

Article Vegetation Fraction Images Derived from PROBA-V Data for Rapid Assessment of Annual Croplands in Brazil

Egidio Arai^{1,*}, Edson Eyji Sano², Andeise Cerqueira Dutra¹, Henrique Luis Godinho Cassol¹, Tânia Beatriz Hoffmann¹ and Yosio Edemir Shimabukuro¹

- ¹ Instituto Nacional de Pesquisas Espaciais (INPE), 12.227-010 São José dos Campos, Brazil; andeise.dutra@inpe.br (A.C.D.); henrique@dsr.inpe.br (H.L.G.C.); tania.hoffmann@inpe.br (T.B.H.); yosio@dsr.inpe.br (Y.E.S.)
- ² Embrapa Cerrados, 73301-970 Planaltina, Brazil; edson.sano@embrapa.br
- * Correspondence: egidio@dsr.inpe.br; Tel.: +55-12-3208-6431

Received: 20 November 2019; Accept: 3 January 2020; Published: 3 April 2020



Abstract: This paper presents a new method for rapid assessment of the extent of annual croplands in Brazil. The proposed method applies a linear spectral mixing model (LSMM) to PROBA-V time series images to derive vegetation, soil, and shade fraction images for regional analysis. We used S10-TOC (10 days synthesis, 1 km spatial resolution, and top-of-canopy) products for Brazil and S5-TOC (five days synthesis, 100 m spatial resolution, and top-of-canopy) products for Mato Grosso State (Brazilian Legal Amazon). Using the time series of the vegetation fraction images of the whole year (2015 in this case), only one mosaic composed with maximum values of vegetation fraction was generated, allowing detecting and mapping semi-automatically the areas occupied by annual crops during the year. The results (100 m spatial resolution map) for the Mato Grosso State were compared with existing global datasets (Finer Resolution Observation and Monitoring-Global Land Cover (FROM-GLC) and Global Food Security—Support Analyses Data (GFSAD30)). Visually those maps present a good agreement, but the area estimated are not comparable since the agricultural class definition are different for those maps. In addition, we found 11.8 million ha of agricultural areas in the entire Brazilian territory. The area estimation for the Mato Grosso State was 3.4 million ha for 1 km dataset and 5.3 million ha for 100 m dataset. This difference is due to the spatial resolution of the PROBA-V datasets used. A coefficient of determination of 0.82 was found between PROBA-V 100 m and Landsat-8 OLI area estimations for the Mato Grosso State. Therefore, the proposed method is suitable for detecting and mapping annual croplands distribution operationally using PROBA-V datasets for regional analysis.

Keywords: linear spectral mixing model; Mato Grosso State; cropland mapping; maximum fraction values mosaic

1. Introduction

Information on the extent, distribution, and dynamics of croplands have been identified as key issues for food security scenarios and economic and policy implications [1–3]. The sustainable natural resources management policies demand for the development of operational cropland mapping and monitoring methods capable of providing accurate and up-to-date spatial information [4,5]. In this context, remote sensing techniques have provided tools for identifying croplands over large areas based on attributes such as long-term historical, frequent, and freely available data that make remotely sensed observations attractive [6–8].



Brazil is a country with high capacity to increase agricultural production. In the past decades, the agricultural sector has grown rapidly because of the technological advances, commodities price increase, and global demand rising [9,10]. With this, research has been conducted on mapping croplands [11] and other land use and land cover (LULC) classes in Brazil at a national scale, as the MapBiomas project [12], or even at a local scale [7,13,14], or at global scale as the Finer Resolution Observation and Monitoring of Global Land Cover (FROM–GLC) [15] and the Global Food Security Support Analysis Data (GFSAD30) [16] provided at certain years. Most of these studies have used a combination of either medium spatial resolution and relatively long re-visiting capability satellite data (e.g., Landsat satellite) or moderate spatial resolution and high temporal resolution imageries [e.g., moderate resolution imaging spectroradiometer (MODIS) sensor onboard Terra and Aqua platforms].

One potentially limiting factor of the existing approaches is the relatively long time re-visiting data, for example, the limited number of Landsat records [17,18] that also can present cloud contamination, especially in tropical areas [19,20]. On the other hand, coarse and moderate spatial resolution sensors are limited whenever the object of interest is smaller than their pixel size. Moreover, mapping croplands is challenging in Brazil, especially because of the diversity of crops, varying planting dates, agricultural landscape fragmentation, spatial heterogeneity, and mosaic of croplands and natural grasslands. In addition, there is a lack of operational approaches to map the spatial distribution of croplands in Brazil automatically and rapidly. For instance, many classification algorithms and methods developed are labor-intensive and time-consuming from the computational point of view to obtain optimized parameters and results, besides the demand for large training datasets [17,18]. Thus, the development of operational approaches is required to subsidize decision-makings of crop planting dates as well as to reduce uncertainties of crop yield estimations [7].

The linear spectral mixing model (LSMM) is a technique used to analyze the smallest element inside of a pixel, under the assumption that a pixel corresponds to a linear spectral combination of its elements [21]. The LSMM generates fraction images containing the proportion of the elements, usually vegetation, soil, and shade, which, in turn, can be used for mapping LULC changes based on the following characteristics: (a) vegetation fraction images highlight the forest cover conditions and allow differentiation between forest and non-forest areas, similar to the vegetation indices such as the Normalized Difference Vegetation Index (NDVI) or Enhanced Vegetation Index (EVI); (b) soil fraction images highlight areas with high reflectance values such as bare soil and clear-cuts [22–24]; and (c) shade fraction images highlight areas with low reflectance values such as water, shadow, and burned areas [25–27].

Croplands show the highest proportions of vegetation fraction, allowing it to be separated from the other land use classes. For instance, the use of year-based fraction images mosaics can reduce data dimensionality by selecting cloud-free pixels to generate mosaics of maximum fractions of vegetation, soil, and shade. Taking advantage of the temporal resolution of PROBA-V satellite, the objective of this paper is to present a new method based on PROBA-V fraction images derived from LSMM to assess the spatial distribution of annual croplands in Brazil and in Mato Grosso State rapidly. The PROBA-V is a mini-satellite developed as continuity for the 15 years old SPOT-VEGETATION mission.

