SOCIEDAD LATINOAMERICANA EN PERCEPCIÓN REMOTA
Y SISTEMAS DE INFORMACIÓN ESPACIAL
XXI SIMPÓSIO INTERNACIONAL SELPER
4-8 de novembro 2024 - Belém, Brasil

DYNAMICS OF SMALLHOLDER AGRICULTURAL PRODUCTION IN THE MUNICIPALITY OF BREVES, PARÁ STATE, BRAZIL

Victória Beatriz Soares [1]; Gustavo Bayma [2;] Édson Luis Bolfe [3];
Taya Cristo Parreiras [4]; Victória H. Matusevicius e de Castro [5]; Edson Eyji Sano [6]

[1] Pontifical Catholic University of Campinas - PUCC, Faculty of Geography,
Campinas-SP, Brazil. victoria.leandro@colaborador.embrapa.br
https://orcid.org/0009-0006-6176-2545

[2] Embrapa Meio Ambiente, Jaguariúna-SP, Brazil. gustavo.bayma@embrapa.br
https://orcid.org/0000-0001-5312-6609

[3] Embrapa Agricultura Digital, Campinas-SP, Brazil. edson.bolfe@embrapa.br
https://orcid.org/0000-0001-7777-2445

[4] Campinas State University - UNICAMP, Postgraduate Program in Geography,
Campinas-SP, Brazil. tayacristo1@gmail.com
https://orcid.org/ 0000-0003-2621-7745

[5] Campinas State University - UNICAMP, Faculty of Agricultural Engineering,
Campinas-SP, Brazil. vhmatusevicius@gmail.com
https://orcid.org/0009-0009-6801-201X

[6] Embrapa Cerrados, Brasília-DF, Brazil. edson.sano@embrapa.br
https://orcid.org/0000-0001-5760-556X

RESUMEN. Este estudio se centró en evaluar la dinámica de las áreas agrícolas de pequeños agricultores entre 2019 y 2023 en Breves (PA), utilizando técnicas de geoprocesamiento y teledetección. La clasificación de las imágenes Sentinel-2 resultó en excelentes tasas de precisión, con una precisión global del 99% y un coeficiente Kappa de 0,9764 para distinguir entre bosque y no bosque, y una precisión del 89,7% y un coeficiente Kappa de 0,8554 para distinguir entre diferentes tipos de cobertura del suelo. El análisis reveló un aumento significativo de las superficies de vegetación secundaria y suelo expuesto, mientras que las superficies de bosque, pastizales y agua permanecieron relativamente estables durante todo el periodo de análisis. La transición entre las clases de cobertura del suelo indicó una conversión de las áreas de bosque a vegetación secundaria y luego a suelo expuesto, lo que posiblemente refleja la tala de árboles y las actividades agrícolas a pequeña escala. Estos resultados ponen de relieve la importancia de la teledetección para caracterizar y supervisar la producción agrícola a pequeña escala, proporcionando valiosos conocimientos para el desarrollo rural sostenible.

Palabras-clave: Amazonía; Uso y cobertura del suelo; Transición agrícola; Random Forest.

ABSTRACT. This study focused on assessing the dynamics of smallholder agricultural areas between 2019 and 2023 in Breves, Pará State, Brazil, using geographical information systems and remote sensing techniques. The classification of Sentinel-2 images resulted in an overall accuracy of 99% and a Kappa coefficient of 0.98 for distinguishing between forest and non-forest, and an accuracy of 89.7% and a Kappa coefficient of 0.86 for distinguishing between different types of land cover. The analysis revealed a significant increase in the areas of secondary vegetation and bare soil, while the areas of forest, pasture, and water remained relatively stable throughout the period of analysis. The transition between land cover classes indicated a conversion of areas from forest to secondary vegetation and then to bare soil, possibly reflecting logging and small-scale farming activities. These results highlight the importance of remote sensing in characterizing and monitoring small-scale agricultural production, providing valuable insights for sustainable rural development.

Keywords: Amazonia; Land use and cover; Agricultural transition; Random Forest.

