Agriculture and Forest Plantation - Appendix

Collection 10

General coordinator
Eliseu Weber

Team
Kênia Samara Mourão Santos
Paulo Domingos Pires Teixeira Junior
Clebson Silva
Rafael Paranhos

Document Overview

This appendix details the methods used for the Agriculture and Forest Plantation products in MapBiomas Brazil Collection 10 agriculture classes land use land cover (LULC) map.

The MapBiomas organization on GitHub has repositories for all the network's initiatives and modules. The repositories for 'Agriculture' and 'Forest Plantation' for MapBiomas Brazil Collection 10 are available at:

- Agriculture: https://github.com/mapbiomas/brazil-agriculture
- Forest Plantation: https://github.com/mapbiomas/brazil-forest-plantation

1 Overview of the classification method

The mapping of Agriculture and Forest Plantation in the MapBiomas project began with the project's inception in 2015, with the challenge of producing maps quickly, efficiently, inexpensively, and annually. From Collection 1 (2015) to Collection 10 (2024), methodological and technological improvements were made to the process, leading to an increase in the number of mapped classes, as illustrated in Figure 1.

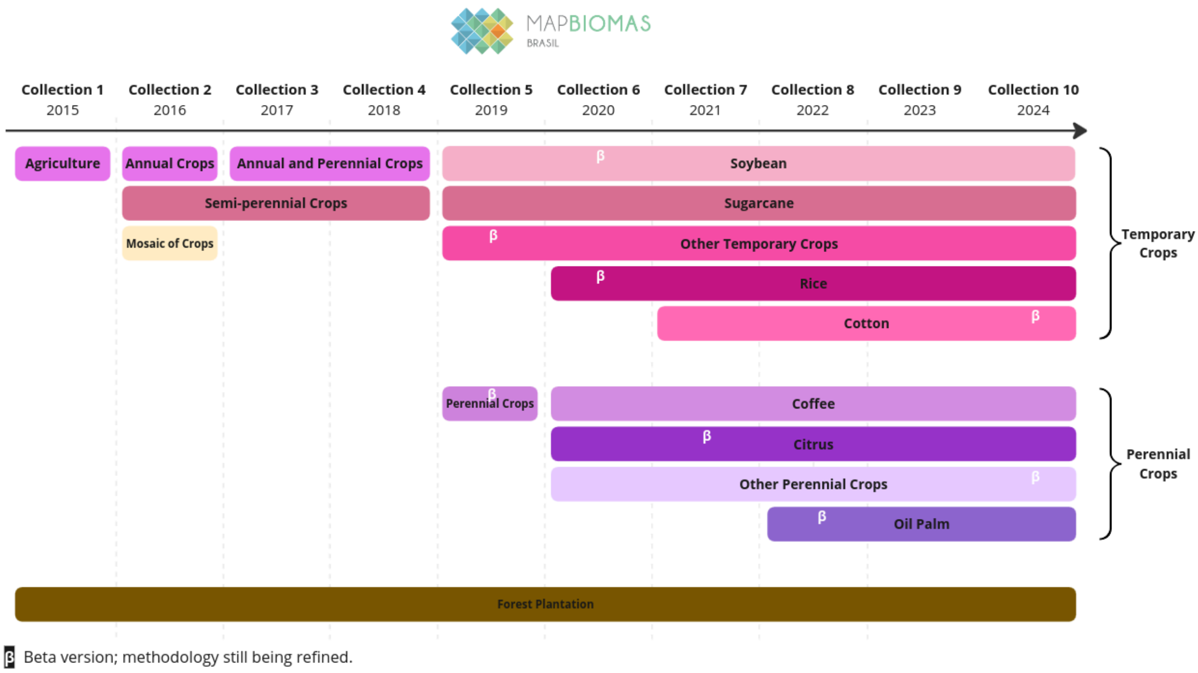


Figure 1. Evolution of mapped classes for agriculture and forest plantation.

In Collection 9, Remap Geotecnologia assumed the mapping of the aforementioned classes, as well as the Irrigated Agriculture classes and other products in the Agriculture Module. Initially, no new classes were added to the map; however, there were improvements in the process and in the quality of the maps, such as the increase in the mapped area for rice in Brazil, due to the combination of Random Forest and U-Net models.

In Collection 10, new updates were made focusing on the classes mapped using Deep Learning models, especially Oil Palm, Citrus and Rice.

2 Classification

In general, the use of supervised classification via machine learning algorithms has adopted the procedure illustrated in Figure 2.

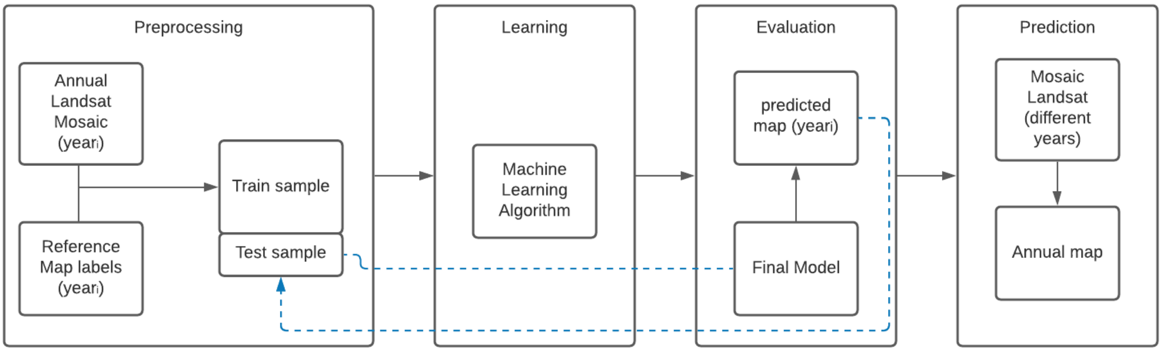


Figure 2. Supervised learning workflow in the context of image classification.

The preprocessing step and prediction were the same for both algorithms used in 'Agriculture' and 'Forest Plantation' mapping (*i.e.*, Random Forest and Convolutional Neural Network). The learning and evaluation steps were specific to each of the algorithms. The annual rice, citrus and oil palm maps were generated using a convolutional neural network (*i.e.*, U-Net), while the maps for the other classes were obtained using Random Forest.

2.1 Landsat image mosaics

As illustrated in Figure 3, the Collection 2 Landsat series offers a comprehensive array of images captured between 1985 and 2024. It is noteworthy that the launch of Landsat 8 and Landsat 9 has led to a significant increase in the number of available images, thereby enhancing the probability of obtaining cloud-free mosaics.

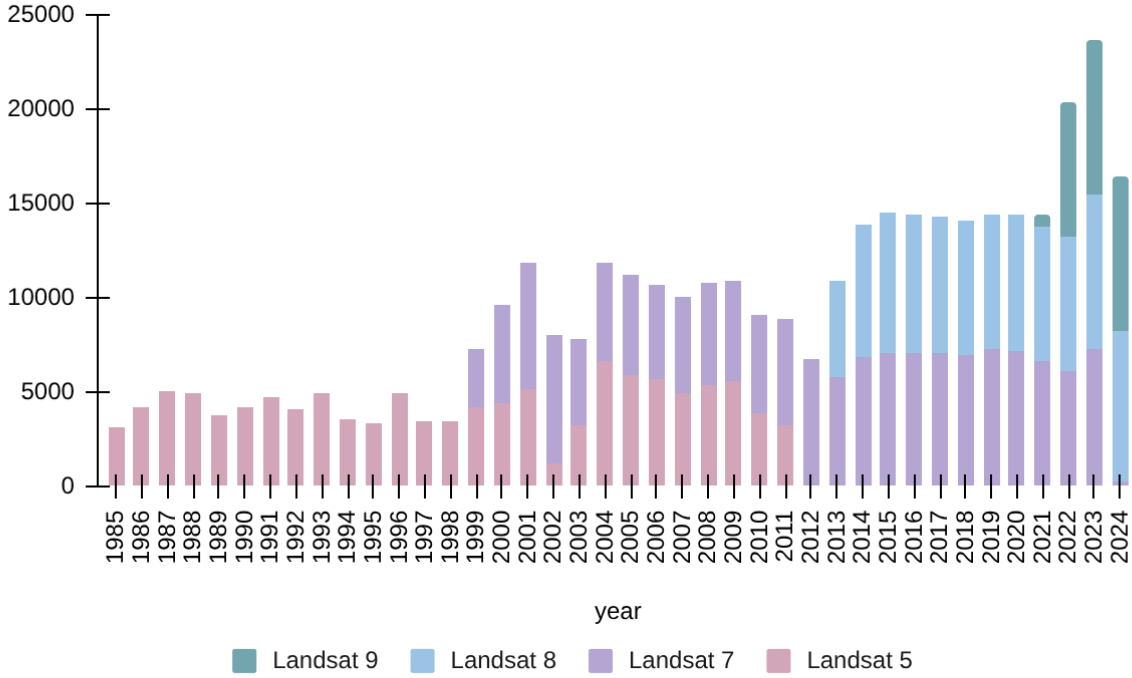


Figure 3. The number of available TOA Landsat Collection 2 images covering the Brazilian territory from 1985 to 2024.

2.2 Definition of the temporal period

To define the best period to compose the mosaics used in the supervised classification of 'Agriculture' and 'Forest Plantation', the seasonal characteristics of each agriculture class were considered to better distinguish the class of interest from the remaining land cover and land use classes. For instance, for different types of agricultural crops and for different regions in Brazil, the growing season can cover different periods of the year, predominantly during the wet season, as is the case in most regions of Brazil. According to phenological developments of the different types of cultures, we can note that for mapping annual crops, the Landsat mosaics require images that span the period from October to March. For semi-perennial and perennial crops, we can use images collected throughout most of the year or all year.

Additionally, given that the Landsat mosaics are composed of images taken during the growing season, which for most regions in Brazil coincides with the wet period and consequently has a high incidence of clouds, it may be necessary to use images from the same period in previous years or temporal filters to address the issue of missing images in the time series.

2.2.1 Cotton, Soybean, and Other Temporary Crops

In the MapBiomas project, the 'Temporary Crops' correspond to those cultivated during the summer season. Consequently, the mapping methodologies need to consider this period to collect the images to build the mosaic. In addition, it is important to note that among the crop types mapped by the MapBiomas project, the cotton class (*beta version*), mainly occurs in the 'second crop cycle', in the largest cotton-producing region in Brazil (Mato Grosso) (ABRAPA, 2020). Thus, as MapBiomas only maps agriculture classes of the first crop cycle, the cotton mapping from MapBiomas Collection 10 will reflect only the areas where cotton is the main or only crop cultivated in the crop year, not the total planted area of the cotton-producing region.

To obtain a one-by-one-year growing season calendar, a time series of EVI2 data calculated from MODIS was smoothed using the Fourier Transform to minimize variations. The full methodology used is described in Teixeira Jr. et al. (2023). Figure 4 shows an example of the methodology used to obtain the annual vegetative peaks per Landsat scene.

