SEMIAUTOMATIC AI-BASED REMOTE SENSING CLASSIFICATION TO DELIVER AN EU LANDMARK 2020 DEFORESTATION MAP FOR BRAZIL

Luciana S. Soler[1], Baggio Luiz de C. e Silva[1], Juliana S. Mota[1], Cassiano G. Messias[1], João.F. S. Kneipp C. Pinto[1], Deborah C. Lima[1], Raíssa Teixeira[1], Igor Perez Cunha[1], Jefferson Jesus de Souza[1], Patrícia Trindade[1], Rodrigo S. do Carmo[1], Filipe G. Dias[1], Ulisses Bezerra[1], Sérgio Dousseau[1], Magog Carvalho[1], Andressa Ferla[1] , Maurício Trancho[1], Karine R. Ferreira[1], Gilberto R. Queiroz[1], Silvana Amaral[1], Cláudio A. Almeida[1]

[1]INPE, National Institute for Space Research, Av. dos Astronautas 1754, São José dos Campos, 12227-010 Brazil
luciana.soler@inpe.br, baggio.silva@inpe.br, juliana.mota@inpe.br, cassiano.messias@inpe.br, joao.kneipp@inpe.br, deborah.lima@inpe.br, raissa.teixeira@inpe.br, igor.cunha@inpe.br, jefferson.jesus@inpe.br, patrícia.trindade@inpe.br, rod19.silva@gmail.com, filipe.dias@inpe.br, ulisses.alencar17@gmail.com, sergiodousseau@gmail.com, magogaraujo@gmail.com, andressaferla94@gmail.com, mauriciotrancho@gmail.com, karine.ferreira@inpe.br, gilberto.queiroz@inpe.br, silvana.amaral@inpe.br, claudio.almeida@inpe.br

ABSTRACT

The expertise of INPE in mapping deforestation is worldwide recognized, and it has been crucial in the development of civil agreements and public policies of deforestation control. A recent and polemic undergoing policy in the new Regulation (EU) 2023/1115 on deforestation-free products, which shall demand a fine and trustful landmark map of deforestation for every country that trades commodities within EU countries. In order to support such a policy, the Brazilian government has set an agreement that INPE shall produce a Landmark deforestation map of Brazil, based on Prodes historical series and supplemented by deforested patterns retrieved using automatized methods of artificial intelligence (AI) up until 31st December 2020. The methodology adopts non-supervised and a supervised AI-based approaches, both using segmentation strategies under distinct protocols. Pilot areas were strategically chosen in all biomes. Both methods have been successful in detecting most deforestation patterns in all biomes, but final outputs still require auditing to reach similar Prodes accuracy.

KEY WORDS — Zero deforestation, Brazilian biomes, artificial intelligence, market regulation, Prodes.

1. INTRODUCTION

Since the late 1980s deforestation in the Brazilian Amazon has been monitored by Prodes (Brazilian Annual Monitoring of Natural Vegetation Removal), using mainly Landsat images. In 2024, Prodes has reached a milestone as the current methodology now adopts Sentinel-2 images organized in data cubes used as input to I.A. time series classification methods . Despite Prodes Amazon history has set 6,25 ha as the minimum mapped area, and also as the threshold adopted to calculate annual rates and total increments, in practical terms this threshold is 1 ha, being such data available to Prodes and all other Brazilian biomes [1]. In addition, Deter (Daily Monitoring of Suppression and Degradation of Native Vegetation) monitors the region daily with a minimum area mapped of 3 ha. Data available after 2016 adopts images of 60 m thanks to WFI and AWFI sensors on board of CBERS 4, CBERS 4A, and Amazonia-1 satellites. Deter has also been expanded to Pantanal biome and to non-forest areas in the Brazilian Amazon [2,3].

The competence and long-term expertise of INPE (National Institute for Space Research) in mapping deforestation is worldwide recognized, and it has been crucial in dampening deforestation rates more than four times from 2004 to 2010 [4,5]. Despite the denial of Prodes deforestation rates by some Brazilian governments in the past, not by chance, in 2007 an article in Science stated that PRODES was “the envy of the world” [6]. Official deforestation data from Prodes and DETER have been consolidated also as referenced data to run zero deforestation and market regulation policies. Key policies implemented adopting Prodes as the ground truth for deforestation mapping are the Soy Moratorium and the Zero-Deforestation Cattle Agreements [7,8]. Both initiatives were extended to the Cerrado biome where cattle ranchers are voluntarily engaging certified markets stamping free from deforestation products, so they can keep governmental subsidies.

