

EXPLORING THE HARMONIZED LANDSAT SENTINEL (HLS) DATACUBE TO MAP AN AGRICULTURAL LANDSCAPE IN THE BRAZILIAN SAVANNA

T. C. Parreiras^{1*}, E. L. Bolfe^{1,2}, E. S. Sano³, D. C. Victoria², I. D. Sanches⁴, L. E. Vicente⁵

¹ State University of Campinas - Institute of Geosciences, Campinas-SP, Brazil. tayacristo1@gmail.com

² Brazilian Agricultural Research Corporation – Embrapa Agricultura Digital, Campinas-SP, Brazil. (edson.bolfe, daniel.victoria)@embrapa.br

³ Brazilian Agricultural Research Corporation – Embrapa Cerrados, Brasília-DF, Brazil. edson.sano@embrapa.br

⁴ Earth Observation and Geoinformatics Division, National Institute for Space Research, São José dos Campos-SP, Brazil. ieda.sanches@inpe.br

⁵ Brazilian Agricultural Research Corporation - Embrapa Meio Ambiente, Jaguariúna-SP, Brazil. luiz.vicente@embrapa.br

Commission III, WG III/10

KEYWORDS: Agriculture, Cerrado Biome, Harmonized Landsat Sentinel, Classification, Random Forest.

ABSTRACT:

Brazil has established itself as one of the world leaders in food production. Different types of remote sensing mapping techniques have been undertaken to support rural planning in the country. However, due to the complex dynamics of Brazilian agriculture, especially in the Cerrado biome (tropical savanna), there is a need for more feasible crop discrimination and monitoring initiatives, which require a consistent time series of remote sensing data at medium meter and potentially up to 3 day Landsat 8 and Sentinel-2 satellite time series, minimizing the cloud cover limitations for rainfed agricultural monitoring. This paper aims to explore the potential of the Harmonized Landsat 8 Sentinel-2 (HLS) data cube to map agricultural landscapes in the Brazilian Cerrado. The HLS multispectral bands from 27 scenes with less than 10% cloud cover, from October 2020 to September 2021, encompassing one entire crop growing season, were processed by the Random Forest algorithm to produce a map with four land use/cover classes (annual crops, sugarcane, renovated sugarcane fields, cultivated pastures, and native Cerrado). We performed accuracy assessment through 10-fold cross-validation and confusion matrix analyses. The results showed a high level of overall accuracy and Kappa coefficient, both with 99%, as well as high user's and producer's accuracies of at least 99%. The HLS dataset has been continuously improved, showing very promising results for rainfed agricultural mapping and monitoring.

1. INTRODUCTION

Global consumption of food, water, and energy has grown exponentially in recent years. This increasing demand is placing pressure on natural resources of countries with high potential for agricultural production, which is the case of Brazil. The country has good technical and agronomic conditions to produce grain, meat, and biofuel with high quality and sustainability. We expect that the grain production will rise from the current 251 million tons in 2019/2020 (65 million hectares) to approximately 318 million tons in 2029/2030 (76 million hectares), a 27% increase (Bolfe et al., 2020).

Nowadays, the Brazilian tropical savanna (Cerrado biome) is the most important agricultural frontier in the country. This biome occupies approximately 23% of the Brazilian territory, covering an area of approximately 2 million km² in the central part of the country (Pereira et al., 2020). It is partially distributed in the states of Bahia, Goiás, Maranhão, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Paraná, Piauí, Rondônia, São Paulo, Tocantins, and the Federal District.

Because of its complex and increasingly dynamic agricultural production, remote sensing analysis is fundamental for mapping, modelling, and monitoring the processes of agricultural expansion, retraction, conversion, intensification, and diversification (Sano et al., 2019). Indeed, different land use and land cover (LULC) mapping initiatives in the Brazilian Cerrado have supported rural planning in this biome (Arantes et al., 2016; Noojipady et al., 2017; Sano et al., 2019). These

initiatives are mostly based on the analysis of the Landsat satellite data. However, the image acquisition mode of the Landsat, which is based on 16-day repeat pass is often limited for accurate LULC mapping and monitoring over rainfed agricultural lands (Prudente et al., 2020).

The constellation of two Sentinel-2 satellites, launched in 2015 and 2017, operates, in some extent, with similar spectral characteristics of the Landsat 8. However, Sentinel-2 provides better spatial resolution (10 m) in the visible (400–720 nm) and near-infrared (720–1100 nm) bands and higher temporal resolution (10 days or 5 days if the two satellites are combined). The open-access policy of Landsat and Sentinel-2 multispectral data allows the synergistic use of the Harmonized Landsat 8 and Sentinel-2 (HLS) data sets (Claverie et al., 2018). This creates unprecedented opportunities for timely and accurate observation of the dynamics of the Earth's surface. The multisensor analysis of satellite data expands the possibility of obtaining more accurate discrimination of different LULC types over agricultural lands (Bègué et al., 2018).