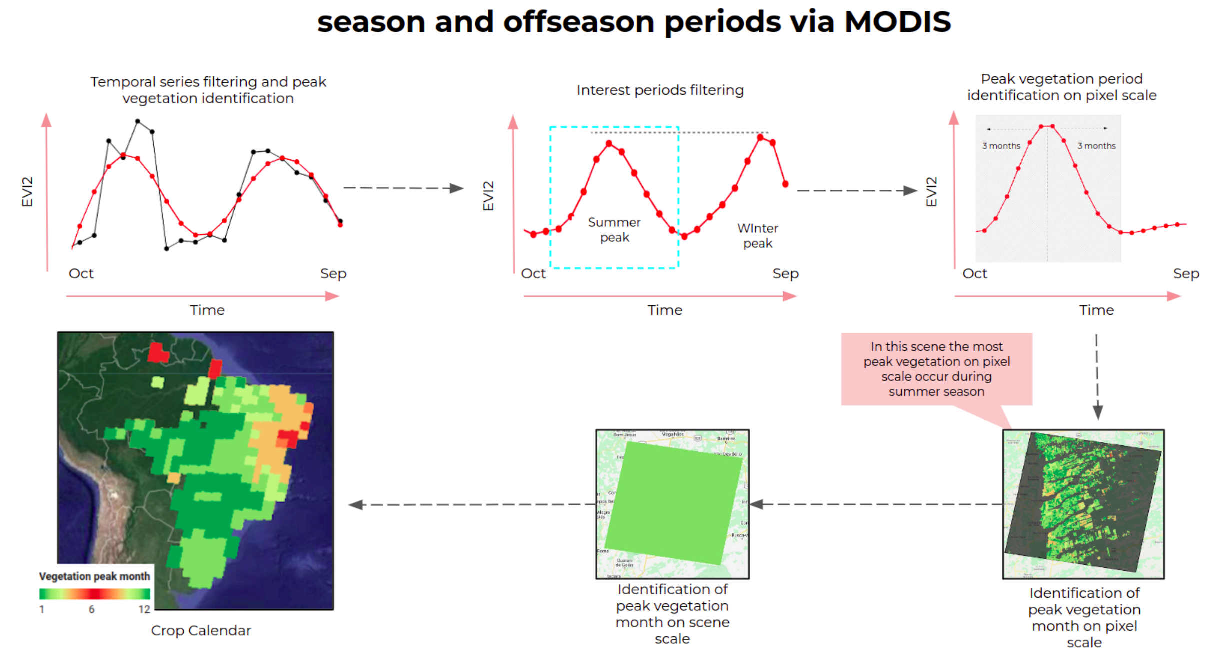


Figure 4. Scheme to obtain the vegetation peak month, year by year, for Landsat scene. Source: Teixeira Jr. et al., (2023).

Thus, the seasonal mosaics for soybean, cotton, and other temporary crops were based on the peak vegetation agricultural crop rotation month information, according to Table 2. The ‘growing season’ was defined as the period between 3 months before and 3 months after (+3/-3) the peak vegetation month, ‘off-season’ as between 5 months before and 3 months before (-5/-3), and ‘annual’ as between the peak month and 12 months after (+0/+12).

Table 2. Periods used to select the selection of mosaic images of cotton, soybean, and other temporary crops in Collection 9.

Period	Start	End
growing season	vegetation peak month -3 months	vegetation peak month -+3 months
off-season	vegetation peak month - 3 months	vegetation peak month - 5 months
annual	vegetation peak month	vegetation peak month + 12 months

2.2.2 Sugar cane

For the sugar cane class, Landsat mosaics were created to highlight intra-annual variations based on bimonthly compositions for the entire country, which were used to select the images according to the periods presented in Table 3.

Table 3. Periods used to select the mosaic images of sugar cane in Collection 10.

Period	Start	End
growing season 1	12/01/year-1	01/31/year
growing season 2	02/01/year	03/31/year
growing season 3	10/01/year	11/30/year
off-season 1	04/01/year	05/31/year
off-season 2	06/01/year	07/31/year
off-season 3	08/01/year	09/30/year

2.2.3 Rice

In Collection 10, the rice class was improved, with the addition of the new states to the map: Pará, Mato Grosso do Sul and Goiás. The selection of images was made based on the growing season period according to the year of mapping carried out in each state (Table 4).

Table 4. Periods used for the selection of mosaic images of rice in Collection 10.

State	Start growing season	End growing season	Start off-season	End off-season
Tocantins - TO	04/01/year	07/30/year	08/01/year-1	11/01/year-1
Rio Grande do Sul - RS	12/01/year-1	03/01/year	-	-
Santa Catarina - SC				
Paraná - PR	10/01/year-1	04/30/year	-	-
Goiás - GO	04/01/year	07/30/year	08/01/year-1	11/01/year-1

Pará - PA	08/01/year-1	01/30/year	05/01/year-1	08/01/year-1
Mato Grosso do Sul - MS	10/01/year-1	04/30/year	01/01/year	07/30/year

2.2.4 Perennial Crop

Since Collection 8, the 'Perennial Crops' classes were divided into four subclasses: coffee, citrus, oil palm, and other perennial crops. The last one doesn't distinguish between types of crops. For all 'Perennial Crops' classes, a median of annual mosaic (*i.e.*, 01-01-year to 12-31-year) was obtained.

2.2.5 Forest Plantation

For the Forest Plantation class, two periods were defined to compose the Landsat mosaics. These periods cover from January to January of the next year and are presented in Table 6.

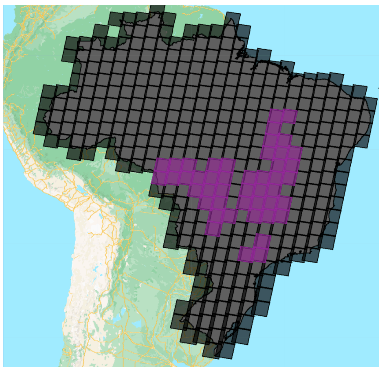
Table 6. Periods used to select mosaic images of "Forest Plantation' in Collection 10.

Period	**Start**	**End**
P1	01/01/year	07/01/year
P2	07/01/year	01/01/year+1

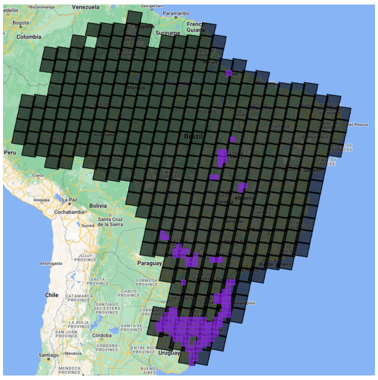
2.3 Definition of regions for classification

The 'Agriculture' and 'Forest Plantation' are heterogeneously distributed in the Brazilian biomes. Therefore, Landsat scenes were selected in regions with the highest occurrence of each class according to the reference maps. Figure 8 illustrates the scenes chosen for each land use class.

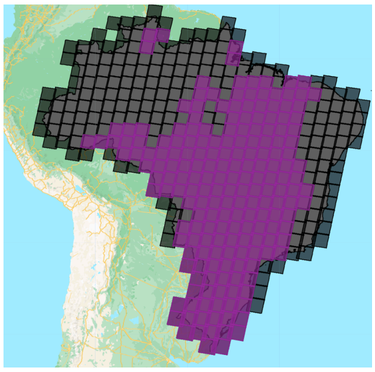
Cotton



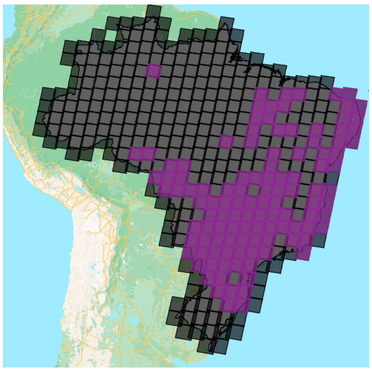
Rice



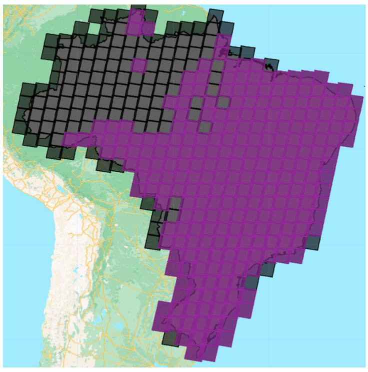
Soybean



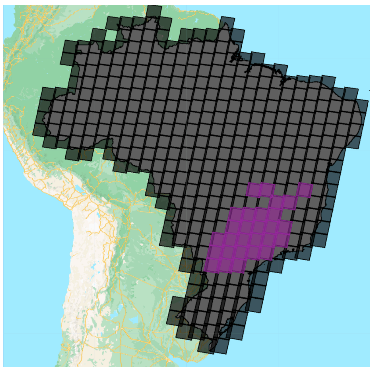
Sugar cane



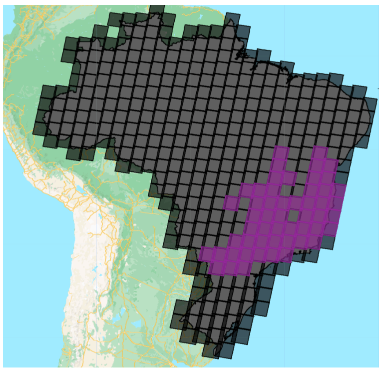
Other Temporary Crop



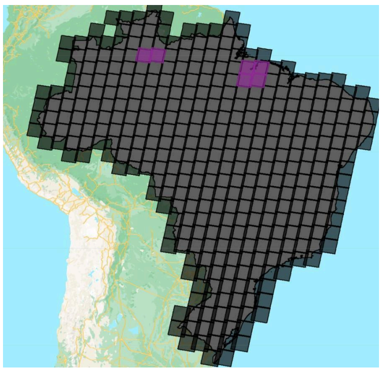
Citrus



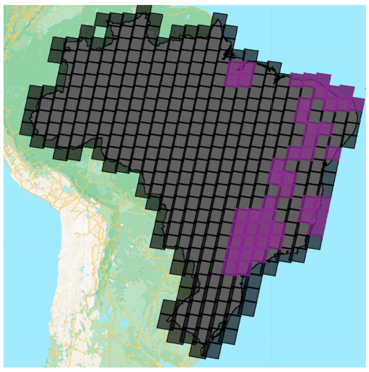
Coffee



Oil Palm



Other Perennial Crop



Forest Plantation

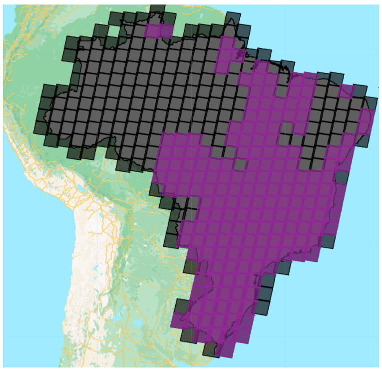


Figure 8. Selected scenes of Landsat series to classify maps by land use class.

2.4 Feature space

In Collection 10, most of the feature space was used based on the improvements made in Collection 8. Since Collection 9, the rice class has been improved with the selection of other bands to enable the map of new regions.

Table 7. Feature space and algorithm used for each class mapped in Collection 10.

Classes	Region	Bands	Indexes	Metrics	Algorithm
Cotton	as Figure 8	BLUE, GREEN, RED, NIR, SWIR1, SWIR2	SAVI, CAI, EVI2, NDWI, LAI	median, mean, max, min, stdDev, 80th percentile, 20th percentile, and CEI (NDWI, NIR, EVI2)	Random Forest
Soybean	as Figure 8	BLUE, GREEN, RED, NIR, SWIR1, SWIR2	SAVI, CAI, EVI2, NDWI, LAI	median, mean, max, min, stdDev, 80th percentile, 20th percentile, and CEI (NDWI, NIR, EVI2)	Random Forest
Other Temporary Crop	as Figure 8	BLUE, GREEN, RED, NIR, SWIR1, SWIR2	SAVI, CAI, EVI2, NDWI, LAI	median, mean, max, min, stdDev, 80th percentile, 20th percentile, and CEI (NDWI, NIR, EVI2)	Random Forest
Sugar Cane	as Figure 8	BLUE, GREEN, RED, NIR, SWIR1, SWIR2	NDVI, NDWI	median	Random Forest
Rice	Tocantins	SWIR1, SWIR2	EVI2, NDWI	CEI (EVI2), CEI (NDWI)	U-Net + Random Forest
	Santa Catarina	SWIR2	EVI2, NDWI	CEI (EVI2), CEI (NDWI)	
	Paraná	SWIR1, SWIR2	EVI2, NDWI	CEI (EVI2)	
	Rio Grande do Sul	SWIR1, SWIR2, TIR1	EVI2	CEI (EVI2)	
Coffee	as Figure 8	BLUE, GREEN, RED, NIR, SWIR1, SWIR2	EVI2, NDWI	median, mean, max, min, stdDev, 80th percentile, 20th	Random Forest

				percentile, and quality mosaic (qmo)	
Citrus	as Figure 8	RED, NIR, SWIR1	-	median	U-Net
Oil Palm	as Figure 8		-	median	U-Net
Other Perennial Crop	as Figure 8	BLUE, GREEN, RED, NIR, SWIR1, SWIR2	NDVI	median, max, min, stdDev, 20th percentile, and quality mosaic (NDVI)	Random Forest
Forest Plantation	as Figure 8	BLUE, GREEN, RED, NIR, SWIR1, SWIR2	EVI2, MNDWI, LAI	median, mean, max, min, stdDev, 80th percentile, and quality mosaic (qmo)	Random Forest

Legend: SAVI (Soil-adjusted Vegetation Index), CAI (Cellulose Absorption Index), EVI2 (Enhanced Vegetation Index 2), NIR (Near Infrared), SWIR (Short Wavelength Infrared), NDWI (Normalized Difference Water Index), LAI (Leaf Area Index), CEI (Crop Enhancement Index).

