LAND USE AND LAND COVER CLASSIFICATION USING A SAR OPTICAL CLOUD COMPUTER APPROACH IN SOUTHERN OF RORAIMA

Victor Hugo Rohden Prudente[1], Nildson Rodrigues de França e Silva[2,3], Andre Dalla Bernardina Garcia[2], Lucas Volochen Oldoni[2], Haron Abrahim Magalhães Xaud[4], Maristela Ramalho Xaud[4], Marcos Adami[,2], Ieda Del'Arco Sanches[2]

[1]School for Environment and Sustainability, University of Michigan, Ann Arbor, MI 484109, United States of America and victor.rohden@yahoo.com; [2]Posgraduate Program in Remote Sensing, Earth Observation and Geoinformatics Division, National Institute for Space Research, Sao Jose dos Campos, SP 12227-010, Brazil and {nildson.silva; andre.garcia; lucas.oldoni; marcos.adami; ieda.sanches}@inpe.br; [3]Agrosatelite Applied Geotechnology, SC-401, 8600 - Corporate Park, Florianopolis, SC 88050-000, Brazil and nildson.silva@agrosatelite.com.br; [4] Brazilian Agricultural Research Corporation, Embrapa Roraima, BR-174, km 8, Distrito Industrial, CEP 69.301-970 – Boa Vista RR, Brazil, {haron.xaud; maristela.xaud}@embrapa.br

ABSTRACT

Our goal in this study was to perform a LULC classification for the southern part of Roraima state. This area has a highly frequent cloud cover and a lack of LULC information. We used a SAR-Optical multisensory methodology, with a cloud computing process, to be able to classify all the areas, with less computational effort and in less time. Our results show an Overall Accuracy of 92.61%, with Users' and Producers' Accuracy (UA and PA), around 90% for all ten classes. Also, this approach identified important classes for the region, such as perennial crops and conversion areas.

Key words — *microwave data, random forest, google engine, amazon region.*

1. INTRODUCTION

Increasing food production meanwhile minimizing the environmental impacts is one of the biggest challenges to be faced by the agriculture sector, government, and researchers. Part of this food demand is met by agriculture expansion, resulting in severe environmental impacts, and contributing to global climatic changes [1]. In this sense, continuous Land Use and Land Cover (LULC) mapping is fundamental to land use management and to understanding the environmental effects at local, regional, and global scales [2]. Thus, Remote Sensing (RS) technology is widely utilized for synoptic and continuous LULC monitoring, allowing the identification of the LULC Changes (LULCC) [3].

Even though LULC map information is highly important for the management of tropical areas, there is a lack of information for some regions in the Brazilian Amazon. For example, the Savannas and Campinarana areas in Roraima are not considered yet in programs designed for forest monitoring despite their ecological importance. Besides, in this region, the use of optical sensors suffers negative impacts due to a high cloud cover frequency [4]. To work around the shortage of optical data, the use of microwave sensors (Synthetic Aperture Radar) appears as a possible workaround. SAR sensors are less influenced by cloud cover frequency when compared with optical data [3]. Methods that integrate optical and SAR data have been explored in LULC studies [5], [6].

However, the quantity of data generated demands high computation power and large space for storage. A way to overcome this is by using cloud processing platforms [7]. To decrease data volume, it is possible to use different metrics to explore the temporal variation from the remote sensing data. This statical transformation of an image time series is called multi-temporal metrics, which can be very useful for LULC characterization [6].

On the SAR-Optical approach, Random Forest (RF) is one of the most commonly non-parametric classifiers used to provide LULC classification. The RF classifier is highlighted due to its robustness and capability to hold a high number of variables and high data dimensionality [2], [5].

In this context, this study aims to classify the LULC in a tropical area, in the south of Roraima state, in the Brazilian Amazon region. In our study, we used RF and combined the Sentinel-1 SAR and Sentinel-2 MSI (Multispectral Instrument) optical images, acquired in different periods during the year 2019.

