Image Classification Methods Assessment for Identification of Small-Scale Agriculture in Brazilian Amazon

Flávia Domingos Pacheco, Maíra Ramalho Matias, Gabriel Máximo da Silva,
Anielli Rosane de Souza, Yosio Edemir Shimabukuro, Maria Isabel Sobral Escada
Earth Observation and Geoinformatics Division – DIOTG
Brazilian National Institute for Space Research
São José dos Campos, Brazil
e-mail: {flavia.pacheco; maira.matias; gabriel.maximo; anielli.souza; yosio.shimabukuro; isabel.escada}@inpe.br

***Abstract*—This paper aims to test different methods for image classification focusing on small-scale agriculture in the region of Mocajuba and Cametá, municipalities in the Northeast of Pará state, Brazil. It is an important land use class, always ignored by Land-Use and Land-Cover monitoring systems because of its small size and variable spectral signature. We used an image from the PlanetScope Surface Reflectance Mosaics (Analysis Ready) with spatial resolution of 4.77 meters and 4 spectral bands (red, green, blue and infra-red). After proceeding with a multiresolution segmentation to identify image objects, two object-oriented classification algorithms were tested: Adapted Nearest-neighbor and C5.0 Decision trees algorithms. We selected 122 random points using the images available on Google Earth Pro as reference to assess the accuracy of classifications. Afterwards, confusion matrices were generated. Both methods showed similar overall accuracy and kappa value. However, C5.0 Decision trees reached a higher producer's accuracy to small-scale agriculture (75%) in comparison to Adapted Nearest-neighbor (65%). The average size of the small-scale agriculture segments estimated was less than 1 ha in both maps, showing the need to carry out studies on scales of greater detail, preferably with images of high spatial resolution to represent these systems properly. In this study, C5.0 Decision trees had the best result, representing the most suitable method for mapping small-scale agriculture in Brazilian Amazon.**

Keywords - digital image processing; segmentation; land use; land cover; smallholders; planetscope.

I. INTRODUCTION

Small-scale agriculture is a rather important land use activity that takes place in small properties, providing income for families and food for local population [1]. Although we have few definitions about it, these producers are somehow invisible to public policy and they have difficulty in accessing credits for their production [2].

Despite of its importance, small-scale agriculture is not always properly represented in the maps produced by Land-Use and Land-Cover (LULC) monitoring systems of Brazilian Amazon, such as TerraClass [3] and MapBiomas [4]. This class is included in generic mixed classes of LULC, that comprises undistinguishable features due to their small size. Besides using medium spatial resolution images, the mapping scale of these monitoring systems makes this land use undistinguishable, so they are included in mixed classes that embraces agriculture, pasture, secondary vegetation, etc.

Identifying small-scale agriculture with satellite images is a challenge and demands new approaches. Only by knowing about the presence of this type of agriculture system and its spatial pattern, cartographic representation can be complete and small-scale agriculture can take its part in policies, since it makes possible to understand how this system works, who are the agents involved on it and so on [5]. Also, it is important to highlight that a large part of the food products that supply cities come from family farming included in small-scale agriculture. Therefore, we reiterate the importance of studying small-scale agriculture in the Amazon as a way of demonstrating quantitative and qualitative data of this important land use to subsidize public policies for the regional economy.

Classifying an image requires high analyst experience and good background knowledge about the region of classification. Moreover, combining diverse parameters in supervised classification can improve the methods and provide good results [6]. Therefore, it is important to address the specific characteristics regarding small-scale agriculture in Brazilian Amazon. Especially for Amazon, scenes from optical sensors face problems with high proportion of cloud cover, which is worsened by low temporal resolution of some types of images [7]. Also, there is a significant confusion between small-scale agriculture with pasture and secondary vegetation, once they show similar spectral responses [8]. However, those LULC classes vary in shape and size [5].

To overcome these issues, different methods have been used to collect meaningful and useful information from image processing, GIS and modeling [9]. To consider shape information, the analyst may select Object-Based Image Analysis (OBIA) methods for image classification [10][11]. These methods address spectral and shape attributes using image segmentation. In addition, an alternative for high spatial resolution images to reduce cloud cover is to make a mosaic of images with the best pixels found in a certain period.

Overall, considering the gap of information and challenges, there is still a lot to do when identifying small-scale agriculture using remote sensing techniques. To assess how different methods perform, this work proposes to classify small-scale agriculture in an area that includes part of two municipalities in the Baixo Tocantins region (Brazilian Amazon), testing different image classification

 ISBN: 978-1-61208-871-6

algorithms based on OBIA and high spatial resolution image, the PlanetScope product of Surface Reflectance Mosaic.

We organized this paper in four sessions. In Section 2, we provide the state of the art regarding small-scale agriculture mapping, we discuss related work, existing solutions and limitations. In Section 3, we describe the PlanetScope scene, discussing its characteristics, as well as the study area. We explain all the methods applied, from image segmentation to classification and accuracy assessment. In Section 4, we show and discuss the results for each classification method, comparing their performance mainly related to small-scale agriculture class. In Section 5, we give our final remarks, highlighting pros and cons of the tested methods and future work opportunities for this matter.