2.5 Classification algorithm, training samples and parameters

2.5.1 Reference Maps

The reference maps used to obtain samples to train the classifier are shown in Table 8.

Table 8. Reference maps used in the Random Forest classification for the classes ‘Agriculture’ and ‘Forest Plantation’ in Collection 10.

Class	Landsat time series	Number of training samples	Sampling Approach	Rule	Type	Year of acquisition	Reference
Soybean (2000 - 2022)	Normalized	10,000	Stratified	-	stable samples	2021	Song et al. (2021)
Soybean (1985 - 1999)	L5 TOA	10,000	Simple	-	stable samples	2000	Song et al. (2021)
Sugar cane	TOA	10,000		-	annual samples	2020	Terraclass
Rice	TOA	-	-	-	chips	2017-2020	Agência Nacional de

							Águas (ANA, 2020) and Companhia Nacional de Abastecimento (Conab)
Cotton	Normalized	10,000	Stratified	-	stable samples	+1/-1 year window from target year	MapBiomas Collecion 8
Other Temporary Crop	Normalized	10,000	Stratified	-	stable samples	+1/-1 year window from target year	MapBiomas Collecion 8
Coffee	Normalized	10,000	Stratified		stable samples	2015, 2016, 2017, 2018, 2019	Companhia Nacional de Abastecimento (Conab)
Citrus	TOA	-	-	-	chips	2020	MapBiomas
Oil Palm	TOA	-	-	-	chips	2020	-
Other Perennial Crop	Normalized	5,000	-	-	-	2016	Quarta comunicação nacional do Brasil à UNFCCC
Forest Plantation	Normalized	10,000	Stratified	-	stable samples	2012 - 2014	Global Forest Watch, Transparent World (2015)

2.5.2 Random Forest

As shown in Table 8, for the classes mapped by the Random Forest algorithm (Breiman, 2001), the process steps are: a) Initially, an annual Landsat mosaic is created, according to the period of the year (i.e. growing season or off-season), specific for each class; b) bands are composed with specific metrics for each class; c) simple or stratified random sampling is performed based on the reference map; d) the samples are used to train the classifier; e) classify the classes of interest. The results of this process are annual maps of interest classes. To reduce noise and inconsistencies, the maps obtained after the classification undergo spatial and temporal post-processing and are then integrated into the other MapBiomas themes. An important observation is that the annual mosaic used in the training process must be from the same year as the reference map. An example of Random Forest classification is presented in Figure 9.

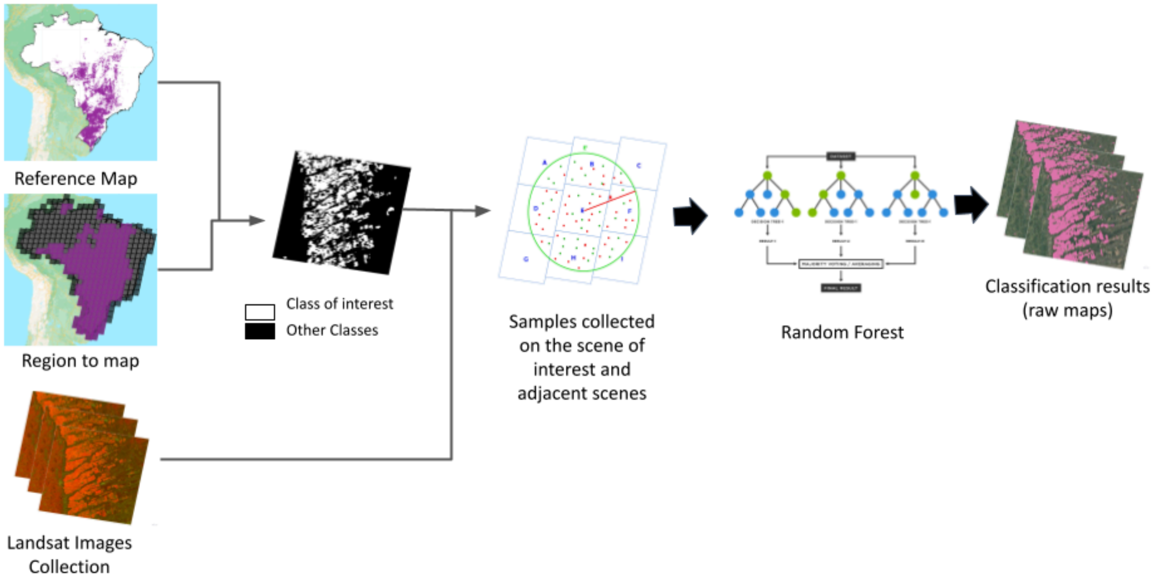


Figure 9. Flowchart of agriculture and Forest Plantation classification using the Random Forest algorithm.

The classes mapped using the Random Forest algorithm were Soybean, Sugar Cane, Cotton, Rice, Other Temporary Crops, Coffee, Other Perennial Crops, and Forest Plantation. All models were trained with 100 trees and the default values for the other parameters.

2.5.2.1 Simple Sampling

The acquisition of training samples was performed by each Landsat scene. In addition to the samples collected in the target scenes, samples from adjacent scenes (blue contour) were included inside a buffer (red fill contour), in which the center of that radius corresponds to the center of the target scene (red), as shown in Figure 10.

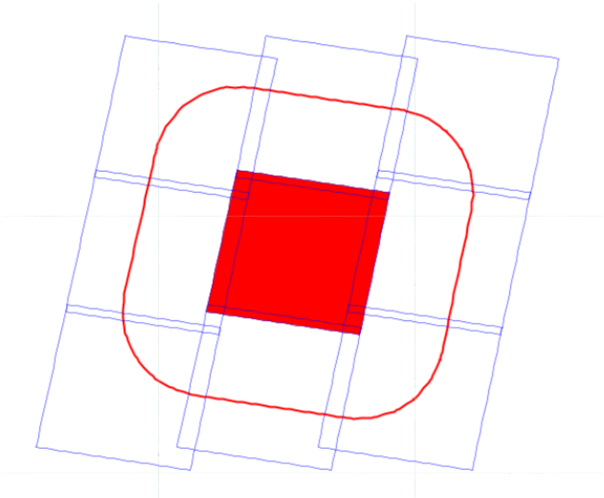


Figure 10. Scheme for sample acquisition for the regionalized training of the Random Forest classifier in 'Agriculture' and 'Forest Plantation'.

The classification result based on a reference map was used to support the subsequent training and classification procedure of previous years up to the year with the available reference map. The Random Forest training scheme was then used to classify the subsequent years for which a reference map was unavailable.

2.5.2.2 Stratified Sampling

The quality of training samples has been pointed out as one of the ways to increase the accuracy of remote sensing image classifications, as well as the algorithm's performance and accurate input data (LI et al., 2021; ZHU et al., 2016). Sampling methods commonly used for supervised classification (such as simple sampling) may often not consider the spatial distribution of the targets of interest in a scene, resulting in unbalanced samples between classes. Thus, a stratified sampling approach aimed to balance the sample distribution between the interest and non-interest targets (LI et al., 2021).

The difference between stratified and simple sampling is that, in the traditional method (simple sampling), the training samples are randomly distributed, considering the whole Landsat scene boundary. In the stratified sampling, the number of training samples is distributed according to the percentage area of the class of interest, resulting in a balanced distribution of samples as shown in Figure 11.

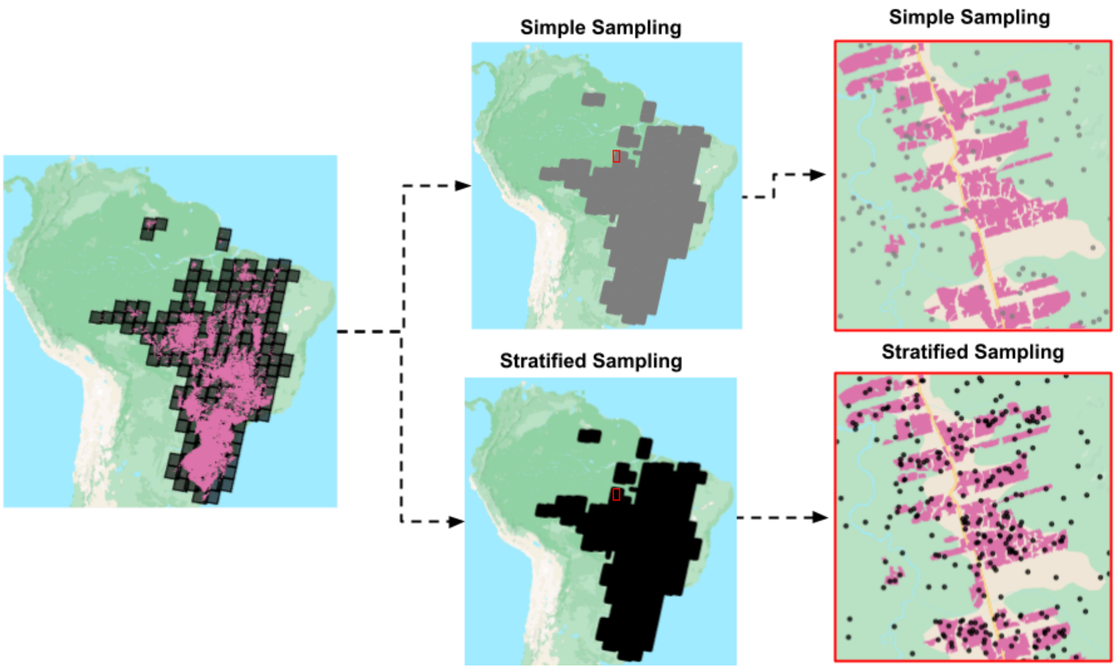


Figure 11 - Difference between sampling approaches. a) simple sampling, and b) stratified sampling.

2.6 Deep Learning

For the mapping of Rice, Citrus and Palm Oil classes, an adaptation of the U-Net convolutional neural network (RONNEBERGER et al., 2015) was used. Unlike machine learning algorithms that classify each pixel based on its spectral response, this architecture considers the context of the pixels. This architecture is illustrated in Figure 12.

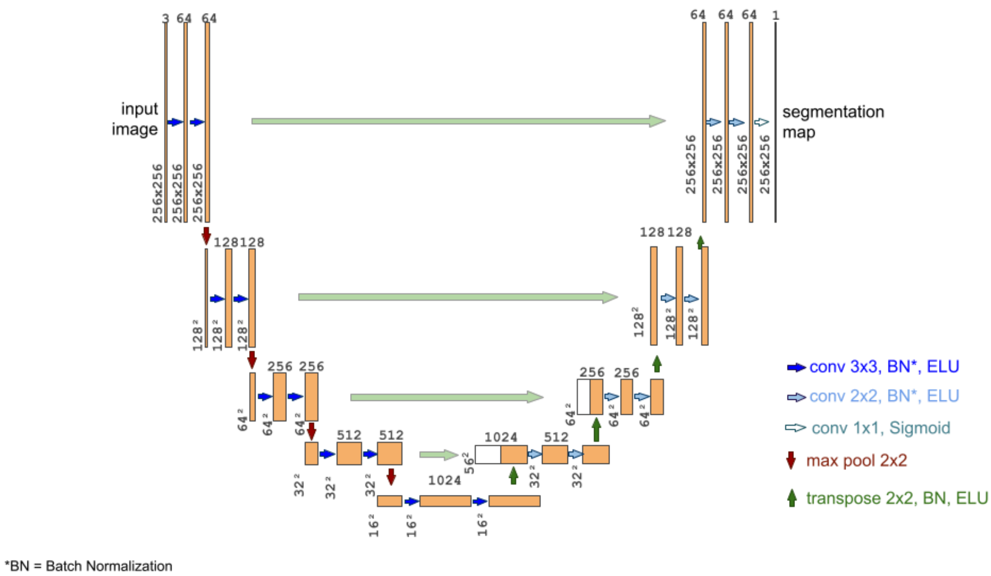


Figure 12 - Adapted U-Net convolutional neural network, with its layers and connections, used for the mapping of rice and citrus.