2. MATERIAL AND METHODS

2.1. Study area

Our study area comprehends five municipalities (Rorainópolis, Caracaraí, São Luiz, São João da Baliza, and Caroebe) located in the southern part of Roraima (Figure 1). In this area is predominant the pasture and the small agriculture areas. Besides, this region has been monitored by the Brazilian Agricultural Research Corporation, Embrapa Roraima projects, as TERRAMZ (*Conhecimento Compartilhado para Gestão Territorial Local na Amazônia*).

In this area, there are three groups of natural vegetation formations in our study area: rainforest, campina-

campinarana, and savannas (Figure 1). Rainforest is divided into Seasonal Semideciduous Forest, Seasonal Forest, Open, and Dense Ombrophilous Forest. Campina-campinarana is formed by campinas (small shrubs) and campinarana. Savannas are predominated by grasslands with few shrubs. Some regions show ecologic tension, where the transition between two or more types of vegetation occurs, such as Savannas and Forests, and Campinarana and Forest [2].

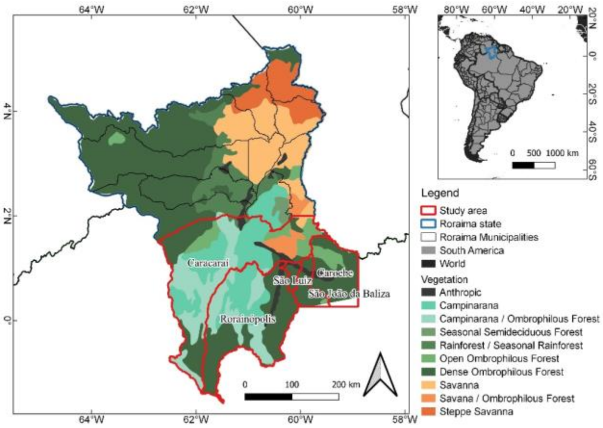

Figure 1. Location of the study area and Roraima natural formations according to IBGE. Source: Adapted from IBGE [8]**.**

2.2. Remote Sensing

We used data from Sentinel-2/MSI and Sentinel-1/SAR data. For Sentinel-2, we used the data from the spectral bands and the NDVI, and the Land Surface Water Index (LSWI), vegetation Indexes (VIs). For Sentinel-1 SAR it was used the IW mode and Ground Range Detected (GRD) data types, with polarization, are VH and VV. Also, we used the ratio between the VH and VV polarizations. To summarize the image collections, we used standard deviation, minimum, maximum, and median metrics for Sentinel-1 and Sentinel-2 data. For Sentinel-1, the metrics were calculated separately for each satellite (Sentinel-1A and Sentinel-1B).

2.3. Field data

Accurate and representative field information is essential for LULC and LULCC classifications, so it is crucial to choose proper periods for field data collection. We performed fieldwork in Roraima in August-September 2019, which correspond to the crop season and the end of the rainy season, allowing us to collect field data from all representative LULC classes, including rainfed crops. We collected the LULC data for classification training purposes, along the roadsides [9], using the Locus Map Pro applications. Our data represents 10 LULC classes: forest, savanna, campinarana, water, sand/rock, annual crops, perennial crops, pasture, conversion, and impermeable (see [6] for details and pictures)

2.4. Classification

In our approach, Sentinel-1 was used for three different periods; P1: January to April; P2: May to August; and P3: September to December, for each year (2017, 2018, and 2019). P2 represents the rainy season and P1 and P3 represent the dry season in Roraima. Due to the low number of cloud-free pixels [4], we adopted to use of Sentinel-2 data for the entire year and not only for each period. In total, there are 216 bands (layers), 144 for Sentinel-1/SAR and 72 for Sentinel-2, as input for the classification process.

The GEE platform [7] was used for the LULC classification process, along with the RF algorithm. Inside GEE, we filtered all Sentinel-2 and Sentinel-1 images available from Roraima. Also, we used all images with less than 50% of cloud cover. To remove the clouds and shadows, we filter using the Bits and Sentinel-2 cloud probabilistic [10], with a threshold of 65%.