2. Materials and Methods

2.1. Study Areas

The first study area, selected for regional analysis of PROBA-V 1 km data set, was the entire territory of Brazil (over 8.5 million km²), which produces most of the world's major crop types, namely soybean, sugarcane, maize, rice, and wheat [28–30]. According to the Municipality Agricultural Production (PAM) data from the Brazilian Institute of Geography and Statistics (IBGE) [31], the Midwest is the most important region in the country in terms of planted area of annual and perennial crops (32%), followed by South (27%), Southeast (21%), Northeast (16%), and North (5%) (Figure 1).

In terms of percentages of planted area by crop types cultivated in Brazil, soybean is the largest crop (42%), followed by maize (21%), sugarcane (14%), bean (4%), wheat (3%), rice (3%), and others (11%).



Figure 1. Location of Brazil in the American continent (inset figure) and location of Mato Grosso (MT) State in Brazil. Brazil and Mato Grosso State correspond to the two selected study areas of this paper. The country is divided into five regions: North, Northeast, Midwest, South, and Southeast. Pie graphs show the percentages of six major national crop types in Brazil by region. Source: IBGE [31].

It has been estimated that row cropping area in Brazil occupies around 50 million ha, depending on the study e.g., [11]. Most of these areas are extensive (field size often more than one thousand ha) and are concentrated in three regions (South, Southeast, and Midwest), where topography is flat and precipitation in the rainy season is sufficient for rainfed agriculture. Row cropping areas in other Brazilian regions that are environmentally sensitive (North, Brazilian Amazon) or present low levels of precipitation (Northeast, except the western part) and hilly topography (mainly in the coastal region) are mostly found in small-size farms (family farming, up to four fiscal modules, one fiscal module = 5 to 110 ha, depending upon the municipality, according to the Federal Law n. 11.326/2006). Figure 2 shows some examples of the Landsat-8 RGB color composites of bands 4 (red; 0.64–0.67 μ m), 5 (near infrared; 0.85–0.88 μ m), and 6 (middle infrared; 1.57–1.65 μ m) from different regions of Brazil with dominant large-scale grain production. In these images, we can note the presence of large fields, often larger than the 10 x 10 km size (10 thousand ha).



Figure 2. Cont.



Figure 2. RBG color composites of Landsat-8 scenes (bands 4, 5, and 6; year = 2015) from the following representative sites of Brazil in terms of large-scale grain production: (**a**) municipality of Formosa do Rio Preto, Bahia State, Northeast region; (**b**) municipality of Sorriso, Mato Grosso State, Midwest region; (**c**) municipalities from Jataí/GO; and (**d**) municipality of Assis Chateaubriand, Paraná State and surroundings.

The second study area, selected for local analysis of PROBA-V 100 m and for evaluation purposes, was the Mato Grosso State, the third largest state in Brazil (over 900,000 km²) and located in the Midwest region. Mato Grosso presents the largest cropland areas in Brazil and encompasses three biomes: Amazonia (54%), Cerrado (40%), and Pantanal (6%) [32]. The rainforest is found in the northern part of the state and currently it is part of the "arc of deforestation" of the Brazilian Legal Amazon [33].

One of the major challenges for mapping croplands on a continental scale accurately, such as Brazil, is related to the wide and diverse crop calendars (Figure 3). For each crop, the period of seeding is different depending on the region, especially because of the varying precipitation patterns. For instance, soybean has two cultivation periods in the North region, one during the wet season (October to February) and another during the dry season (May to June), and one cultivation period in the other regions, which extends from September to January, depending on the latitude. Mato Grosso State often presents two planting periods during the same year of crop calendar as part of agricultural intensification policy [7]. The implication is that most of the tropical crop fields in Brazil, including Mato Grosso State, are found in the wet season, period in which cloud cover is persistent, reducing the capacity of optical sensors to map agricultural areas.



Figure 3. Brazil's crop calendar of five main annual crops (soybean, maize, bean, wheat, and rice) by region. Source: adapted from [31,34].

2.2. Calibration and Validation Data Set

For this study, we used, as calibration data set, the S10-TOC (1 km spatial resolution composites of 10 days global synthesis, and top-of-canopy) products for the entire Brazilian territory (total of 36 images/year, each image composed of 32 tiles) for regional analysis, and the S5-TOC (100 m spatial resolution composites, 5 days global composites, and top-of-canopy) products for the Mato Grosso State (total of 72 images/year, each image composed of 4 tiles) for local analysis. The PROBA-V was launched on May 7, 2013 with 100 m spatial resolution at nadir and 350 m in every field of view [35]. The sensor has spectral channels similar to the four-band SPOT-VEGETATION instrument: blue (0.438–0.486 μ m); red (0.615–0.696 μ m); near infrared (0.772–0.914 μ m); and medium infrared (1.564–1.634 μ m) [36].

A set of 47 Landsat-8 OLI path/rows from the entire Mato Grosso State, obtained in the EarthExplorer user interface developed by the United States Geological Survey (USGS), was selected for validation purposes of PROBA-V 100 m data. For each confluence of full degree latitude and longitude, we defined a buffer zone of 5 km radius (total of 292 sampling units) where Landsat-8 OLI (Operational Land Imager) color composites of bands 4 (red; $0.64-0.67 \mu m$), 5 (near infrared; $0.85-0.88 \mu m$), and 6 (middle infrared; $1.57-1.65 \mu m$) were clipped. Within an average of six images available in January, February, and March, we chose the one with lowest cloud cover condition and then with the highest radiometric quality based on the visual inspection of these three bands.

