# Agriculture and Forest Plantation - Appendix

## Collection 9

**General coordinator**
Eliseu Weber

**Team**
Kênia Samara Mourão Santos
Paulo Domingos Pires Teixeira Junior

## 1 Overview of the classification method

The mapping of Agriculture and Forestry 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 9 (2023), methodological and technological improvements were made to the process, leading to an increase in the number of mapped classes, as illustrated in Figure 1.

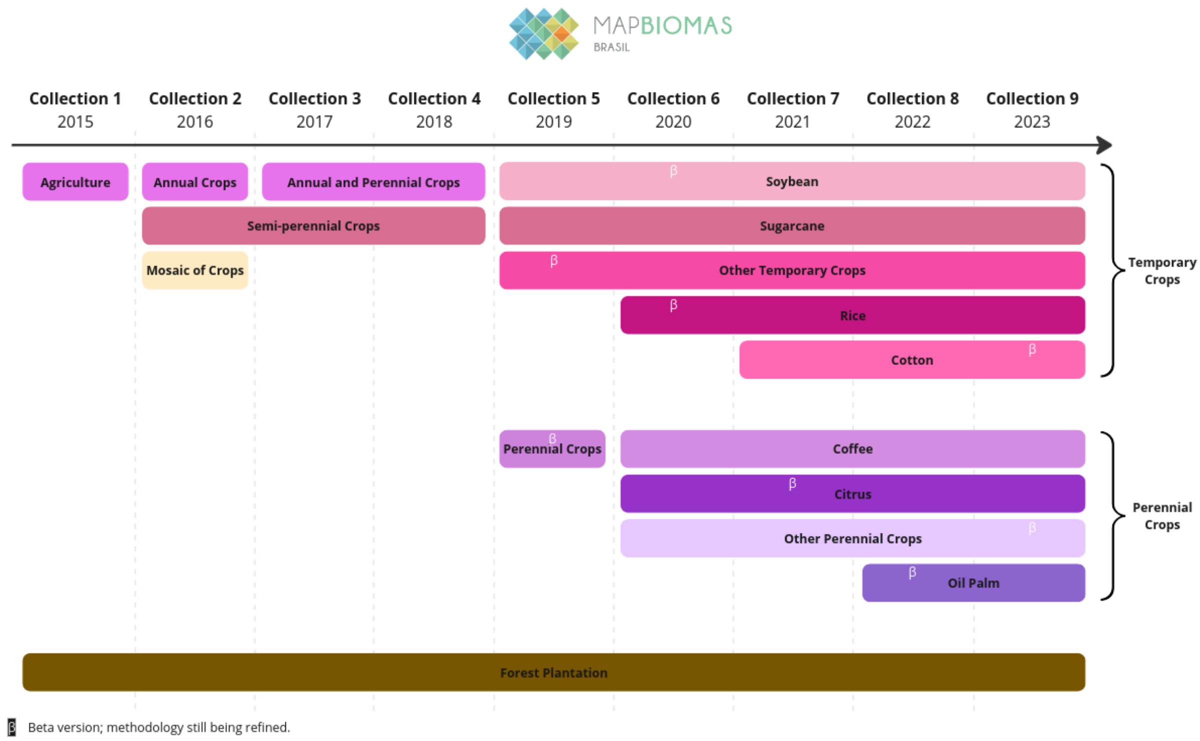



**Figure 1**. Evolution of mapped classes for agriculture and forestry.

In Collection 9, Remap Geotecnologia assumed the mapping of those classes, as well as the classes in the Irrigated 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 Unet models.

## 2 Classification

The MapBiomas-brazil account in GitHub has all the scripts used to classify 'Agriculture' and 'Forest Plantation' classes in MapBiomas Collection 9. The repository links are:

- Agriculture:

  https://github.com/mapbiomas-brazil/agriculture/tree/mapbiomas90

- Forest Plantation:

https://github.com/mapbiomas-brazil/forest-plantation/tree/mapbiomas90

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

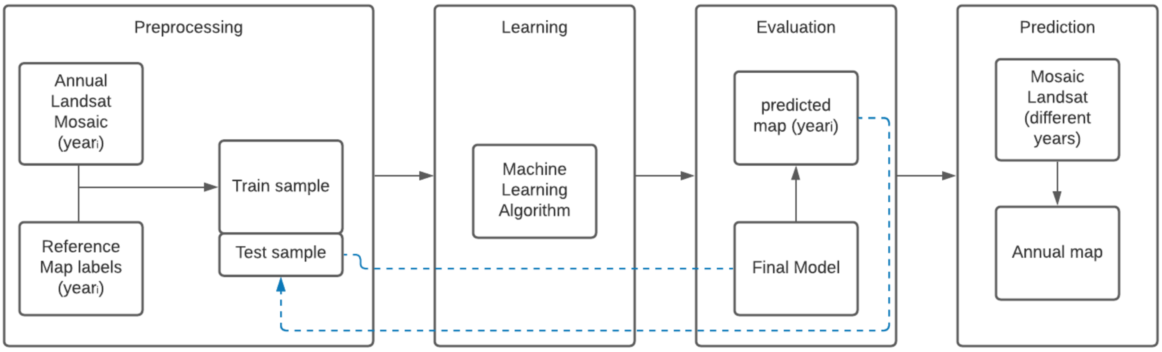



**Figure 2**. Supervised learning workflow in context of image classification.

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 according to each of the algorithms. The annual rice, citrus and oil palm maps were generated using a convolutional neural network (*i.e*. U-Net) and the other classes were obtained using Random Forest.

## 2.1 Landsat image mosaics

Figure 3 shows the availability of images from the Collection 2 Landsat series between 1985 and 2023. It is worth mentioning the greater number of images available from the launch of Landsat 8 and Landsat 9, which provides a greater probability of obtaining cloud-free mosaics.

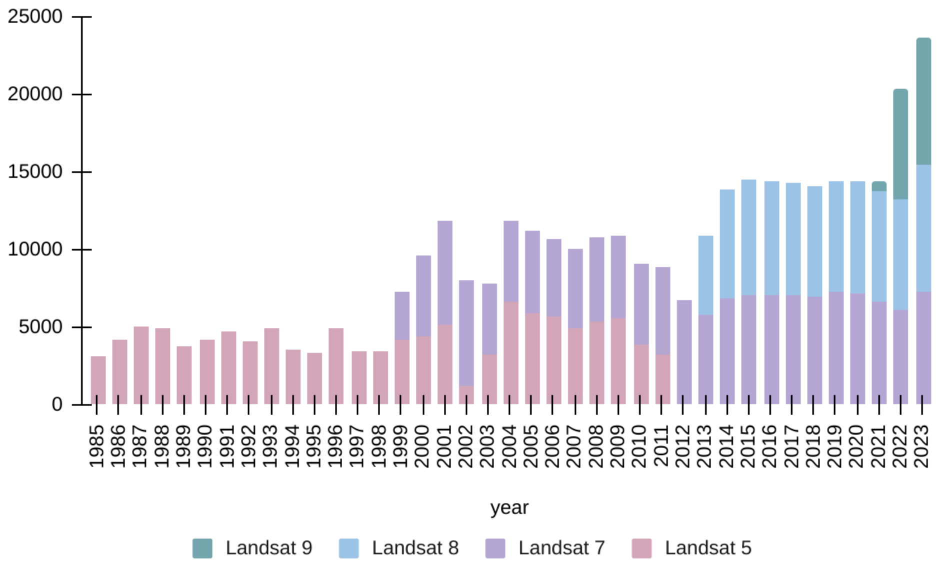



**Figure 3.** The number of available TOA Landsat Collection 2 images covering the Brazilian territory from 1985 to 2022.

## 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 agriculture crops and for different regions in Brazil, the growing season can cover different periods of the year, predominantly during the wet period, in most Brazil regions. According to phenological developments of the different types of cultures, we can note that for mapping annual crops, the Landsat mosaics require images that cover the period from October to March, while for semi perennial and perennial crops, we can use images collected throughout most of the year or all year.

Additionally, because 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 is essential to use images from the same period in previous years 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 in the summer period 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 period called 'second crop', in the largest cotton producing region in

Brazil (Mato Grosso) (ABRAPA, 2020). Thus, as MapBiomas only maps agriculture classes of the first crop period, the cotton mapping from MapBiomas Collection 9 will not reflect the total planted area of the cotton-producing regions.

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 can be found in Teixeira et al. (2023).

An example of the methodology to obtain the annual vegetative peaks per Landsat scene is shown in Figure 4.