For Collection 10, the preparation of the images for training the model was updated. Specifically for the Rice, Palm Oil and Citrus classes, a new normalization procedure based on reference areas was applied to the period Landsat mosaics (Table 4) used for training. This method primarily aims to standardize the spectral values, attenuate information outside its range, and ensure a consistent representation of the spectral distribution of the class of interest in the 0 to 1 range.

The normalization process involves the following special steps:

- Defining essential parameters, such as the year of analysis, the date period for the image search, the maximum limit of cloud cover allowed, the output spatial resolution, and the percentiles (e.g., 1% and 99%) that will be used for normalizing the spectral values.
- Uploading the necessary geographic data, including the reference areas (polygons representing the class of interest), and creating a binary image (raster) from these areas, where the corresponding pixels are assigned a value of 1 and the rest 0.
- Processing and compositing of collections of Landsat images (from satellites 5, 7, 8 and 9), that are filtered by period, location and cloud boundary. Auxiliary functions are applied to mask unwanted pixels (such as clouds, shadows, snow and water) and standardize spectral band names. Then, a median composition of the filtered image collection is created to reduce the impact of noise or masked pixels.
- The crucial normalization step is applied: the percentile values (lower and upper) are calculated exclusively from the pixels located within the reference areas in the composite image. Next, the values of all the pixels in the selected spectral bands are scaled to the range 0 to 1, where the lower percentile value is mapped to 0 and the upper percentile value is mapped to 1. Values outside this range are limited to 0 or 1.

After preparing and normalizing the period mosaics, training data were generated from them, consisting of mosaic blocks and the reference map masks for each class.

2.6.1 Rice

The delimitation of the mapping area was based on the map of irrigated rice in Brazil published by the National Water Agency (ANA) and the National Supply Company (Conab) in 2020. Images were selected based on the growing season according to the year of mapping carried out in each state. The reference map was divided into blocks of 0.5 x 0.5 degrees (~300 thousand ha each). The blocks used for rice mapping and training were those that overlapped the reference map and the states of interest, as illustrated in Figure 13.

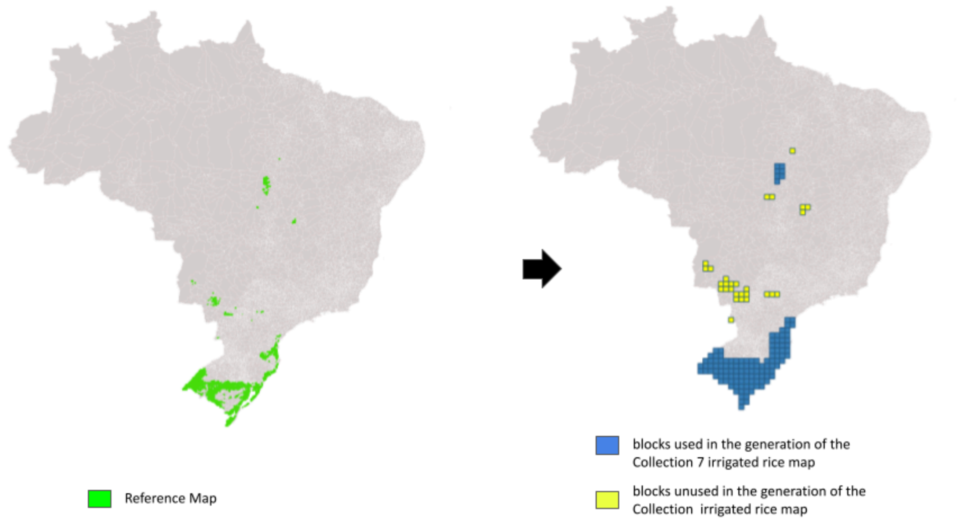


Figure 13. Study area used for the mapping of irrigated rice in the MapBiomas Project.

From the reference map and the annual Landsat mosaics, training samples were created, consisting of pairs of blocks of the annual mosaic (from the reference year) and in the mask of the reference map for this same block. A sample U-Net entry training example is shown in Figure 14.

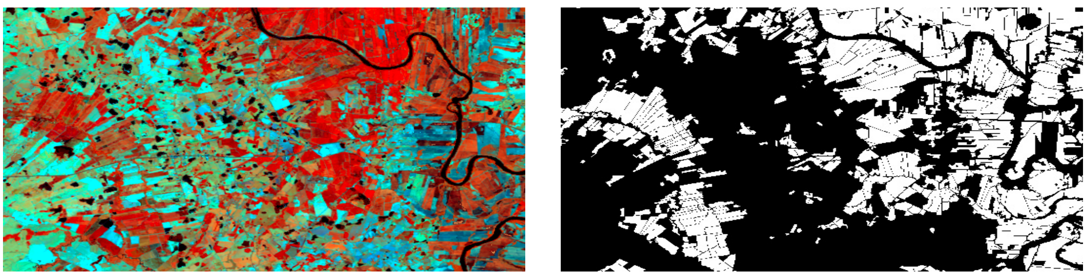


Figure 14. Example of U-Net samples for mapping rice.

The test data were used for the accuracy analysis of the trained model. The final model (*i.e.*, the one with the best results) was used in the process of classification of irrigated rice in different states for each year of the series (1985-2024).

2.6.2 Citrus

The Citrus map was performed, similar to rice, using a neural network based on the U-Net architecture. Reference data for training were generated by visual interpretation of Sentinel and Landsat images for the year 2020.

2.6.3 Oil Palm

The Oil Palm map was performed, similar to rice, using a neural network based on the U-Net architecture. Reference data for training were generated by visual interpretation of Sentinel and Landsat images for the year 2020.

3 Post-classification

Temporal and spatial filters were applied to remove noise and classification errors.

3.1 Spatial filter

The minimum connected pixels filter was applied to most classes, except those mapped with U-Net, because the result of the semantic segmentation showed little or no spatial noise. This spatial filter removed groups of 6 or fewer pixels from the class of interest or the "others" class (Figure 15).

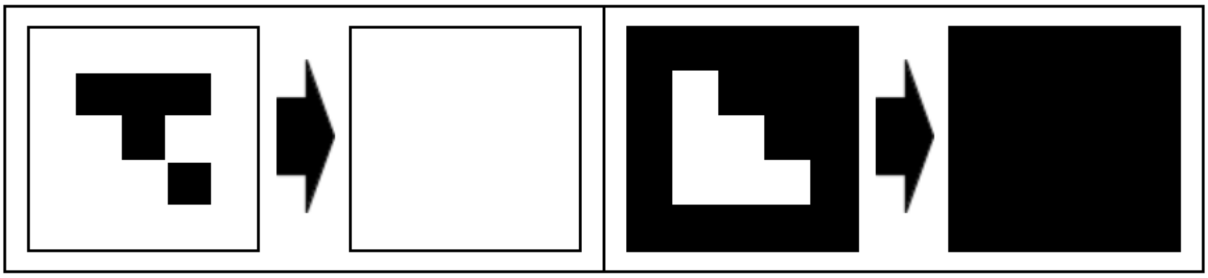


Figure 15. Example of the minimum connected pixels spatial filter. The image on the left shows an exclusion of pixels of the interest class (in black). The image on the right shows an inclusion of pixels of "other class" (in white) to the interest class.

3.2 Temporal filter

In general, two temporal window filters were applied: 3 years with 2 years threshold or 5 years with 3 years threshold. The 3-year window excludes the center year when neither of the adjacent years are of interest class and includes the center year when both adjacent years are of interest class (Figure 16). The 5-year window excludes the center year when no more than 2 years are of interest class and includes when at least 3 adjacent years are of interest class (Figure 17).

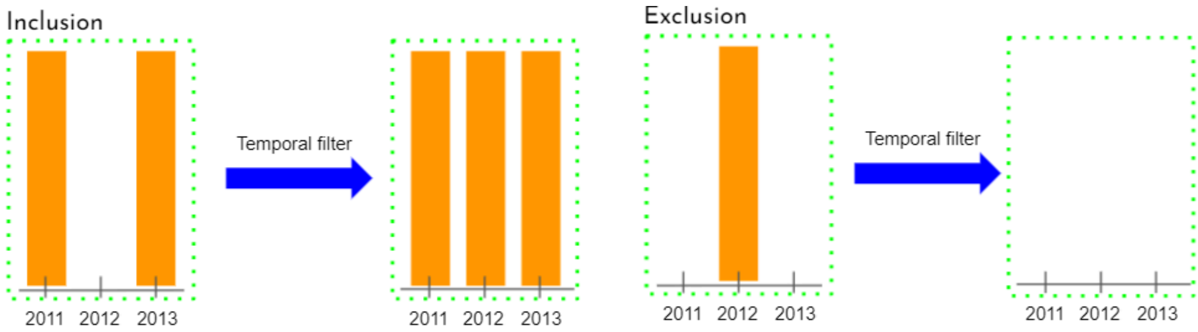


Figure 16. 3-year temporal window filter: The orange bars represent pixels of the mapped class (interest class). The exclusion filter changes a pixel to the "others" class when the same pixel was not of interest class in the adjacent years. The inclusion filter changes a pixel to the interest class when the same pixel was of the interest class in the adjacent years.

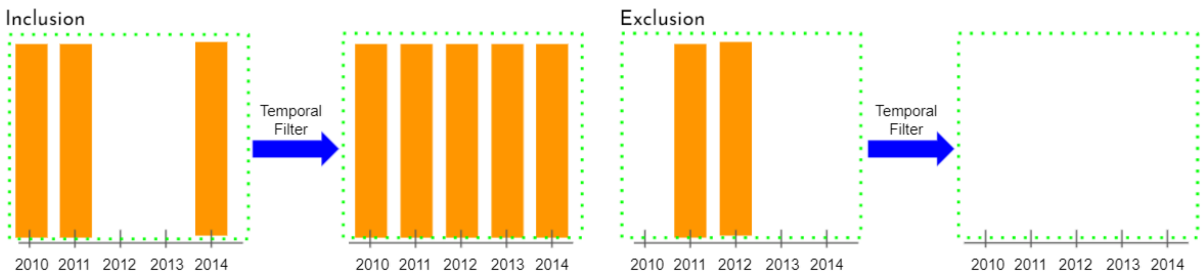


Figure 17. 5-year temporal window filter: The orange bars represent pixels of the mapped class (interest class). The exclusion filter changes a pixel to the "others" class when no more than 1 year is of the interest class. The inclusion filter changes a pixel to the interest class when at least 3 adjacent years are of the interest class.

In addition, for all agriculture classes, the first year of the series (*i.e.,* 1985), pixels were excluded when, in the following year, they were not classified, and included when, in the following year, they were. No temporal filter was applied to the last year of the time series (*i.e.,* 2024).

3.2.1 Soybean, Cotton, and Other Temporary Crops

For MapBiomas Collection 10, the soybean, cotton, and Other Temporary Crops classes followed a unified process (Figure 18).

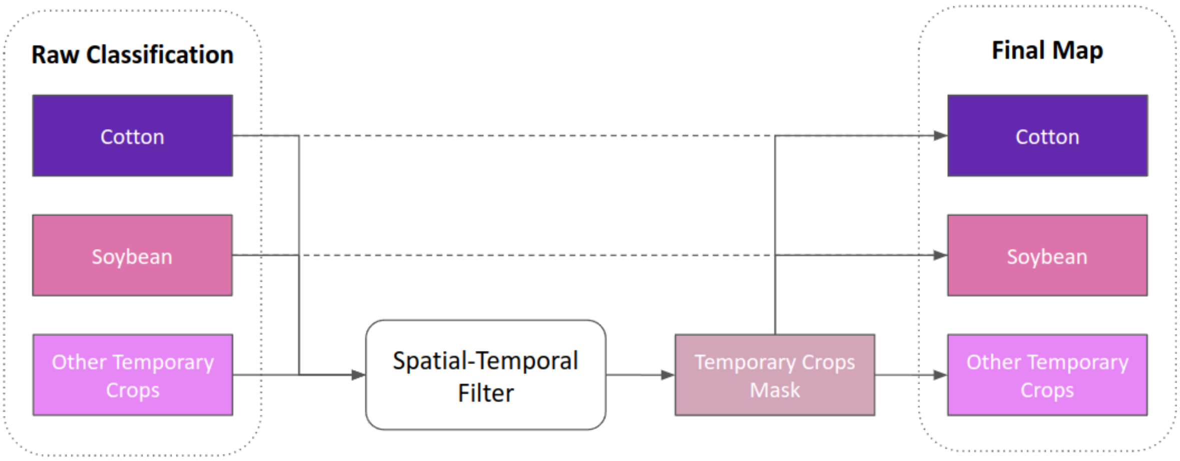


Figure 18: Soybean, cotton, and Other Temporary Crops classification and post-classification Flowchart.

