

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/343222791>

# APPLYING A PHENOLOGICAL OBJECT-BASED IMAGE ANALYSIS (PHENOBIA) FOR AGRICULTURAL LAND CLASSIFICATION: A STUDY CASE IN THE BRAZILIAN CERRADO

Conference Paper · July 2020

CITATIONS

0

READS

70

13 authors, including:



**Hugo Bendini**

National Institute for Space Research, Brazil

40 PUBLICATIONS 102 CITATIONS

SEE PROFILE



**Leila Fonseca**

Brazilian Institute for Space Research (INPE), Brazil

92 PUBLICATIONS 917 CITATIONS

SEE PROFILE



**Anderson Reis Soares**

National Institute for Space Research, Brazil

26 PUBLICATIONS 16 CITATIONS

SEE PROFILE



**Philippe Rufin**

Humboldt-Universität zu Berlin

20 PUBLICATIONS 266 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Mapping and analyzing spatial and temporal pattern of insect disturbances using Landsat time series [View project](#)



GeoDMA - Geographic Data Mining Analyst [View project](#)

# APPLYING A PHENOLOGICAL OBJECT-BASED IMAGE ANALYSIS (PHENOBIA) FOR AGRICULTURAL LAND CLASSIFICATION: A STUDY CASE IN THE BRAZILIAN CERRADO

*Hugo N. Bendini<sup>1</sup>; Leila M. G. Fonseca<sup>1</sup>; Anderson R. Soares<sup>1</sup>; Philippe Rufin<sup>2,3</sup>; Marcel Schwieder<sup>2</sup>; Marcos A. Rodrigues<sup>1</sup>; Raian V. Mareto<sup>1</sup>; Thales S. Körting<sup>1</sup>; Pedro J. Leitão<sup>2,4</sup>; Ieda D. A. Sanches<sup>1</sup>; Patrick Hostert<sup>2,3</sup>*

1 Brazilian National Institute for Space Research (INPE); {hugo.bendini, anderson.soares, marcos.almeida, raian.mareto, thales.korting, ieda.sanches, leila.fonseca}@inpe.br

2 Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, Berlin 10099, Germany; {philippe.rufin, marcel.schwieder, p.leitao, patrick.hostert}@geo.hu-berlin.de

3 Integrative Research Institute on Transformations of Human-Environment Systems (IRI THESys), Humboldt University zu Berlin, Unter den Linden 6, Berlin 10099, Germany

4 Department Landscape Ecology and Environmental Systems Analysis, Technische Universität Braunschweig, Langer Kamp 19c, 38106 Braunschweig, Germany

\* Correspondence: hnbendini@gmail.com; Tel.: +55 12-98806-6386

## ABSTRACT

Mapping agriculture with high accuracy is important to generate reliable information about crop production. Pixel-based methods still present problems with noise and usually require post-processing approaches to reach satisfactory results. Object-based Image Analysis (OBIA) enable the detection of homogeneous objects in remote sensing images based on spectral similarity. However, traditional OBIA does not consider the multi-temporal characteristics of land cover or land use, such as agriculture. The objective of this study is to evaluate a phenological object-based approach with dense Landsat image time series for mapping agriculture in different level of detail in the Brazilian Cerrado. We derived pixel-wise EVI fitted time series with 8-day temporal resolution and applied multi-resolution segmentation using all image bands to incorporate the influence of space and time. Then we generated phenological metrics and applied OBIA of agricultural lands in Brazil using a hierarchical classification scheme. The overall accuracies for each hierarchical level were around 90%, and the spatial consistency of the generated maps is promising.

**Index Terms**— Big data, Time-series mining, Phenometrics, OBIA.

## 1. INTRODUCTION

Brazil's agricultural frontier is considered one of the largest in the world. Accurate maps are important to generate reliable information in the context of energy, water, and land systems interactions. Pixel-based image classification methods still present problems with noise and usually require post-classification processing with temporal rules to