2.3. Ancillary Data Set

The national land use and land cover map from 2015 (collection 3.1), produced by the MapBiomas Project [12] and the Agricultural Census statistical data obtained by IBGE in 2017 [31] were selected to subsidize the interpretation of PROBA-V data set. The MapBiomas Project is an initiative of several Brazilian institutions, including universities, private companies, and non-governmental organizations to automatically classify, on a yearly basis, the major LULC classes of the entire Brazilian territory. It is based on cloud computing of the Landsat time series available in the Google Earth Engine platform.

The Agricultural Census obtains different types of agricultural data at the municipality level, including annual and perennial croplands, pasturelands, and agroindustry, approximately every ten years. The last Census was released in 2017 (preliminary data, under review) and refers to the crop calendar of October 2016–September 2017 [34].

2.4. Methodological Overview

The method is summarized in the flowchart shown in Figure 4 and involves five main steps: (1) selection of the annual PROBA-V time series images (1km and 100m datasets); (2) generation of the median PROBA-V image mosaic from the time series of four spectral bands; (3) definition of the endmembers (vegetation, soil, and shade) spectral profiles by visual inspection of median image mosaic and running the unmixing algorithm for all PROBA-V (1 km and 100 m datasets), generating the corresponding vegetation, soil, and shade fraction images using the linear spectral mixing model; (4) generation of the maximum value images, only one for the corresponding endmembers (vegetation, soil and shade); segmentation and classification of the maximum value mosaic of vegetation fraction that highlights (brighter areas) the annual cropland, using the unsupervised classification algorithm; (5) visual comparison of the obtained results with the existing global maps and between PROBA-V 1km and 100m datasets, and validation of the results using Landsat OLI (30 m) images.



Figure 4. General methodological flowchart for annual cropland classification and evaluation.

2.4.1. Calibration

The first step of calibration procedure consisted of preparing the mosaic of the tiles in the HDF format and converting to the geocoded Geotiff format for each date obtained from PROBA-V. The reflectance provided in the data are presented as digital count numbers (DN), and in order to obtain reflectance values, a conversion using the offset and scale parameters for the PROBA-V dataset were applied, represented by 0.0 and 2000 respectively [35].

Clouds and corresponding shadows were masked using the Status Map (SM) file information. The SM dataset is presented in numeric values (bits) that summarize the pixel condition for each image. A bit can only assume binary values, i.e., 0 and 1, and considering only the three bits contained in the SM, they provide five different combinations (000, 001, 010, 011, and 100) that correspond to a particular indicator of the pixel condition. These five different combinations can be described as: Clear, Shadow, Undefined, Cloud, and Ice [35]. Thus, we created a mask for each scene containing all the pixel indicators to preserve only the "Clear" pixel in our dataset.

The next step, it was the application of the LSMM, which assumes that pixel values are linear combinations of reflectance from different components, called endmembers Equation (1):

$$R_{i} = \sum_{j=i}^{n} f_{i} r_{i,j} + \varepsilon_{i}$$
⁽¹⁾

where: R_i = the spectral reflectance in the *ith* spectral band; $r_{i,j}$ = the spectral reflectance of the *jth* component in spectral band *ith* (endmember); f_j = the proportion of the *jth* component within the pixel; i = 1, n = 4 (number of spectral bands used); j = 1, m = 3 (number of considered components); and ε_i = the residue for the *ith* spectral band.

After several tests, the median value of all images for the four spectral bands was computed using the PROBA-V 1km dataset for the year 2015, generating a single image mosaic that was used only to extract the spectral reflectance values of the endmembers by visual inspection. The median has the advantage of not being strongly affected by the extreme values (outliers). In this study, we considered three endmembers (vegetation, soil, and shade) for the PROBA-V 1 km and 100 m datasets, following the methodology suggested by [37] (Figure 5a). Using the generated median image mosaic, the three pure pixels (endmembers) of green vegetation, soil, and shade were selected by visual inspection of the mosaic. The spectral values in each spectral band are presented in Table 1 and the theoretical spectral responses of the green vegetation, soil, and shade are presented in Figure 5b. Then we applied the LSMM Equation (1) in the PROBA-V 1 km and 100 m annual time-series.



Figure 5. (a) Conceptual composition of a pixel with different elements inside a pixel; (b) the subsequent theoretical spectral response of its elements to obtain the endmembers; and (c) representation of temporal behavior of vegetation fraction images and variation of the proportions through the period.

		Endmembers		
	Spectral Bands	Soil	Vegetation	Shade
Spectral reflectance factor	447–493 nm 610–690 nm 777–893 nm 1570–1650 nm	0.0710 0.1850 0.3185 0.4205	0.0355 0.0505 0.8155 0.3605	0.0365 0.0330 0.0215 0.0110

Table 1. Spectral reflectance values for the three endmembers extracted from the median image mosaic described previously.

As a result of the LSMM, fraction images were obtained for each component considered in the pixel, which represents the proportion in the original data [38–40] by numeric values between 0 and 1. Then, we generated 36 and 72 sets of vegetation, soil, and shade fraction images for 1 km and 100 datasets, respectively. Following, using the results of fraction image series derived from the LSMM, the maximum proportion values of the respective fractions were computed in PROBA-V 1 km, reducing the time series into a single image for each fraction, as illustrated in Figure 5c. We built similar composites for the Mato Grosso State using 100 m dataset to allow comparison of the classification results obtained with different spatial resolutions.

Figure 6 shows the maximum proportion values composite of vegetation fraction image enhancing the agricultural areas in Brazil and in Mato Grosso State for the year 2015. It can be noted that the agricultural areas are highlighted by the brighter gray tones in both 1 km and 100 m composites, allowing the classification of these areas using a given classification algorithm. Moreover, using the maximum proportion value obtained in the entire year, each pixel can have a variation in the extraction date, which is represented by the vegetative peak considered in the "Clear" pixel interval of our dataset.



Figure 6. Maximum vegetation fraction mosaic derived from the linear spectral mixing model (LSMM) of PROBA-V mosaics datasets. (a) Brazil, PROBA-V 1 km mosaic; (b) Mato Grosso, PROBA-V 1 km mosaic; and (c) Mato Grosso, PROBA-V 100 m mosaic.