The classification was initially performed with distinct classes for cotton, soybean, and other temporary crops based on a reference map. However, during the temporal filtering stage, these classes were aggregated and treated as a single 'Temporary Crops' class. The temporal filter was applied using a three-year moving window with a minimum threshold of two years.

The filtered result was an annual mask that indicated the area used for ‘Temporary Crops’ in general. As a final step, the raw classifications of cotton and soybean were masked by the ‘Temporary Crops’ mask, resulting in the final maps for those classes. The remaining area in the ‘Temporary Crops’ mask was considered as Other Temporary Crops. This method maintained the temporal stability of the ‘Temporary Crops’ areas, while the annual crop variation during the growing season period was preserved.

3.2.2 Rice

The filters for the rice class were regionalized, based on knowledge of the different regional dynamics of this land use. For all of Brazil, a temporal filter with a 5-year window and 3-year threshold was applied. A second temporal filter with a 3-year window and 2-years threshold was applied in all regions with the exception of southern Brazil. This exception is due to the known agricultural dynamics in rice-producing areas of southern Brazil, which often involve crop rotation and would otherwise be omitted..

3.2.4 Sugar cane

For sugarcane post-processing, four temporal filters were used:

1) Temporal filter using 3 years with 2 years threshold applied only on the initial edge year (1986). In the initial year (1985) and final year (2024) no temporal filter was applied.

2) Temporal filter using 5 years with 3 years threshold for the period 1987 to 2023.

3) Temporal filter using 5 years with 2 years threshold applied only on the final edge year (2024).

4) Temporal filter using 3 years with 2 years threshold applied to all series, except to the edge years (1986-2024) to ensure temporal consistency.

3.2.6 Citrus

As with the coffee class, for the citrus class the same filters were applied for the edge years of the series (1986 and 2023) as for the other years, plus a time consistency filter as follows:

1) Temporal filter using 3 years with 2 years threshold applied only on the edge years (1986 and 2023). In the initial year (1985) and final year (2024) no temporal filter was applied.

2) Temporal filter using 5 years with 3 years threshold for the period 1987 to 2022.

3) A fill filter was also applied to convert pixels that were not classified as citrus during a period when these pixels were classified as citrus.

3.2.7 Coffee

For post-processing of the coffee class three temporal filters were used:

1) Temporal filter using 3 years with 2 years threshold applied only on the edge years (1986 and 2023). In the initial year (1985) and final year (2024) no temporal filter was applied.

2) Temporal filter using 5 years with 3 years threshold for the period 1987 to 2022.

3) We also defined as pixels of the coffee class, starting in 2016, those pixels that were classified as coffee in any of the mappings of the last 5 years.

3.2.8 Oil Palm

For post-processing of the oil palm class three temporal filters were used:

1) Temporal filter using 7 years with 3 years threshold applied from 1986 to 2000.

2) Temporal filter using 5 years with 3 years threshold for the period 2000 to 2023.

3) And, a Temporal filter using 3 years with 2 years threshold was applied for the last years, from 2019 to 2024.

3.2.9 Other Perennial Crop

For Other Perennial Crop, a temporal filter using 5 years with 3 years threshold was used, in addition to a filter to remove intervals of the class of interest with less than 5 consecutive years. Therefore, a 6-year window was utilized: the year of interest and 1 year before and 4 years after the year of interest.

3.2.10 Forest Plantation

The same temporal filters applied to the coffee class were used for the 'Forest Plantation' class:

1) Temporal filter using 3 years with 2 years threshold applied only on the edge years (1986 and 2023). In the initial year (1985) and final year (2024) no temporal filter was applied.

2) Temporal filter using 5 years with 3 years threshold for the period 1987 to 2022.

3) Another consideration was at the end of the series. When the trees are cut down, it may take a while for them to grow back, during which time the classifier cannot identify them as 'Forest Plantation'. Since it takes 3 to 5 years for 'Forest Plantation' to become identifiable again, pixels from 2016 to 2021 were converted to 'Forest Plantation' when they were of this class in the 3 years before (2013 to 2015). Figure 19 illustrates this filter.

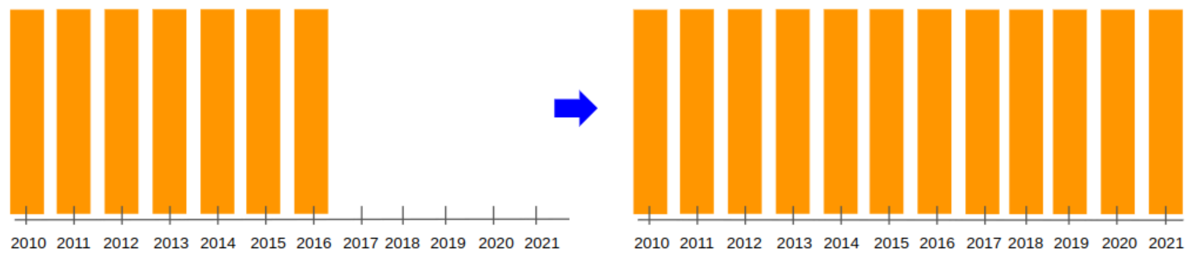


Figure 19. The temporal filter was applied in the last years of the 'Forest Plantation' series.

4) Another temporal filter was applied to fill longer intervals of non-occurrence of 'Forest Plantation' when it was a forest plantation some year in the past and it became again years after, like the example in Figure 20.

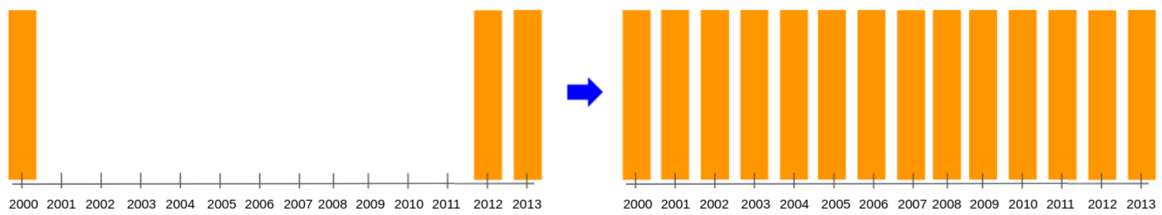


Figure 20. Temporal filter that converted longer intervals into 'Forest Plantation' when it was in the past and became again years after.

4 Integration with biomes and themes

After classifying the 'Agriculture' and 'Forest Plantation' themes, they were integrated with the other land use and land cover classes to create the final MapBiomas Collection 10 maps. This integration was based on the hierarchical order of the classes. The integration process enhances the quality of the 'Agriculture' and 'Forest Plantation' maps by eliminating certain commission errors.

A notable improvement in Collection 9 was the introduction of post-integration filters for the Agriculture and Forest Plantation classes. Figure 21 illustrates examples of these filters applied to the Temporary Crops and Forest Plantation classes.

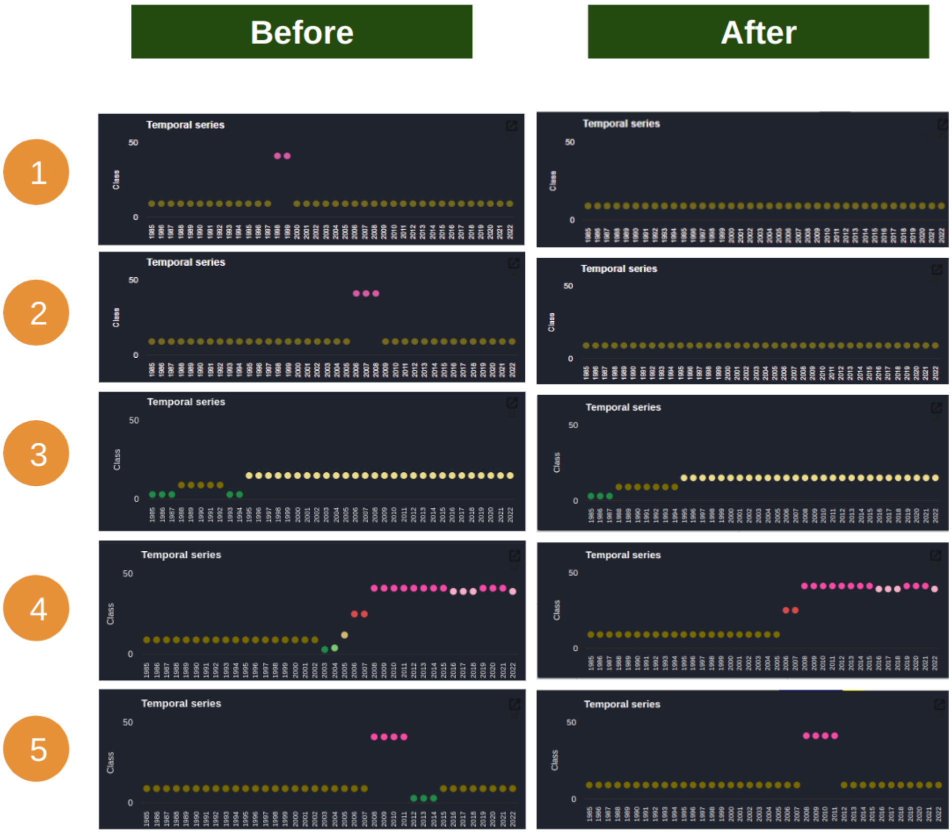


Figure 21. Examples of integration filters for Agriculture and Forest Plantation themes.

Examples 1 and 2 illustrate the filter's function, excluding agricultural occurrences that appear in three or fewer years, which were considered noise in the classification. Examples 3, 4, and 5 demonstrate how forest formation and savanna classes are remapped to forest plantation when these classes appear after a forest plantation and before other anthropic classes, such as pasture. The filter identifies forest formation as a misclassification of forest plantation, a common error due to their similar spectral responses. The filter's impact is particularly noticeable when analyzing the transition of classes from natural to anthropic.

5 Accuracy analysis

5.1 Comparison with Independent Validation Points

The accuracy analysis was produced using independent validation points provided by the *Laboratório de Processamento de Imagens e Geoprocessamento* (LAPIG) of the Goias Federal University (UFG). LAPIG points were collected only for the level 3 classes of 'Forest Plantation', 'Perennial Crops', and 'Temporary Crops', without distinction between the crops that compose these classes. For this reason, we aggregate all perennial classes (coffee, citrus, and Other Perennial Crops) into 'Perennial Crops' and all temporary classes (soybean, sugarcane, rice, cotton, and Other Temporary Crops) into 'Temporary Crops' to evaluate the accuracy using LAPIG points. LAPIG points used for the accuracy assessment are shown in Figure 22. MapBiomas' complete accuracy metrics can be found here:
https://brasil.mapbiomas.org/en/analise-de-acuracia/

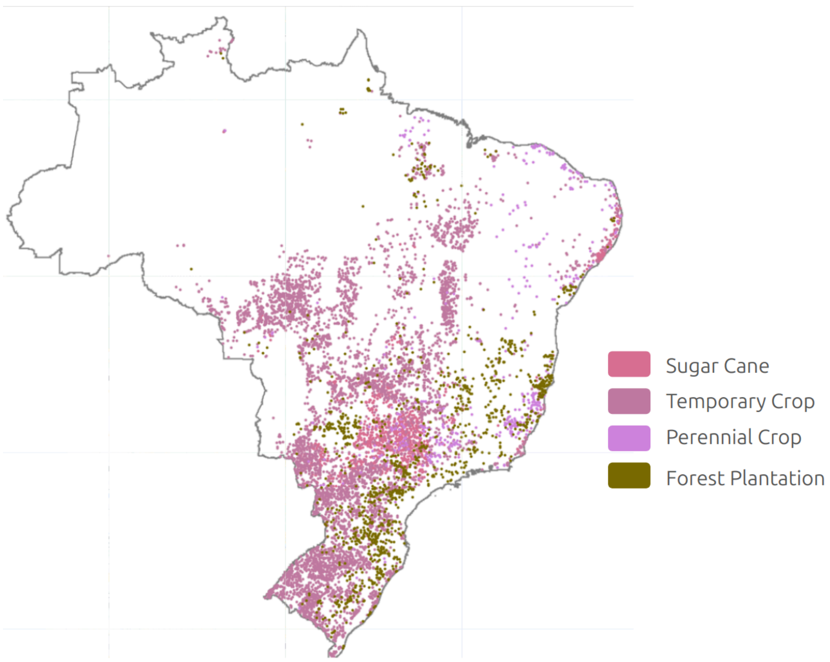


Figure 22. LAPIG points that were used for the accuracy assessment of 'Temporary Crops', 'Perennial Crops', and 'Forest Plantation' classes.

