*, 120–130. [CrossRef]
67. Yang, J.; Weisberg, P.J.; Bristow, N.A. Landsat remote sensing approaches for monitoring long-term tree cover dynamics in semi-arid woodlands: Comparison of vegetation indices and spectral mixture analysis. *Remote Sens. Environ.* **2012**, *119*, 62–71. [CrossRef]
68. Souza, C.M.; Roberts, D.A.; Cochrane, M.A. Combining spectral and spatial information to map canopy damage from selective logging and forest fires. *Remote Sens. Environ.* **2005**, *98*, 329–343. [CrossRef]
69. Bullock, E.L.; Woodcock, C.E.; Souza, C.; Olofsson, P. Satellite-based estimates reveal widespread forest degradation in the Amazon. *Glob. Chang. Biol.* **2020**, *26*, 2956–2969. [CrossRef]
70. Naidoo, L.; Mathieu, R.; Main, R.; Wessels, K.; Asner, G.P. L-band synthetic aperture radar imagery performs better than optical datasets at retrieving woody fractional cover in deciduous, dry savannahs. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 54–64. [CrossRef]
71. Lee, J.; Cardille, J.A.; Coe, M.T. Agricultural expansion in Mato Grosso from 1986–2000: A Bayesian time series approach to tracking past land cover change. *Remote Sens.* **2020**, *12*, 688. [CrossRef]
72. Fortin, J.A.; Cardille, J.A.; Perez, E. Multi-sensor detection of forest-cover change across 45 years in Mato Grosso, Brazil. *Remote Sens. Environ.* **2020**, *238*, 111266. [CrossRef]