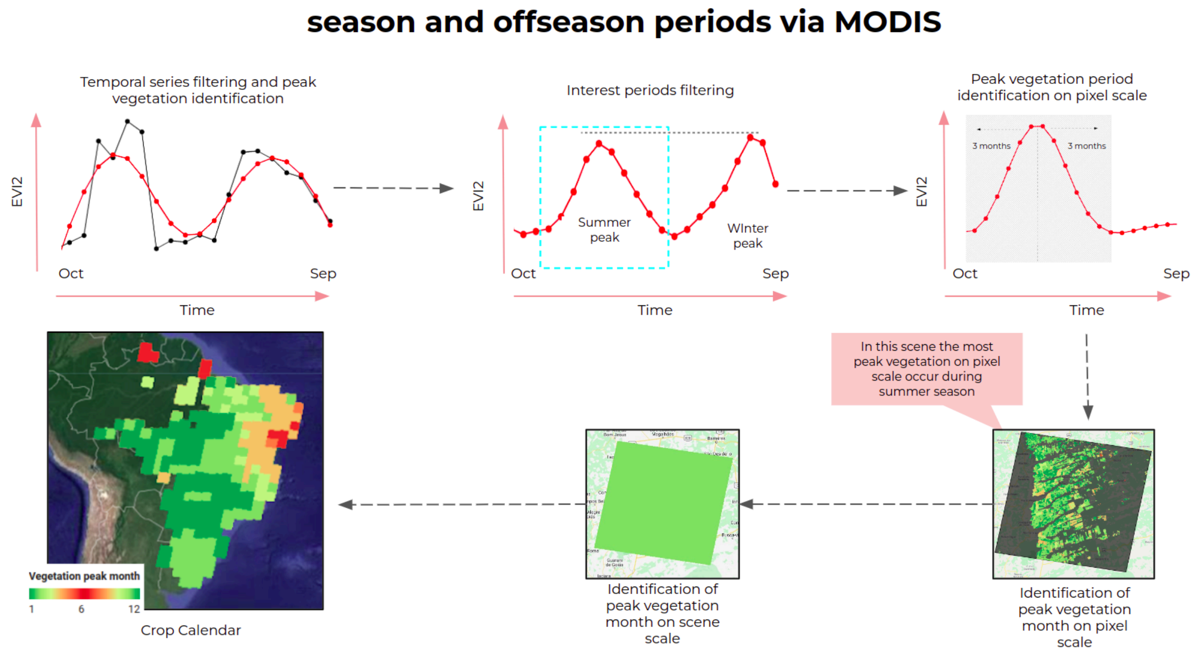



**Figure 4.** Scheme to obtain vegetation peak month, year by year, to Landsat scene. Source: Teixeira et al. (2023).

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. Was defined the 'growing season' as 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 |

#### 1.1.1.1 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 for selecting the selection of mosaic images of sugar cane in Collection 9.

| 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 |

#### 1.1.1.1 Rice

In Collection 9, 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 9.

| 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** | 10/01/year-1 | 04/01/year | 01/10/year-1 | 01/01/year |

| | | | | |
|---|---|---|---|---|
| **Santa Catarina - SC** | | | | |
| **Paraná - PR** | 10/01/year-1 | 04/30/year | 01/01/year | 07/30/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 |

### 2.2.2 Perennial Crop

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

### 2.2.3 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 9.

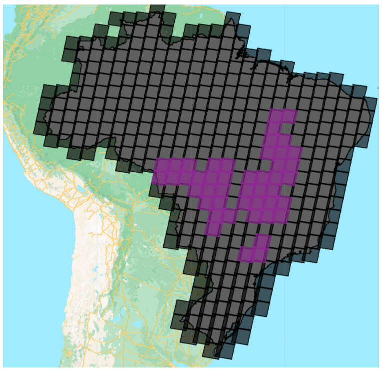
| Period | Start | End |
|---|---|---|
| **P1** | 01/01/year | 07/01/year |
| **P2** | 07/01/year | 01/01/year+1 |

### 2.2.4 Definition of regions for classification

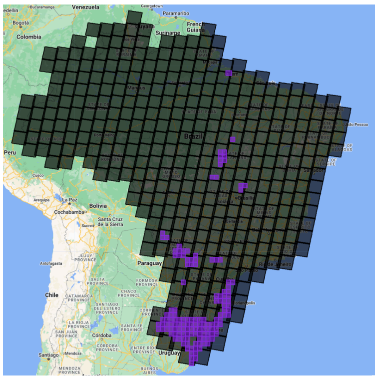
The 'Agriculture' and 'Forest Plantation' are heterogeneously distributed in the Brazilian biomes. Therefore, Landsat scenes were selected in regions with the highest occurrence of each class according to the reference maps. Figure 8 illustrates the scenes chosen for each land use class

Due to an error in Collection 9 integration, the citrus area inside the Atlantic Forest biome was omitted and is included in the mosaic of uses class. This misrepresented area corresponds to approximately 100.000 ha. This problem will be corrected in Collection 10.

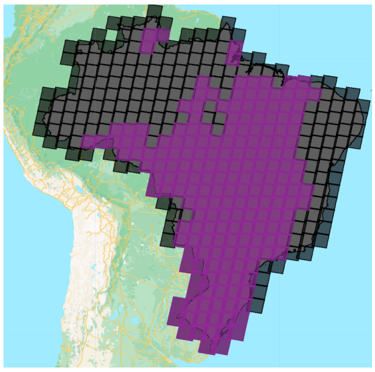
**cotton**



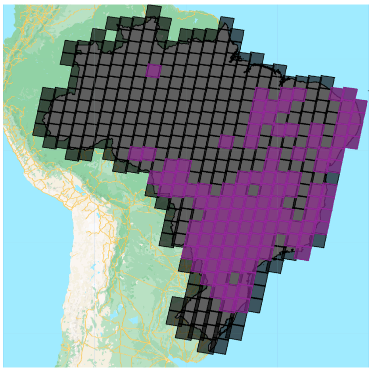
**Rice**



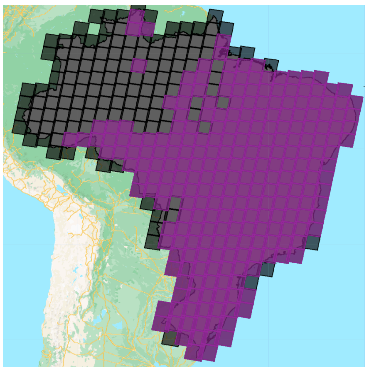
**Soybean**



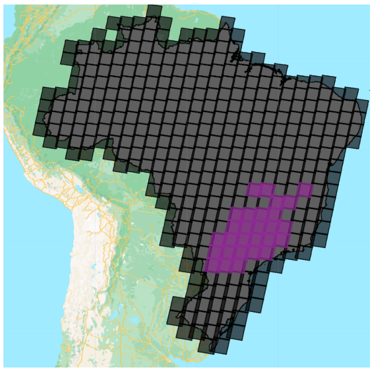
**Sugar cane**



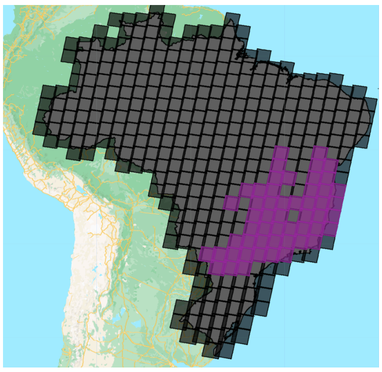
**Other Temporary Crop**



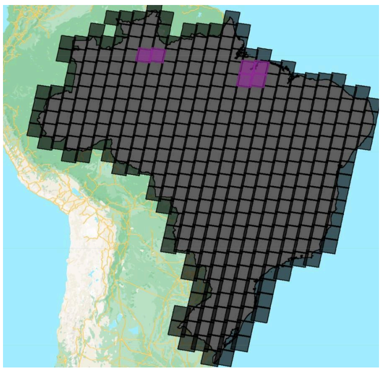
**Citrus**



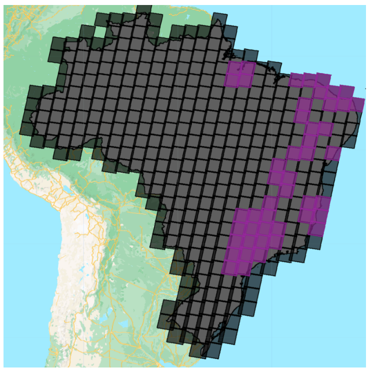
**Coffee**



**Oil Palm**



**Other Perennial Crop**



**Forest Plantation**

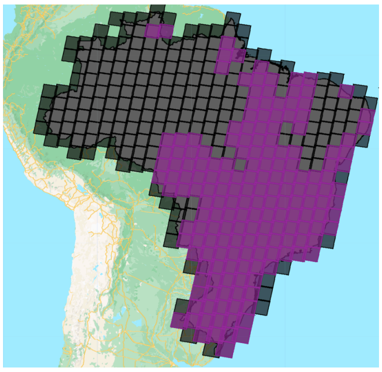


**Figure 8.** Selected scenes of Landsat series to classify maps by land use class.

### 2.3 Feature space

In Collection 9, most of the feature space was used based on the improvements made in Collection 8, however, the rice class was 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 9.