5.2 Comparison with others maps

5.2.1 TerraClass

TerraClass Project aims to demonstrate the dynamics of deforestation in the Amazon and Cerrado biomes, identifying land use and land cover types following deforestation events detected by the PRODES project, which has been conducted by the National Institute for Space Research (INPE) since 1988. Utilizing remote sensing and geoprocessing techniques, maps of land use and land cover for these biomes are generated to identify changes in deforested areas. This information is publicly available and can be accessed through the institute's website. It serves as an important reference for comparison with those produced by the MapBiomas project.

To provide the analysis, the years 2020 and 2022 were used because both Amazon and Cerrado have compatible maps with the same classes, which were comparable with MapBiomas classes, as shown in Table 9.

Table 9. Compatibility of classes from TerraClass 2020 and MapBiomas Collection 9.

Class	TerraClass (2020 and 2022)	MapBiomas
Sugar Cane	13	20
Temporary Crop	14, 15	41, 62, 39, 40
Perennial Crop	12	48, 46, 47 e 35
Forest Plantation	9	9

5.2.1.1 Amazon 2022

The compatible maps of MapBiomas and TerraClass are shown in Figure 28.

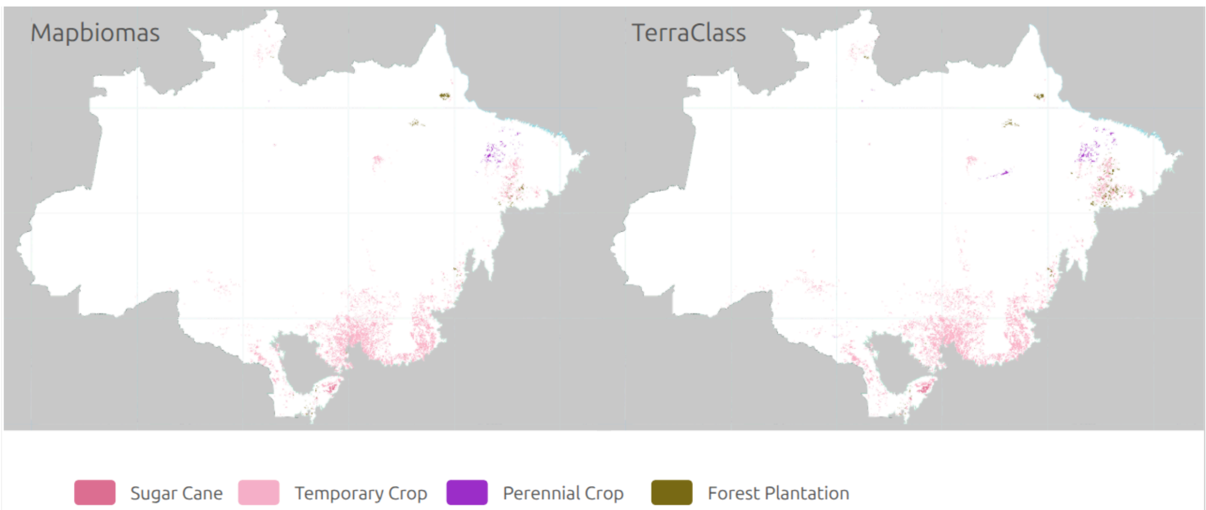


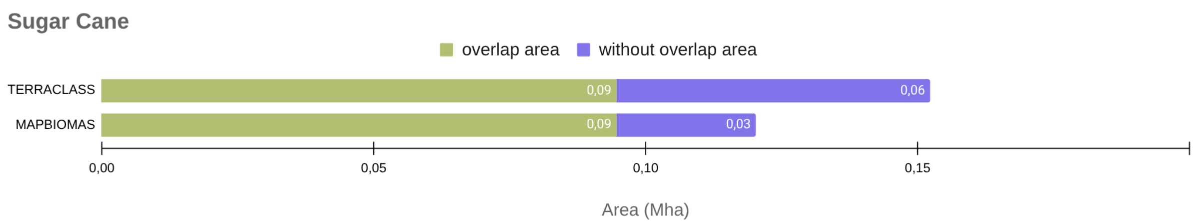
Figure 28. Compatible maps of MapBiomas and TerraClass maps to the base year 2022.

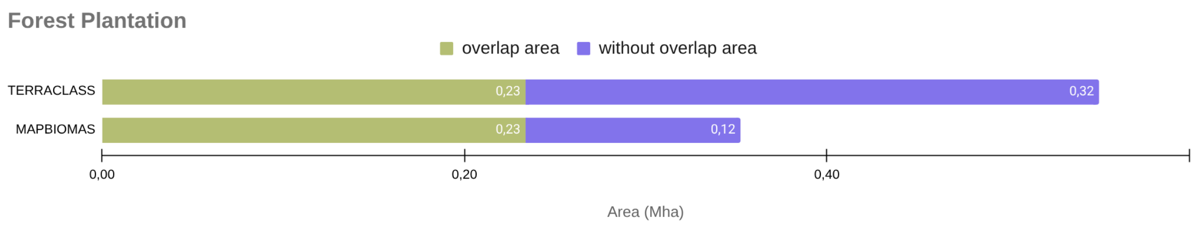
5.2.1.1.1 Area comparison

The pixel area between TerraClass and MapBiomas in the Amazon biome shows the difference between the sources. For the Amazon biome all the classes presents higher values than MapBiomas, as demonstrated in Figure 29.

Figure 29. Area per classes from MapBiomas and TerraClass maps to base year 2022.

The overlap area between the maps is shown in Figure 30.





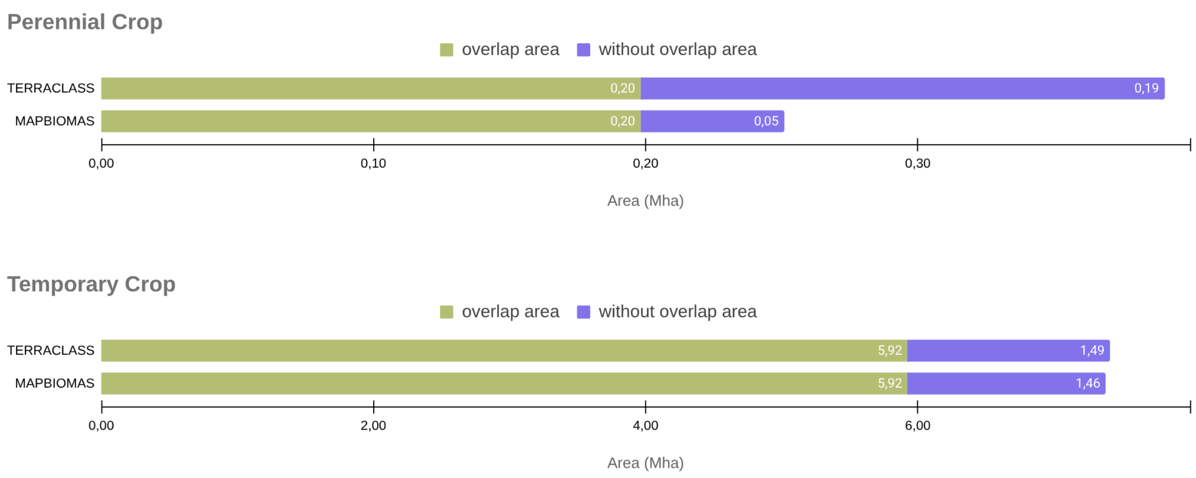


Figure 30. Overlap area per classes from MapBiomas and TerraClass project to 2022.

In most classes, the mapping from TerraClass obtained larger area values. Part of this difference may be due to the type of methodology adopted, which involves visual interpretation to group pixels into classes. This differs from MapBiomas's pixel-by-pixel classification, which contributes to a more pixelated or noisy map.

For Sugar Cane, the TerraClass project mapped 150,000 hectares and MapBiomas mapped 120,000 hectares. This represents a difference of 30,000 hectares more mapped by TerraClass. The Forest Plantation class shows a similar trend, with 550,000 hectares mapped by the TerraClass project and 350,000 hectares by the MapBiomas project. The Perennial Crop class showed 390,000 hectares mapped by TerraClass and 250,000 hectares mapped by MapBiomas. MapBiomas showed similar area values for the Temporary Crop class, resulting in a mapping of 7.38 million hectares in 2020 compared to the 7.41 million hectares mapped by TerraClass in the same year.

5.2.1.1.2 Accuracy Analysis

After making the classes compatible between MapBiomas and TerraClass, the MapBiomas validation points were used to compare the accuracies between the two maps. The result is shown in Figure 31.

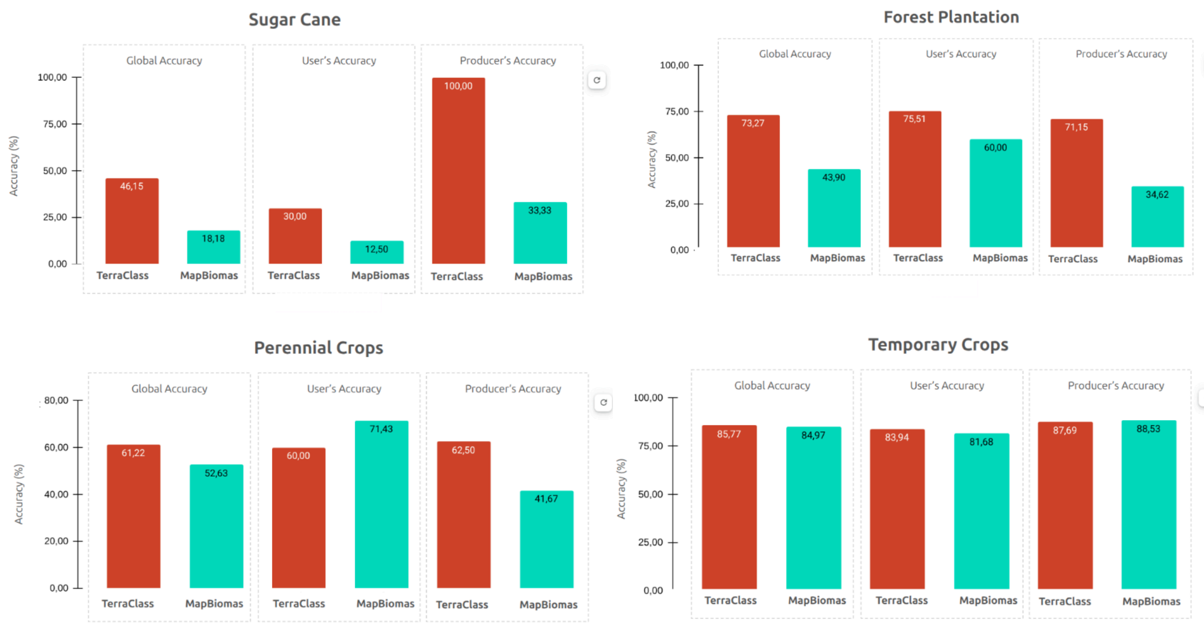


Figure 31. Accuracy analysis with Lapig validation points between MapBiomas and TerraClass project to 2022.

For all classes, TerraClass showed higher producer accuracy values compared to MapBiomas, except for Temporary Crops. This reinforces the trend towards a larger area mapped by TerraClass, as shown in the previous topic. When analyzing user accuracy, MapBiomas shows higher values for Perennial Crops. Overall, the differences between global accuracy for the two maps for the Temporary Crop class are only 0.81%, indicating a great deal of similarity in the quality of the maps.