II. State of the art

In Brazilian Amazon several studies on agriculture have been carried out with remote sensing techniques. However, most of them addresses large-scale agriculture. Few studies can be found related to small-scale agriculture. On the other hand, there are plenty of techniques that can be tested for mapping this land use class. Therefore, the main contribution of this study relies on testing and evaluating techniques capable of detecting this type of agriculture, which is largely invisible, despite its importance to society, environment and economy. By doing so, we explore the challenges, potentials and constraints of mapping small-scale agriculture in Brazilian Amazon.

When searching for techniques, we can find plenty of options of classification algorithms that can be used for agriculture classification. Some authors use traditional approaches and algorithms to identify large-scale agriculture, such as Maximum likelihood (ML) and ISODATA [12][13]. Those algorithms are pixel-wise and depend uniquely on spectral response of the targets. However, small-scale agriculture is often composed of more than one crop types at different stages of growth, which results frequently in a spectral mixture at the pixel level.

By using ML classification, a study carried out in Ethiopia could not discriminate crop types not even when the crop classes were groupings. To improve the result, the authors applied a second approach by using a neural network for sub-pixel classification of the image [13]. Another study carried out in Brazilian Amazon used ML to map smallholders and reported that this algorithm is not appropriate for it, the results showed a very low accuracy (8%) for small-scale agriculture [6]. The authors also performed a classification by segmenting the image and applying Adapted Nearest Neighbor method. In this case, they were able to reach a higher accuracy (64%) for detecting small-scale agriculture.

Segmentation and object-based analysis are broadly used in many studies [10][14]-[17]. The main gain of using this technique for small-scale agriculture detection is because the segmentation allows the use of more features, such as shape, texture and so on, rather than only spectral ones. Once small-scale agriculture has specific shape and texture, and spectral mixture, an object-based analysis unfolds as a key technique [6]. For some authors, working with object is also an advantage due to the ease of interpretation, for the features correspond to elements of landscape [15].

Lastly, it is important to mention the image spatial resolution. Although [18] do not work directly with LULC classification methods, their research is based on remote sensing data and techniques and the analysis of small-scale agriculture intensification. The authors used Landsat and MODIS imagery, which have a spatial resolution of 30 and 250 m, respectively. However, the use of higher spatial resolution images is more suitable for small farms, where cultivation takes place in areas smaller than 1ha, as in Northeast of Pará state [6]. In that sense, many authors used RapidEye imagery, which have a spatial resolution of 5 m [6][12][19]. Furthermore, the red-edge band from RapidEye imagery is adequate to discriminate the stage of vegetation, supporting small-scale agriculture mapping [20]-[22].

Overall, we can notice that frequently small-scale agriculture is not considered in LULC mapping and there are only a few researches regarding this matter. In the researches we reviewed we could commonly observe the use of the combination of different techniques: ML and neural networks [13]; multiresolution segmentation and adapted nearest neighbor [6]; and segmentation and random forest algorithm [14][15]. In some studies, we can observe that the authors adapted and tested techniques used to large-scale agriculture, but considering the unique features of small-scale agriculture in Amazon.

III. Material and methods

PlanetScope product of Surface Reflectance Mosaic is a free access and analysis-ready level product. This product is a Level-1 processing data, including geometric and atmospheric correction. Atmospheric parameters are estimated from an external data source, such as MODIS [23]. The availability of Planet's monthly mosaics comprises an initiative of Norway's International Climate and Forest Initiative (NICFI), which aims to provide universal access to monitoring the tropics through high spatial resolution satellite imagery, to support efforts to stop the deforestation of the world's rainforests [23].

The scene is a monthly composition of pixels acquired during the month of September, 2020. This mosaic has spatial resolution of 4.77 m and has four spectral bands: red, green, blue and near-infrared [23]. The image covers part of Mocajuba and Cametá municipalities, in the Northeast of Pará state, Brazil (Figure 1). We selected this study area because it is a hotspot of small-scale agriculture in Pará State [10]. This region is historically occupied by many smallholders that are responsible to provide food for local and regional markets, besides self-consumption [24]. In both municipalities, the main crop of small-scale agriculture is cassava (Table 1), which is planted exclusively in farms smaller than 10 ha. Açai, black pepper and cocoa are important crops, but they are not exclusive to small-scale farms, they are also cultivated in farms with area up to 200 ha [25].

Furthermore, this region is a hotspot of secondary vegetation [26], that can be an indicator of shifting

 ISBN: 978-1-61208-871-6

cultivation, usually performed in Amazon by smallholders that leave part of the land to fallow and return their cultivation afterwards, reincorporating the nutrients and minerals to the soil [27]. The main river body are Tocantins River and its tributaries.

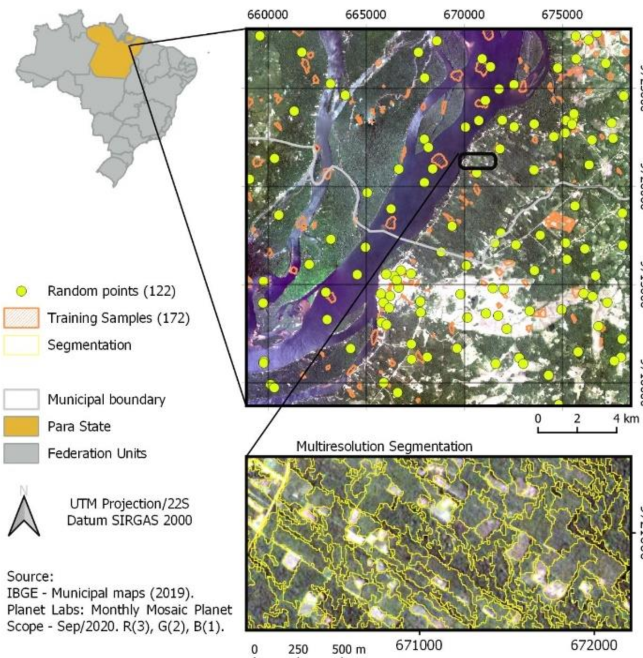

Figure 1. Study area (Mocajuba and Cametá municipalities, Pará, Brazil), training samples, random validation points, and segmentation example.