After, we used a stratified sampling approach to sample a maximum of 5000 pixels per class. For this set, we randomly split 70% for training and 30% for the validation process. For the RF, we set up 100 the number of trees, and the other parameters were left as default at GEE. The classification results were generated with 20 meters of pixel size. As post-classification, we used a mode filter (kernel in circle and radius equal to 1) in a post-classification step to smooth the results and avoid isolating misclassified pixels. To access the accuracy classification, from the confusion matrix, we extracted the overall accuracy (OA), user's accuracy (UA), and producer's accuracy (PA) [11].

3. RESULTS

Our approach achieved an OA of 92.61%. According to the confusion matrix, Table 1, most classes have UA and PA near or higher than 90%. The lower accuracy was found for the annual crop class (UA = 41.1%), with confusion with the pasture class

	1	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	**UA**
1	1505	0	22	1	6	0	0	19	13	0	**96.1**
2	0	71	0	0	0	0	0	0	2	0	**97.3**
3	51	0	1362	13	23	0	0	3	35	1	**91.5**
4	0	0	16	1498	1	0	0	0	0	0	**98.9**
5	12	0	31	0	1345	0	0	90	39	0	**88.7**
6	1	0	7	16	3	239	0	1	2	2	**88.2**
7	0	0	0	0	35	0	30	4	4	0	**41.1**
8	10	0	1	0	137	0	0	1367	0	0	**90.2**
9	2	0	6	0	44	2	0	10	1242	0	**95.1**
10	0	0	2	0	5	2	0	3	0	144	**92.3**
PA	**95.2**	**100**	**94.1**	**98**	**84.1**	**98.4**	**100**	**91.2**	**92.9**	**98**	

Table 1. Confusion matrix (number of pixels) and Users (UA) and Produces (PA) Accuracies (in percentage) for the LULC classification. 1: Forest, 2: savanna, 3: campinarana, 4: water, 5: pasture, 6: sand/rock, 7: annual crops, 8: perennial crops, 9: conversion, 10: impermeable.

This higher commission error means that our approach overestimates the annual crop area. The omission error for Annual crops was small (high producer accuracy) than the commission error. Almost all our reference data (field data) are correctly classified as annual crops.

According to the variable importance (Figure 2), to achieve this result, the more important features are from optical data. NDVI max and Band 11 (Swir) percentile 75% and median have the more important contribution. Other metrics from LSWI, B5 (Vegetation Red Edge), B12 (Swir), B3 (Green), and B4 (Red), appear among the 10 more important variables.

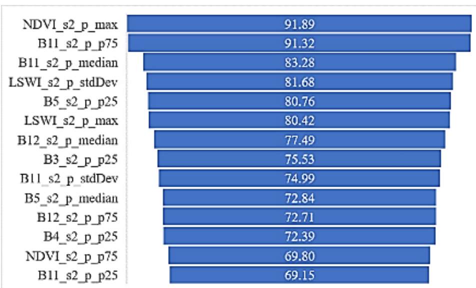

Figure 2. List of the 10 more important variables for LULC classification for the year 2019.

Rainforest class is predominant in our LULC classification (Figure 3 and Figure 4) for all five municipalities. Campinarana is more frequent in Rorainópolis and Caracarai municipalities. Pasture, conversion areas, and perennial crops are the predominant anthropic classes (Figure 5). Caracaraí has the biggest conversion, pasture, and perennial crop amount area (Figure 5). Rorainópolis is the second one in the area for the same classes.