In this context, on May 31st, 2023 the European Parliament and the Council of the European Union deliberated the new Regulation (EU) 2023/1115 on deforestation-free products [9] which is expected to bring down greenhouse gas emissions and biodiversity loss. As the abovementioned Brazilian policies, this new EU Regulation shall demand a fine and trustful Landmark map of deforestation for every country that trades commodities within EU countries. In the case of Brazil, after disputing

opinions among delegates from the EU and the Brazilian Ministry of Agriculture, Livestock and Food Supply (MAPA) regarding why JRC maps (https://www.wri.org/initiatives/global-forest-watch) had been chosen instead of Prodes maps, an overall agreement was set. Therefore, it has been decided that a Landmark deforestation map of Brazil, based on Prodes would be chosen (EU Landmark 2020), but EU delegates required INPE to fulfill the observable deforested patterns that may occur outside the Prodes calendar up until December 2020.

This work presents the under construction methodology adopted to build the EU Landmark 2020, a map of deforestation to attend the Regulation (EU) 2023/1115 on deforestation-free products and preliminary results in pilot areas. The term deforestation here also encloses non-forest natural vegetation removal that is currently detected by Prodes in all Brazilian biomes.

2. MATERIAL AND METHODS

The EU Landmark 2020 will adopt Sentinel-2/MSI mosaics and time series from Analysis Ready Data (ARDs) retrieved from the Brazil Data Cube (BDC) [10]. Its minimum mapped area is 0,5 ha, and the Prodes mask is adopted as a true and only the remaining areas are classified. The first tests are being performed in 19 pilot areas of all Brazilian biomes (Figure 1).

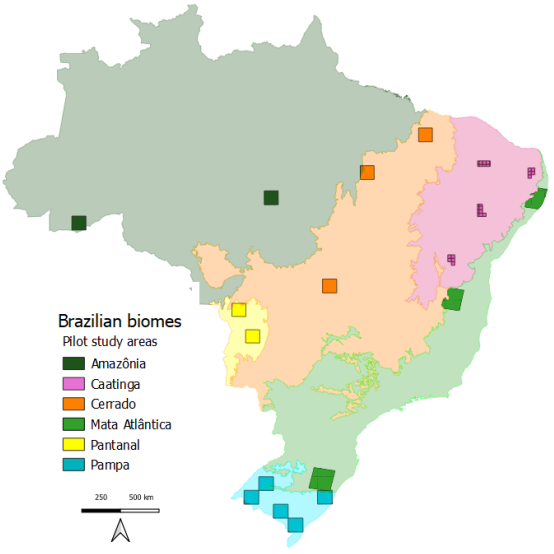


Figure 1. Study areas.

The pilot areas were selected considering two basic aspects: their representativeness of deforested patches and the level of complexity to distinguish these patterns. This was based on Prodes experience that label each scene/tile according to the time and visual interpretation difficulties in accomplishing the mapping. Figure 1 illustrates the location and dimensions of all pilot areas.

2.1. Challenges and AI-based approaches

The methodology adopted non-supervised methods in all pilot areas. Because this method has already been tested to large areas in Cerrado to distinguish secondary vegetation from natural pasture, it was here tested to all biomes. The supervised approach was adopted in some areas, but as the use of time series had not yet been implemented, it was considered a complementary method to the overall building of the EU Landmark 2020. The biggest challenges are mostly two: 1) The non-observable areas outside the monitoring calendar (i.e January-June and October-December), when deforestation still occurs but is hardly seen through cloudiness; 2) The georeferenced alignment between the >30-year Prodes historical series build upon Landsat series (30 m) and the AI classification maps based on Sentinel-2 images with 10 m of spatial resolution. In the latter, the main issues were bordering pixels of the Prodes mask that did not exactly match the new borders of Sentinel-2 AI classification maps. We then assumed that total deforestation figures based on the new classified images AI-based would ramp as observed in several Landsat vs Sentinel comparisons in the literature.