TABLE I. MAIN CROPS OF SMALL-SCALE AGRICULTURE IN CAMETÁ AND MOCAJUBA, CONSIDERING PROPERTIES UP TO 10 HA (2017)

Crop	Harvested area (ha)
Cassava	4,507
Açai	1,676
Black pepper	413
Cocoa	402
Corn	319

Source: [25].

We used two different supervised methods: Adapted Nearest-neighbor (NN) [28] and C5.0 Decision trees, an improved version of C4.5 [29]. The main steps applied for each approach were: a) multiresolution segmentation and parameters definition; b) training sampling; c) classification; and d) assessment of accuracy. We chose LULC classes considering visual surveys and landscape descriptions in similar studies developed at the same location (Table 2).

TABLE II. LAND-USE AND LAND-COVER CLASSES

Classes	Description
Water	Water bodies: rivers, lagoons, etc.
Forest	Natural vegetation with predominance of trees
Secondary vegetation	Natural vegetation in regeneration emerged from previously deforested areas, with trees, shrubs and herbs
Urban areas	Built-up areas with population clusters: city, village and community
Pasture	Predominance of herbaceous and grassy vegetation, it may occur also sparse shrub vegetation and few arboreal individuals
Small-scale agriculture	Small agriculture lands with mainly annual crops
Others	Aggregate of land use and land cover, such as rocky outcrops, sand banks
Non observed	Clouds and cloud shadows

Source: adapted from [6].

A. *Multiresolution segmentation*

The multiresolution segmentation was generated by eCognition 9.0.1 [30]. This kind of segmentation creates initially an 1-pixel-sized object and merges the neighbor pixels with similar features consecutively into the object [31]. When observing similar features, the algorithm considers a combination of spectral and shape criteria. The analyst can weigh the priority criteria when creating the segments. We repeatedly tested different values for segmentation parameters, until the algorithm could create appropriate objects for small-scale agriculture. We found the values of the following parameter more adequate:

- Image layer weights = 1, 1, 1, 1. It ranges from 0 to 1. The same importance was assigned to all bands (R, G, B, NIR);
- Scale parameter = 60. It defines the size of polygons;
- Shape = 0.7. It defines the weight the shape criteria must have for segmentation. It ranges from 0 to 1, the higher its value, the lower the influence of the color. The chosen value prioritizes the shape over color;
- Compactness = 0.5. It defines the weight of compactness criteria. It ranges from 0 to 1, the higher the value, the more compact image objects may be, that is the borders are closer to the center of the segment.

B. *Sampling design*

This step was performed using simple random sampling method throughout the scene. The sample units were objects obtained from the multiresolution segmentation. We selected the same objects/segments as training samples for both methods to keep them uniform. In total, we selected 172 objects for all the 8 classes. We sampled 20 objects for each class described in Table 2, except *Cloud shadows,* with 12 samples. We collected fewer samples of *Cloud shadows* because this class was less representative in terms of area.

Note that we collected *Clouds* and *Cloud shadows* samples individually. After performing the classification, we merged those features into *Non observed* class.

Overall, this step involved the random selection of segments from different shapes and sizes in each class. Water sample segments presented a mean area of 12.83 ha ± 8.60 ha for these specific samples, compared to other classes (Table 3). On the other hand, small-scale agriculture presented a mean area of 0.84 ha ± 0.36 ha for these samples. As we standardized only the number of samples, this variation in sample size might affect the final classification according to spectral mixture features of the pixels, and it may impact the accuracy. The main class for this study is small-scale agriculture and the random samples selected in this class presented similar sizes. As we see in Figure 2, this is a size pattern for this class in the Amazon for both algorithms, so the samples represented well the classification for this region.

TABLE III. DESCRIPTIVE STATISCS FROM SAMPLING POLYGONS, IN HECTARE

Classes	μ	σ	σ²
Water	12.83	8.60	73.87
Forest	3.86	3.25	10.56
Secondary Vegetation	6.05	2.63	6.90
Urban Area	1.44	0.94	0.89
Pasture	3.09	2.32	5.37
Small-scale agriculture	0.84	0.36	0.13
Others	2.72	2.74	7.49

Units: hectare; μ = polygon mean area; σ = standard deviation; σ^2 = variance.

C. *Object-based image analysis*

OBIA is an alternative to pixel-to-pixel approaches as it relies on identifying regions of the image: it uses segments to extract neighborhood, spectral and spatial features, composing the feature space [10][11]. This method has shown to be more suitable to identify small-scale agriculture and to distinguish this land use from other classes [6]. Therefore, we ran two OBIA methods using the same image, samples and accuracy points. Both methods differ when it comes to classifying an object: the first one considers the nearest neighbor in the feature space, while the latter one uses multiple decision trees to identify the proper class, as explained in the next sections.