The single mosaic derived from vegetation fraction images time series were then processed by the segmentation and ISOSEG unsupervised classification techniques. Image segmentation is a technique to group the data, in which only contiguous regions with similar spectral characteristics are joined. The segmentation approach used in this work is based on a region growing technique. For this, two

parameters are defined by the user to define regions that are used in the classification procedure: (1) similarity threshold that is the Euclidean distance between the mean digital number of two regions, that can be grouped together; and (2) area threshold that is the minimum area to be considered as a region, set as the number of pixels [41,42].

Then, the segmented images are classified using ISOSEG [41] (Figure 7e,f), a region classifier algorithm that is based on clustering techniques. This unsupervised algorithm uses the covariance matrix and the mean of the regions to estimate the centers of the classes. The user defines an acceptance threshold (for each sensor data used), the maximum allowed Mahalanobis distance that a mean digital number may be from the center of a class, to be considered as belonging to that class. After the classification process (using the 75% threshold value), some classes were manually assigned to the corresponding cropland class of the predefined legend.



Figure 7. Representative process of the methodological flowchart. (a) Example of RGB color composites (bands SWIR, NIR, and red; date = 2015/12/11) of the PROBA-V 100 m, containing clouds; (b) PROBA-V 100 m after masked using SM dataset; (c) vegetation fraction image; (d) single maximum proportion value mosaic (year = 2015); (e) image segmentation results; and (f) croplands classification by ISOSEG technique.

The segmentation and classification processes were carried out on both PROBA-V 1 km and 100 m mosaics in the SPRING 5.5 Geographical Information System (GIS) software [43]. In this study, after several empirical tests, we fixed the best value of similarity and area in 8 and 10 pixels for the PROBA-V 100 m and 8 and 5 pixels for PROBA-V 1 km, respectively. These thresholds are the only two input parameters required by the SPRING 5.5 software.

2.4.2. Validation

The validation of the area estimation occupied by the annual croplands in the Mato Grosso State was based on the orthorectified Landsat-8 Operational Land Imager (OLI) scenes from January to March of 2015. We considered a systematic sampling strategy based on a regular 0.5×0.5 degree, full latitude and longitude grid (Figure 8). Within a total of 298 sampling units, six polygons were disregarded since there was no cloud-free image available. This type of sampling design has been

used, for example, by [44] in the context of the Food and Agriculture Organization's Forest Resources Assessment 2010 Remote Sensing Survey [45].

Landsat-8 OLI bands 4, 5, and 6 were selected for image segmentation also using the region growing algorithm available in the SPRING 5.5 geographical information system (GIS) software [43] (see details of this image segmentation approach in the Section 2.4.1 Calibration). After several empirical tests, we set the similarity as 25. The minimum area was defined as 67 pixels, corresponding to approximately 6 ha in the 30 meter spatial resolution of Landsat imageries, which is likely the minimum mapping area in the PROBA-V 100 m scenes. The segments were exported in shapefile format and interpreted visually on a computer screen by overlaying them on the corresponding R4 G5 B6 color composites in the ArcGIS 9.2 software. An interpretation key was prepared to visually identify the crop plantations in the Mato Grosso State (Figure 9) and to differentiate them from the other representative LULC classes (mainly primary forest, Cerrado shrubland, and planted pasture). Examples of the Landsat-8 OLI color composites and corresponding results of image segmentation and interpretation are shown in Figure 10.



Figure 8. Sampling design (5 km buffer zone in the confluence of regular 0.5×0.5 degree, full latitude and longitude grid) for validation of PROBA-V's estimates of croplands in the Mato Grosso State. Total number of samples = 292. Polygons inside the map correspond to the municipality boundaries. Red-colored samples were excluded from the analysis because of the unavailability of cloud-free Landsat scenes.

The segments related to annual and semi-perennial croplands (soybean, maize, cotton, and sugarcane) of each validation samples were selected for area estimation using the Calculate Geometry function available in the ArcGIS 10.1 software. These areas estimations were compared with those obtained by the PROBA-V 100 m data in the same sampling units.

Land Use and Land Cover Class	Interpretation Key	Example of RGB Color Composite	
Cropland	Color: reddish, light green Texture: smooth Geometry: regular Municipality: Sapezal, Mato Grosso Overpass: March 30, 2015		

	Color: pink/light blue	and the second
	Texture: medium	
Cultivated pastureland	Geometry: regular	
Cultivated pastuleland	Municipality: Feliz Natal, Mato	
	Grosso	
	Overpass: January 4, 2015	
	Color: pink	
	Texture: intermediate	
Shruhland	Geometry: undefined	And the second second
Shi ubland	Municipality: Sapezal, Mato	
	Grosso	
	Overpass: March 30, 2015	
	Color: green	
	Texture: intermediate	
Forestland	Geometry: undefined	
Forestiand	Municipality: Feliz Natal, Mato	and the first
	Grosso	
	Overpass: January 4, 2015	X





Figure 10. Examples of Landsat-8 OLI color composite (**a**–**c**) and corresponding image segmentation and interpretation (**d**–**f**) in the municipalities of Sapezal/Campos de Júlio, Brasnorte, and Nova Mutum, respectively. Yellow lines represent the segments while red polygons represent the annual croplands.

3. Results

Figure 11 shows the map of annual croplands in Brazil in 2015, obtained by the PROBA-V 1 km mosaic. We found a total of 11.8 million ha of cropland areas in the entire national territory. The classification of agricultural areas in the Mato Grosso State using PROBA-V 100 m showed a total of 5.3 million ha (Figure 12), which was nearly half of estimation obtained by previous studies ([11]: 9.7 million hectares in 2014; [12]: 10.2 million hectares (annual and perennial crops) in 2015; [7]: 11.1 million hectares in 2015; and [34]: 9.7 million hectares in 2017).