| Classes | Region | Bands | Indexes | Metrics | Algorithm |
|---|---|---|---|---|---|
| **Cotton** | as table x | 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 table x | 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 table x | 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 table x | 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 table x | 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 table x | RED, NIR, SWIR1 | - | median | U-Net |
| **Oil Palm** | as table x | | - | median | U-Net |
| **Other Perennial Crop** | as table x | BLUE, GREEN, RED, NIR, SWIR1, SWIR2 | NDVI | median, max, min, stdDev, 20th percentile, and quality mosaic (NDVI) | Random Forest |
| **Forest Plantation** | as table x | BLUE, GREEN, RED, NIR, SWIR1, SWIR2 | EVI2, MNDWI, LAI | median, mean, max, min, stdDev, 80th percentile, and quality mosaic (qmo) | Random Forest |

## 2.4 Classification algorithm, training samples and parameters

### 2.4.1 Reference Maps

The reference maps used to obtain samples to train the classifier are shown in Table 15.

**Table 8.** Reference maps used in the Random Forest classification for the classes 'Agriculture' and 'Forest Plantation' in Collection 9.

| 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) |

### 2.4.2 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 and off-season), specific for each class; b) bands are compose 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 the process are annual maps of interest classes. To reduce the amount of noise and inconsistencies, the maps obtained after the classification undergo spatial and temporal post-processing and then are integrated into the other themes of MapBiomas. An important observation is that the annual mosaic used in the training process must be from the same year as the reference map used. An example of Random Forest classification is presented in Figure 9.

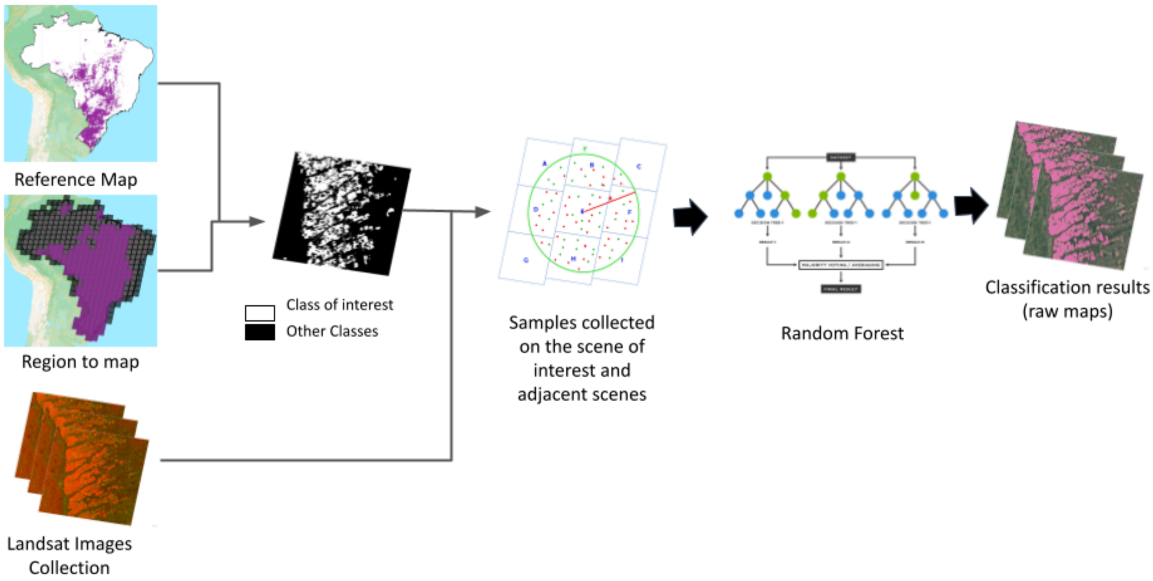



**Figure 9.** Flowchart of agriculture and Forest Plantation classification using Random Forest algorithm.

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

### 2.4.2.1 Simple Sampling

The acquisition of training samples was performed by each Landsat scene. In addition to the samples collected in the target scenes, samples collected in 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.

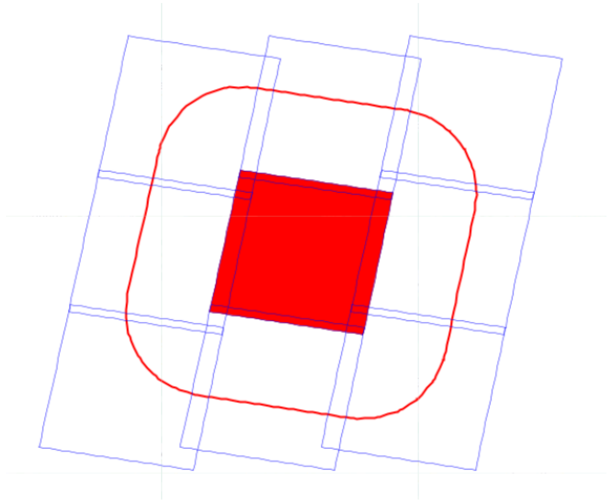


**Figure 10.** Scheme for sample acquisition for the regionalized training of the Random Forest classifier in 'Agriculture' and 'Forest Plantation'.

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 used to classify the subsequent years in which a reference map was not available.

#### 2.4.2.2 Stratified Sampling

The quality of training samples has been related 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). The sampling methods commonly used for supervised classification (such as simple sampling), may often not consider the spatial distribution of the targets of interest in the 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 distributed randomly, considering the whole Landsat scene boundary, while in the stratified sampling method, the distribution of the number of training samples is weighted by the percentage area of the class of interest, obtaining a balanced distribution of samples as is shown by Figure 11.

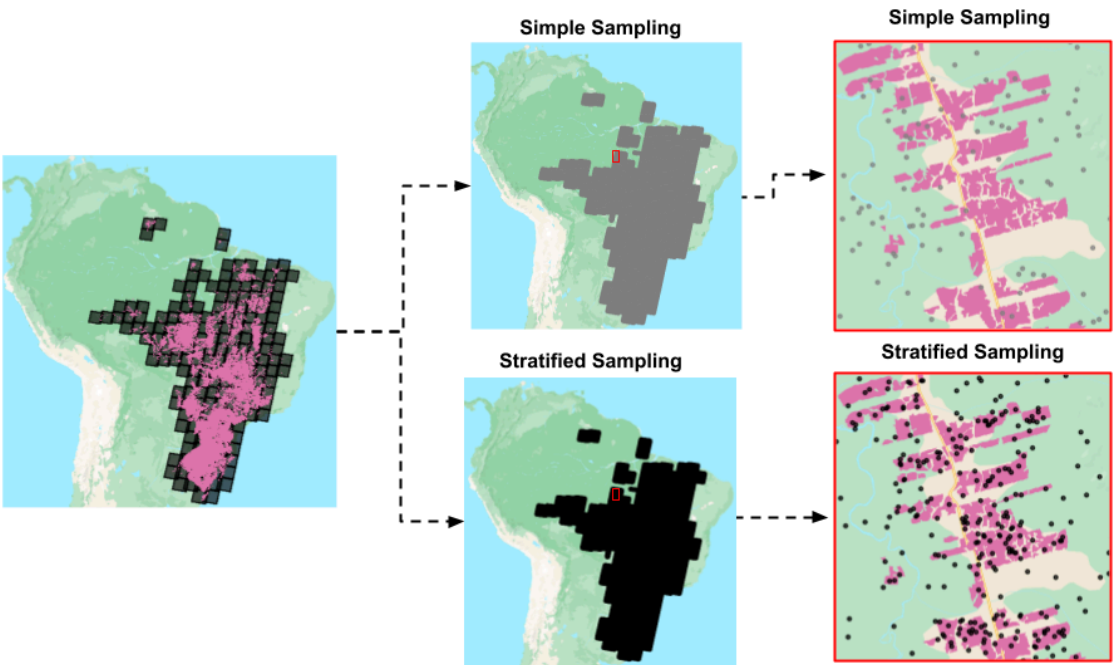



**Figure 11** - Difference between sampling approaches. a) simple sampling, and b) stratified sampling.

### 2.5 Deep Learning

For the mapping of Rice, Citrus and Oil Pam, an adaptation of the U-Net convolutional neural network (RONNEBERGER et al., 2015) was used. Unlike machine learning algorithms that classify each pixel considering the spectral response for each pixel, this architecture uses the context in which the pixels are. This architecture is illustrated in Figure 12.