The high persistence of clouds in tropical regions (Prudente et al., 2020), the high level of spectral mixing mainly in the coarse-resolution sensors, and the difficulty of obtaining field data because of the poor road network and conditions are the main sources of classification errors. Researchers have developed and tested different classification strategies and routines to minimize the effects of these sources of errors. Among these strategies, we find, in literature, the use of

*Corresponding author

different vegetation indices (Kuchler et al., 2020), phenology metrics (Werner et al., 2019), and machine learning algorithms such as the Random Forest (RF), Support Vector Machine (SVM), neural networks, and decision trees, which have presented promising results (Chen et al., 2018; Picoli et al., 2018). These and other strategies have resulted in increasingly detailed and accurate mappings and can be easily applied to the HLS data. However, detailed agricultural monitoring over large areas by combining high temporal and spatial resolution data is still a challenge. The multisensor approach involving data harmonization has great potential to overcome this issue (Bègué et al., 2018; Hao et al., 2019; Bolton et al., 2020).

The use of multisensor digital image classification techniques to improve the accuracy of agricultural mapping and monitoring in different countries is increasing. Some studies have demonstrated the potential of the harmonized series for different applications, including crop type mapping (Hao et al., 2019; Dong et al., 2020; Gao et al., 2020), crop phenology (Nguyen et al., 2020), irrigated areas (Bolognesi et al., 2020), intensification of pasture lands (Griffiths et al., 2020), and vegetation characterization (Bolton et al., 2020).

In Brazil, time series of remote sensing images also have been used to study the LULC dynamics (Picoli et al., 2018; Bendini et al., 2019) with many investigations combining different image analysis methods to improve the agricultural mapping, modelling, and monitoring (Werner et al., 2019; Kuchler et al., 2020). To date, few efforts have used the harmonized Landsat 8 and Sentinel-2 (HLS) multispectral images in the Cerrado biome. Since 2020, NASA is making HLS data available for the Cerrado region, especially after 2021, with the release of the V2.0 collection (NASA, 2021). This paper aims to explore the potential of HLS time series to map LULC in a watershed located in the Brazilian Cerrado with high agricultural production, mostly rainfed grains and sugarcane.

2. MATERIAL AND METHODS

2.1 Study area

The study area corresponds to the Santa Barbara watershed that occupies an area of 1539 km² in the southern part of the Goiás State (Figure 1). It encompasses part of the municipalities of Goiatuba, Bom Jesus de Goiás, Joviânia, and Vicentinópolis, an important agricultural production region in the Brazilian Cerrado. The altitude varies between 452 m and 824 m, with an average of 637 m. The topography is dominantly flat (typical slope: 0–4%) (Fioreze et al., 2010). The climate is tropical, with hot and humid summers, and dry winters – Aw in the Köppen's climate classification system (Alvares et al., 2014). The average annual precipitation is 1520 mm, concentrated between December and March (Fioreze et al., 2010).

Reddish, deep, and well-developed Oxisols (*Latossolo Vermelho* in the Brazilian System of Soil Classification) is the dominant type of soil in this watershed. The flat relief favours the introduction of irrigated agriculture, mostly by the center-pivot irrigation system (Fioreze and Oliveira, 2010). Although the region is marked by sugarcane production, we also find large-scale cultivation of annual crops, mainly soybean and maize. Crops such as soybeans, maize, and sorghum represent between 65% and 75% of the total agricultural production of the municipalities that cover partially the study area. Sugarcane corresponds to 27% of the harvested area in the municipality of Goiatuba, and 23% in the municipality of Bom Jesus de Goiás (IBGE, 2021).

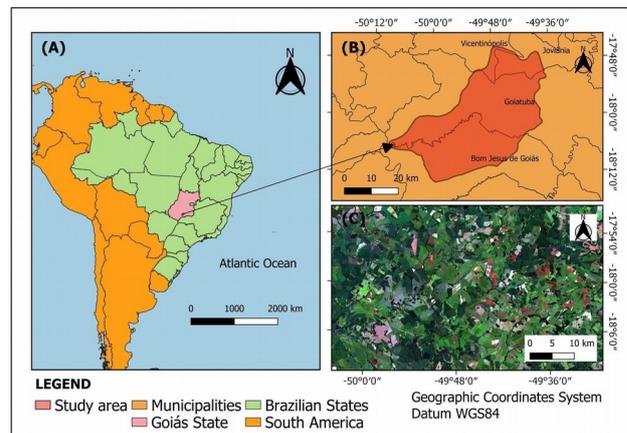


Figure 1. Santa Barbara watershed (study area) in the Goiás State, Brazil. The true-color composition (Red-Blue-Green) (C) is part of the Harmonized Landsat 8 Sentinel-2 collection from 21 April 2021. Base maps were obtained from the Brazilian Institute of Geography and Statistics (IBGE) and Goiás State Geoinformation System (SIEG).