RESUMO. Este estudo teve como objetivo avaliar a dinâmica das áreas agrícolas de pequenos produtores entre 2019 e 2023 em Breves (PA), utilizando técnicas de geoprocessamento e sensoriamento remoto. A classificação das imagens Sentinel-2 resultou em excelentes taxas de precisão, com uma precisão geral de 99% e um coeficiente Kappa de 0,98 para distinguir entre floresta e não floresta, e uma precisão de 89,7% e um coeficiente Kappa de 0,86 para distinguir entre diferentes tipos de cobertura do solo. A análise revelou um aumento significativo das áreas de vegetação secundária e de solo exposto, enquanto as áreas de floresta, prados e água se mantiveram relativamente estáveis ao longo do período de análise. A transição entre classes de ocupação do solo indicou uma conversão de áreas florestais em vegetação secundária e depois em solo exposto, possivelmente refletindo o abate de árvores e atividades agrícolas de pequena escala. Estes resultados destacam a importância do sensoriamento remoto para caracterizar e monitorar a produção agrícola de pequena escala, fornecendo informações valiosas para o desenvolvimento rural sustentável.

Palavras-chave: Amazônia; Uso e cobertura do solo; Transição agrícola; Random Forest.

1. INTRODUCTION

Smallholder agricultural production is present in all of Brazil's biomes and plays a fundamental role in ensuring the food and nutritional security of the Brazilian population (EMBRAPA, 2017). According to data from the 2017 Agricultural Census, 77% of Brazilian agricultural establishments are classified as family farms and employ more than 10 million people, which corresponds to 67% of the country's agricultural workforce (IBGE, 2017a).

Despite the importance of smallholder agricultural production, the most advanced types of agricultural technology tend to reach large-scale farms more effectively, while small and medium-scale farms have numerous difficulties in accessing digital technologies (Bolfe et al., 2020). In this sense, smallholder agricultural production is closer to agro-ecological principles because it is characterized by smaller plots of land, the use of local or nearby inputs, its own labor force and a short marketing chain (Altieri, 2004). In addition, many small-scale farms are closely linked to the production of ecosystem services, such as promoting biodiversity, preventing erosion and reusing waste to improve soil fertility (Silva *et al.*, 2021), providing benefits to the community (Schmidt., 2016).

For this reason, public projects and actions that focus on the production profile are increasingly necessary, such as the “Semear Digital” project, sponsored by the São Paulo Research Foundation (FAPESP) and coordinated by the Brazilian Agricultural Research Corporation (Embrapa Agricultura Digital), which aims to bring digital inclusion to small and medium-sized rural farmers through the establishment of Agrotechnological Districts (DATs) (CCD-AD Semear, 2022). In Brazil, smallholder agricultural production evolves in different ways depending on regional characteristics, including spatial distribution, variation in the products produced, and the techniques adopted (EMBRAPA, 2017). In Breves (PA) municipality, one of the Semear Digital´s DAT and the study area of this research, the smallholder agricultural production is diversified, including temporary and perennial crops such as açaí, cassava, and banana (IBGE, 2024). The production system for some crops involves the use of controlled fire in areas averaging 1 to 3 hectares to clean and prepare the soil for planting subsistence crops. After the harvest, the farmers often move to other areas and repeat the process. This process results in areas of “capoeira”, characterized by secondary vegetation, mainly grasses and shrubs, which are considered partially degraded.

Inspecting the Earth from space is critical for determining the long-term impact of human activity on natural resources. The information collected by remote sensing satellites for mapping the Earth's characteristics and infrastructure has become dynamic. Change detection of land use and land cover processes is very important, as it allows the analysis of changes that occur in different time periods (Mohanrajan; Loganathan; Manoharan., 2020). In this context, the aim of the study was to assess the dynamics of agricultural areas between 2019 and 2023, using geoprocessing and remote sensing techniques, with a focus on generating information to support the actions of local public actors in relation to more sustainable rural development.

2. MATERIAL AND METHODS

2.1 Study area

The Amazon estuary is formed by the Amazon and the Tocantins- Araguaia River basins and includes the Marajó Island, the largest fluvial archipelago in the world (Gonçalves *et al.*, 2015). Among the 17 municipalities of this region, we find the municipality of Breves in the Pará State (our study area), located between the municipalities of Melgaço and Bagre. Breves covers an area of 9549.52 km². The local landscape consists of an ecosystem composed mainly of “várzea”, “igapós”, “campos”, and “terra firme”. The region is composed of several interconnected islands surrounded by streams, holes, and channels. These channels serve as waterways for the Amazon and Tocantins rivers (IBGE, 2024). For this research, we selected three family farming areas (AOIs) in this municipality.