C.1 Adapted Nearest-neighbor (NN)

This method is an adapted version of nearest neighbor, which considers not only the spectral features, but also other features related to the object [28]. The algorithm formulates a feature space considering the attributes including all segments, then it searches for the closest sample and assigns that class to the segment [30]. The analyst determines the correspondent features. To proceed with classification, we chose the features according to [6]: a) Spectral attributes: brightness, mean and standard deviation of each band, b) Object attributes: shape index. In total, we used 13 attributes.

C.2 Decision trees (C5.0)

We used the objects from multiresolution segmentation to perform a feature extraction from the original image by Geographical Data Mining Analyst (GeoDMA), an open-source plug-in available for TerraView 5.5.1 [32]. In total, the algorithm generated 103 features, considering both spectral and spatial features, e.g., mean, mode, and maximum values of each band, polygon angle, shape index, compactness etc. After finishing the feature extraction, we collected the training samples and ran a boosting C5.0 Decision trees classification.

This algorithm generates a pre-set number of decision trees from the sample features, which is applied when classifying the segments. In total, the algorithm generated 99 trees independently. The final classification for each segment is the one that was assigned by the most of the trees [33].

Among the 103 features, the algorithm highlights the ones that were more used when classifying an object. The main features, that showed 100% usage in this classification were:

- Band 2 (Green): maximum value and band ratio;
- Band 3 (Blue): median, mode, dissimilarity and contrast;
- Band 4 (Near-infrared): mean, median and band ratio.

Band ratio is the contribution of the given band to the region. Contrast is the measure of the intensity contrast between a pixel and its southeast neighbor over the object, aka Sum of Squares Variance. Dissimilarity is the measure of how different the elements of the Gray-level co-occurrence matrix are from each other [34].

Regarding object features, the more important ones, which had ca. 60% of usage, were: Perimeter, compacity, radius and bounding box area.

D. *Assessment of accuracy*

To assess the accuracy of both classifications, 122 points were randomly distributed throughout the scene, representing all classes. For each class, we collected 20 random points, except for *Pasture*. Note that we had only 2 points for *Pasture* due its low scene cover.

We used those points as test samples once the ground truth was assigned by inspecting the actual land use and land cover with Google Earth Pro 2020 images. Then, we used those points to validate NN and C5.0 classifications. From that point on, confusion matrix was created for each classification and we computed the accuracy and kappa index. These information were the basis for identifying the main confusion occurring to small-scale agriculture areas and the overall performance of each method, leading to the most suitable one for small-scale agricultural mapping [35][36].

IV. RESULTS AND DISCUSSION

Classification maps are showed on Figure 2. Small-scale agriculture is more present in the upland region, even though that both riverine and upland population are acknowledged as important agents involved into this land use activity [6]. As stated before, high spatial resolution sensors are more adequate to improve classification accuracy due to the small-scale agriculture's size: our results presented mean area of 0.97 ha ± 0.69 ha for NN and 0.70 ha ± 0.39 ha for C5.0 Decision trees (Table 4). Considering that TerraClass maps have a minimum mapping area of 6.25 ha [6], it can not identify and map properly small-scale agriculture. That explains why this class is not explicitly visible in current LULC monitoring systems.

 ISBN: 978-1-61208-871-6

Descriptive statistics indicate similar classification area in both algorithms for the classes of water, forest, secondary vegetation, urban area, and others. The study area shows great forest and secondary vegetation cover in mainland and at the islands, covering ca. 60% of the scene in both maps. Water covers around 18% of the area.

As presented in confusion matrix (Table 5), both algorithms presented similar overall accuracy (NN = 75% and C5.0 Decision trees = 73%) and kappa values. On the other hand, C5.0 Decision trees algorithm found better results when mapping small-scale agriculture (75%), compared to NN (65%). This performance of Adapted Nearest-neighbor algorithm is corroborated with other studies that found around 62% of producer's accuracy for small-scale agriculture carried out in the same region of Brazilian Amazon [6].

NN may be overclassifying small-scale agriculture, representing 9.3% of the mapping area. For this classification, commission error was 19%, which indicates that a significant number of polygons were classified by mistake as small-scale agriculture, increasing the area of this class. These classification errors occurred due to confusion, especially with secondary vegetation, forest, and others.

For C5.0 Decision trees, there was no commission error for small-scale agriculture class, which represents 4.8% of the mapping area. In other words, C5.0 Decision trees is more conservative for mapping small-scale agriculture and did not included other classes in small-scale agriculture by mistake as NN did.

Both algorithms showed the same omission errors for small-scale agriculture regarding secondary vegetation (15%) and pasture (10%) classes. NN also showed omission errors for small-scale agriculture with the class others (10%).

According to the literature, similar spectral attributes may affect the classification of small-scale agriculture, once this class has similar spectral responses to other classes, such as pasture [8] and secondary vegetation [5].

Small-scale agriculture and pasture differ in terms of size and shape [5]. Although the last one showed a mean area of 1.77 ha in both maps, the only area in the scene, identified as pasture by the analysts, had actually 100 ha. Yet, this single pasture area was segmented into many smaller objects by the multiresolution segmentation, once covered by grassland in different stages, for instance, clean pasture, shrubby pasture. When segmenting an image with parameters adjusted to small-scale agriculture objects, it is necessary to resort the spatial resolution with better detail, which influences other targets segmentation. In this case, although the targets differ when it comes to object features, there was still confusion among them. Anyhow, the confusion with pasture might also have been influenced by the small number of samplings due to the lack of other pasture areas in this scene. To better represent these classes we should separate them in two classes, clean and shrubby pasture. But due to the low presence of pasture on the scene, this step was not feasible. However, we highlight the confusion and the need to separate this land use in two distinct classes in case this methodology is applied in an area where pasture is more significant.