Figure 11. Map of annual croplands in Brazil, derived from the PROBA-V 1 km mosaic (**a**). Inset maps correspond to the five major regions of Brazil (**b**) and the six Brazilian biomes (**c**).



Figure 12. Map of annual croplands in Mato Grosso State, derived from the PROBA-V 100 m. Divisions in the map correspond to the microregions (regions that aggregate a varying number of municipalities with high levels of economic and social similarity). The five microregions highlighted in the map correspond to the microregions with the highest extents of annual croplands in 2015.

4. Discussion

4.1. PROBA-V Calibration Data Set

Compared with other results derived from different products (census and Landsat-like products), the estimation of the cultivated areas from PROBA-V 1 km for the entire national territory was quite low (11.8 million ha). According to the [11], the national area occupied by annual croplands was 46.1 million ha in 2014. The estimation from the MapBiomas project, which uses 30 m resolution data [12], was 46.9 million ha in 2015, while the last Census of Agriculture [34] estimated 52.9 million ha in 2017. This does not mean that the proposed method is wrong or inaccurate, only shows a typical limitation of global or large-scale products that use coarse resolutions [46,47]. Besides, official area estimates cannot be compared if we consider the existence of double-cropped areas, since these areas are counted twice [11].

Nevertheless, the PROBA-V 1 km data were able to depict the most important regions of Brazil in terms of rainfed grain productions, namely: historically well-established agricultural producers in the states of Rio Grande do Sul and Paraná; mid-1980 expansion in Mato Grosso do Sul, Goiás, Mato Grosso, and Bahia States; and the south of Piauí and Maranhão States, the most recent agricultural expansion in the country locally known as the MATOPIBA region, a continuous region encompassing the states of Maranhão, Tocantins, Piauí, and Bahia. Reference [9] also highlighted these areas in their analysis of patterns of Brazilian agriculture dynamics. Besides, there are several studies in literature discussing the advance of grain production in Brazil, from south/southeast to the Cerrado (tropical savanna, central part of the country) and then to the Amazon (rainforest, north of Brazil). These studies focus mainly on the advances of grain production in the following regions: southwest of Goiás State (municipalities of Jataí and Rio Verde), which was one of the first region to be settled by intensive, heavily-mechanized agriculture in the Cerrado biome [48]; western of Bahia State (municipality of Luís Eduardo Magalhães), a well-known producer of soybean, cotton, and maize over the very extensive and flat terrain known as "chapadão" [49]; central part of Mato Grosso State (municipality of Lucas

do Rio Verde and Sinop, along BR-163 highway), an environment-sensitive transition region between Cerrado and Amazon ecotones in the Mato Grosso State [50,51]; and MATOPIBA region (municipalities of Alto Parnaíba, Tasso Fragoso, and Balsas), where land prices are relatively low, and the climate and topographic conditions are favorable for rainfed agriculture [52].

Although our results agree with others cropland maps in terms of the major spatial distribution pattern, discrepancies are observed in some areas, especially in São Paulo State, that may be affected by the semi-perennial crops. São Paulo State (Southeast region) has been the major sugarcane Brazilian producer [31,34], however, fields of this crop type usually have a vegetative cycle that exceeds twelve months, not computed considering our one-year calendar (January to December). Besides, sugarcane crops do not present relatively high brightness differences in the vegetation fraction image, if compared with others crops types and different spectral responses, making it difficult to be detected and separated by the proposed method. According to the Agricultural Census [34], there are approximately 9.1 million hectares of sugarcane in Brazil.

These discrepancies noted here have two main reasons. The first is due to the longer growing season, which allows farmers to adopt two commercial cropping per year (double cropping, usually soybean followed by maize) [53]. The second reason is due to the limitation in relation to the low spatial resolution, which is better for discriminating larger farmers than smaller ones. For instance, smaller cultivated areas are usually formed by land-use mosaics (e.g., agroforestry, pasture, and subsistence crops), which cannot be discriminated at 100-m or 1-km resolution remote sensing data. Thus, the proposed approach is limited to detect double cropped areas since, in the peak of wet season (January to March), most of the soybeans are already harvested and maize is still in the beginning of the growing cycle (only cotton is in the flowering stage). Nevertheless, the top five most productive microregions (state-based political regions encompassing varying number of municipalities with relatively similar social and economic aspects) detected by the PROBA-V 100 m matched the ones detected by the Agricultural Census [34] (Table 2). The exception was the Primavera do Leste, ranked as fifth most productive microregion by the PROBA-V 100 m and as seventh by the Agricultural Census (data not shown).

PROBA-V 100 m			Agricultural Census		
Rank	Microregion	Area (ha)	Rank	Microregion	Area (ha)
1	Alto Teles Pires	1,258,684	1	Alto Teles Pires	2,244,905
2	Parecis	767,372	2	Parecis	1,309,408
3	Canarana	487,669	3	Canarana	1,007,393
4	Sinop	464,582	4	Sinop	812,524
5	Primavera do Leste	345,399	5	Arinos	697,565

Table 2. Comparison between area estimation of annual croplands in the Mato Grosso State obtained by PROBA-V 100 m mosaic (year: 2015) and Agricultural Census (year: 2016/2017).

Comparing our results from existing global datasets, as the FROM-GLC [15] and the GFSAD30 [16] (Figure 13), we found several differences on cropland area estimation. The FROM-GLS estimates an area of 27 million ha, while the GFSAD30 estimates 237 million ha in Mato Grosso State, overestimating all the results from the Brazilian initiatives cited before. Both methods (FROM-GLC and GFSAD30) are based on 30 m spatial resolution data set and the supervised algorithms are trained through representative samples. This means that the training dataset can be large if landscape is complex [16]. The GFSAD30 data consider almost all non-forest areas as cropland while in the FROM-GLC project the croplands varies a lot from bare field, seeding and crop growing to harvesting (e.g., the "bare farmland" class), and cannot be exactly identified as crop in that year. In this sense, our study advances on cropland estimations since the approach attempts to identify only planted areas, that is, areas with some crop plantation at least once a year, without using any training dataset related to the supervised classifications.