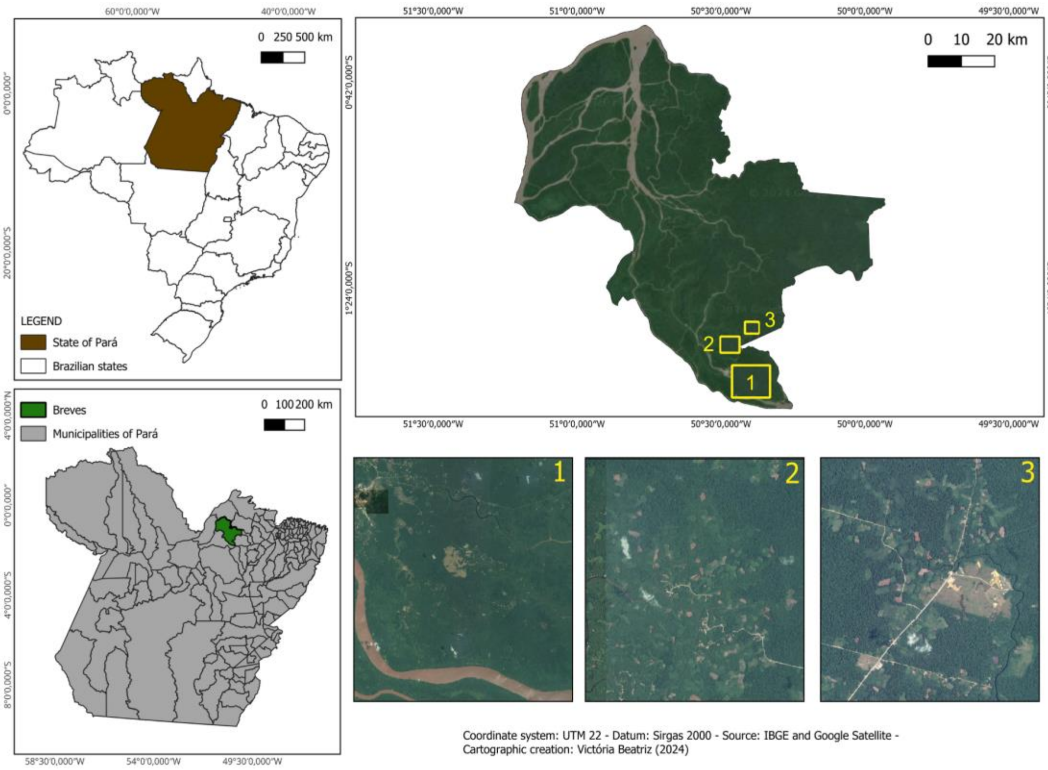

Figure 1. Location of study areas (AOIs) 1, 2 and 3 in the municipality of Breves, Pará State. Area 1 - 6 km southeast of the city of Breves. Area 2 -3 km west of the PA-159 at Retiro Duarte. Area 3 - near the Tucanu Açu Bridge, along the PA-159 highway.

2.2 Orbital image data processing

The first step was to select images from the Sentinel-2 satellite in the Google Earth Engine (GEE) platform. The blue (~ 493 nm), green (~ 560 nm), red (~ 665 nm), near infrared (~ 833 nm), and shortwave infrared 1 and 2 (~ 1610 and ~ 2190 nm) bands were used, with a spatial resolution of 20 m and a radiometric resolution of 12 bits (ESA, 2022). The annual images were assembled by selecting pixels from the first half of each year to reduce the presence of clouds. Thus, images representing the years between 2019 and 2023 were assembled.

The selected images were segmented in the GEE platform. Segmentation is a process that allows the identification and delineation of groups with similar spectral features in the images (Gonzalez; Woods, 2008). The segmentation method used in this study was the Simple Non-Iterative Clustering (SNIC) (Achanta.; Süsstrunk, 2017). The SNIC was chosen mainly because of its ability to enforce good connectivity between pixels from the beginning of the segmentation process. The average segment size in pixels was set to 16 to ensure a sufficient resolution for processing. A connectivity of 8 was chosen to define the relationship between pixels, which means that each pixel is connected to eight surrounding neighbors. The neighborhood size was set to 32, allowing a more comprehensive analysis of each point in the image. Finally, no specific seeds were defined, meaning that the algorithm does not use predetermined starting points for segmentation or analysis.