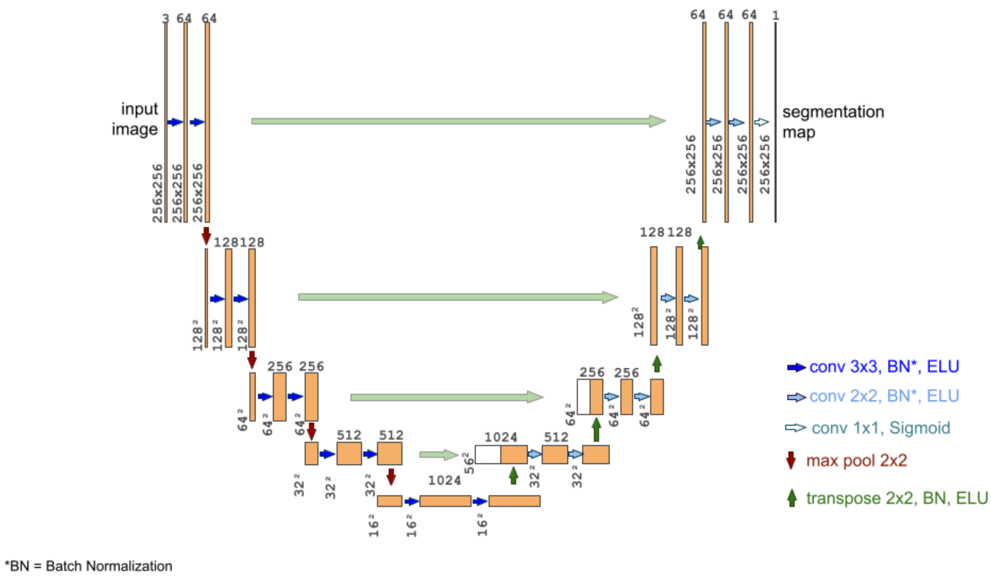



**Figure 12** - Adapted U-Net convolutional neural network, with its layers and connections, used for the mapping of rice and citrus.

### 2.5.1 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. The selection of images was made 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.

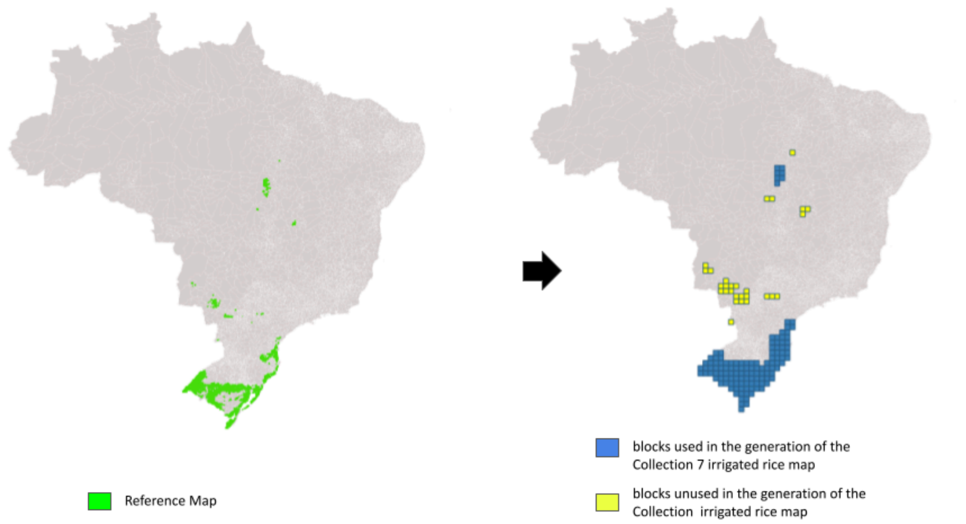



**Figure 13.** Study area used for the mapping of irrigated rice in the MapBiomas Project.

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.

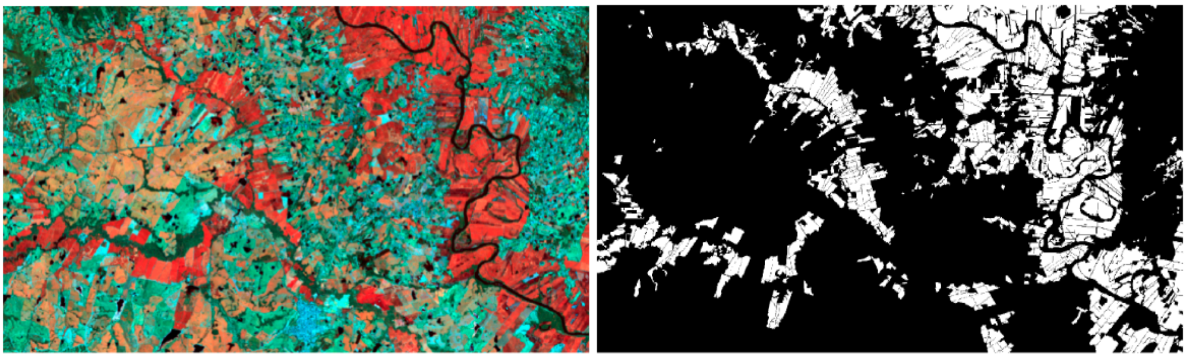


**Figure 14**. Example of U-Net samples to mapping rice.

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 (1985-2022).

#### 2.5.2 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.

#### 2.5.3 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 filter of minimum connected pixels was applied in most classes, except on the classes mapped with U-Net, because the result of the semantic segmentation showed little or no spatial noise. This spatial filter removed groups of pixels with 6 or fewer pixels of the interest class or the “others” class (Figure 15).

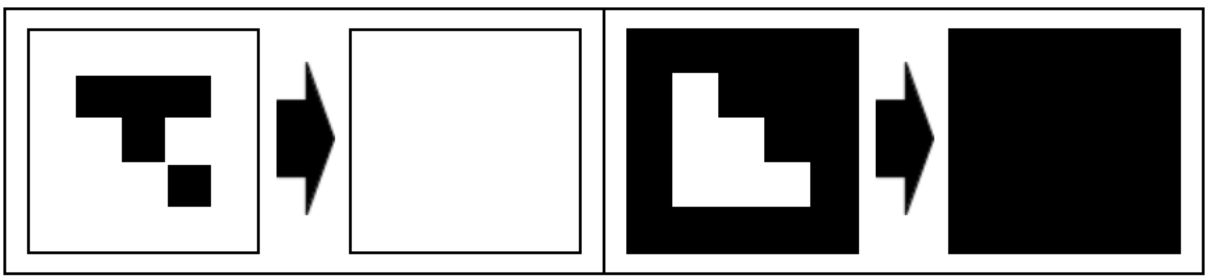


**Figure 15.** Example of the minimum connected pixels spatial filter. The image on the left shows an exclusion of pixels of the interest class (in black). The image on the right shows an inclusion of pixels of “other class” (in white) to the interest class.

### 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 none 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).

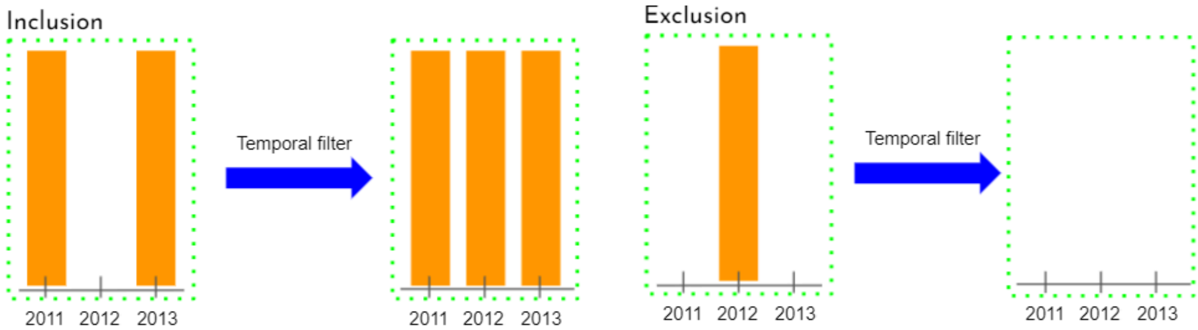



**Figure 16.** 3-year temporal window filter: The orange bars represent pixels of the mapped class (interest class). The exclusion filter changes a pixel to the "others" class when the same pixel was not of interest class in the adjacent years. The inclusion filter changes a pixel to the interest class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