Figure 13. The Mato Grosso State cropland classification derived from PROBA-V 100 m spatial resolution data and comparison of the cropland maps in 2015. Note: The images have different scale bars in this figure.

PROBA-V 1 km mosaic, in relation to the PROBA-V 100 m, presented even more underestimated total area of annual croplands in Mato Grosso (3.4 million ha). Most of these differences were due to the presence of small areas (often smaller than the pixel size of PROBA-V 1 km) classified as annual crops only in the PROBA-V 100 m mosaic (Figure 14). These areas might be related to the small-scale maize production. As pointed out by [9,11], maize in Brazil is found in almost all municipalities of Brazil in small extents, since this crop is one of the most important national products of subsistence agriculture.



Figure 14. (**a**) The 2015 Mato Grosso State cropland classification derived from PROBA-V data and the comparison with 1 km and 100 m spatial resolutions. Details in (**b**) and (**c**).

4.2. Landsat-8 Validation Data Set

Within the 288 sampling units, we found 111.2 thousand ha of annual croplands according to the PROBA-V 100 m estimation in Mato Grosso State and 209.1 thousand ha according to the Landsat-8 OLI estimation, that is, an underestimation of 53% by the PROBA-V 100 m data set. The relation between Landsat-8 OLI estimation of area occupied by annual croplands and the corresponding estimation by the PROBA-V 100 m is shown in Figure 15. A set of 73 sampling units presented no annual croplands in both Landsat-8 OLI and PROBA-V datasets. The coefficient of determination was 0.82 and the linear regression equation was: y = 1.7131 x + 64.806. We suggest using this equation to correct the estimation of PROBA-V 100 m to overcome the underestimation trend presented by this study. In this case, the total area of annual croplands, estimated as 5.3 million ha will increase to 9.0 million ha, which is much closer to the 9.8 million ha estimated by the Agricultural Census of 2017 [34].



Figure 15. Relation between annual croplands estimated by Landsat-8 OLI and corresponding estimation by PROBA-V 100 m in the 288 sampling units. Continuous line refers to the linear regression line. Dashed line refers to the 1:1 line.

5. Conclusions

The maximum proportion values mosaic obtained from vegetation fraction images highlighting the agricultural areas in Brazil, allowed us to map the main annual cropping areas quickly. However, because of the concept used for generating the vegetation mosaics, it was not possible to discriminate the crop types. The generation of maximum fraction values minimizes the effect of cloud cover in the images, as observed by [54] who considered 16-day, MODIS NDVI composites from the rainy season to map crops in Canada.

This study showed that the bigger the pixel size of the data set, the lower the mapping accuracy of annual croplands in Brazil, since small patches of agricultural lands are not depicted, leading to underestimations. At this time, we disregarded the effects of local variations in the crop management (e.g., different planting dates, single or double cropping) in our data analysis.

Our results also showed underestimation of annual croplands both in Brazil and in Mato Grosso State, as compared with previous studies based either on Agricultural Census or on Landsat-like and time series of MODIS data, but agreed in terms of the major spatial distribution pattern. Despite the underestimation, which can be related to the coarse spatial resolution, varying periods in the crop calendar, and different cycle crops, the proposed method has the potential to be used operationally for detecting and mapping agricultural areas rapidly at global, regional, and local scales. The resulting maps are quite important for government and non-governmental organizations for planning, management, and monitoring the grain production in tropical environments, since the method is conservative because it only identifies areas that had at least one crop in the analyzed period.

Recommendations for further studies involving this proposed method are: to fix the crop calendar from October (previous year) to September, as adopted by the Brazilian Agricultural Census; to test different crop-dependent calendars, aiming to reduce uncertainties related to double-cropping areas; and to extract statistics from linear spectral mixing model to highlight semi-perennial crops as sugarcane.

Author Contributions: Processing data and formal analysis, E.A. and E.E.S.; writing and revision, E.A., E.E.S., A.C.D., H.L.G.C., and Y.E.S.; graphics and figures preparation, E.E.S. and A.C.D.; data collect, E.A and T.B.H.; conceptualization and experimental design, Y.E.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: This study was supported by the São Paulo Research Council (FAPESP-Grant 19806-3/2016 and by the Brazilian Research Council (CNPQ-Grant 301190/2013-5). H.L.G.C acknowledges the grant received from the FAPESP (2018/14423-4). The authors would like to thank the anonymous reviewers for their valuable comments and suggestions that greatly improved the quality of the technical and presentation form of our manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Thenkabail, P.S.; Wu, Z. An automated cropland classification algorithm (ACCA) for Tajikistan by combining Landsat, MODIS, and secondary data. *Remote Sens.* **2012**, *4*, 2890–2918. [CrossRef]
- See, L.; Fritz, S.; You, L.; Ramankutty, N.; Herrero, M.; Justice, C.; Becker-Reshef, I.; Thornton, P.; Erb, K.; Gong, P.; et al. Improved global cropland data as an essential ingredient for food security. *Glob. Food Sec.* 2015, 4, 37–45. [CrossRef]
- 3. Jayne, T.S.; Chamberlin, J.; Headey, D.D. Land pressures, the evolution of farming systems, and development strategies in Africa: A synthesis. *Food Policy* **2014**, *48*, 1–17. [CrossRef]
- Matton, N.; Canto, G.; Waldner, F.; Valero, S.; Morin, D.; Inglada, J.; Arias, M.; Bontemps, S.; Koetz, B.; Defourny, P. An automated method for annual cropland mapping along the season for various globally-distributed agrosystems using high spatial and temporal resolution time series. *Remote Sens.* 2015, 7, 13208–13232. [CrossRef]
- 5. Lambert, M.-J.; Waldner, F.; Defourny, P. Cropland mapping over Sahelian and Sudanian agrosystems: A knowledge-based approach using PROBA-V time series at 100-m. *Remote Sens.* **2016**, *8*, 232. [CrossRef]
- Gao, F.; Anderson, M.C.; Zhang, X.; Yang, Z.; Alfieri, J.G.; Kustas, W.P.; Mueller, R.; Johnson, D.M.; Prueger, J.H. Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. *Remote Sens. Environ.* 2017, 188, 9–25. [CrossRef]
- Chen, Y.; Lu, D.; Moran, E.; Batistella, M.; Dutra, L.V.; Sanches, I.D.; Silva, R.F.B.; Huang, J.; Luiz, A.J.B.; Oliveira, M.A.F. Mapping croplands, cropping patterns, and crop types using MODIS time-series data. *Int. J. Appl. Earth Obs. Geoinf.* 2018, 69, 133–147. [CrossRef]
- 8. Phalke, A.R.; Özdoğan, M. Large area cropland extent mapping with Landsat data and a generalized classifier. *Remote Sens. Environ.* **2018**, *219*, 180–195. [CrossRef]
- 9. Dias, L.C.P.; Pimenta, F.M.; Santos, A.B.; Costa, M.H.; Ladle, R.J. Patterns of land use, extensification, and intensification of Brazilian agriculture. *Glob. Chang. Biol.* **2016**, *22*, 2887–2903. [CrossRef]
- 10. 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]
- 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] [PubMed]
- 12. MapBiomas. Mapeamento Anual da Cobertura e Uso do Solo no Brasil. 2019. Available online: http://mapbiomas.org (accessed on 12 September 2019).
- 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]