To train the Random Forest classifier, the AOIs from 2019 and 2023 were classified using visual interpretation. An interpretation key was created to analyze the Sentinel-2 RGB color composites of bands 4, 3, and 2 as well as bands 5, 3, and 6 (Figure 2). Forest areas appear dark green and with a rough texture, which indicates dense vegetation canopy. Secondary vegetation appears in a lighter shade of green and with low roughness, presenting a “smoother” appearance. The bare soil, considered here to be indicative of smallholder agricultural production, is purple colored, which indicates the presence of post-burn dry matter. Finally, the water body in this region appears in colors closer to purple, with variations within this palette, indicating that it is a river with deposits of organic matter (Moraes, 2021).

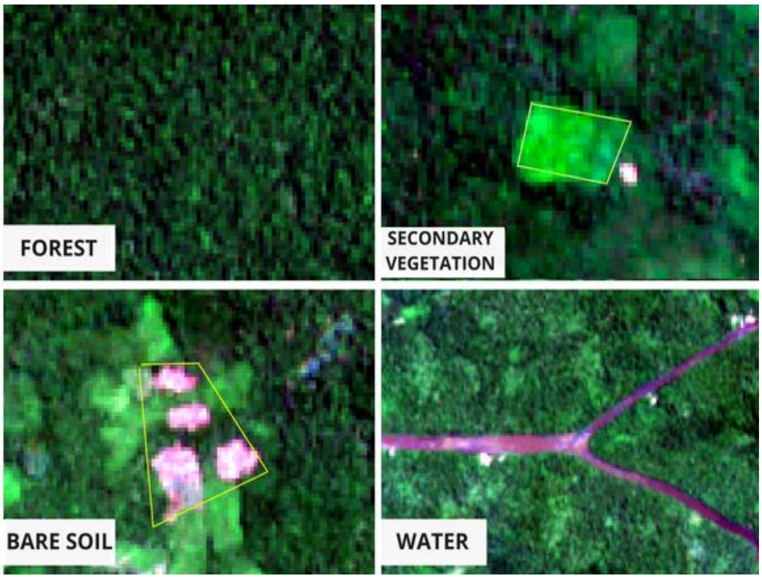

Figure 2. Interpretation key for delineating the main land use and land cover classes in Breves municipality, Pará State.

Based on the 2019 and 2023 visually classified segments, a sampling process was performed. In order to ensure that the sample set could be used in all years of the analysis, a previous step in preparing a set of possible samples involved identifying stable areas between 2019 and 2023, i.e., those that did not show a significant change during the period. The sample design was carried out with three assumptions: i) a minimum number of 50 samples from each class to validate the models, according to Congalton e Green (2019); ii) a sample volume within what is considered ideal for machine learning models, according to Maxwell; Warner, e Fang (2018), i.e. between 100 and 1000 times the number of input variables; and iii) a maximum non-information rate of 50%. This assumption ensures that the spatial representativeness of the majority class (in this case, forest) is maintained, while seeking greater model robustness with respect to minority classes. Thus, 400 sample segments of forest, 200 of secondary vegetation and water, 167 of bare soil and 100 of pasture were selected, distributed between the three AOIs. In order to fulfill the aforementioned assumptions, the decision was made to randomly select three sample pixels for each segment, increasing the size of the sample set but maintaining the spatial representativeness of the classes. Next, the values of each band (B2, B3, B4, B8, B11, and B12) were extracted from each sample for the year 2023, generating the data set containing the explanatory variables (bands) and the response variables (classes).

The final step was to divide the database into training and validation subsets at a ratio of 70%/30% and to perform supervised classification using the Random Forest (RF) algorithm, with k-fold cross-validation of k = 5 and 10 replicates. This was carried out with tools from the RF and caret packages in the R environment. After automatic classification of AOIs images from 2019 to 2023, the values for overall accuracy, precision, Kappa index, confusion matrix, and omission and commission statistics were used to assess the model performance. To minimize the salt and pepper effect, mode filtering with a 3 x 3 window was used in the final maps for levels 1 and 2. In addition, 50 samples from each class were selected in the 2019 images to assess the accuracy of the model's temporal transferability.

3. RESULTS AND DISCUSSION

3.1. Model performance results

The model's performance was tested at two levels. At Level 1, the classification sought to discriminate between forest (F) and non-forest (NF). The model achieved an overall accuracy of 99% and a Kappa coefficient of 0.98, indicating excellent agreement between the model's predictions and the actual values. The McNemar's test showed a p-value of 0.18, indicating that there is no significant difference between the classification errors, while the p-value [Acc > NIR] was less than 2×10^{-16}, confirming that the accuracy of the model is significantly higher than the uninformed accuracy level (NIR). The specificity was 0.995 and the sensitivity was 0.982. Garcia, Vilella e Rizzo (2013) performed a hierarchical classification in the Amazon and found an accuracy of 93% for the classification between forest, semi-forest, and water.