reach higher accuracies and satisfactory spatial consistency. To consider the context and the hierarchical structure in a classification process, OBIA (Object-based Image Analysis) techniques are a useful tool, as it considers neighborhood relationships, texture, form and compacity, based on segmentation and feature extraction [1]. However, traditional OBIA approaches do not consider the multi-temporal characteristics of the different types of land cover or land use, such as agriculture. According to [2], including spatiotemporal heterogeneity and correlation, may possibly enhance the performance of classification approaches. Recently, Belgiu and Csillik [3] compared the performances of pixel and object-based approaches for satellite image time series classification using Sentinel-2 imagery. The results indicated that OBIA outperformed pixel-wise classification. Despite the reported advantages, its adaptation to satellite image time series (SITS) remains an issue due to the need to incorporate the temporal variability of the observed targets. Recent studies showed the potential of the phenological metrics (phenometrics) derived from dense satellite image time series for crop classification in the Brazilian Cerrado [4]. Very few studies, though, applied OBIA combined with phenological metrics [5, 6]. They showed the potential of object-oriented approaches based on phenological parameters for rice mapping. However, they derived phenometrics from coarse resolution imagery (250m), and found that mixed pixel is a problem [5]. They pointed out that there is a need to investigate the application of the method with dense and higher spatial resolution SITS. Exploring these approaches and considering large, diverse and dynamic agricultural regions such as the Brazilian Cerrado offers opportunities to improve current mapping approaches. The objective of this study is to evaluate a phenological object-based approach with dense Landsat

image time series for mapping agriculture in different level of detail in the Brazilian Cerrado.

## 2. METHODOLOGY

### 2.1. Study area and field work

The method was evaluated in a study area in the west of Bahia State, within the Brazilian Cerrado biome (Figure 1). This area consists mostly of large-scale and market-oriented agriculture and belongs to the most recent agricultural frontier of the Cerrado - the Matopiba region. We collected data at different thematic detail to test the approach using a hierarchical classification scheme. Field surveys were conducted during the 2015–2016 cropping seasons, with the “Rally da Safra 2016” team, which is a project that visits several farms on Brazil to evaluate crop yield during the growing peak. We registered the sample locations for each class using a GPS and collected information about crop types and crop rotations with farmers. Field boundaries were digitized to obtain a polygon database with 774 labelled samples. We used a hierarchical classification approach by which the first level classes domains are isolated, and land cover is classified by correspondence at level 2. Then we isolated the segments classified as “Annual crop” and used it to classify land use by correspondence for each domain for the subsequent nomenclature level. The land cover classes (Level 2) are based on the nomenclature of the Systematic Survey of Agricultural Production of the Brazilian Institute of Geography and Statistics [7]. The natural vegetation classes consist of three main Cerrado physiognomies, based on the definition of Ribeiro and Walter (2008): forest, savanna and natural grasslands [8]. The Crop Group (Level 3) is defined by the main agricultural practices in the Cerrado region [4].

### 2.2. Satellite data

We used all available Landsat images between April 2015 and April 2017, applying a weighted ensemble of Radial Basis Function (RBF) convolution filters as a kernel smoother to fill data gaps such as cloud cover and Scan Line Corrector (SLC)-off data [9, 10, 4]. Through this approach, we created a dense Enhanced Vegetation Index (EVI) data cube with an 8-day temporal resolution and derived a set of 30 phenometrics [11, 12]. Phenometrics included day-of-the-year (DOY) of start, mid, end, length of crop seasons and phenological proxies like peak, base value, seasonal amplitude or rate of increase, decrease [11] and the polar features, which are based on the representation of the time series by projecting the values onto angles in the interval  $[0, 2\pi]$  [12].

### 2.3. Multiresolution segmentation

Multiresolution Segmentation (MRS) algorithm [13] is an image segmentation approach that aims to minimize the variability of a segment, relying on the potential of the local variance [14]. We performed the segmentation of a dense EVI datacube, considering all the bands with the same

weight with the following parameters: scale: 300, shape: 0.3 and compactness: 0.4. The “eCognition” software [15] was used to perform segmentation and to extract feature from the segments.

### 2.4. Feature extraction and selection

The object-based metrics minimum, maximum, mean, standard deviation and Grey Level Co-occurrence Matrix (GLCM) textures were derived from each EVI band and phenometrics. Shape-based features such as area, length, perimeter, length/width ratio, compactness, were also extracted resulting in 770 attributes for each segment/object. We proceeded feature selection using the Correlation-based Feature Subset Selection (CfsSubsetEval), which evaluates the individual predictive ability of each feature along with the degree of redundancy between them [16] and selects a subset of 59 attributes.