2.2. Non-supervised and supervised approaches

First, we adopted a method developed by [10] based on two unsupervised clustering algorithms: Self-Organizing Map (SOM) and agglomerative hierarchical clustering. SOM is an unsupervised learning method based on competitive learning that reduces a high-dimensional feature input space to a low-dimensional feature output space. A dataset can be mapped and represented by a set of neurons by using weight vectors. After that, hierarchical clustering is used with the help of binary trees called dendrograms, which support deciding whether clusters should be merged or split. The spectral bands and indexes of the Sentinel-2 time series adopted in the first tests were B04, B8A, B11, NDVI, NBR, NDWI, AWEINSH, and a texture image based on the average of B8A. Sentinel-2/MSI data was retrieved from the BDC. In this method, there is no need for sampling classes.

The second approach adopts image segmentation based on the Simple Non-Iterative Clustering (SNIC) segmentation method [11]. After SNIC, the Random Forest machine learning algorithm is applied, using Google Earth Engine. The innovation here is the adoption of fraction images generated by a Spectral Mixture Model (SMM) using Sentinel-2 B03, B04, and B08 bands at 10 m resolution. The method was extended from pilot areas to two ecoregions in the Amazon using a mosaic of Sentinel-2 images with tiles ranging from June 1st to 31st December 2020. In this method, samples for Random Forest are retrieved on Soil fraction image and NDVI, using the clusters resulting from SNIC

segmentation. Thus, the final classification is performed at the cluster level instead of pixel level.

2.3. Auditing procedure based on visual interpretation

In order to guarantee the current Prodes level of accuracy to the EU Landmark 2020, an auditing process through visual interpretation adopted: 1) Sentinel-2 time series from data cubes to check omitted deforestation due to gaps in the time series; 2) color composition of Sentinel-1/SAR (R-VV,G-VH,B-VV/VH) to solve remaining recurrence of clouds; 3) Landsat images from the 80´s to solve misclassifications between ancient deforestation not yet in Prodes from false targets listed in the gray lines of Table 1. Auditing also fixed border issues of Prodes mask due to historical analogic and digital data that compose the historical data.

Class name	Description
Abiotic	Rock outcrops, non-vegetated natural areas
Non-forest	Unmapped non-forest natural areas
Deforestation landmark	Deforested patterns not accounted in Prodes 2020 masks
Border	Bordering deforested pixels of Prodes mask
Flooded forest	Flooded areas misclassified as deforestation
Natural forest loss	Primary forest areas lost by natural causes

Table 1. Classes adopted in the auditing using AI-based outputs to build the EU Landmark 2020 map.

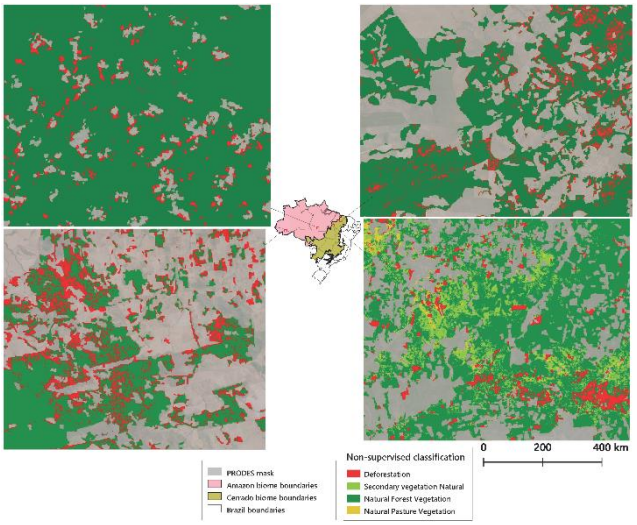


Figure 2. Tiles classified in the Brazilian Amazon and Cerrado biomes using the non-supervised approach.