At Level 2, four classes were evaluated: water (W), secondary vegetation (SV), pasture (P), and bare soil (BS). The overall accuracy of the model was 89.7% and the Kappa coefficient was 0.86, indicating an excellent level of agreement. For this level, McNemar's test was not applicable (NA), but the p-value [Acc > NIR] was less than 2.2×10^{-16}, which shows that the performance was better than the NIR index. The confusion matrix and class accuracies are shown in Table 1. Garcia, Vilella e Rizzo (2013) carried out a hierarchical classification in the

Amazon and found an accuracy of 86% for the classification of Forest, Savannah Formation, Flooded Areas, Secondary Vegetation, Agriculture, and Bare Soil.

	Water	Secondary vegetation	Pasture	Bare soil	Total	Commission error	Omission error
Water	197 (100%)	0	0	0	197	0	0
Secondary vegetation	0	183 (91%)	12	6	201	0.05	0.09
Pasture	0	2	44 (80%)	9	55	0.02	0.38
Bare soil	0	16	15	99 (76%)	130	0.06	0.13
Total	197	201	71	114	583	-	-

Table 1. Level 2 confusion matrix of land use land cover classes.

In terms of temporal transferability of the classification model, the validation carried out with samples obtained in 2019 revealed that the method was efficient, even with a difference of five years. The confusion matrix indicated an accuracy of 89%. The model performed worse when classifying pastures and bare soil, with commission and omission errors of 20% and 5%, and 12% and 6%, respectively. These results demonstrate high accuracy in classifying F and NF areas, and excellent performance in distinguishing between the types of land cover investigated in this study.

3.2. Analysis of land use and land cover dynamics

SV areas changed substantially, starting at 3,795 hectares, in 2019, and increasing to 4,258 hectares, in 2023. Similarly, the BS class, used in this study as an indicator of smallholder agricultural production, was 594 hectares, in 2019, and nearly doubled to 1,077 hectares, in 2023. Forest, pasture, and water areas remained relatively stable over the period of analysis (Figure 3).

In general, the F class showed little variation in area, while SV showed an increase in the three studied areas. The area of SV in study area 1 decreased from 1,939 ha, in 2019, to 1317 ha, in 2021, increased to 2,872 ha, in 2022, and had an area of 2,320 ha, in 2023. In study area 2, the area of SV was 1,335 ha, in 2019, decreasing to 1,167 ha, in 2021, and was 1,404 ha, in 2023. With a similar pattern of variation, the study area 3 showed 524 ha, decreased to 472 ha in 2021, and increased to 534 ha in 2023. The BS class showed a similar pattern in the three study areas. In study area 1, it increased from 333 ha, in 2019, to 451 ha, in 2020, decreased to 305 ha in 2021 and increased to 659 ha in 2023. In the study area 2, it increased from 182 ha to 306 ha, between 2019 and 2020, decreased to 270 ha in 2022 and increased to 278 in 2023. In the study area 3, it increased from 79 ha, in 2019, to 94 ha, in 2021, and increased again to 140 ha in 2023. P and W classes showed stable dynamics and no major changes (Figure 4).