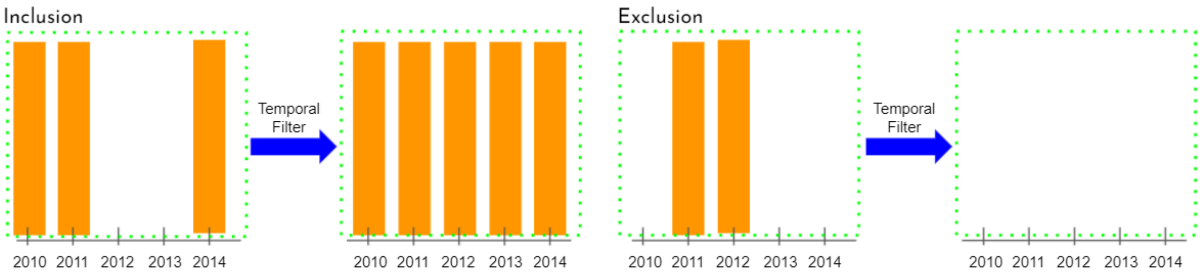



**Figure 17.** 5-year temporal window filter: The orange bars represent pixels of the mapped class (interest class). The exclusion filter changes a pixel to the "others" class when no more than 1 another year is of the interest class. The inclusion filter changes a pixel to the interest class when at 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. For the last year of the time series (*i.e.* 2023), no temporal filter was applied.

### 3.2.1 Soybean, Cotton, and Other Temporary Crops

For MapBiomas Collection 9, the soybean, cotton, and Other Temporary Crops classes followed a unified process that was proposed in the last Collection (Figure 18).

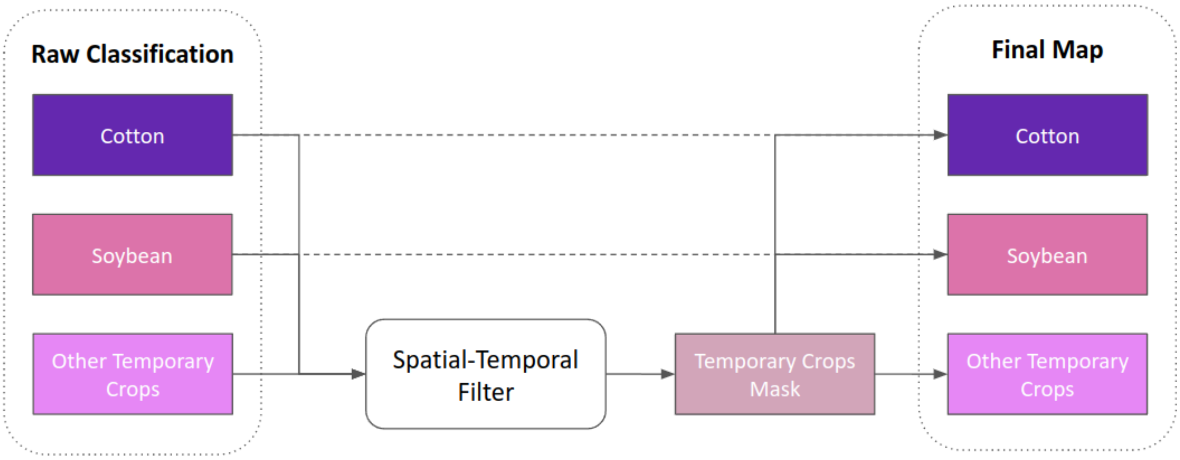



**Figure 18**: Soybean, cotton, and Other Temporary Crops classification and post-classification Flowchart.

The classification was made using a reference map with class distinction for cotton, soybean, and Other Temporary Crops. However, the raw classification results were not all used in the filters as separate classes, but as a unified 'Temporary Crops' class. The temporal filter used a 3-year window with 2 years threshold.

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. In this way, it was possible to maintain 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 Brazil, a temporal filter with a 5 years window and 3 years threshold was applied. A second temporal filter with a 3 years window and 2 years threshold was applied in all regions with the exception of southern Brazil. The exception is due to the known dynamics of agriculture in rice producing areas in southern Brazil, that often involve crop rotation, which would be omitted otherwise.

### 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 (2023) no temporal filter was applied.

2) Temporal filter using 5 years with 3 years threshold for the period 1987 to 2021.

3) Temporal filter using 5 years with 2 years threshold applied only on the final edge year (2023).

4) Temporal filter using 3 years with 2 years threshold applied to all series, except to the edge years (1986-2022) 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 2021) 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 2021). In the initial year (1985) and final year (2022) no temporal filter was applied.

2) Temporal filter using 5 years with 3 years threshold for the period 1987 to 2020.

3) A fill filter was also applied to convert pixels that were not classified as citrus between 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 2021). In the initial year (1985) and final year (2022) no temporal filter was applied.

2) Temporal filter using 5 years with 3 years threshold for the period 1987 to 2020.

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 2021.

3) And, a Temporal filter using 3 years with 2 years threshold was applied for the last years, from 2019 to 2022.

### 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

For the 'Forest Plantation' class, the same temporal filters of the coffee class was applied:

1) Temporal filter using 3 years with 2 years threshold applied only on the edge years (1986 and 2021). In the initial year (1985) and final year (2022) no temporal filter was applied.

2) Temporal filter using 5 years with 3 years threshold for the period 1987 to 2020.

3) Another consideration was at the end of the series. When the trees are cut, it may take a while for them to grow again and then the classifier can't identify them as 'Forest Plantation'. Since it takes 3 to 5 years for 'Forest Plantation' to become identifiable again, to solve this situation 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.

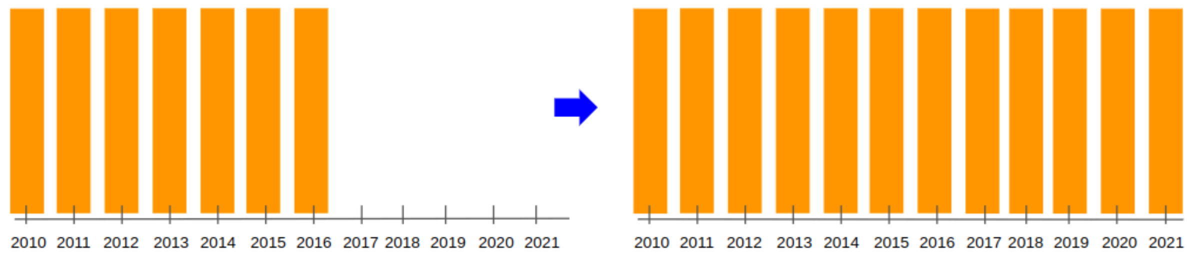



**Figure 19.** Temporal filter was applied in the last years of the 'Forest Plantation' series.

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

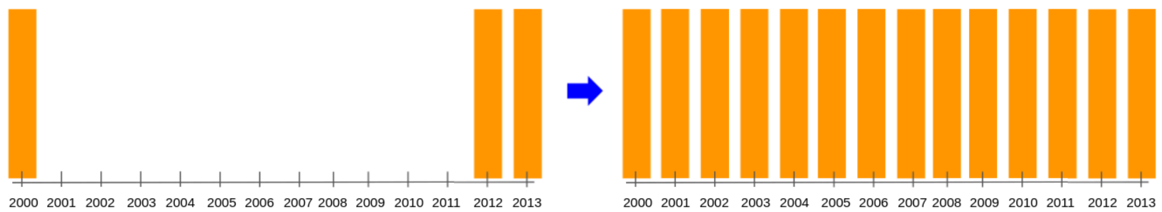



**Figure 20.** Temporal filter that converted longer intervals into 'Forest Plantation' when it was in the past and became again years after.

## 4 Integration with biomes and themes

After classifying the 'Agriculture' and 'Forest Plantation' themes, they were integrated with other land use and land cover classes to create the final maps of MapBiomas Collection 9. 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.