Regarding the confusion with secondary vegetation, once again the spectral features are the main reason for it, as well as the different stages of secondary vegetation, that poses similar stratum as agriculture. The mean size of secondary vegetation polygons was 1.71 ± 1.34 for NN and 1.58 ± 1.18 for C5.0. Even though this confusion is not opportune, it is important to address that secondary vegetation can be part of small-scale agriculture production system by forming biomass while the land is under fallow [37]. In this case, secondary vegetation poses as an asset, once its function is to ensure the land fertility [27].

Overall, the results for small-scale agriculture were adequate and despite the different accuracies, both methods showed limitations when differentiating this class from pasture and secondary vegetation.

V. Conclusion and Future Work

As the first challenge faced in this study, we can highlight the successful attempted of using the same training and validation samples for the classification and evaluation steps, to promote an adequate comparison between the algorithms tested. Using the same samples in different software and algorithms is not always possible or easy. For instance, C5.0 automatically generates the confusion matrix for the classification, but does not point out the samples used in the validation process. For this reason, we collected randomly distributed points in the image to evaluate the two classifications. We can summarize our main findings as following:

- Considering all small-scale agriculture identification challenges, C5.0 Decision trees results were able to reach a higher producer's accuracy compared to NN method. Note that both methods were run relying upon different magnitude of features: C5.0 Decision tree identify automatically 103 features, however the ones that mattered the most for the classification were 33 of them, which were used at least in 50% of the decision trees. On the other hand, NN used 13 features pointed out by the analyst. So, while the latter requests that the analyst decides how many and which features are going to be inserted in the feature space, C5.0 Decision trees use as default 103 features and saves the analyst from selecting the most suitable features for small-scale agriculture.

- Therefore, C5.0 showed greater results, representing the most suitable method for mapping small-scale agriculture in the study area. Nevertheless, we recommend carrying out more studies over larger areas to identify the best attributes for classifying small-scale agriculture and overcome misclassification errors, as well as in other mixed land cover types and landscape diffuse patterns.

 ISBN: 978-1-61208-871-6

Regarding future work, we have the following remarks:

- We recommend investigating which features are more significant for the identification of small-scale agriculture by C5.0. We suggest a systematically removal of features at the classification level and performing a sensitive analysis;

- We believe that temporal analysis can be explored in future work. The inclusion of the temporal component coupled with machine learning and deep learning techniques may contribute for selecting other important variables for small-scale agriculture classification. Additionally, the use of these methods may contribute for the advancement of studies linked to agricultural intensification and fallowing practices in shifting cultivation agriculture, widely used in the context of small-scale agriculture in the Brazilian Amazon. Then, the inclusion of the time component will also be important to test whether it is relevant or not to use phenological metrics for agriculture with large species diversity that does not present regular crop cycles.

- Also, we strongly recommend testing different sampling design to test better results and perform a sensitive analysis.

TABLE IV. DESCRIPTIVE STATISTICS FROM IMAGE CLASSIFICATION ACCORDING TO LAND USE AND LAND COVER, IN HECTARE

Classes	Adapted Nearest-neighbor					C5.0 Decision trees				
	μ	*σ*	*σ²*	*Total*	*%*	*μ*	*σ*	*σ²*	*Total*	*%*
Water	5.69	5.62	31.62	7,082.61	18.66	5.92	5.75	33.01	6,885.70	18.14
Forest	1.55	1.18	1.40	11,317.54	29.82	1.62	1.26	1.58	10,477.20	27.60
Secondary vegetation	1.71	1.27	1.61	11,182.62	29.46	1.58	1.18	1.39	11,934.05	31.44
Urban Area	0.69	0.60	0.36	633.66	1.67	0.70	0.63	0.40	537.14	1.42
Pasture	1.77	1.34	1.79	718.84	1.89	1.77	1.10	1.21	2,340.14	6.17
Small-scale agriculture	0.97	0.69	0.48	3,526.97	9.29	0.70	0.39	0.15	1,837.95	4.84
Others	1.22	1.37	1.88	3,493.84	9.20	1.38	1.38	1.92	3,943.92	10.39
Total	-	-	-	37,956.08	100	-	-	-	37,956.08	100

Units: hectare; μ = polygon mean area; σ = standard deviation; σ² = variance.