2.2 Harmonized Landsat 8 Sentinel-2 (HLS) characteristics

The HLS data made available by NASA with a regular grid of 109.8 km × 109.8 km tiles are converted into the Universal Transverse Mercator (UTM) projection system, datum WGS84, spatial resolution of 30 m, and radiometric resolution of 16 bits. The Land Surface Reflectance Code (LaSRC) algorithm is used for atmospheric correction. HLS products provide per-pixel masks of cloud, cloud shadow, snow, and water. For spatial co-registration, the Automated Registration and Orthorectification Package (AROP) is used, with resampling of the Sentinel-2 bands, from 10 m, 20 m, and 60 m to 30 m. The HLS30 (Harmonized Landsat 8 Sentinel-2 data with 30-meter spatial resolution) uses the Band Pass Adjustment procedure, a band-by-band process in which the Landsat spectral bands are used as a reference to perform linear adjustment of equivalent Sentinel bands. Except for the near-infrared (B8) band and for the three red-edge (B5, B6, and B7) bands, all other Sentinel bands have equivalent Landsat bands. The bidirectional reflectance distribution function (BRDF) is normalized to minimize the differences between the sensor and solar viewing angles using the c-factor technique and global coefficients (Claverie et al., 2018).

In September 2021, NASA released the new HLS30 data collection, version 2.0, following the version 1.5 (NASA, 2021). By the time this study was conducted, HLS collection 2.0 was under active development, with new data for the Brazilian Cerrado being released regularly, especially from January 2021.

2.3 Image acquisition and processing

In this study, we used the following harmonized products: Sentinel-2 Multispectral Instrument (MSI) surface reflectance 30 m (called S30) and the Landsat 8 Operational Land Imager (OLI) surface reflectance and top-of-atmosphere (TOA) brightness 30 m over the Sentinel-2 tilling system (called L30), from collections 1.5 and 2.0.

The HLS product description can be accessed at the following website: <https://hls.gsfc.nasa.gov/products-description/>. We downloaded all data from the Earth Data portal (<https://search.earthdata.nasa.gov/>). With the launch of version

2.0, the availability of images for the Cerrado region increased significantly, with the temporal resolution reaching up to 3 days. In our data set, we considered a threshold of 10% of cloud cover over the study area and absence of cirrus as the main requirements for selecting the images. The Cerrado agricultural calendar for the first crop goes mostly from October to March, comprising the rainy season. However, due to the longer sugarcane cycle (semi perennial crop), with harvest starting in August, the images were taken for a period of one year, between October 2020 and September 2021.

2.4 Datasets

We masked clouds and water bodies using the F-mask product, while urban areas, main roads, and center-pivots were masked manually. The HLS30 (S30 + L30) dataset was composed of the following multispectral bands: blue (490 nm), green (560 nm), red (665 nm), near-infrared (NIR, 840 nm), and two shortwave infrared channels (SWIR, 1610 and 2190 nm). We stacked all bands into a single file and clipped to the area of interest.

2.5 LULC classes and sampling data

In this study, we considered the following representative LULC classes from the study area, which were selected based on the agricultural production data (IBGE, 2021) and the time series of multispectral images: cultivated pastures, native vegetation, sugarcane, and annual crops. The HLS time series showed the presence of two well-defined spectral patterns of the sugarcane, one corresponding to the regular, productive sugarcane, and the other corresponding to renovated fields. After successive harvests, sugarcane presents a gradual loss in the yield, so that the area needs to be renewed before new sowing (Rudorff et al., 2010). The most common leguminous species used for renovation are the *Crotalaria juncea*, soybean, sorghum and millet.

The sampling regions were delimited with the support of QGIS 3.8 software and based on the analysis of one color composite HLS image per month, considering the following parameters of photo-interpretation: geometry, color, texture, and context. We also analyzed the Normalized Difference Vegetation Index (NDVI) profiles derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) data to double check the selected LULC class. These profiles are produced by the Brazilian Agricultural Research Corporation (Embrapa) and can be accessed through the Vegetation Temporal Analysis System (SATVeg) web tool (Embrapa, 2021). Using the R statistical package, version 4.1.2, the polygons were converted into points, from which the band attributes were extracted at the pixel level. Table 1 describes the number of sampled pixels and the corresponding total sampled area. Examples of NDVI (2020–2021) temporal profiles for each class are shown in Figure 2.

2.6 Classification and validation

The classification with the RF algorithm was performed using the R statistical package, version 4.1.2 and the randomForest package (Liaw and Wiener, 2002). The RF is an algorithm to classify images using an ensemble of decision trees. It is based on the Gini index in which the independence of the trees is a key factor for obtaining better performance. The RF is widely used in crop mapping and yield estimation because of its high efficiency to deal with large sampling databases (Fang et al., 2020).

The sample dataset was randomly split into two subsets, one for training, with 70% of the data, and another for validation, with 30%. In order to perform RF classification, the number of trees (nTree), and the number of variables in each node (mTry) need to be adjusted. After evaluating the error variation in relation to the number of trees, the nTree was adjusted to 500 while the mTry was adjusted through 10-fold cross-validation and defined as 8.

The caret package (Kuhn, 2008) was used to calculate the confusion matrix between the observed and predicted values in the validation subset, from which the overall accuracy and Kappa coefficient values were obtained, in addition to the user's and producer's accuracies (UA and PA, respectively). According to Landis and Koch (1977), Kappa coefficient provides the level of agreement between the observed and the predicted classes. It can be classified as poor (≤ 0.00), slight (0.00–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), and almost perfect (0.81–1.00). The UA metric indicates the probability of a classified pixel actually represents the correct category on the ground, and the PA is a measure of omission error, indicating how well a certain area can be classified (Congalton, 1991). To assess the model's robustness, we run a 10-fold cross-validation with the dataset split randomly into training (70%) and test (30%) subsets.