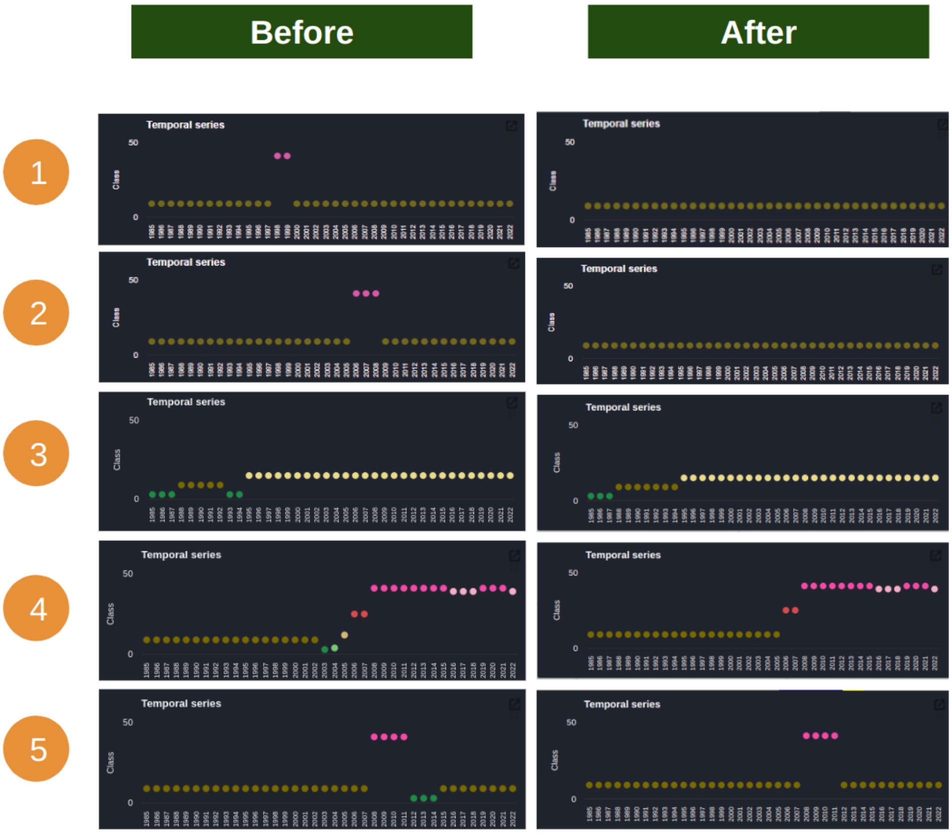



**Figure 21.** Examples of integration filters for Agriculture and Forest Plantation themes.

Examples 1 and 2 illustrate the filter's function excluding agricultural occurrences when they appear in three or fewer years, which were considered noise's classification. Examples 3, 4, and 5 demonstrate the remapping of forest formation and savanna classes 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.

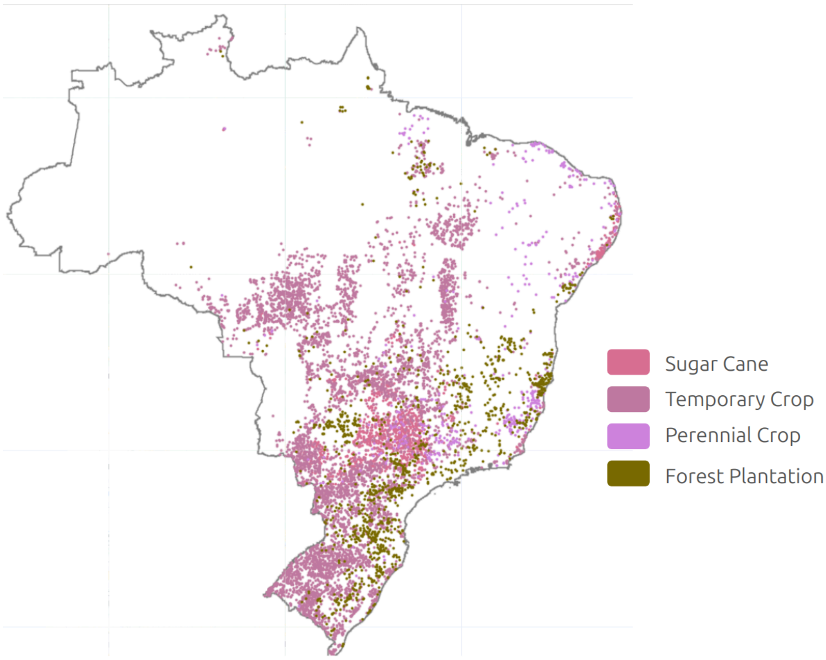



**Figure 22.** LAPIG points that were used for the accuracy assessment of 'Temporary Crops', 'Perennial Crops', and 'Forest Plantation' classes.

#### 5.1.1 Temporary Crops

The results of the accuracy assessment of the 'Temporary Crops' class are presented in Figures 23 and 24, covering the period from 1985 until 2021.

Figure 23 shows the producer's accuracy of the 'Temporary Crop' class in the Collection 9. This statistical metric indicates the portion of the pixels of 'Temporary Crop' were correctly classified. In this sense, for 'Temporary Crop' the producer's accuracy informs that over all time-series, more than 80% of 'Temporary Crop' areas were classified by MapBiomas as 'Temporary Crop'.

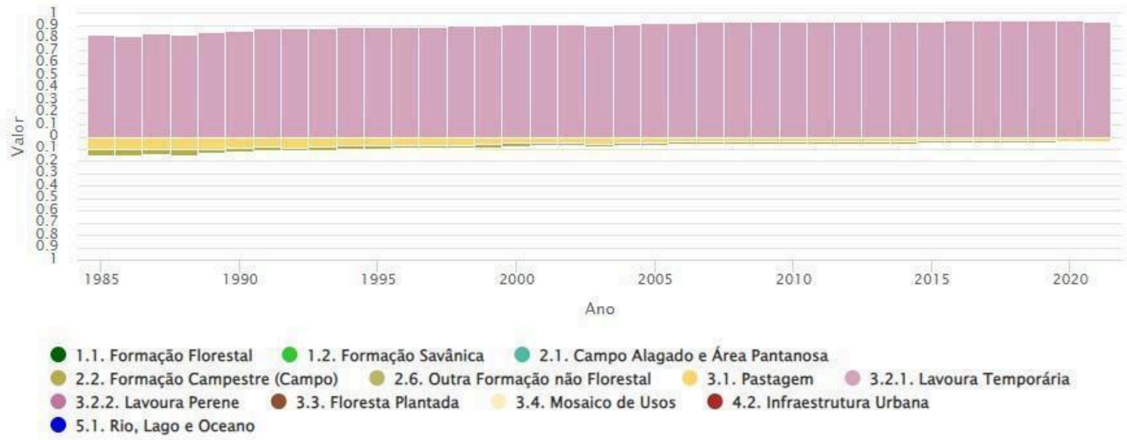



**Figure 23.** Producer's accuracy of the 'Temporary Crop' class in Collections 9.

Figure 24 presents the user's accuracy of the 'Temporary Crop' class in Collections 9. This statistical metric indicates the portion of the map pixels that were correctly classified. For 'Temporary Crop' in this Collection, the user's accuracy shows that from 2000, between 70-90% of the area mapped as 'Temporary Crop' is in fact 'Temporary Crop'. However, from 1985 until 2000 we can see lower user accuracy values, indicating that an important portion of the area mapped as 'Temporary Crop' is pasture and grassland.

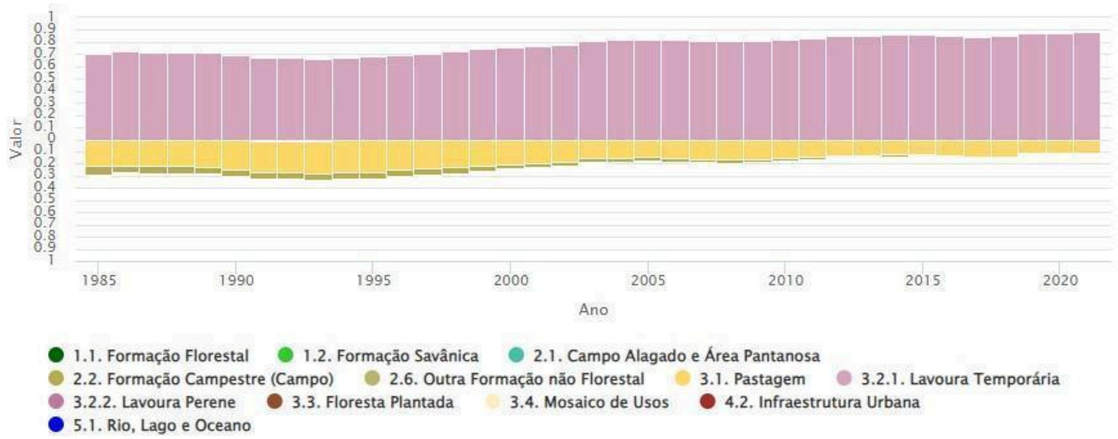



**Figure 24.** User's accuracy of the 'Temporary Crop' class in Collections 9.

### 5.1.2 Perennial Crops

The producer's accuracy for 'Perennial Crop' classes, for Collection 9, is presented by Figure 25. This metric indicates the omission of the 'Perennial Crop' classification, where we can observe higher omission at the beginning of the time-series, while from the 2000 until the end of the time-series, the mapping could identify more areas of 'Perennial Crop' classes as 'Perennial Crop'. In addition, it is important to note that the most part of these areas of 'Perennial Crop' that were not identified in the MapBiomas as 'Perennial Crop' were classified as pasture or natural forests.