5.2.1.2 Cerrado 2022

The compatible maps of MapBiomas and TerraClass are shown in Figure 32.

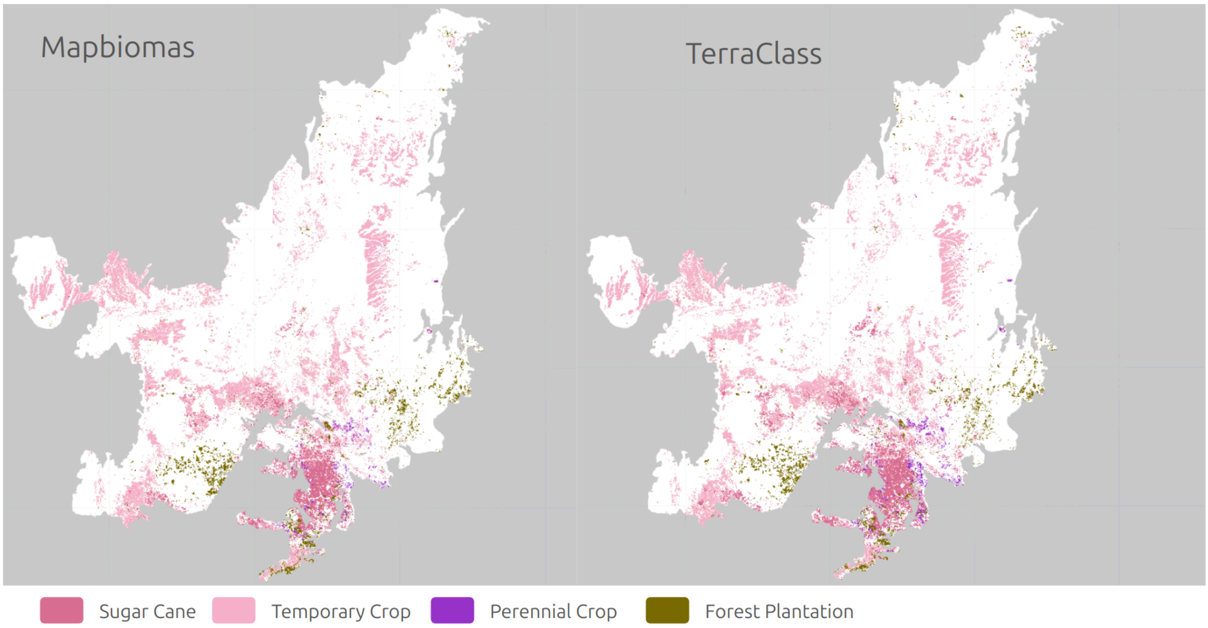


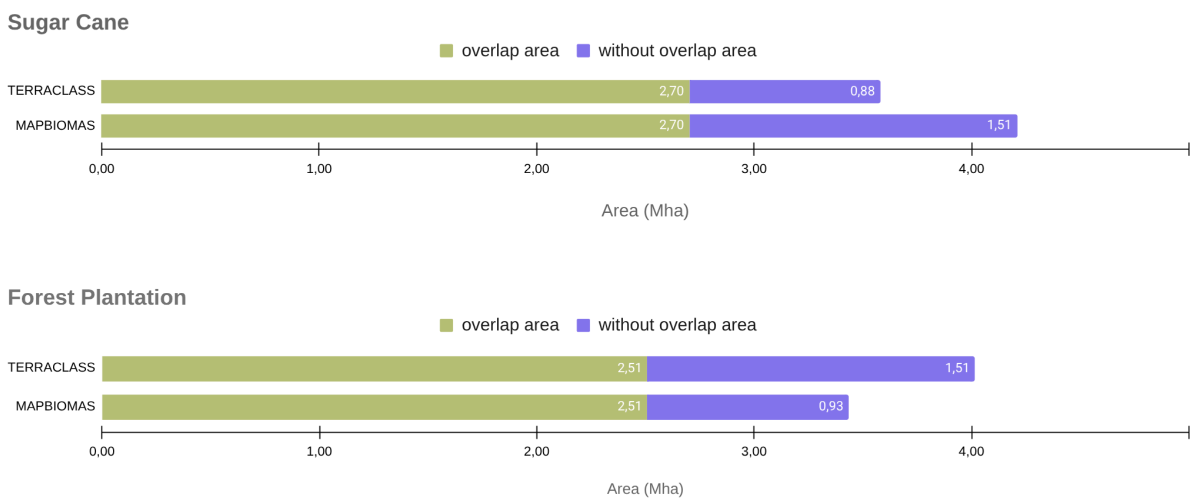
Figure 32. Compatible maps of MapBiomas and TerraClass maps to base year 2022.

5.2.1.2.1 Area comparison

The pixel area between TerraClass and MapBiomas to the Cerrado biome shows the difference between the sources. For the classes Temporary Crops, Forest Plantation, and Perennial Crop, TerraClass presents higher values than MapBiomas. Otherwise, MapBiomas presented a Sugar Cane area higher than TerraClass, as demonstrated in Figure 33.

Figure 33. Area per classes from MapBiomas and TerraClass maps to base year 2022.

The overlap area between the maps is shown in Figure 34.



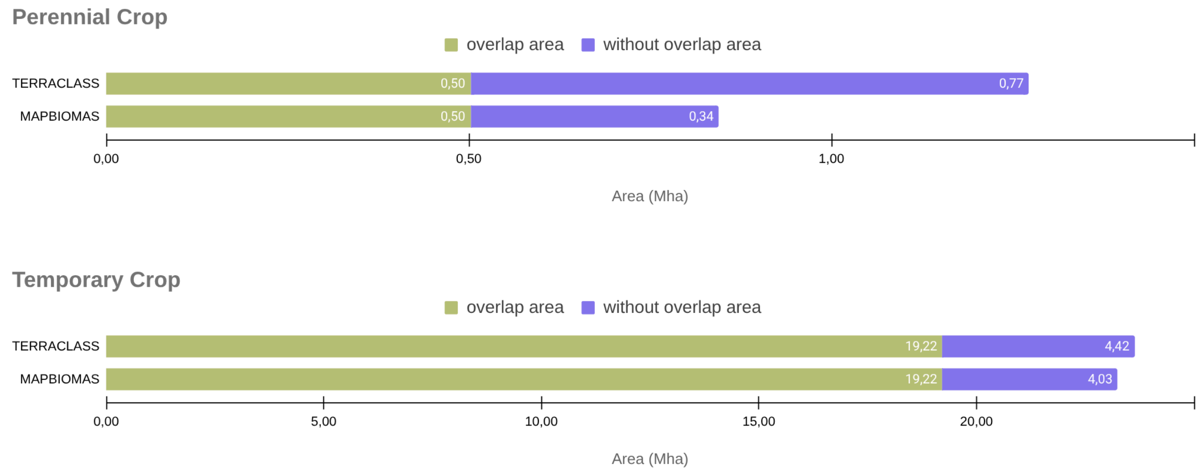


Figure 34. Overlap area per classes from MapBiomas and TerraClass project to 2022.

Unlike the pattern found in the Amazon biomes, in the Cerrado, there is a lot of overlap between the mappings from TerraClass and MapBiomas. The areas obtained for the Temporary Crop class were very similar, with more than 81% of both maps having overlapping areas. The remaining areas were those mapped by one but not the other,, i.e., a disagreement in allocation.

5.2.1.2.2 Accuracy analysis

The result of the comparison between the accuracies obtained from the MapBiomas validation points applied to both compatible maps is shown in Figure 35.

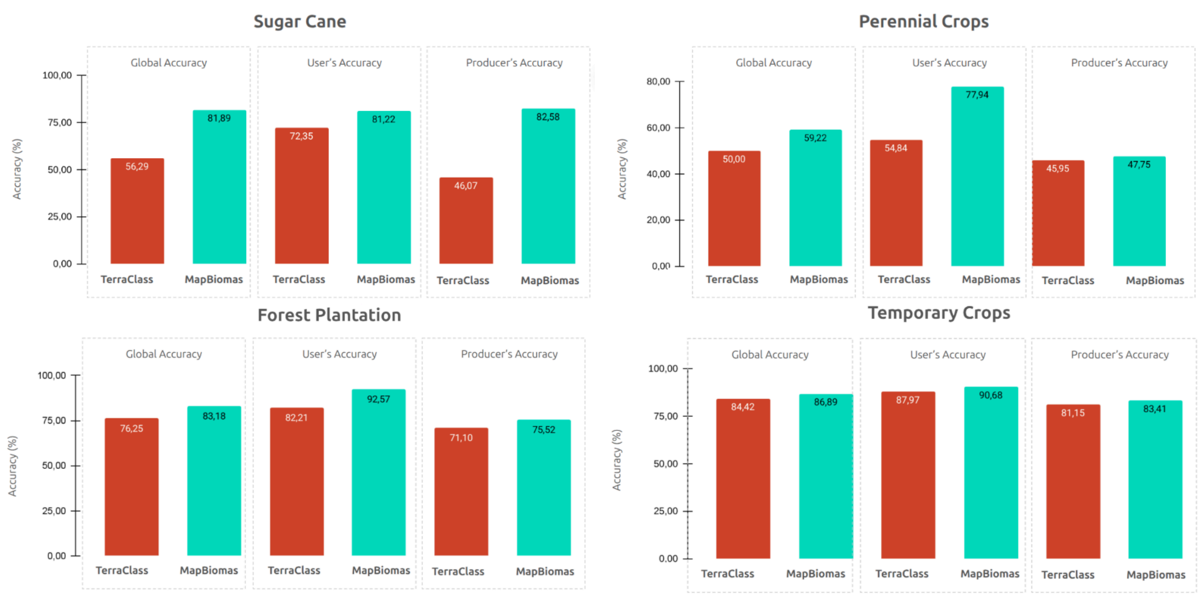


Figure 35. Accuracy analysis with Lapig validation points between MapBiomas and TerraClass project to 2022.

5.2.2 Brazil’s Fourth National Communication- QCN - 2016

The Fourth National Communication to the United Nations Framework Convention on Climate Change (UNFCC) refers to the fulfillment of Brazil's commitment to the global climate agenda. In the context of national communications, a national inventory of anthropogenic emissions by sources and anthropogenic removals by sinks of all greenhouse gases (GHG) is periodically compiled. From 1990 to 2016, this inventory was carried out following the “2006 Intergovernmental Panel on Climate Change (IPCC) Guidelines for National Greenhouse Gas Emissions Inventories. One of the results of the Fourth National Communication (QCN) was the Land Use and Land Cover (LULC) map, which serves as a basis for calculating emissions. This map has been made public and is an important source for comparing results with those obtained by MapBiomas, since calculating greenhouse gas emissions is one of the project's main objectives.

The comparison between the two sources to the base year 2016 was carried out similarly to the comparison between MapBiomas and TerraClass data. Figure 36 shows the compatibility between the classes used in the analysis.

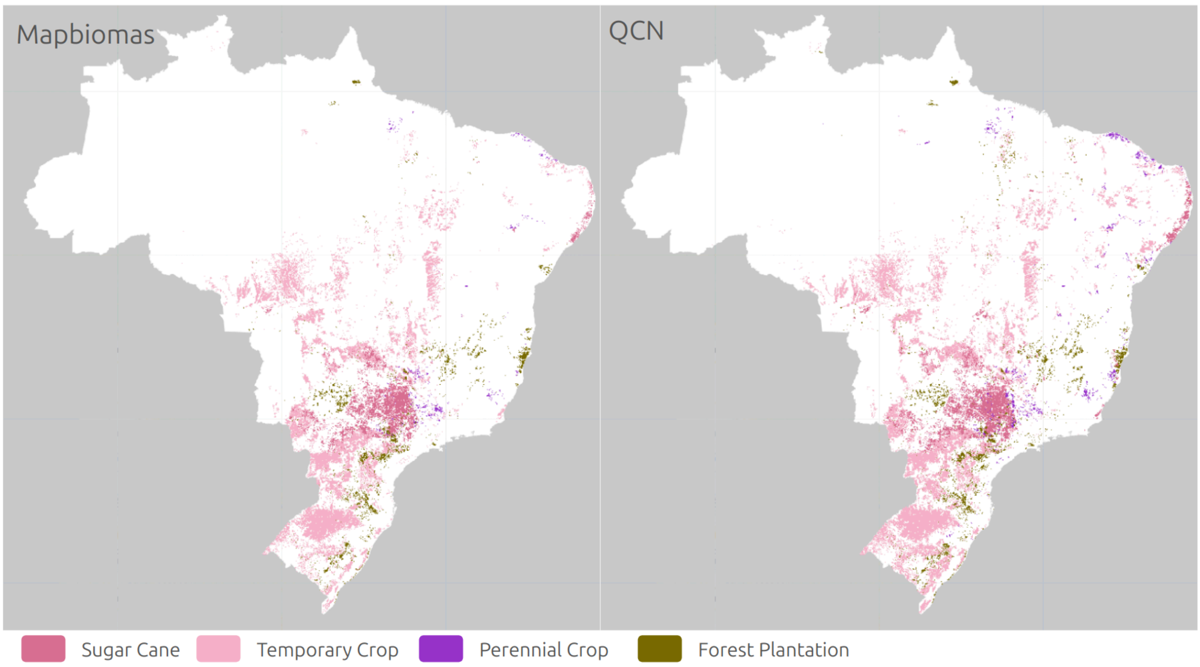


Figure 36. Compatible maps of MapBiomas and Fourth National Communication to the United Nations Framework Convention on Climate Change (UNFCC) maps to base year 2016.