TABLE V. CONFUSION MATRIX FOR ADAPTED NEAREST-NEIGHBOR AND C5.0 DECISION TREES ALGORITHMS

Adapted Nearest-neighbor

	%	*Reference*							
		(A)	(B)	(C)	(D)	(E)	(F)	(G)	User's accuracy
Classification	(A) Water	**100**	0	0	0	0	0	0	100
	(B) Forest	0	**55**	35	0	0	0	0	61
	(C) Secondary vegetation	0	40	**60**	0	0	15	0	52
	(D) Urban area	0	0	0	**95**	0	0	0	100
	(E) Pasture	0	0	0	0	**50**	10	10	20
	(F) Small-scale agriculture	0	5	5	0	50	**65**	0	81
	(G) Others	0	0	0	5	0	10	**90**	86
	Producer's accuracy	100	55	60	95	50	65	90	
	Samples	20	20	20	20	2	20	20	
	Kappa = 0,70				*Overall accuracy = 75%*				

C5.0 Decision trees

	%	*Reference*							
		(A)	(B)	(C)	(D)	(E)	(F)	(G)	User's accuracy
Classification	(A)	**90**	0	0	0	0	0	0	100
	(B)	0	**50**	30	0	0	0	0	63
	(C)	0	45	**60**	0	0	15	0	50
	(D)	0	0	0	**85**	0	0	5	94
	(E)	0	0	10	15	**100**	10	20	15
	(F)	0	0	0	0	0	**75**	0	100
	(G)	10	5	0	0	0	0	**75**	83
	Prod. acc.	90	50	60	85	100	75	75	
	Samples	20	20	20	20	2	20	20	
	Kappa = 0,68				*Overall accuracy = 73%*				

 ISBN: 978-1-61208-871-6

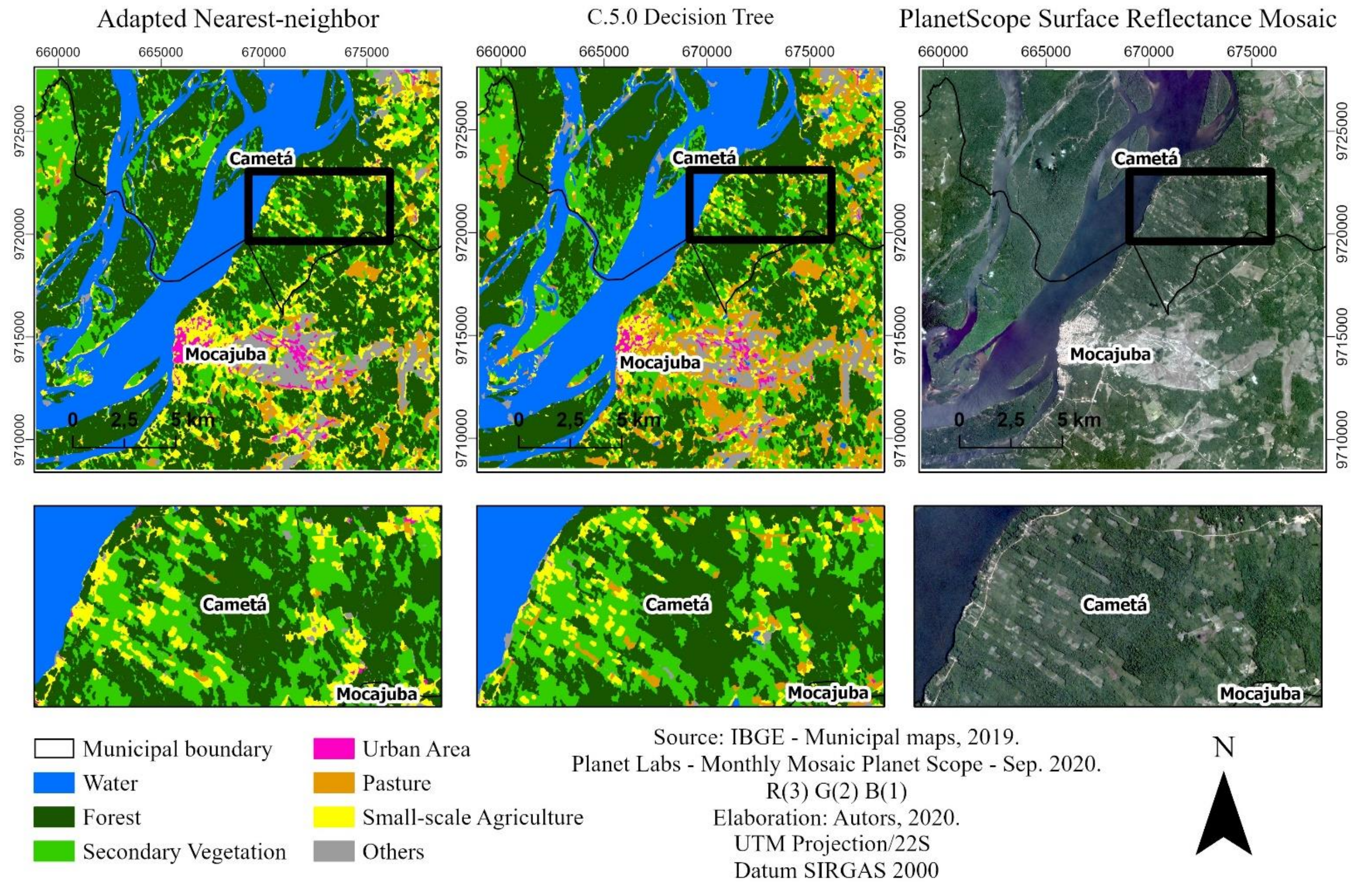

Figure 2. Land-Use and Land-Cover classification using NN and C5.0 Decision trees methods to identify small-scale agriculture.

ACKNOWLEDGMENT

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001" and in part by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). We thank the Brazilian National Institute for Space Research (INPE) for supporting the authors. Finally, we thank all anonymous reviewers for constructive suggestions on an earlier version of this paper.