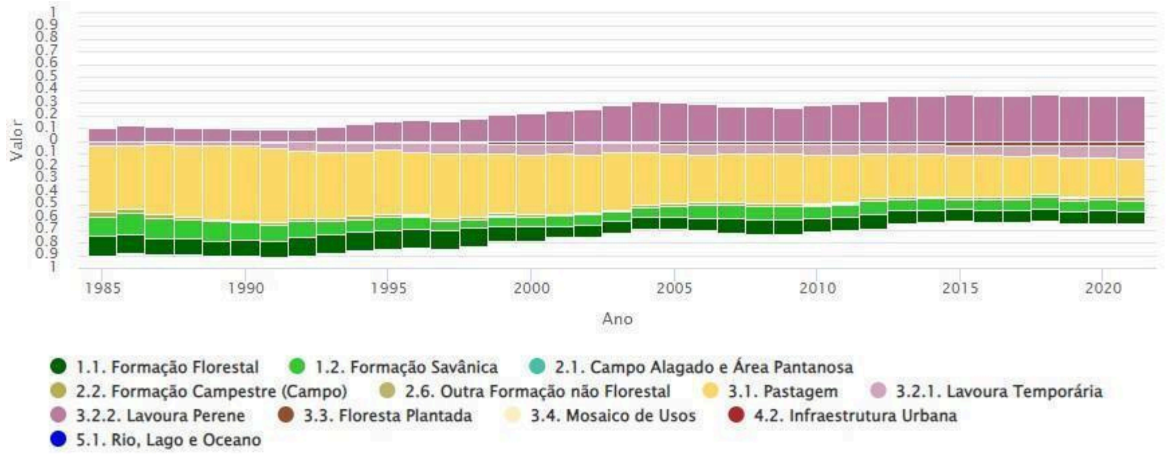



**Figure 25.** Producer's accuracy of the 'Perennial Crops' class in Collections 9.

Figure 26 presents the user's accuracy of the 'Perennial Crop' class in Collections 9. This metric indicates that at the beginning of the time-series a lower portion of the area mapped as 'Perennial Crop' is in fact 'Perennial Crop', while a higher portion of the area mapped as 'Perennial Crop' is in fact 'Perennial Crop' is pasture and natural forests. However, this metric improved over time, showing better agreements between the LAPIG points and the 'Perennial Crop'.

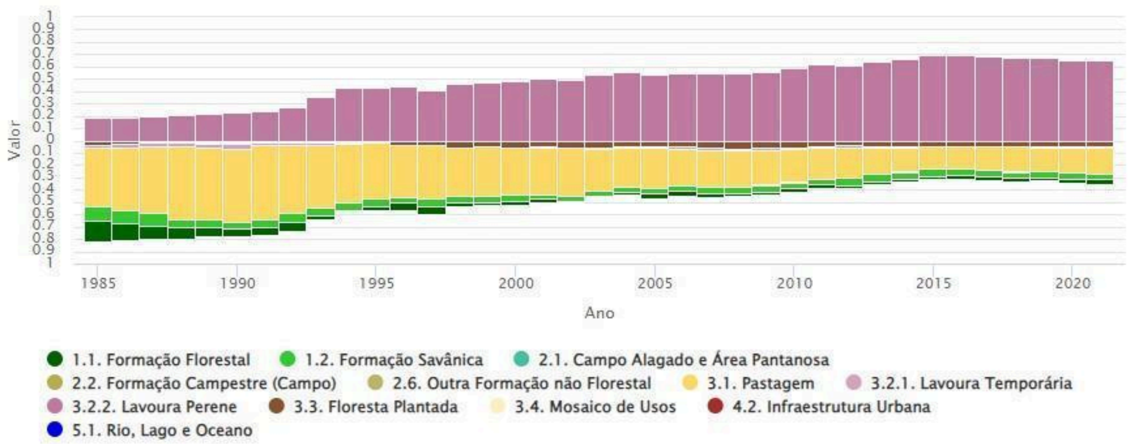



**Figure 26.** User's accuracy of the 'Perennial Crop' class in Collections 9.

**5.1.3 Forest Plantation**

Figure 27 presents the results of producer's accuracy for the 'Forest Plantation' class for Collections 9. This metric indicates the amount of the area mapped as 'Forest Plantation' is 'Forest Plantation' and the amount of the 'Forest Plantation' area was not classified as 'Forest Plantation'. Thus, it is possibly observed that in the beginning of the time-series, from 1985-1993, a higher portion of the 'Forest Plantation' area was classified as Mosaic of Uses and natural forests. However, this amount has decreased since 1993, showing that the map improved the identification of the 'Forest Plantation'.

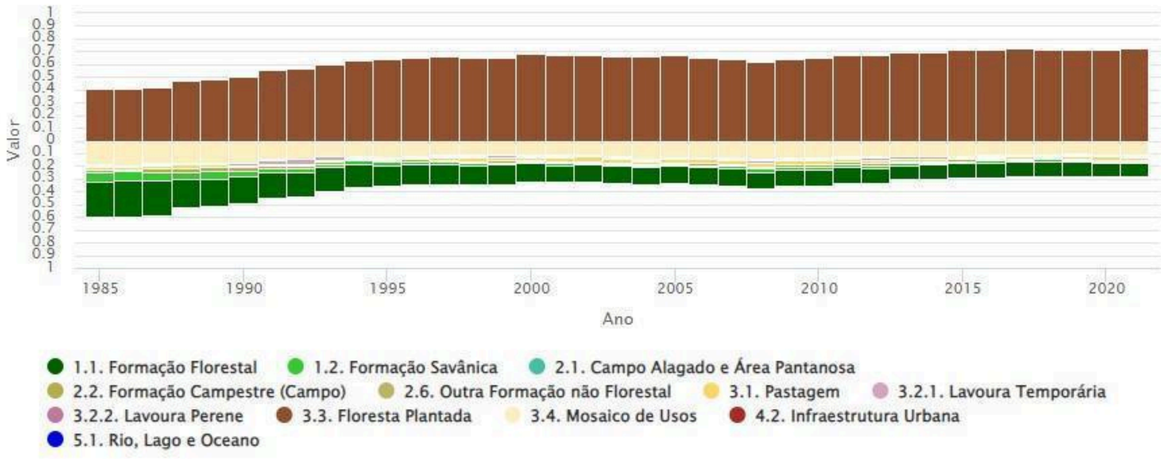



**Figure 27.** Producer's accuracy of the 'Forest Plantation' class in Collections 9.

## 5.2 Comparison with others maps

### 5.2.1 TerraClass

The aim of TerraClass Project is 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 maps produced by the MapBiomas project.

To provide the analysis, the year 2020 was 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 | MapBiomas 2020 |
|---|---|---|
| *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 2020

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

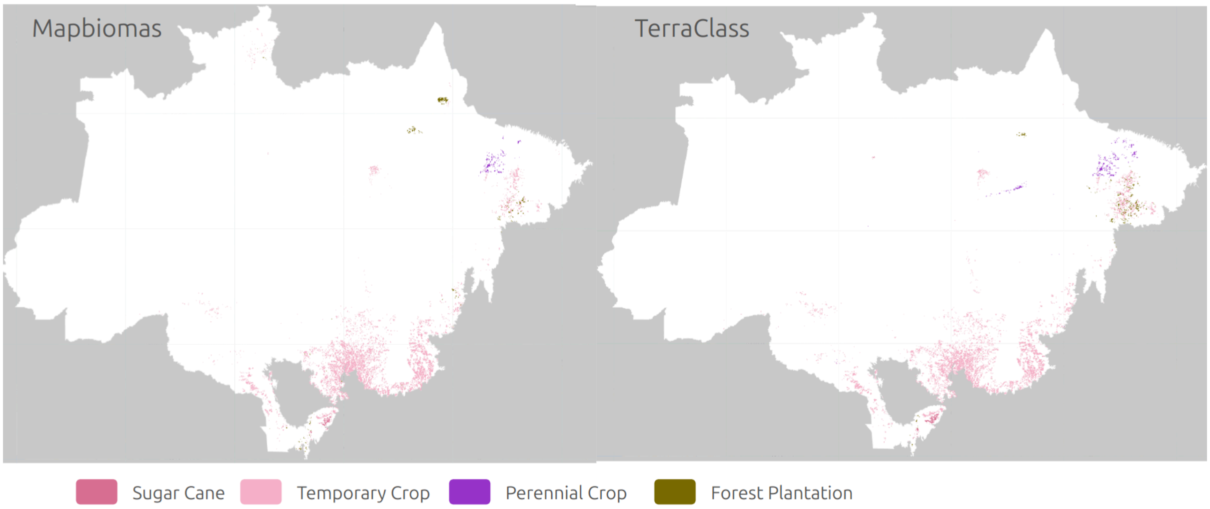



**Figure 28.** Compatible maps of MapBiomas and TerraClass maps to base year 2020.

#### 5.2.1.1.1 Area comparison

The pixel area between TerraClass and MapBiomas to the Amazon biome shows the difference between the sources.  For the classes Sugar Cane, Forest Plantation, and Perennial Crop, TerraClass presents higher values than MapBiomas. Otherwise, MapBiomas presented a Temporary Crop area 1.22Mha bigger than TerraClass, as demonstrated in Figure 29.