5.2.2.1 Area comparison

The area presented by the QCN shows higher values than those obtained in the MapBiomas project for the four classes analyzed, as shown in Figure 37.

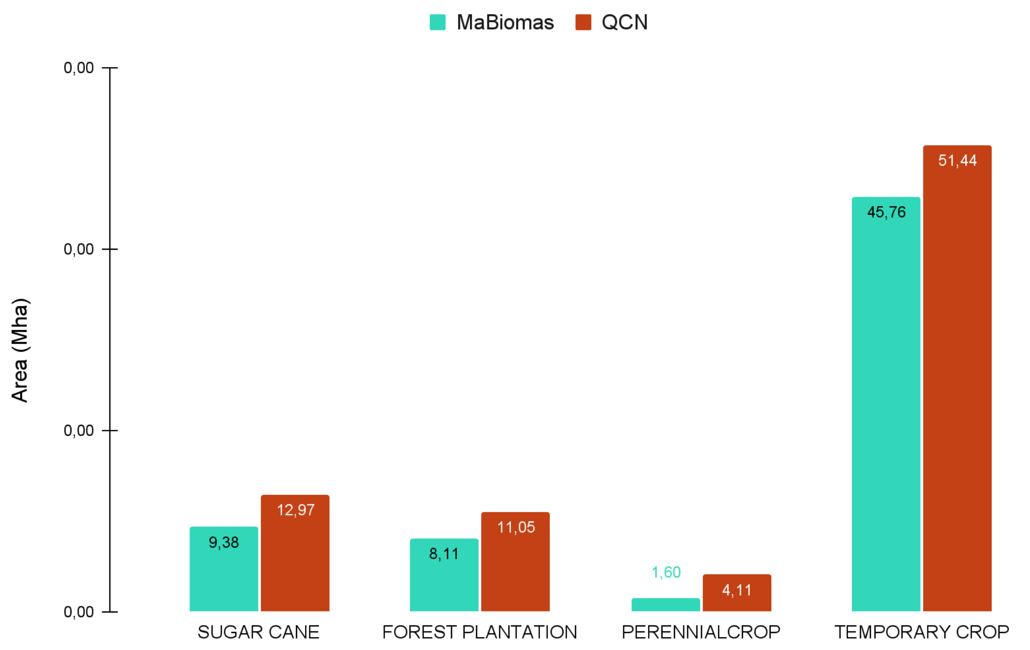
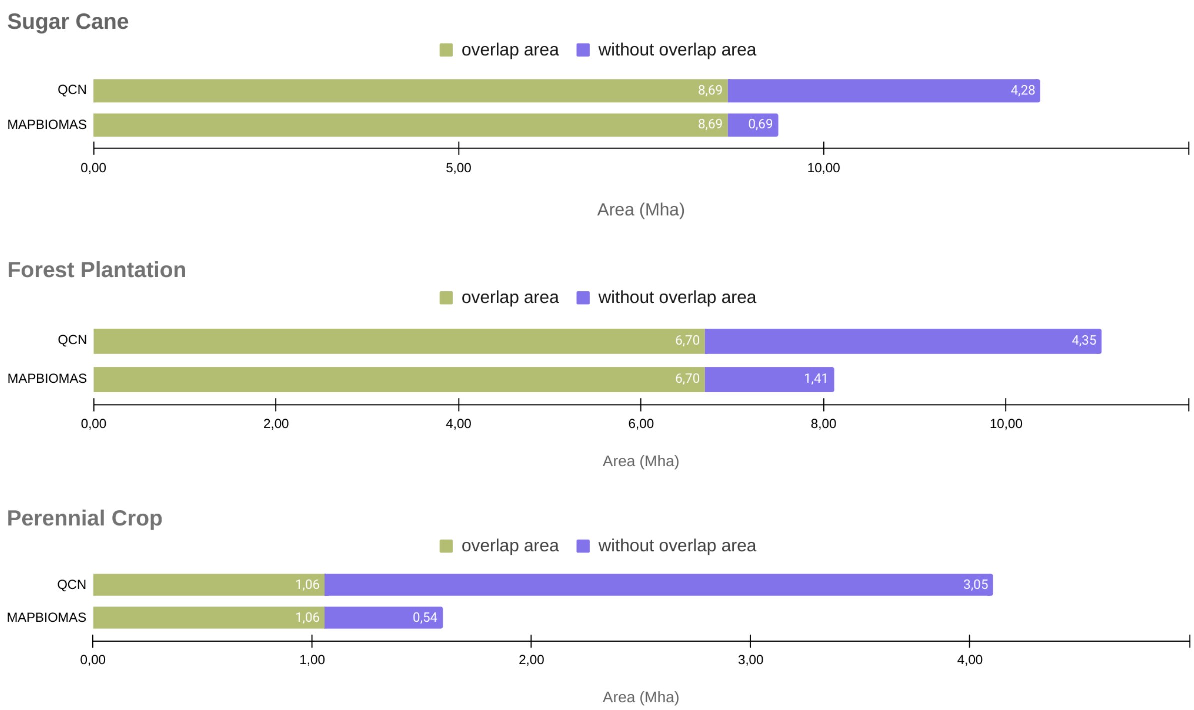


Figure 37. Area per classes from MapBiomas and QCN maps to base year of 2016.



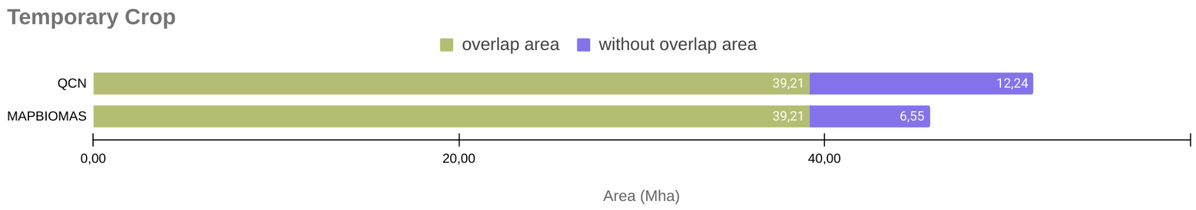


Figure 38. Overlap area per classes from MapBiomas and QCN project to 2016.

The mapping from QCN obtained larger area values. Same as TerraClass, part of this difference may be due to the type of methodology adopted, which is mapping via visual interpretation, where pixels are grouped in a way that covers the entire class. This is different from MapBiomas's pixel-by-pixel classification, which results in a more pixelated or noisy map.

For Sugar Cane, the QCN project mapped 12.97 million hectares and MapBiomas mapped 9.38 million hectares in 2016. This represents a difference of 3.59 million hectares more mapped by QCN in Brazil. The Forest Plantation class shows a similar trend, with 11.05 million hectares mapped by the QCN and 8.11 million hectares by the MapBiomas project. The Perennial Crop class showed 4.11 million hectares mapped by QCN and 1.6 million hectares mapped by MapBiomas. And QCN showed higher area values for the Temporary Crop class, resulting in a mapping of 51.44 million hectares compared to 45.76 million hectares mapped by MapBiomas in the same year.

5.2.2.2 Accuracy analysis

By using the same MapBiomas validation points applied to the QCN maps, the results shown in Figure 39 were obtained.

Figure 39. Accuracy analysis with Lapig validation points between MapBiomas and QCN project to 2016.

For all analyzed classes, QCN maps showed higher producer accuracy results than MapBiomas, which corroborates the result of larger mapped areas shown in Figure 37. However, MapBiomas showed higher results in terms of user accuracy, indicating high reliability in the results obtained by the project when mapping the classes. Overall, the results showed similar accuracy results for the classes, except for Perennial Crops, which showed large differences in mapping. Temporary crops show very similar results in terms of overall, producer, and user accuracy. Of the more than 45 million hectares mapped by the MapBiomas project, 39.21 million hectares show an overlapping area with the QCN mapping results, corresponding to 86% of the temporary crops map.

These results suggest that MapBiomas and the other land use and land cover mapping sources present very similar results, indicating that the MapBiomas project's methodology is promising for land use and land cover mapping

References

AGÊNCIA NACIONAL DE ÁGUAS (ANA). Levantamento da Agência Nacional de Águas - Brasília: ANA, 2020.

AGÊNCIA NACIONAL DE ÁGUAS (ANA). Mapeamento do arroz irrigado no Brasil. Brasília: ANA & Conab, 2020.

ASSOCIAÇÃO BRASILEIRA DOS PRODUTORES DE ALGODÃO (ABRAPA). Algodão no Brasil. 2020. Available at: <https://www.abrapa.com.br/Paginas/dados/algodao-no-brasil.aspx>

Breiman, Leo. Random forests. Machine learning, v. 45, n. 1, p. 5-32, 2001.

CONAB. COMPANHIA NACIONAL DE ABASTECIMENTO. Mapeamento Agrícolas. Available at: <https://www.conab.gov.br/info-agro/safras/mapeamentos-agricolas?start=20>.

Global Forest Watch Transparent World. Tree Plantations. 2015. Accessed through Global Forest Watch in January 2017. Available at: <www.globalforestwatch.org>.

Li, J.; Tian, L.; Wang, Y.; Jin, S.; Li, T.; Hou, X.; Optimal sampling strategy of water quality monitoring at high dynamic lakes: A remote sensing and spatial simulated annealing integrated approach. Science of The Total Environment, 10, 146113, 2021.

Quarta Comunicação Nacional Do Brasil À Convenção-Quadro Das Nações Unidas Sobre Mudança Do Clima (UNFCC). Available at: <https://www.gov.br/mcti/pt-br/centrais-de-conteudo/publicacoes-mcti/quarta-comunicacao-nacional-do-brasil-a-unfccc/sumario_executivo_4cn_brasil_web.pdf >

Ronneberger, O.; Fischer, P.; Brox, T.. U-Net: Convolutional networks for biomedical image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention; Springer: Berlin, Germany, 2015; pp. 234–241, 2015.

Song, X. P., Hansen, M. C., Potapov, P., Adusei, B., Pickering, J., Adami, M., ... & Tyukavina, A. Massive soybean expansion in South America since 2000 and implications for conservation. Nature sustainability, 4(9), 784-792, 2021.

Teixeira Jr. P. D. P; et al. Obtenção De Calendário Agrícola Baseado Em Série Temporal De Índice De Vegetação. In: Anais Do Xx Simpósio Brasileiro De Sensoriamento Remoto. Florianópolis. Anais eletrônicos... São José dos Campos, INPE, 2023. Disponível em: <https://proceedings.science/sbsr-2023/trabalhos/obtencao-de-calendario-agricola-baseado-em-serie-temporal-de-indice-de-vegetacao?lang=en> Accessed on: 15 Jul. 2024.

Zhu, Z., Gallant, A. L., Woodcock, C. E., Pengra, B., Olofsson, P., Loveland, T. R., ... & Auch, R. F.. Optimizing selection of training and auxiliary data for operational land cover classification for the LCMAP initiative. ISPRS Journal of Photogrammetry and Remote Sensing, 122, 206-221, 2016.