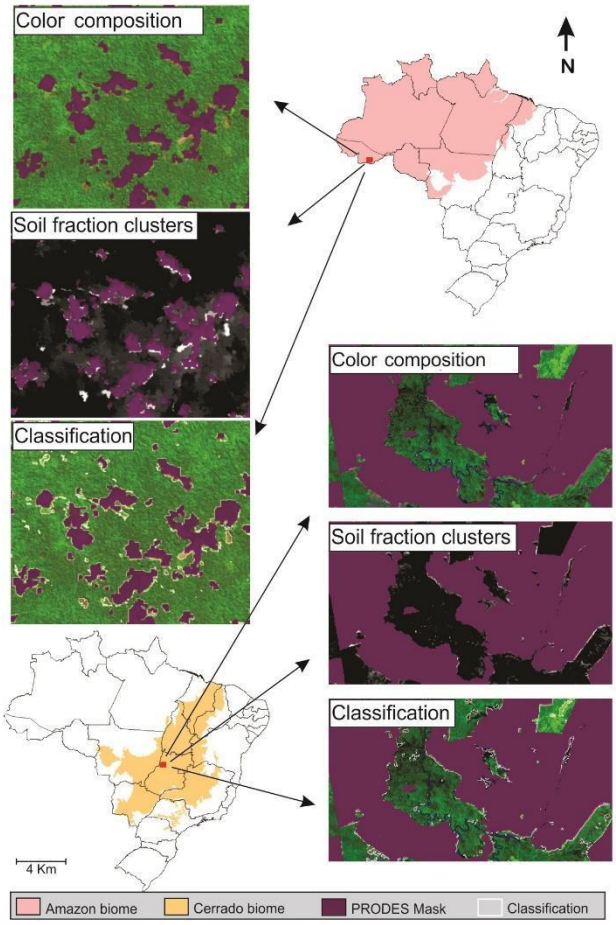


Figure 3. Tiles classified in the Brazilian Amazon biome using the supervised approach using SNIC and SSM.

3. RESULTS AND DISCUSSION

The non-supervised classification approach (Figure 2) detected deforestation from non-deforested areas with a visual accuracy close to Prodes´s. In the Cerrado biome, secondary vegetation detection has shown reasonable results, but still a quite considerable number of misclassification. This method was successful within the period not covered by the Prodes 2020 calendar, due to the temporal series of Sentinel-2 images in data cubes previously treated for cloud removal. However, the final output still requires some final auditing to match with the Prodes deforestation mask from previous years. Despite that, this approach considerably decreases the time spent by interpreters and auditors in delivering the final EU Landmark 2020 deforestation map.

Regarding the supervised classification approach, the SNIC segmentation method was very efficient in separating detailed features between deforested and non-deforested areas (Figure 3). This is likely because SNIC considers not only the distance of pixel values but also the spatial distance and texture. Another advantage of this method is the short processing time when compared to the HCA method for example. The SSM method has shown promising results in

classifying deforested borders and small polygons of deforestation seen in Sentinel-2 images of 10 m resolution, that were outside Prodes 2020 mask. We expected a boom in small deforested areas in the Amazon when using Sentinel-2 images, which did not occur in the tests performed. This was because the total deforested area was balanced between the rough trace of borders resulting from visual interpretation and Landsat images compensate for the fine trace of SNIC-SMM method using Sentinel-2 images.

4. CONCLUSIONS AND NEXT STEPS

The preliminary results are promising as they show that the majority of deforested patches are detected with a rather precise delimitation than Prodes in most cases, and the overall mapping time is indeed diminished. However, to reach overall Prodes accuracy, the method still requires quite some auditing time throughout visual interpretation. This is especially true in more complex biomes regarding vegetation seasonality and variety of forest/non-forest types such as Caatinga, Pampa, Pantanal and Cerrado biomes, i.e., in most Brazilian biomes. Therefore, a considerable number of tests are needed with extent to other regions, added to different bands and indexes combination and further analysis of time series of false targets. Considering all these, together with continuous algorithm improvement and computational intelligence, we expect to obtain more confident outputs that allow a more efficient mapping process, with less need for auditing.

The challenge to build a Landmark deforestation map to 2020 using AI-based models has also motivated BiomasBR program to reach a new paradigm of mapping the suppression of natural vegetation in Brazil. Thus, we expect that at the end of this journey INPE´s team will be ready to incorporate these methods in the regular monitoring projects to all biomes.