Class	Sampled pixels	Sampled area (km ²)
Sugarcane	2276	2.0
Annual crops	2113	1.97
Cultivated pasture	2145	1.91
Cerrado	2393	1.96
Renovated *	2246	1.96
Total	11173	9.8

Table 1. Number of sampled pixels and corresponding sampled area, per class. *Renovated sugarcane fields.

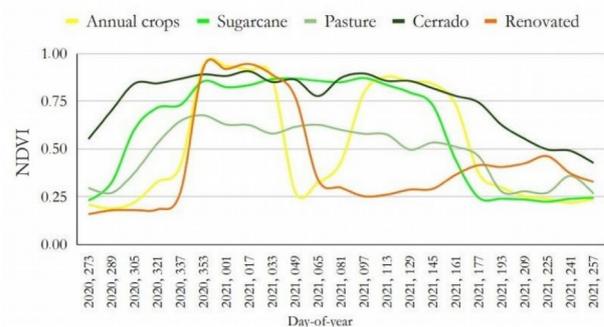


Figure 2. Representative temporal profile of the MODIS NDVI for each LULC class of the Santa Barbara watershed, Goiás State, over the 2020–2021 season. X-axis corresponds to the Julian days.

3. RESULTS

3.1 Number of available cloud-free HLS30 images

During the 2020–2021 season, we selected 27 nearly cloud-free HLS30 overpasses (Table 2), which is a remarkable achievement. For comparison purposes, we carried out a search in the Earth Explorer platform to verify the number of Landsat Operational Land Imager (OLI) overpasses available in the study area for the same time period considered in this study (October 2020 to September 2021), using a cloud-cover threshold of 10%. Only 9 out of 24 scenes were available. However, despite the high number of nearly cloud-free HLS30 images, none of them was obtained in December, January, or

February, which correspond to the local rainy season. In a study developed by Parente et al. (2017), 80% of the Brazilian territory presented less than 12 cloud-free OLI images in 2015, which undermines any analysis involving monitoring of annual crop plantation and production.

The balance between the spatial and temporal resolutions of time series is one of the biggest challenges in agricultural monitoring based on satellite data. While MODIS sensor provides a high-frequency revisiting rate (1–2 days) and coarse resolution data (250 m), Landsat 8 OLI (and, more recently, Landsat 9) provides images with 30 m spatial resolution, but with coarser temporal resolution (16-day), which can result in gaps, especially in the rainy season (Prudente et al., 2020; Silva Júnior et al., 2020).

Data	Tile	Date	Bands
HLS	T22KEF T22KFE T22KFF	2020-10-04	S30 (2, 3, 4, 8A, 11 and 12); L30 (2, 3, 4, 5, 6 and 7)
		2020-10-14	
		2020-11-28	
		2021-03-28	
		2021-04-21	
		2021-04-22	
		2021-05-02	
		2021-05-27	
		2021-06-16	
		2021-06-21	
		2021-06-24	
		2021-06-26	
		2021-07-01	
		2021-07-06	
		2021-07-10	
		2021-07-16	
		2021-07-21	
		2021-07-26	
		2021-07-31	
		2021-08-11	
		2021-08-15	
		2021-08-20	
		2021-08-25	
		2021-08-27	
		2021-09-04	
		2021-09-19	
2021-09-24			

Table 2. Detailed information about the Harmonized Landsat 8 Sentinel-2 (HLS) data used in this study.

The Multispectral Instrument (MSI), onboard the Sentinel-2A and Sentinel-2B satellites, indeed provide improved spatial (10 and 20 m), temporal (5 days), and spectral (13 bands) resolutions in relation to the Landsat OLI. Nevertheless, its computational demand is relatively high, making its application in large-scale monitoring more time-consuming. In this sense, the ready-to-use HLS30 freely available data can balance the high processing time demanding issue of Sentinel-2 datasets and the relatively low temporal resolution of Landsat datasets. Another advantage of the HLS30 is the possibility of having a higher number of nearly cloud-free images in comparison with the datasets involving only Sentinel-2 and Landsat 8 images.

3.2 Classification results and accuracy assessment

Figure 3 shows the result of the RF classification for the study area while Table 3 shows the total area mapped by each LULC class. It can be noted the predominance of agricultural activity in the study area. Approximately 89% of the watershed are destined for the production of grains, sugarcane, irrigated

agriculture, and cultivated pastures. The native Cerrado occupies 12% of the watershed, distributed mainly along the watercourses as riparian forests. A set of 62 center-pivots were visually identified, occupying an area of 4397 ha.