- Zhu, C.; Lu, D.; Victoria, D.; Dutra, L. Mapping fractional cropland distribution in Mato Grosso, Brazil using time series MODIS Enhanced Vegetation Index and Landsat Thematic Mapper data. *Remote Sens.* 2015, *8*, 22. [CrossRef]
- 15. Gong, P.; Wang, J.; Yu, L.; Zhao, Y.; Zhao, Y.; Liang, L.; Niu, Z.; Huang, X.; Fu, H.; Liu, S.; et al. Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *Int. J. Remote Sens.* **2013**, *34*, 2607–2654. [CrossRef]
- 16. Zhong, Y.; Giri, C.; Thenkabail, P.S.; Teluguntla, P.; Congalton, R.G.; Yadav, K.; Oliphant, A.; Xiong, J.; Vijay, V.; Poehnelt, J. NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) Global Food Security-Support Analysis Data (GFSAD) @ 30-m for South America: Cropland Extent Product (GFSAD30SACE); NASA EOSDIS Land Processes DAAC: Sioux Falls, SD, USA, 2017.
- 17. Xu, Y.; Yu, L.; Zhao, F.R.; Cai, X.; Zhao, J.; Lu, H.; Gong, P. Tracking annual cropland changes from 1984 to 2016 using time-series Landsat images with a change-detection and post-classification approach: Experiments from three sites in Africa. *Remote Sens. Environ.* **2018**, *218*, 13–31. [CrossRef]
- 18. Hao, P.; Löw, F.; Biradar, C. Annual cropland mapping using reference Landsat time series—A case study in central Asia. *Remote Sens.* **2018**, *10*, 2057. [CrossRef]
- 19. Asner, G.P. Cloud cover in Landsat observations of the Brazilian Amazon. *Int. J. Remote Sens.* 2001, 22, 3855–3862. [CrossRef]
- 20. Sano, E.E.; Ferreira, L.G.; Asner, G.P.; Steinke, E.T. Spatial and temporal probabilities of obtaining cloud-free Landsat images over the Brazilian tropical savanna. *Int. J. Remote Sens.* **2007**, *28*, 2739–2752. [CrossRef]
- 21. 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]
- 22. Asner, G.P.; Knapp, D.E.; Broadbent, E.N.; Oliveira, P.J.; Keller, M.; Silva, J.N. Selective logging in the Brazilian Amazon. *Science* 2005, *310*, 480–482. [CrossRef] [PubMed]
- 23. Anwar, S.; Stein, A. Detection and Spatial Analysis of Selective Logging with Geometrically Corrected Landsat Images. *Int. J. Remote Sens.* 2012, *33*, 7820–7843. [CrossRef]
- 24. Shimabukuro, Y.E.; Arai, E.; Duarte, V.; Jorge, A.; Santos, E.G.D.; 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]
- 25. Quintano, C.; Fernández-Manso, A.; Fernández-Manso, O.; Shimabukuro, Y.E. Mapping Burned Areas in Mediterranean Countries Using Spectral Mixture Analysis from a Uni-Temporal Perspective. *Int. J. Remote Sens.* **2006**, *27*, 645–662. [CrossRef]
- 26. Shimabukuro, Y.E.; Arai, E.; Anderson, L.O.; Aragão, L.E.O.C.; Duarte, V. Mapping degraded forest areas caused by fires during the year 2010 in Mato Grosso state, Brazilian Legal Amazon using Landsat-5 TM fraction images. *Rev. Bras. Cart* **2017**, *69*, 23–32.
- 27. Shimabukuro, Y.E.; Duarte, V.; Arai, E.; Freitas, R.M.; Lima, A.; Valeriano, D.M.; Brown, I.F.; Maldonado, M.L.R. Fraction images derived from Terra MODIS data for mapping burnt areas in Brazilian Amazonia. *Int. J. Remote Sens.* **2009**, *30*, 1537–1546. [CrossRef]
- Roy, E.D.; Richards, P.D.; Martinelli, L.A.; Coletta, L.D.; 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]
- 29. FAO; FAOSTAT. Food and Agriculture Data. Available online: http://faostat.fao.org (accessed on 10 June 2019).
- 30. Vale, P.; Gibbs, H.; Vale, R.; Christie, M.; Florence, E.; Munger, J.; Sabaini, D. The expansion of intensive beef farming to the Brazilian Amazon. *Glob. Environ. Chang.* **2019**, *57*, 101922. [CrossRef]
- 31. IBGE. Produção Agrícola Municipal. Rio de Janeiro: IBGE, 2018. Available online: https://sidra.ibge.gov.br/ pesquisa/pam/tabelas (accessed on 17 June 2019).
- 32. IBGE. Mapas de biomas e de vegetação. Instituto Brasileiro de Geografia e Estatística. Available online: http://geoftp.ibge.gov.br/informacoes_ambientais/vegetacao/mapas/unidades_da_federacao/mt_vegetacao.pdf (accessed on 3 December 2019).
- 33. Durieux, L.; Machado, L.A.T.; Laurent, H. The impact of deforestation on cloud cover over the Amazon arc of deforestation. *Remote Sens. Environ.* **2003**, *86*, 132–140. [CrossRef]