### 2.5. Classification and Validation

We used a hierarchical classification with three levels, from land cover classes to classes representing cropping practices and evaluated results by overall accuracies and class f1 score using a Monte Carlo simulation. After the time series feature extraction, we used our field database to train a Random Forest (RF) [17] model and obtained a classifier for each nomenclature level. We empirically determined the parameters *mtry* of 5 and *ntrees* of 600. The “randomForest” package in R was used for our classification tasks. The final maps were validated based on Monte Carlo simulation where, for each model, 100 simulations were carried out by randomly selecting 70% of the samples to train and 30% for validation. For each subdivision, a confusion matrix was calculated, and the average confusion matrix was used to derive the overall accuracy and the class f1-scores for each model.

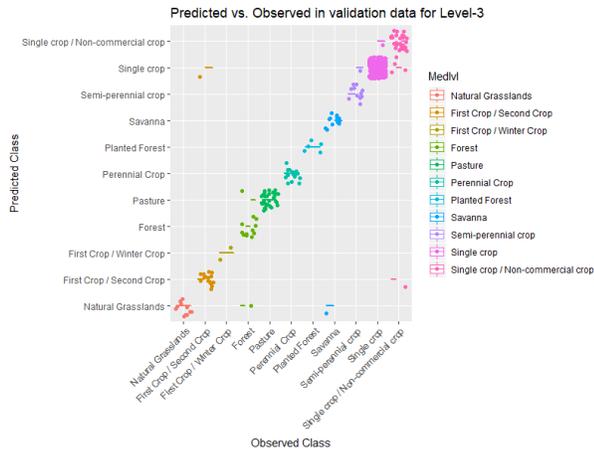
## 3. RESULTS AND DISCUSSIONS

The overall accuracies for Level 1, 2 and 3 were respectively 91.1%, 90.1% and 87.4%. Table 1 presents the class wise f1-scores, considering the hierarchical and non-hierarchical approaches.

**Table 1** - Class wise f1-scores, considering the hierarchical and non-hierarchical approaches on Level 3 classification.

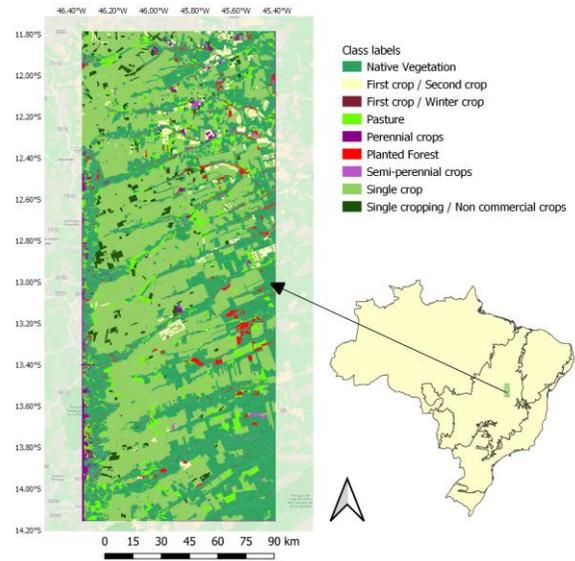
	Hierarchical	Non-hierarchical
First crop/Second crop	0.7156766	0.6974652
Single crop	0.9512889	0.9157691
Single/Non-commercial	0.4527556	0.4422892
Natural grasslands	0.510509	0.4237434
Forest	0.6066552	0.451746
Pasture	0.8091056	0.5204885
Perennial crops	0.9760317	0.9007172
Planted forest	0.9583333	0.877551
Savanna	0.6815478	0.628766
Semi-perennial crops	0.8238095	0.5947261

We can see that the hierarchical classification improved the accuracies substantially, especially for semi-perennial crops and pastures. This reinforces the findings of Neves et al. (2019) and Lebourgeois (2017) et al., who showed the potential of hierarchical classification in comparison to traditional non-hierarchical approaches [18, 19]. This can be explained also by the fact that some misclassifications were avoided by splitting the higher-level classes on the first level of classification.