5. ACKNOWLEDGDMENTS

Thanks to the National Council of Technological and Scientific Development - CNPq project number 422354/2023-6 (Monitoramento e avisos de mudanças de cobertura da terra nos Biomas Brasileiros – capacitação e semiautomatização do programa BiomasBR), supported by the National Institute for Space Research (INPE). Thanks to the World Bank Group for the financial support.

6. REFERENCES

[1] Assis, L. F. F. G., Ferreira, K. R., Vinhas, L., Maurano, L., Almeida, C., Carvalho, A., Rodrigues, J., Maciel, A., Camargo, C. : TerraBrasilis: 2019. A Spatial Data Analytics Infrastructure for Large-Scale Thematic Mapping. *ISPRS Int. Journal of Geo-Information. 8(11), 513* https://doi.org/10.3390/ijgi8110513

[2] INPE - Instituto Nacional de Pesquisas Espaciais. Dados do Projeto Prodes e Deter. Plataforma TerraBrasilis. Acesso em: 31/10/2024. Disponível em: http://terrabrasilis.dpi.inpe.br/.

[3] C. G. Messias, J. F. S. K. C. Pinto, V. L. Camilotti et al. DETER Monitoring on Non-Forest Vegetation in the Brazilian Amazon. Outubro de 2024. Revista Brasileira de Cartografia, 76:2024. DOI: 10.14393/rbcv76n0a-72531. Licença: CC BY 4.0.

[4] C.G. Messias, D. Silva., M. Brasil, T. Souza, C. Almeida. Analysis of deforestation rates and their drivers in the Brazilian Legal Amazon during the last three decades. *Raega - O Espaço Geográfico em Análise*. Available online: https://revistas.ufpr.br/raega, 2011.

[5] G. Câmara. A pesquisa espacial no Brasil: 50 anos de Inpe (1961-2011). *REVISTA USP*, São Paulo, n.89, pp. 234-243, March/April 2011.

[6] H. E. Kintisch. Carbon Emissions: Improved Monitoring of Rainforests Helps Pierce Haze of Deforestation. *Science*, 316, 5824, 536–537, 2007.

[7] K. Gibbs. et al., 2015. Brazil's Soy Moratorium. *Science* 347, 377-378.doi:10.1126/science.aaa0181, 2015.

[8] M.E. Skidmore, F. Moffette, L. Rausch, M. Christie, J. Munger, H.K.Gibbs. Cattle ranchers and deforestation in the Brazilian Amazon: Production, location, and policies. Global Environmental Change, 68, 102280. doi.org/10.1016/J.GLOENVCHA.2021.102280, 2021.

[9] European Union, EU. Regulation (EU) 2023/1115 of the European Parliament and of the Council of 31 May 2023 on the making available on the Union market and the export from the Union of certain commodities and products associated with deforestation and forest degradation and repealing Regulation (EU) No 995/2010. EUR-Lex, Access to European Union law. Disponível em: https://eurlex.europa.eu/eli/reg/2023/1115/oj. Acesso em: 31/10/2024.

[10] K.R. Ferreira, G.R. Queiroz, L. Vinhas, R.F.B. Marujo., R.E.O Simoes, M.C.A. Picoli, G.Camara, R. Cartaxo, V.C.F. Gomes, L.A. Santos, and *et al*. Earth observation data cubes for Brazil: Requirements, methodology and products. *Remote Sensing*, 12(24), 2020.

[11] B.L.C. Silva, F.C. Souza, K.R. Ferreira, G.R. Queiroz, L.A. Santos. Spatiotemporal segmentation of satellite image time series using self-organizing map, *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, V-3-2022, 255–261, https://doi.org/10.5194/isprs-annals-V-3-2022-255-2022, 2022.

[12] R. Achanta, S. Süsstrunk. Superpixels and Polygons Using Simple Non-iterative Clustering, 2017 *IEEE Conference on Computer Vision and Pattern Recognition* (CVPR), Honolulu, HI, USA, pp. 4895-4904, doi: 10.1109/CVPR.2017.520, 2017.