There is a predominance of double cropping system rather than the single cropping system, as showed by the two well-defined peaks of NDVI values during the rainy season. In the case of the renovated sugarcane fields, there is only one growing peak in the rainy season, since the sugarcane will be planted for the next productive cycle (Figure 2). Although we expected misclassification between sugarcane and cultivated pasture, as they are both grasses, we found a quite different spectral responses throughout the year. Cultivated pastures are much more sensitive to the strong climate seasonality of the Cerrado than sugarcane plantations.

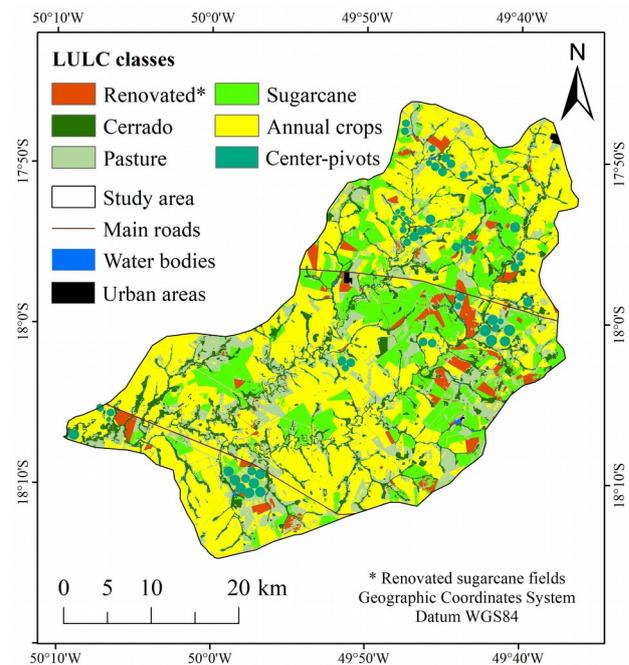


Figure 3. Final map produced through the Random Forest classification with the Harmonized Landsat 8 Sentinel-2 multispectral dataset, showing the target classes and masked areas.

Classes	Area (km ²)	Area (%)
Sugarcane	317.74	20.76
Annual crops	694.03	45.35
Cultivated pasture	290.77	19.00
Cerrado	163.43	10.67
Renovated*	64.58	4.22

Table 3. Total area mapped with Harmonized Landsat 8 Sentinel-2 (HLS30) images and Random Forest classifier, per class. *Renovated sugarcane fields.

The RF model achieved an overall accuracy and a Kappa index of 0.99. The 10-fold cross-validation assured the performance by showing the same coefficient for both metrics (0.99), indicating an almost perfect level of agreement between estimated and predicted data. Table 4 shows the confusion matrix along with the UA and PA for each class. The results show that the HLS time series was very efficient in capturing the differences of the spectral response variations among the LULCs, producing only three occurrences of misclassifications in more than 3000 pixels in the validation subset.

Although comparing the results from other studies is not straightforward, as it involves different scales, areas, and products, the results obtained by our study are very promising. Bendini et al. (2019) analyzed the Landsat 7 and Landsat 8 time series converted into a vegetation index from different agricultural lands of the Brazilian Cerrado by a temporal interpolation technique and RF classifier. They achieved an accuracy higher than 0.88 involving more than 10 LULC classes. Oldoni et al. (2019), using metrics derived from Landsat 8 multispectral bands and RF classifier, obtained a Kappa index of 0.72 and an overall accuracy of 76%. These authors analyzed seven LULC classes in the Paraná State. Souza Júnior et al. (2020) mapped three decades of LULC classes for the entire Brazilian territory using a hierarchical classification of the Landsat image collection available in the Google Earth Engine platform. The accuracy of the RF-based classification ranged from 75% to 95%, depending on the biome. At the first level of the hierarchical classification, the authors observed accuracy of 81% for the Cerrado biome.

Class	Sug.	Ann.	Past.	Cer.	Ren.
Sug.	633	0	0	0	0
Ann.	1	680	0	0	0
Past.	0	0	641	0	0
Cer.	2	0	0	711	0
Ren.	0	0	0	0	669
UA	100	99.85	100	99.72	100
PA	99.52	100	100	100	100

Table 4. Confusion matrix and user's (UA) and producer's accuracy (PA) for each class obtained with the Random Forest classification applied to the Harmonized Landsat 8 Sentinel-2 multispectral dataset. Sug. = sugarcane; Ann. = annual crops; Past. = pastures; Cer. = Cerrado native vegetation; Ren. = renovated sugarcane fields.

As already discussed, the 30-meter temporal resolution is one of the main advantages of the Landsat multispectral images because it allows a relatively detailed mapping at regional scales, with lower computational cost. However, the low revisiting frequency of the Landsat does not permit the acquisition of consistent time series, often requiring the use of interpolation techniques to reconstitute highly contaminated images by clouds. Silva Júnior et al. (2020) compared the performance of MODIS, Landsat 8 OLI, and Sentinel-2 MSI sensors in mapping soybean production in the Mato Grosso State. The Landsat images resulted in the worst performance because of their low number of available images. By adopting a 3% threshold for cloud cover, Oldoni et al. (2019) obtained 11 Landsat 8 OLI images for two consecutive agricultural years (2015–2016 and 2016–2017) in a study carried out in the Paraná State. Montibeller et al. (2019) also pointed out the low temporal resolution of the Landsat 8 OLI as the main source of misclassifications.