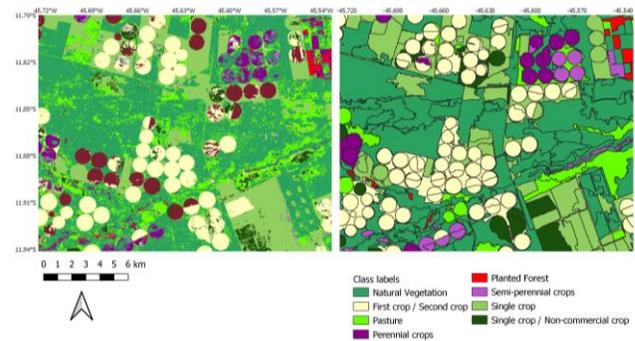
**Fig. 1.** Confusion matrix obtained by the best model over the Monte Carlo simulation, with the hierarchical classification on Level 3.

High rates of confusion were detected for the single/non-commercial crops. This class was confused with either single crops when the second season was not detected in TIMESAT, or with double crops because of the spectral-temporal similarities between some non-commercial crops like sorghum and millet with maize [4]. Misclassifications were observed between pastures and natural grasslands as well. Based on the feature selection results we found that that base value was important for the separation of the distinct agriculture classes [4], which was also identified by Bendini et al. (2019). Agreeing with Singha et al., (2019), the GLCM metrics were not too important for the classification [6], showing that the temporal behavior is more relevant. The mean EVI values were also significant, especially during the peak of crop seasons.

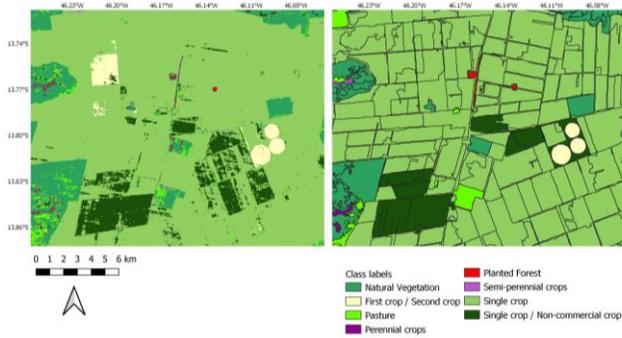


**Fig. 2.** Classification results obtained by the best model over the Monte Carlo simulation, with the hierarchical classification on Level 3 (Overall accuracy of 87.4%).

The map of Figure 2 presents the spatial patterns of the classes on the study area, showing that single-cropping system is being widely used in this region, which agrees with the results of Spera et al. (2016) which showed that, in 2015, 85% of the large-scale agriculture in the Matopiba region was based on single-cropping systems (Spera et al., 2016) [20]. Double-cropping systems occur mostly in irrigated areas. Perennial crops were mapped, being represented mostly by irrigated coffee crops. This agrees with findings of Bendini et al. (2019) [4].



**Fig. 3.a.** Comparison between pixel-based classification results obtained by Bendini et al. (2019) and the present approach results. Figure 3 shows the difference between a pixel-wise classification [4] on the same area and period, and the use of the PHENOBIA approach, where the spatial consistency of the maps is more homogeneous, and the crop fields are well delineated.



**Fig. 4.b.** Comparison between pixel-based classification results obtained by Bendini et al. (2019) and the present approach results.

#### 4. FINAL CONSIDERATIONS AND OUTLOOKS

The objective of this study is to evaluate a phenological object-based approach with dense Landsat image time series for mapping agriculture in different level of detail in the Brazilian Cerrado. We presented an integration of OBIA and a hierarchical classification approach using phenometrics to perform classify SITS, here called PHENOBIA. This work is promising because of the improvements obtained in crop mapping in the Brazilian Cerrado. The overall accuracies for each hierarchical level were around 90%. The spatial consistency of the generated maps was high. However, there is still a gap to be explored in terms of developing robust methods to generate spatiotemporal objects and fully explore all the spatiotemporal properties of the time series datacubes.