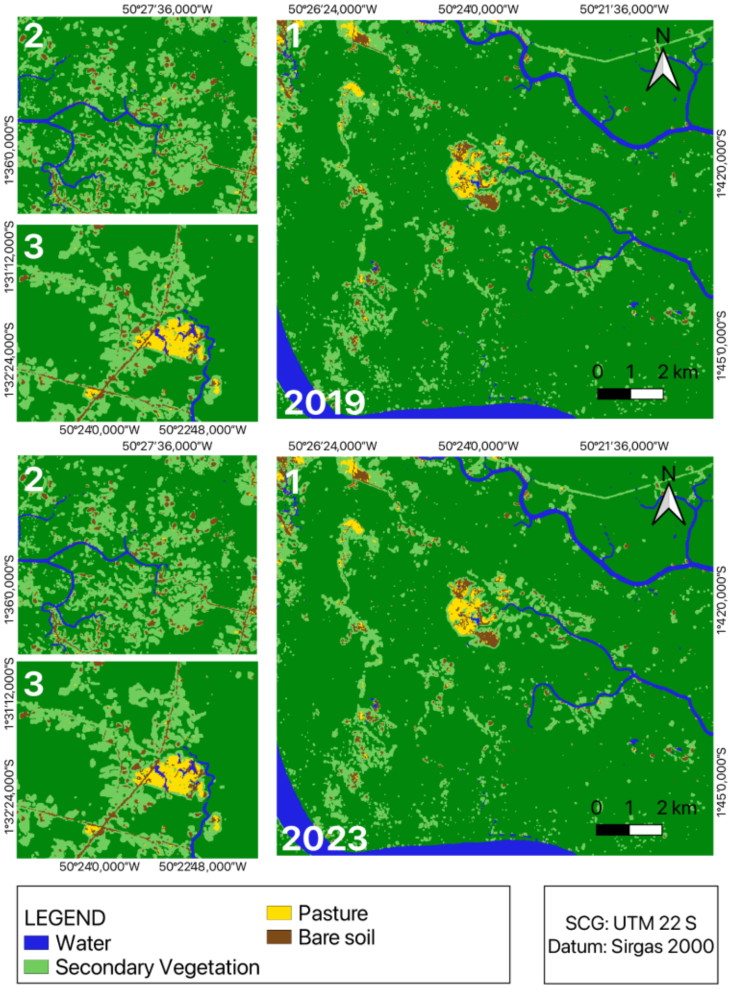

Figure 3. Land use and land cover in 2019 and 2023 in the study areas located in Breves municipality, Pará State.

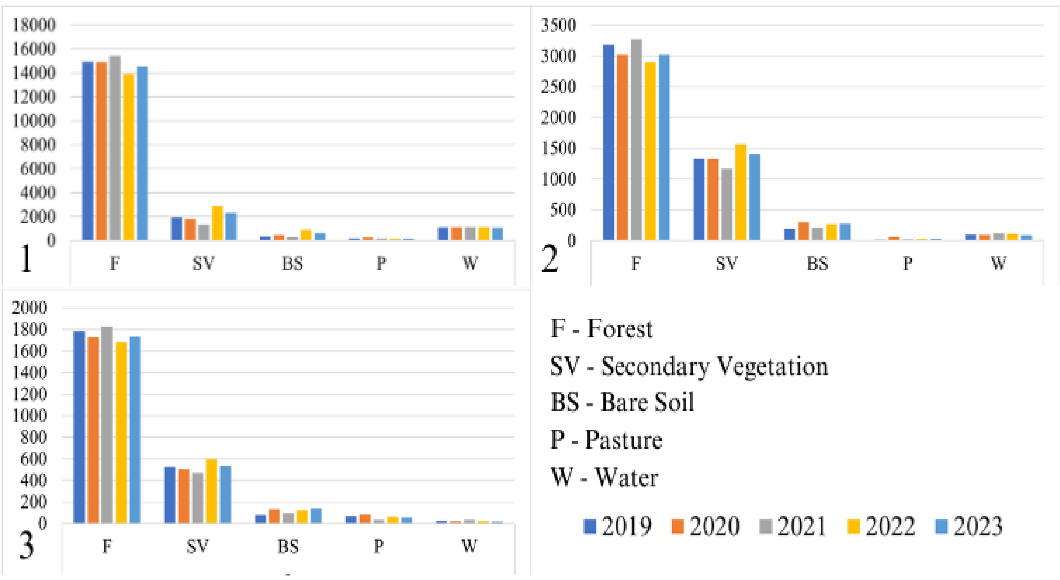

Figure 4. Land use land cover classes areas in the analysis period, in hectares

Figure 5 shows the results for the three areas of interest using a Sankey graph. This type of graph makes the visualization and quantification of areas of transition between classes easier. In the three areas analyzed, there was a conversion of forest areas into secondary vegetation areas, mainly between 2019 and 2020 and in the period between 2021 and 2022. This may be indicative of logging, one of the most important economic activities for the municipality (Alves; Oliveira e Costa, 2022; IBGE, 2017b). In all the transitions between years, it was possible to see the conversion of areas of SV into areas of BS, indicating the use of previously anthropized areas by smallholder agricultural production.