REFERENCES

[1] J. Berdegué and R. Fuentealba, “Latin America: the state of smallholders in agriculture,” IFAD Conference on New Directions for Smallholder Agriculture. Jan, 2011, pp. 1-37. doi: 10.1093/acprof:oso/9780199689347.003.0005.

[2] C. Caetano, “Family farming, the struggle for land in the legal Amazon and the impacts of Law 13.465/2017,” “Agricultura Familiar, a luta pela terra na Amazônia Legal e os Impactos da Lei 13.465/2017,” Revista de Direito Agrário e Agroambiental, vol. 4, pp. 34–55, Jun. 2018.

[3] C. Almeida et al., “High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data,” Acta Amazonica, v. 46, pp. 291 – 302, Sep, 2016, doi:10.1590/1809-4392201505504.

[4] C. Souza et. al., “Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat Archive and Earth Engine,” Remote Sensing, v. 12, issue 17, p. 2735, Aug. 2020, doi:10.3390/rs12172735.

[5] A. R. Souza, Economy and nature: land use and land cover patterns associated with agricultural and extractive activities in communities in the Southwest of Pará, Economia e natureza: padrões de uso e cobertura da terra associados a atividades agropecuárias e extrativistas de comunidades do Sudoeste do Pará, Brazilian National Institute for Space Research, 2016. [Online]. Available from: http://urlib.net/8JMKD3MGP3W34P/3M8SD78. 2021.06.13.

[6] A. R. Souza, M. I. S. Escada, R. Marujo, and A. M. Monteiro, “Cartography of the invisible: unveiling small-scale agriculture with Rapideye Images in the Baixo Tocantins Region, Pará,” “Cartografia do Invisível: Revelando a Agricultura de Pequena Escala com Imagens Rapideye na Região do Baixo Tocantins, Pará,” Revista do Departamento de Geografia, v. 38, pp. 137–153, Dec. 2019, doi:10.11606/rdg.v38i1.151603.

[7] M. Reis, M. I. S. Escada, S. Sant’anna, and L. Dutra, “Land use and land cover trajectory classification and analysis methods in the amazon: implications for forest regeneration studies,” Revista Brasileira de Cartografia, v. 72, p. 1114, Dec. 2020, dx.doi.org/10.14393/rbcv72nespecial50anos-56535

[8] M. Rodrigues, H. Bendini, H. A. Soares, T. S. Körting, and L. Fonseca, “Remote sensing image time series metrics for distinction between pasture and croplands using the Random Forest Classifier,” 2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS), Mar. 2020, pp. 149-154, ISSN: 978-1-7281-4350-7, doi: 10.1109/LAGIRS48042.2020.9165671.

[9] E. M. L. M. Novo, Remote sensing: principle and applications, Sensoriamento remoto: princípios e aplicações, 4th ed., São Paulo: Blucher, 2010.

[10] T. Blaschke, “Object based image analysis for remote sensing,” ISPRS Journal of Photogrammetry and Remote Sensing, vol. 65, n. 1, pp. 2–16, 2010. 10.1016/j.isprsjprs.2009.06.004.

[11] G. J. Hay and G. Castilla, “Object-Based Image Analysis: Strengths, Weaknesses, Opportunities and Threats (SWOT),” International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI-4/C42, pp. 4-5, 2006.

[12] O. Morariu And F. Burescu, “Agricultural land cover classification using RapidEye satellite imagery,” Journal of Young Scientist, vol. 6, pp. 115-122, 2018, ISSN 2284-8012.

[13] J. Delrue et al., “Crop mapping in countries with small-scale farming: a case study for West Shewa, Ethiopia,” International Journal of Remote Sensing, vol. 34, pp. 2566-2582, Dec. 2012, https://doi.org/10.1080/01431161.2012.747016.

[14] M.F.A. Vogels, S.M. Jong, G. Sterk, and E.A. Addink, “Mapping irrigated agriculture in complex landscapes using SPOT6 imagery and object-based image analysis – A case study in the Central Rift Valley, Ethiopia,” Int J Appl Earth Obs Geoinformation, vol. 75, pp. 118-129, Mar. 2019, https://doi.org/10.1016/j.jag.2018.07.019.

[15] L. P. Dutrieux, C. C. Jakovac, S. H. Latifah, and L. Kooistra, “Reconstructing land use history from Landsat time-series: Case study of a swidden agriculture system in Brazil,” International Journal of Applied Earth Observation and Geoinformation, vol. 47, pp.112-124, Nov. 2015, http://dx.doi.org/10.1016/j.jag.2015.11.018.

[16] Y. Inoue et al., “Traceability of slash-and-burn land-use history using optical satellite sensor imagery: a basis for chronosequential assessment of ecosystem carbon stock in Laos,” International Journal of Remote Sensing, vol. 28, n. 24, Dec. 2007, doi: 10.1080/01431160701656323.

[17] H. T. T. Nguyen, T. M. Doan, E. Tomppo, and R. E. McRoberts, “Land Use/Land Cover Mapping Using Multitemporal Sentinel-2 Imagery and Four Classification Methods—A Case Study from Dak Nong, Vietnam,” Remote sensing, vol. 12, n. 9, pp. 1-27, doi:10.3390/rs12091367.

[18] M. Jain, P. Mondal, R. S. DeFries, C. Small, and G. L. Galford, “Mapping cropping intensity of smallholder farms: A comparison of methods using multiple sensors,” Remote Sensing of Environment, vol. 134, pp. 210-223, Jul. 2013, http://dx.doi.org/10.1016/j.rse.2013.02.029.