#### 5. ACKNOWLEDGMENTS

Thanks to the project “Development of systems to prevent forest fires and monitor vegetation cover in the Brazilian Cerrado” (World Bank Project #P143185) – Forest Investment Program (FIP) and the Environmental Monitoring of Brazilian Biomes project (Brazil Data Cube), funded by the Amazon Fund through the financial collaboration of the Brazilian Development Bank (BNDES), the Foundation for Science, Technology and Space Applications (FUNCATE) Process 17.2.0536.1. The authors also thank the São Paulo Research Foundation (FAPESP, grant no 2017/24086-2), and the Brazilian National Council for Scientific and Technological Development (CNPq, grant no 303360/2019-4)

#### 6. REFERENCES

[1] T. Blaschke, “Object based image analysis for remote sensing,” *ISPRS journal of photogrammetry and remote sensing*, vol. 65, no. 1, pp. 2–16, 2010.

[2] Wenqiang Xi, et al. “A spatiotemporal cube model for analyzing satellite image time series: Application to land-cover mapping and change detection”, *Remote Sensing of Environment*, vol. 231, pp. 111212, 2019.

[3] Mariana Belgiu and Ovidiu Csillik, “Sentinel-2 crop-land mapping using pixel-based and object-based time-weighted dynamic time warping analysis,” *Remote sensing of environment*, vol. 204, pp. 509–523, 2018.

[4] Bendini, Hugo et al. Detailed agricultural land classification in the Brazilian cerrado based on phenological information from dense satellite image time series. *International Journal of Applied Earth Observation and Geoinformation*, v. 82, p. 101872, 2019.

[5] Zhang, Meng, and Hui Lin. “Object-based rice mapping using time-series and phenological data.” *Advances in Space Research* 63.1 (2019): 190-202.

[6] Singha, Mrinal and Sarmah, Sangeeta. Incorporating crop phenological trajectory and texture for paddy rice detection with time series MODIS, HJ-1A and ALOS PALSAR imagery. *European Journal of Remote Sensing*, v. 52, n. 1, p. 73-87, 2019.

[7] IBGE, 2016. IBGE: Brazilian Institute Of Geography And Statistics. Available online: (accessed on October). <https://sidra.ibge.gov.br/pesquisa/pam/tabelas>.

[8] Ribeiro, J.F. and Walter, B.M.T., 2008. As principais fitofisionomias do bioma cerrado. In: Sano, S.M., Almeida, S.P., Ribeiro, J.F. (Eds.), *Cerrado: Ecologia E Flora*. Embrapa Cerrados, Brasília, DF, Brazil, pp. 153–212.

[9] Schwieder, M., et al. 2016. Mapping Brazilian savanna vegetation gradients with Landsat time series. *Int. J. Appl. Earth Obs. Geoinf.* 52, 361–370.

[10] Rufin, P. et al. 2019. Mapping Cropping Practices on a National Scale Using Intra-Annual Landsat Time Series Binning. *Remote Sens* 11, 232.

[11] Jönsson, P. and Eklundh, L., 2015. *TIMESAT 3.2 With Parallel Processing Software Manual*. Lund University, Sweden, pp. 22–24.

[12] Körting, T.S., Fonseca, L.M.G. and Câmara, G., 2013. GeoDMA—geographic data mining analyst. *Comput. Geosci.* 57, 133–145.

[13] Baatz, M. and Schäpe, A. (2000): Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. In: Strobl/Blaschke/Griesebner (eds.): *Angewandte Geographische Informationsverarbeitung XII*, Wichmann-Verlag, Heidelberg, 12-23

[14] Drăguț, Lucian et al. (2014). Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing*. 88. 119–127. 10.1016/j.isprsjprs.2013.11.018.

[15] Baatz M, et al., 2004; *eCognition Professional: User guide 4.*; Munich, Germany: Definiens-Imaging

[16] M. A. Hall (1998). *Correlation-based Feature Subset Selection for Machine Learning*. Hamilton, New Zealand.

[17] Breiman, L., 2001. *Mach. Learn.* 45 (5), 5–32.

[18] Neves, A., et al., 2019. “Hierarchical classification of brazilian savanna physiognomies using very high spatialresolution image, superpixel and geobia,” in *IGARSS2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, pp. 3716–3719.

[19] Lebourgeois, V., et al., 2017. A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated Sentinel-2 time series. VHRS and DEM). *Remote Sensing* 9.3, 259.