- 34. IBGE. Censo Agropecuário. Rio de Janeiro: IBGE, 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 15 March 2019).
- 35. Wolters, E.; Dierckx, W.; Iordache, M.-D.; Swinnen, E. PROBA-v products user manual v3.01. Image (rochester, n.y.), n. C, pp. 1–110, 2018. Available online: http://proba-v.vgt.vito.be/sites/proba-v.vgt.vito.be/files/products_user_manual.pdf (accessed on 15 March 2019).
- 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]
- 37. Dutra, A.C.; Shimabukuro, Y.E.; Arai, E. Linear spectral mixing model applied in images from PROBA-V sensor: A spatial multiresolution approach. *RAEGA—O Espaço Geográfico em Análise* **2019**, *46*, 48–62. [CrossRef]
- Settle, J.J.; Drake, N.A. Linear mixing and the estimation of ground cover proportions. *Int. J. Remote Sens.* 1993, 14, 1159–1177. [CrossRef]
- Adams, J.B.; Sabol, D.E.; Kapos, V.; Almeida Filho, R.; Robert, D.A.; Smith, M.O.; Gillespie, A.R. Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon. *Remote Sens. Environ.* 1995, 52, 137–154. [CrossRef]
- 40. Keshava, N.; Mustard, J.F. Spectral Unmixing. IEEE Signal Process. Mag. 2002, 19, 44–57. [CrossRef]
- Bins, L.S.; Erthal, G.J.; Fonseca, L.M.G. Um método de classificação não supervisionada por regiões. In Proceedings of the Sixth Brazilian Symposium on Graphic Computation and Image Processing, Recife, Brazil, 9–18 October 1993; Volume 6, pp. 65–68.
- 42. Ouma, Y.O.; Josaphat, S.S.; Tateishi, R. Multiscale remote sensing data segmentation and post-segmentation change detection based on logical modeling: Theoretical exposition and experimental results for forestland cover change analysis. *Comput. Geosci.* **2008**, *34*, 715–737. [CrossRef]
- 43. Camara, G.; Souza, R.C.M.; Freitas, U.; Garrido, J. SPRING: Integrating remote sensing and GIS by object-oriented data model. *Comput. Graph.* **1996**, *20*, 395–403. [CrossRef]
- 44. Beuchle, R.; Grecchi, R.C.; Shimabukuro, Y.E.; Seliger, R.; Eva, H.D.; Sano, E.; Achard, F. 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]
- 45. Ridder, R.M. Global Forest Resources Assessment 2010: Options and Recommendations for a Global Remote Sensing Survey of Forests; Working Paper 141; FAO: Rome, Italy, 2007; p. 68.
- Waldner, F.; Fritz, S.; Di Gregorio, A.; Defourny, P. Mapping priorities to focus cropland mapping activities: Fitness assessment of existing global, regional and national cropland maps. *Remote Sens.* 2015, 7, 7959–7986. [CrossRef]
- 47. Fritz, S.; See, L.; McCallum, I.; Schill, C.; Obersteiner, M.; van der Velde, M.; Boettcher, H.; Havlík, P.; Achard, F. Highlighting continued uncertainty in global land cover maps for the user community. *Environ. Res. Lett.* 2011, *6*, 044005. [CrossRef]
- Scaramuzza, C.A.M.; Sano, E.E.; Adami, M.; Bolfe, E.L.; Coutinho, A.; Esquerdo, J.C.D.M.; Maurano, L.; Narvaes, I.D.; 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. Cart.* 2017, *69*, 1041–1051.
- Brannstrom, C.; Jepson, W.; Filippi, A.M.; Redo, D.; Xu, Z.; Ganesh, S. Land change in the Brazilian savanna (Cerrado), 1986–2002: Comparative analysis and implications for land-use policy. *Land Use Policy* 2008, 25, 579–595. [CrossRef]
- Morton, D.C.; DeFries, R.S.; Shimabukuro, Y.E.; Anderson, L.O.; Arai, E.; Espirito-Santo, F.B.; Freitas, R.; Morisette, J. Cropland expansion changes deforestation dynamics in the southern Brazilian Amazon. *Proc. Natl. Acad. Sci. USA* 2006, 103, 14637–14641. [CrossRef] [PubMed]
- 51. Gollnow, F.; Lakes, T. Policy change, land use, and agriculture: The case of soy production and cattle ranching in Brazil, 2001–2012. *Appl. Geogr.* **2014**, *55*, 203–211. [CrossRef]
- 52. Araújo, M.L.S.; Sano, E.E.; Bolfe, E.L.; Santos, J.R.N.; Santos, J.S.; Silva, F.B. Spatiotemporal dynamics of soybean crop in the Matopiba region, Brazil (1990–2015). *Land Use Policy* **2019**, *80*, 57–67. [CrossRef]

- Spera, S.A.; Cohn, A.S.; VanWey, L.K.; Mustard, J.F.; Rudorff, B.F.; Risso, J.; Adami, M. Recent cropping frequency, expansion, and abandonment in Mato Grosso, Brazil had selective characteristics. *Environ. Res. Lett.* 2014, 9, 064010. [CrossRef]
- 54. Chen, Y.; Song, X.; Wang, S.; Huang, J.; Mansaray, L.R. Impacts of spatial heterogeneity on crop area mapping in Canada using MODIS data. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 451–461. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).