In this study, we relied on multispectral bands instead of vegetation indices. Montibeller et al. (2019) showed that spectral-temporal profiles using multispectral bands are more efficient than the vegetation indices in distinguishing agricultural classes such as soybean, maize, and sugarcane. Chaves et al. (2019) also reported that the combination of multispectral bands in the shortwave infrared and red-edge wavelengths, in general, improves classification performances. The authors also showed that the machine learning algorithms improved the mapping accuracy. They also highlighted the potential of open-access harmonized images.

Mapping the spatial distribution of agricultural areas and the types and patterns of cropping systems is an important activity for crop yield estimation and LULC planning. They are parts of a critical stage of studies involving remote sensing in agriculture (Bègué et al., 2018; Chen et al., 2018). The acquisition of consistent cloud-free time series is the biggest challenge for the accurate and detailed mapping of agricultural systems in Brazil. The higher the density of the time series, the higher the chances of detecting variations inherent to the crop phenology. This information is essential not only to characterize each crop cycle but also to distinguish crop types and crop management techniques (Prudente et al., 2020; Bègué et al., 2018). The HLS dataset was able to capture these variations even without cloud-free images from December to February.

The HLS data were efficient in mapping LULC classes found in the Santa Barbara watershed located in the southwest of the Brazilian Cerrado with an accuracy of 0.99, an almost perfect level of agreement, and precision of at least 99% in the detection of sugarcane and renovated sugarcane fields, annual crops, native vegetation, and pasturelands. Since 2021, the HLS30 version 2.0 is providing more regular and high-quality data over the Cerrado and the results can be even better for new studies. Although there are still challenges and some issues, especially regarding time gaps over the Cerrado and irregularities in the F-mask, HLS is an important source of open access, multitemporal dataset in which other applications can be explored. These images can be processed through different image classification methods and different enhancement techniques such as the vegetation indices and spectral mixture modeling. Their accuracy and uncertainties should be evaluated in other landscape conditions from the Cerrado biome or other ecosystems.

This study was carried out using training and validation data acquired remotely. Therefore, further investigations using field data as the ground truth to train and validate the classifiers may bring more reliability of such models. The main advantage of the HLS data cube, that is, the high temporal resolution combined with the 30-meter resolution, was demonstrated in this study. The ongoing studies involving other agricultural frontiers in the Brazilian Cerrado will be able to show the potential of this time series to monitor food and energy production in this biome more consistently.

4. CONCLUSIONS

1. The HLS30, multispectral dataset showed high potential for mapping rainfed crop production in the Brazilian savanna.
2. We were able to count on 27 nearly cloud-free HLS overpasses within the period from October 2020 to September 2021, overcoming the low density of time series if only Landsat or Sentinel-2 images are considered.
3. Our results showed the potential of HLS30 to support public policies that rely on accurate LULC maps of the Brazilian savanna.

ACKNOWLEDGEMENTS

We gratefully acknowledge the funding support by the São Paulo Research Foundation (FAPESP) (Grant # 2019/26222, Agricultural mapping in the Cerrado via combined use of multisensor images), and the National Council for Scientific and Technological Development (CNPq) for the research fellowship for the second, third and fifth authors.