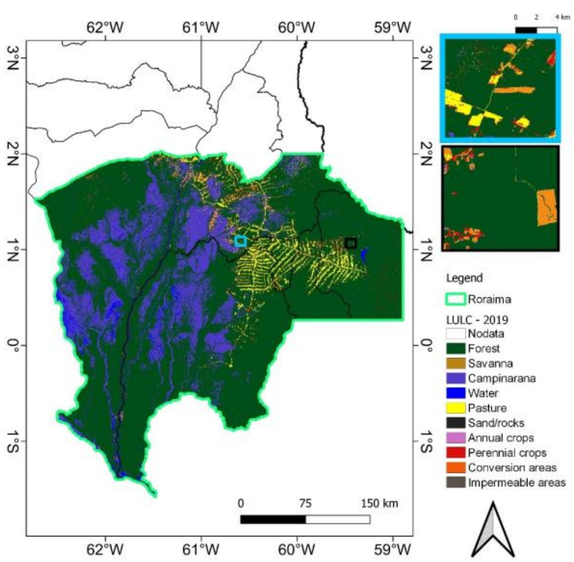

Figure 3. LULC classification for the year 2019, and zoom views for conversion areas.

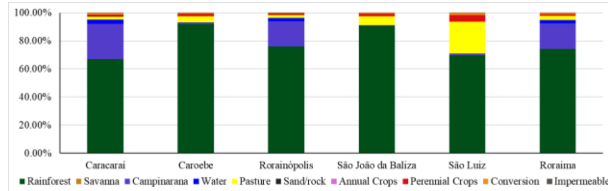

Figure 4. LULC classification percentage for each municipality for the year 2019.

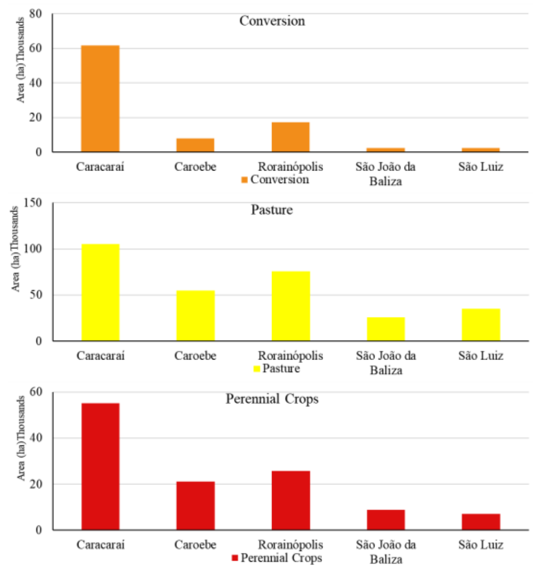

Figure 5. Detailed LULC classification area (in hectares) distribution for conversion, pasture, and perennial crop classes, for each municipality for the year 2019.

4. DISCUSSION

Our results show a way to provide LULC information using a cloud computing platform and multisensor data. This allows us to obtain LUC information with less computational effort in a minimum amount of time.

Although SAR data does not appear among the more important variables in our results, it was verified in previous studies [6] that they contribute to certain classes. Besides, in areas with frequent cloud cover, SAR data could be the only source of information.

Annual crops happen in a few fields in our study. Besides, is more for subsistence and small fields. Due to these factors and the seasonality of the region, some confusion with pasture was expected. The perennial crops class contains mostly palm oil (*dendê*), bananas, and orange fields. Our approach got around 90% of accuracy when identifying this class among the others. This is crucial for the region, where palm oil is increasing in area and socio-economic importance. For bananas and orange fields, the spatial resolution of 20 meters is coarse. These fields are small and a spatial resolution of 10 meters or higher should be used to better results.

We identified areas under conversion, which means that the forest (or other native vegetation) was removed in 2019 and we did some lugging activity or soil turning. In Figure 3 we have some examples of this situation. This information is important to help understand where the is happening natural vegetation loss and, monitor what will happen in the future years in this region.

After an inspection of the LULC classification, it was possible to identify that our approach overestimates the perennial crop areas. However, that was not shown in the accuracy approach due to the roadside example. Seems like the areas with possible forest degradation (near a conversion area) are being classified as perennial crops. Also, for some regions where the cloud cover frequency impacts the number of optical images, seems that our approach could overestimate the pasture and perennial crop classes. Although these points, this approach could be improved and used to help the public and private policies in southern Roraima.