[19] B. Tapsall, P. Milenov, and K. Tasdemir, “Analysis of RapidEye imagery for annual landcover mapping as an aid to European Union (eu) common agricultural policy,” ISPRS TC VII Symposium (100 Years ISPRS), Jul. 2010, pp. 568-573.

[20] S. Adelabu, O. Mutanga, E. Adam, “Evaluating the impact of red-edge band from Rapideye image for classifying insect defoliation levels,” ISPRS Journal of Photogrammetry and Remote Sensing, vol. 95, pp. 34–41, Sep. 2014, https://doi.org/10.1016/j.isprsjprs.2014.05.013.

[21] H. O. Kim, J. M. Yeom, “Effect of red-edge and texture features for object-based paddy rice crop classification using RapidEye multi-spectral satellite image data,” International Journal of Remote Sensing, vol. 35, n. 19, pp. 7046–7068, Jan. 2014, https://doi.org/10.1080/01431161.2014.965285.

[22] J. González-Piqueras et al., “Monitoring crop N status by using red edge-based indices,” Advances in Animal Biosciences, vol. 8, n. 2, pp. 338–342, Jul. 2017, doi:10.1017/S2040470017000243.

[23] Norway’s International Climate and Forest Initiative. *NICFI Data Program*. 2020. [Online]. Available from: assets.planet.com/docs/NICFI_General_FAQs.pdf. 2021.06.13.

[24] I. M. C. C. Cordeiro, L. G. T. Rangel-Vasconcelos, G. Schwartz, and F. A. Costa, Northeast of Pará: overview and sustainable use of secondary forests, Nordeste Paraense: panorama geral e uso sustentável das florestas secundárias. Belém: EDUFRA, 2017. ISBN 978-85-7295-118-0.

[25] Brazilian National Institute of Geography and Statistics. *Agricultural census. Censo agropecuário.* [Online]. Available from: sidra.ibge.gov.br/pesquisa/censo-agropecuario/censo-agropecuario-2017. 2021.06.13.

[26] R. Carvalho, M. Adami, S. Amaral, F. Bezerra, and A. P. Aguiar, “Changes in secondary vegetation dynamics in a context of decreasing deforestation rates in Pará, Brazilian Amazon,” Applied Geography, v. 106, pp. 40-19, May, 2019, doi: 10.1016/j.apgeog.2019.03.001.

[27] F. A. Costa, “Contributions of fallow lands in the Brazilian Amazon to CO2 balance, deforestation and the agrarian economy: Inequalities among competing land use trajectories,” Elementa Science of the Anthropocene, v. 4, pp. 1-24, Oct. 2016, doi: 10.12952/journal.elementa.000133.

[28] A. R. Webb, Statistical Pattern Recognition, 2nd ed., Malvern: John Wiley & Sons, 2002.

[29] J. Quinlan, C4.5: programs for machine learning, San Mateo, CA: Morgan Kaufmann, 1993.

[30] Trimble Germany GmbH. *eCognition Documentation - Classification*. 2019. [Online]. Available from: http://docs.ecognition.com/v9.5.0/eCognition_documentation/User%20Guide%20Developer/6%20About%20Classification.htm. 2021.06.13.

[31] M. Baatz and A. Schape, Multiresolution Segmentation: An Optimization Approach for High Quality Multi-Scale Image Segmentation. 2000. In: J. Strobl, T. Blaschke, and G. Griesbner, Eds., Angewandte Geographische Informations-Verarbeitung, XII, Wichmann Verlag, Karlsruhe, Germany, 12-23.

[32] GeoDMA 2.0.1. São José dos Campos: Brazilian National Institute for Space Research (INPE), 2011. [Online]. Available from: http://www.dpi.inpe.br/geodma. 2021.06.13.

[33] L. Maurano, M. I. S. Escada, and C. Renno, “Spatial patterns of deforestation and estimation of accuracy of PRODES maps for the Brazilian Legal Amazon,” “Padrões espaciais de desmatamento e a estimativa da exatidão dos mapas do PRODES para Amazônia Legal Brasileira,” Ciência Florestal, v. 29, issue 4, pp. 1763-1775, Dec. 2019, doi:10.5902/1980509834380.

[34] T. S. Korting, GeoDMA: a toolbox integrating data mining with object-based and multi-temporal analysis of satellite remotely sensed imagery, Brazilian National Institute for Space Research, 2012. [Online]. Available from: urlib.net/8JMKD3MGP7W/3CCH86S. 2021.06.13.

[35] G. Banko, A review of assessing the accuracy of classifications of remotely sensed data and of methods including remote sensing data in forest inventory, International Institute for Applied Systems Analysis, 1998. [Online]. Available from: https://www.researchgate.net/publication/23738314. 2021.06.13.

[36] R. G. Congalton and K. Gree, Assessing the accuracy of remotely sensed data: principles and practices. 2nd ed., Boca Raton: CRC Press, 2009.

[37] C. C. Jakovac, L. P. Dutrieux, L. Siti, M. Peña-Claros, and F. Bongers, “Spatial and temporal dynamics of shifting cultivation in the middle-Amazonas river: Expansion and intensification,” PLOS ONE, v. 12, n. 7, pp. 1–15, 2017, doi.org/10.1371/journal.pone.0181092.