REFERENCES

- Alvares, C.A., Stape, J.L., Sentelhas, P.C., Gonçalves, J.L.M., Spavorek, G. 2014. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, 22(6), 711–728. <https://doi.org/10.1127/0941-2948/2013/0507>.
- Arantes, A., Ferreira, L., Coe, M. 2016. The seasonal carbon and water balances of the Cerrado environment of Brazil: past, present, and future influences of land cover and land use. *ISPRS Journal of Photogrammetry and Remote Sensing*, 117(7), 66–78. <https://doi.org/10.1016/j.isprsjprs.2016.02.008>.
- Bègué, A., Arvor, D., Bellon, B., Betbeder, J., Albelleyra, D., Ferraz, R.P.D., Lebourgeois, V., Lelong, C., Simões, M., Veron, S.R. 2018. Remote sensing and cropping practices: A review. *Remote Sensing*, 10(1), 99. <https://doi.org/10.3390/rs10010099>.
- Bendini, H.D., Fonseca, L.M.G., Schwieder, M., Kortig, T.S., Rufin, P., Sanches, I.D., Leitão, P.J., Hostert, P. 2019. Detailed agricultural land classification in the Brazilian Cerrado based on phenological information from dense satellite image time series. *International Journal of Applied Observation and Geoinformation*, 82, 101872. <https://doi.org/10.1016/j.jag.2019.05.005>.
- Bolfe, E., Jorge, L., Sanches, I., Luchiani Jr., A., Costa, C., Victoria, D., Inamasu, R., Grego, C., Ferreira, V., Ramirez, A. 2020. Precision and digital agriculture: Adoption of technologies and perception of Brazilian farmers. *Agriculture*, 10(12), 653. <https://doi.org/10.3390/agriculture10120653>.
- Bolognesi, S., Pasolli, E., Belfiori, O., Michele, C., D'Urso, G. 2020. Harmonized Landsat-8 and Sentinel-2 time series data to detect irrigated areas: An application in southern Italy. *Remote Sensing*, 12(8), 1275. <https://doi.org/10.3390/rs12081275>.
- Bolton, D., Gray, J., Melaas, E., Moon, M., Eklundh, L., Friedl, M. 2020. Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. *Remote Sensing of Environment*, 240(4), 111685. <https://doi.org/10.1016/j.rse.2020.111685>.
- Chaves, M.E.D., Picoli, M.C.A., Sanches, I.D. 2020. Recent applications of Landsat 8/OLI and Sentinel-2/MSI for land use and land cover mapping: A systematic review. *Remote Sensing*, 12, 3062. <https://doi.org/10.3390/rs12183062>.
- Chen, Y.L., Lu, D.S., Moran, E., Batistella, M., Dutra, L.V., Sanches, I.D., Silva, R.F.B., Huang, J.F., Luiz, A.J.B., Oliveira, M.A.F. 2018. Mapping croplands, cropping patterns, and crop types using MODIS time-series data. *International Journal of Applied Earth Observation and Geoinformation*, 69, 133–147. <https://doi.org/10.1016/j.jag.2018.03.005>.
- Claverie, M., Ju, J., Masek, J.G., Dungan, J., Vermote, E., Roger, J., Skakun, S., Justice, C. 2018. The harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment*, 219(12), 145–161. <https://doi.org/10.1016/j.rse.2018.09.002>.
- Congalton, R.G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1), 35–46. [https://doi.org/10.1016/0034-4257\(91\)90048-B](https://doi.org/10.1016/0034-4257(91)90048-B).
- Dong, T., Liu, J., Qian, B., He, L., Liu, J., Wang, R., Jing, Q., Champagne, C., McNairn, H., Powers, J., Shi, Y., Chen, J., Shang, J. 2020. Estimating crop biomass using leaf area index derived from Landsat 8 and Sentinel-2 data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 168(10), 236–250. <https://doi.org/10.1016/j.isprsjprs.2020.08.003>.
- Embrapa. 2021. SATVeg - Sistema de Análise Temporal da Vegetação. <https://www.satveg.cnptia.embrapa.br> (10 November 2021).
- Fang, P., Zhang, X., Wei, P., Wang, Y., Xiang, H., Liu, F., Zhao, J. 2020. The classification performance and mechanism of machine learning algorithms in winter wheat mapping using Sentinel-2 10 m resolution imagery. *Applied Sciences*, 10, 5075. <https://doi.org/10.3390/app10155075>.
- Fioreze, A.P., Oliveira, L.C. 2010. Usos dos recursos hídricos na Bacia Hidrográfica do Ribeirão Santa Bárbara, Goiás. *Pesquisa Agropecuária Tropical*, 40(1), 28–35. <https://doi.org/10.5216/pat.v40i1.3869>.
- Fioreze, A.P., Oliveira, L.C., Franco, A.P.B. 2010. Caracterização morfológica da Bacia Hidrográfica do Ribeirão Santa Bárbara, Goiás. *Pesquisa Agropecuária Tropical*, 40(2), 167–173. <https://doi.org/10.5216/pat.v40i2.3931>.
- Gao, F., Anderson, M., Daughtry, C., Karnieli, A., Hively, D., Kustas, W. 2020. A within-season approach for detecting early growth stages in corn and soybean using high temporal and spatial resolution imagery. *Remote Sensing of Environment*, 242(1), 111752. <https://doi.org/10.1016/j.rse.2020.111752>.
- Griffiths, P., Nendel, C., Pickert, J., Hostert, P. 2020. Towards national-scale characterization of grassland use intensity from integrated Sentinel-2 and Landsat time series. *Remote Sensing of Environment*, 238(3), 111124. <https://doi.org/10.1016/j.rse.2019.03.017>.
- Hao, P., Tang, H., Chen, J., Wu, M. 2019. High resolution crop intensity mapping using harmonized Landsat-8 and Sentinel-2 data. *Journal of Integrative Agriculture*, 18(12), 2883–2897. [https://doi.org/10.1016/S2095-3119\(19\)62599-2](https://doi.org/10.1016/S2095-3119(19)62599-2).
- IBGE. 2021. Produção Agrícola Municipal (PAM). <https://dados.gov.br/dataset/pa-producao-agricola-municipal> (10 November 2021).
- Kuchler, P.C., Bègué, A., Simões, M., Gaetano, R., Arvor, D., Ferraz, R.P.D. 2020. Assessing the optimal preprocessing steps of MODIS time series to map cropping systems in Mato Grosso, Brazil. *International Journal of Applied Earth Observation and Geoinformation*, 92, 102150. <https://doi.org/10.1016/j.jag.2020.102150>.
- Kuhn, M. 2008. Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5), 1–26. <https://doi.org/10.18637/jss.v028.i05>.
- Landis, J., Koch, G. 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>.
- Liaw, A., Wiener, M. 2002. Classification and regression by randomForest. *R News*, 2(3), 18–22.