5. CONCLUSION

SAR-optical data, associated with GEE and RF classifier, represents a possibility to get LULC information over large tropical areas. We demonstrate it in this approach, classifying the LULC for southern Roraima, in 2019.

It achieved accuracy near 90% for almost all the classes. Besides, the results brought important information about perennial crops and conversion areas for southern Roraima. However, some improvements are necessary to improve the LULC classification quality. For future works we do recommend providing a classification for each municipality, to try performance the classification in a higher spatial resolution (i.e., 10 meters) and classify different years to analyze the LULC changes locally.

6. ACKNOWLEDGMENTS

This study was funded in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* - Brazil (CAPES) - Finance Code 001. The authors acknowledge the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) for the Research Productivity Fellowship of I.D. Sanches (PQ - 303012/2018-8), and M. Adami (PQ - 306334/ 2020-8) and for V.H.R. Prudente doctorate scholarship. We appreciate the *Embrapa Roraima* fieldwork support by Dr. Haron Xaud and Dr. Maristela Xaud, through TERRAMZ (*Conhecimento Compartilhado para Gestão Territorial Local na Amazônia*) project, and the support from the Agricultural Remote Sensing Laboratory's team at INPE.

7. REFERENCES

[1] X. Deng, C. Zhao, and H. Yan, "Systematic Modeling of Impacts of Land Use and Land Cover Changes on Regional Climate: A Review," *Adv. Meteorol.*, vol. 2013, pp. 1–11, 2013.

[2] J. A. P. Pavanelli, J. R. dos Santos, L. S. Galvão, M. R. Xaud, and H. A. M. Xaud, "PALSAR-2/ALOS-2 and OLI/Landsat-8 data integration for land use and land cover mapping In Northern brazilian Amazon," *Bol. Ciências Geodésicas*, vol. 24, no. 2, pp. 250–269, 2018.

[3] M. A. Wulder *et al.*, "Virtual constellations for global terrestrial monitoring," *Remote Sens. Environ.*, vol. 170, pp. 62–76, 2015.

[4] V. H. R. Prudente, V. S. Martins, D. C. Vieira, N. R. de F. e Silva, M. Adami, and I. D. Sanches, "Limitations of cloud cover for optical remote sensing of agricultural areas across South America," *Remote Sens. Appl. Soc. Environ.*, vol. 20, no. May, p. 100414, Nov. 2020.

[5] N. Clerici, C. A. Valbuena Calderón, and J. M. Posada, "Fusion of Sentinel-1A and Sentinel-2A data for land cover mapping: a case study in the lower Magdalena region, Colombia," *J. Maps*, vol. 13, no. 2, pp. 718–726, Nov. 2017.

[6] V. H. R. Prudente *et al.*, "Multisensor approach to land use and land cover mapping in Brazilian Amazon," *ISPRS J. Photogramm. Remote Sens.*, vol. 189, no. July 2021, pp. 95–109, Jul. 2022.

[7] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.*, vol. 202, pp. 18–27, Dec. 2017.

[8] I. B. de G. e E. IBGE, *Manual Técnico da Vegetação Brasileira*, 2nd ed. Rio de Janeiro, 2012.

[9] F. Waldner *et al.*, "Roadside collection of training data for cropland mapping is viable when environmental and management gradients are surveyed," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 80, no. January, pp. 82–93, Aug. 2019.

[10] Google, "Sentinel-2: Cloud Probability," *Earth Engine Catalog*. [Online]. Available: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_CLOUD_PROBABILITY. [Accessed: 18-Jul-2021].

[11] P. Olofsson, G. M. Foody, M. Herold, S. V. Stehman, C. E. Woodcock, and M. A. Wulder, "Good practices for estimating area and assessing accuracy of land change," *Remote Sens. Environ.*, vol. 148, pp. 42–57, May 2014.