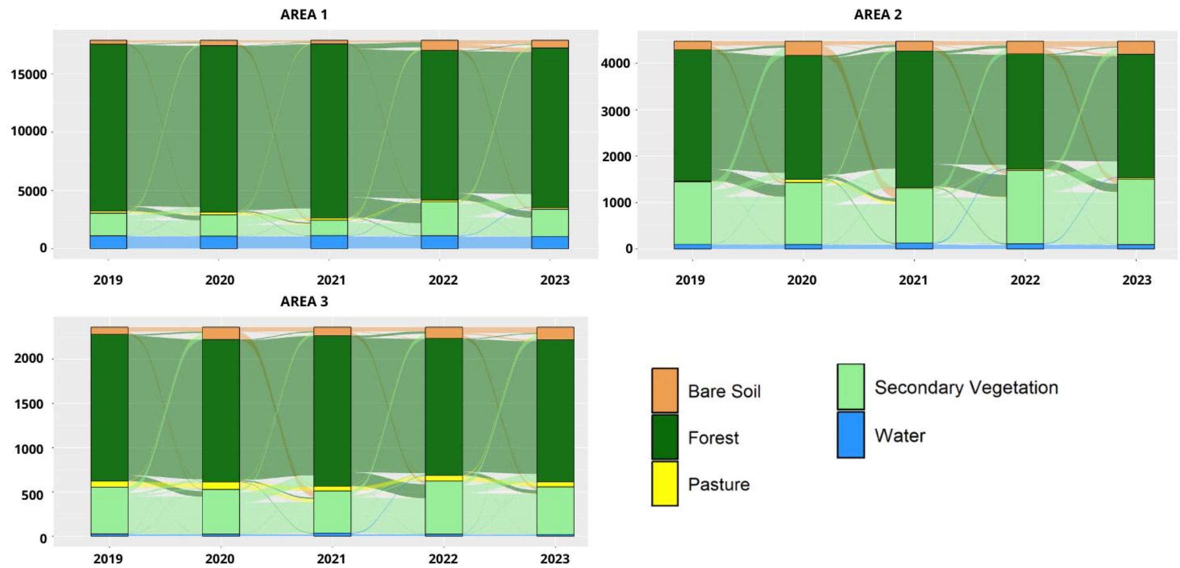

Figure 5. Land use land cover dynamic transition areas in the analysis period, area in hectares.

The transition area analysis for BS in the study area 1 showed that only between 2021 and 2022 the areas of bare soil areas were F (55.9%). In the other years, most BS areas in the current year were anthropogenic areas in the previous year, 74.4%, 92.6% and 92.4% in 2020, 2021 and 2023, respectively. In study area 2, the BS areas were not converted mainly from forest areas, but were predominantly converted from previously anthropized areas, 76.2%, 78.5%, 93.3%, and 83.4% in the years 2020, 2021, 2022, and 2023. In the study area 3, there was also no predominance of conversion of F areas to BS as in study area 2, and the BS areas were predominantly converted from previously anthropized areas, 90.1%, 90.6%, 90.6%, and 98.1% in the years 2020, 2021, 2022, and 2023.

4. CONCLUSIONS

The analysis conducted in this study, using Sentinel-2 images, yielded promising results in distinguishing the indicated classes. At Level 1, the model achieved an overall accuracy of 99% and a Kappa coefficient of 0.98. At Level 2, the accuracy was 89.7% and the Kappa coefficient was 0.86 when distinguishing between the classes of water, secondary vegetation, pasture, and bare soil. These results indicated a notable agreement between the model's predictions and the observed values in practice. The methodology used can contribute to a better characterization of the land use and land cover and to the methodological improvement in mapping the dynamics of smallholder agricultural production in Breves, Pará State.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to the National Council for Scientific and Technological Development (CNPq) for the opportunity provided through the scientific initiation scholarship. Special thanks to Embrapa Agricultura Digital for providing all the necessary facilities to conduct this study. We also acknowledge the São Paulo State Research Foundation (FAPESP), Grant number 2022/09319-9 (Science Center for the Development in Digital Agriculture), and the Coordination for the Improvement of Higher Education Personnel (CAPES) - Funding Code 001.

REFERENCES

ACHANTA, R.; SÜSSTRUNK, S. **Superpixels and polygons using simple non-iterative clustering**. IEEE Conference on Computer Vision and Pattern Recognition, p. 4895-4904, July 2017. Available in: https://ieeexplore.ieee.org/document/8100003. Access: 10 jun 2024.

ALTIERI, M. Agroecologia: conceitos e princípios. Porto Alegre: Editora do UFRGS. 2004. ISBN 85-7025-538-1.

ALVES, L. S. C.; OLIVEIRA, S. A.; COSTA, A. S. **Evaluation of environmental impacts caused by disabled loggers in the municipality of Breves – PA**. Brazilian Journal of Animal and Environmental Research, v. 5, n. 4, p. 3647-3665, 2022. DOI:10.34188/bjaerv5n4-017.