- Montibeller, B., Sanches, I.D.A., Luiz, A.J.B., Gonçalves, F., Aguiar, D.A. 2019. Spectral-temporal profile analysis of maize, soybean and sugarcane on OLI/Landsat 8 data. *Brazilian Journal of Agriculture*, 94(3), 242–258. <https://doi.org/10.37856/bja.v94i3.3612>.
- NASA. 2021. Release of Harmonized Landsat and Sentinel-2 (HLS) Version 2.0. <https://lpdaac.usgs.gov/news/release-of-harmonized-landsat-and-sentinel-2-hls-version-20/> (20 October 2021).
- Nguyen, M.D., Baez-Villanueva, O., Bui, D.D., Nguyen P.T., Ribbe, L. 2020. Harmonization of Landsat and Sentinel-2 for crop monitoring in drought prone areas: case studies of Ninh Thuan (Vietnam) and Bekaa (Lebanon). *Remote Sensing*, 12(2), 281. <https://doi.org/10.3390/rs12020281>.
- Noojipady, P., Morton, D., Macedo, M., Victoria, D., Huang, C., Gibbs, H., Bolfe, E. 2017. Forest carbon emissions from cropland expansion in the Brazilian Cerrado biome. *Environmental Research Letters*, 12(2), 025004. <https://iopscience.iop.org/article/10.1088/1748-9326/aa5986>.
- Oldoni, L.V., Cattani, C.E.V., Mercante, E., Johann, J.A., Antunes, J.F.G., Almeida, L. 2019. Annual cropland mapping using data mining and OLI Landsat-8. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 23(12), 952–958. <http://dx.doi.org/10.1590/1807-1929/agriambi.v23n12p952-958>.
- Parente, L., Ferreira, L., Faria, A., Nogueira, S., Araújo, F., Teixeira, L., Hagens, S. 2017. Monitoring the Brazilian pasturelands: A new mapping approach based on the Landsat 8 spectral and temporal domains. *International Journal of Applied Earth Observation and Geoinformation*, 62, 135–143. <https://doi.org/10.1016/j.jag.2017.06.003>.
- Pereira, P.R.M., Bolfe, E.L., Rodrigues, T.C.S., Sano, E.E. 2020. Dynamics of agricultural expansion in areas of the Brazilian savanna between 2000 and 2019. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, v. XLIII-B3-2020, 1607–1614. <https://doi.org/arch-photogramm-remote-sens-spatial-inf-sci.ne>.
- Picoli, M.C.A., Câmara, G., Sanches, I.D., Simões, R., Carvalho, A., Maciel, A., Coutinho, A., Esquerdo, J., Antunes, J., Begotti, R.A., Arvor, D., Almeida, C. 2018. Big earth observation time series analysis for monitoring Brazilian agriculture. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145, 328–339. <https://doi.org/10.1016/j.isprsjprs.2018.08.007>.
- Prudente, V.H.R., Martins, V.S., Silva, N.R.F., Adami, M., Sanches, I.D. 2020. Limitations of cloud cover for optical remote sensing of agricultural areas across South America. *Remote Sensing Applications: Society and Environment*, 20, 100414. <https://doi.org/10.1016/j.rsase.2020.100414>.
- Rudorff, B.F.T., Aguiar, D.A., Silva, W.F., Sugawara, L.M., Adami, M., Moreira, M.A. 2010. Studies on the rapid expansion of sugarcane for ethanol production in São Paulo State (Brazil) using Landsat data. *Remote Sensing*, 2(4), 1057–1076. <https://doi.org/10.3390/rs2041057>.
- Sano, E., Rosa, R., Scaramuzza, C., Adami, M., Bolfe, E., Coutinho, A., Esquerdo, J., Maurano, L., Narvaes, I., Oliveira, F., Silva, E., Victoria, D., Ferreira, L., Brito, J., Bayma, A., Oliveira, G., Silva, G. 2019. Land use dynamics in the Brazilian Cerrado in the period from 2002 to 2013. *Pesquisa Agropecuária Brasileira*, 54(2), 1–5. <https://doi.org/10.1590/s1678-3921.pab2019.v54.00138>.
- Silva Júnior, C.A., Leonel, A.H.S., Rossi, F.S., Corrêia, W.L.F., Santiago, D.D., Oliveira, J.F., Teodoro, P.E., Lima, M., Capristo-Silva, G.F. 2020. Mapping soybean planting area in midwest Brazil with remotely sensed images and phenology-based algorithm using the Google Earth Engine platform. *Computers and Electronics in Agriculture*, 169, 105194. <https://doi.org/10.1016/j.compag.2019.105194>.
- Souza Júnior, C.M., Shimbo, J.Z., Rosa, M.R., Parente, L.L., Alencar, A.A., et al. 2020. Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat archive and Earth Engine. *Remote Sensing*, 12, 2735. <https://doi.org/10.3390/rs12172735>.
- Werner, J.P.S., Oliveira, S.R.D., Esquerdo, J.C.D.M. 2019. Mapping cotton fields using data mining and MODIS time-series. *International Journal of Remote Sensing*, 41(2), 2454–2476. <https://doi.org/10.1080/01431161.2019.1693072>.