BOLFE, E.; JORGE, L.; SANCHES, I. **Precision and digital agriculture: adoption of technologies and perception of Brazilian farmers.** Agriculture, v. 10, n. 12, 653, 2020. DOI: 10.3390/agriculture10120653.

CCD - AD SEMEAR. Centro de Ciência para o Desenvolvimento em Agricultura Digital. Embrapa - FAPESP, 2022. https://bv.fapesp.br/pt/auxilios/111242/centro-de-ciencia-para-odesenvolvimento-em-agricultura-digital-ccd-adsemear.

CONGALTON, G; GREEN, K. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices (3rd ed.). Boca Raton: CRC Press. 2019. ISBN 9780429052729.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. A real contribuição da agricultura familiar no Brasil. Embrapa, 2017. Available in: https://www.embrapa.br/agropensa/busca-de-noticias/-/noticia/27405640/a-real-contribuicao-da-agricultura-familiar-no-brasil. Acess: 15 jun. 2024.

EUROPEAN SPACE AGENCY. Sentinel online. Available in: https://sentinel.esa.int/web/sentinel/missions. Acess: 15 jun. 2024.

GARCIA, A; VILELA, V; RIZZO, R. **Assessing land use/cover dynamics and exploring drivers in the Amazon arc of deforestation through a hierarchical, multiscale and multi-temporal classification approach**. Remote Sensing Applications: Society and Environment, v. 15, 100223. 2019. DOI: 10.1016/j.rsase.2019.05.002.

GONÇALVES, D. A., S. C. MARQUES, A. L. PRIMO, F. MARTINHO, M. D. B. BORDALO & M. A. PARDAL. **Mesozooplankton biomass and copepod estimated**

production in a temperate estuary (Mondego estuary): effects of processes operating at different timescales. Zoological Studies 54, 12. 2015. DOI: 10.1186/s40555-015-0135-6.

GONZALEZ, R. C.; WOODS, R. Digital Image Processing (3rd ed). Tennessee: Pearson International Edition. 2008.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Censo Agropecuário de 2017. Brasília: IBGE, 2017a. Available in: https://censoagro2017.ibge.gov.br/ Acess:05 jun. 2024.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Histórico do município de Breves. Brasília: IBGE, 2017b. Available in: http://www.ibge.gov.br/cidadesat/topwindow.htm. Acess: 05 jun. 2024.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. O município e o seu contexto. Brasília: IBGE, 2024. Available in: http://www.ibge.gov.br. Acess: 10 jun. 2024.

MAXWELL, A; WARNER, T; FANG, F. **Implementation of machine-learning classification in remote sensing: an applied review**. International Journal of Remote Sensing, v. 39, n. 8, p. 2447-2471, 2018. DOI: 10.1080/01431161.2018.1433343.

MOHANRAJAN, N; LOGANATHAN, A; MANOHARAN, P., 2020. **Survey on Land Use/Land Cover (LU/LC) change analysis in remote sensing and GIS environment: Techniques and Challenges**. Environmental Science and Pollution Research 27, 29900–29926, 2020. DOI: 10.1007/s11356-020-09091-7.

MORAES, E. Sensoriamento Remoto e Geoprocessamento. São José dos Campos: Instituto Nacional de Pesquisas Espaciais (INPE). 2021. Available in: http://www.dsr.inpe.br/DSR/educacao/uso-escolar-sensoriamento-remoto/apresentacoes-2021-arquivos/21jul21_elisabetemoraes.pdf. Acess: 03 jun. 2024.

SILVA, L. F.; MALTEZ, M. A. P. F.; OLIVEIRA, C. E. A.; GUSMÃO, Y. J. P.; SOUZA, M. A.; NASCIMENTO, J. A. C.; OLIVEIRA, C. P.; BUENO, O. C. **Sustainability, family farming and public policies in Brazil: a literature review**. Research, Society and Development, v. 10, n. 4, e42310414220, 2021. DOI: 10.33448/rsd-v10i4.14220.

SCHMIDT, C. M. **Dificuldades, falhas e desafios da ação coletiva na agricultura familiar: um estudo na Associação de Produtores de Corumbataí do Sul-PR**. CBEO - Congresso Brasileiros de Estudos Organizados, 4. Porto Alegre, 2016. Available in: https://anaiscbeo.emnuvens.com.br/cbeo/article/view/219. Acess: 15 jun. 2024