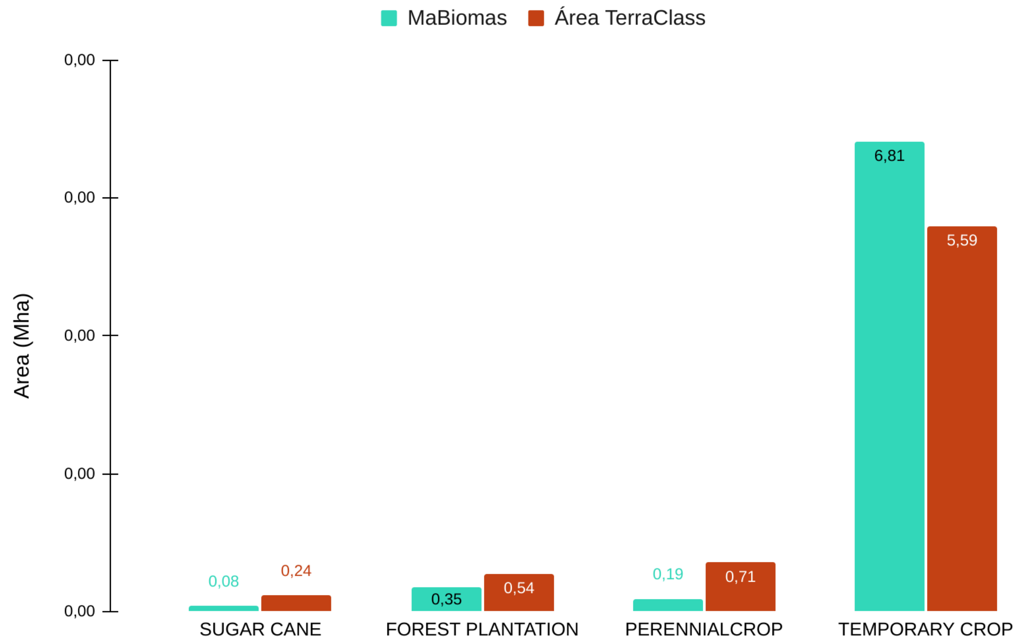
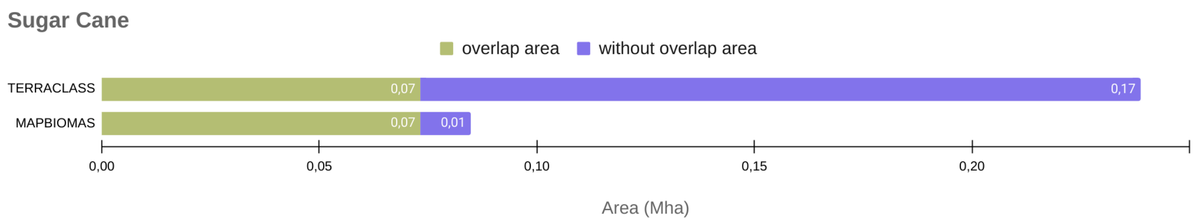
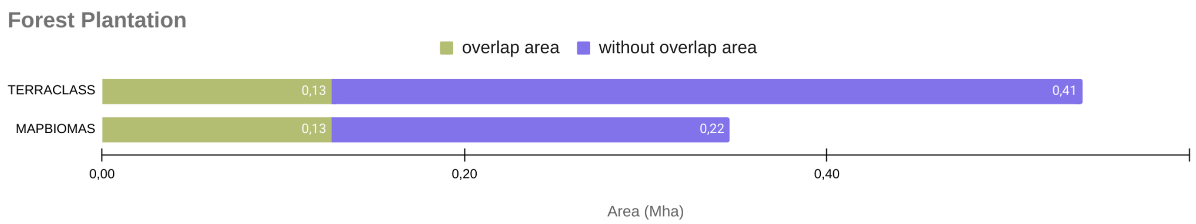



**Figure 29.** Area per classes from MapBiomas and TerraClass maps to base year 2020.

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



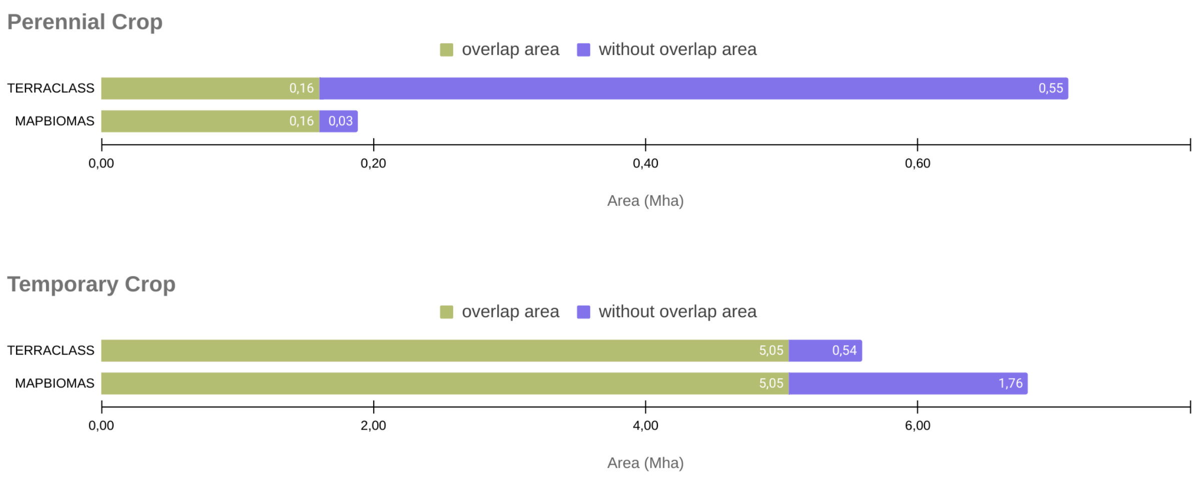



**Figure 30.** Overlap area per classes from MapBiomas and TerraClass project to 2020.

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 is mapping via visual interpretation, where pixels are grouped in a way that covers the entire class, unlike Mapbiomas, which uses pixel-by-pixel classification, which contributes to a more pixelated or noisy map.

For Sugar Cane, the TerraClass project mapped 240,000 hectares and MapBiomas mapped 80,000 hectares. This represents a difference of 160,000 hectares more mapped by TerraClass. The Forest Plantation class shows a similar trend, with 540,000 hectares mapped by the TerraClass project and 350,000 hectares by the MapBiomas project. The Perennial Crop class showed 710,000 hectares mapped by TerraClass and 190,000 hectares mapped by MapBiomas. MapBiomas showed higher area values for the Temporary Crop class, resulting in a mapping of 6.81 million hectares in 2020 compared to 5.59 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 2020.

For Sugar Cane, Forest Plantation and Perennial Crops, TerraClass showed higher producer accuracy values compared to MapBiomas. This reinforces the trend towards a larger area mapped by TerraClass, as shown in the previous topic. When analyzing user accuracy, on the other hand, MapBiomas shows higher values, i.e. when a class is mapped in MapBiomas, there is a greater chance that this class has been mapped correctly, according to the validation points used.

When the Temporary Crop class is analyzed, on the other hand, the values of the two maps are very close, but there is an inversion. TerraClass has higher user accuracy values and MapBiomas has higher Producer accuracy values. Overall, the global accuracy for the two maps for the Temporary Crop class is only 1.33%, indicating a great deal of similarity in the quality of the maps.

#### 5.2.1.2 Cerrado 2020

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

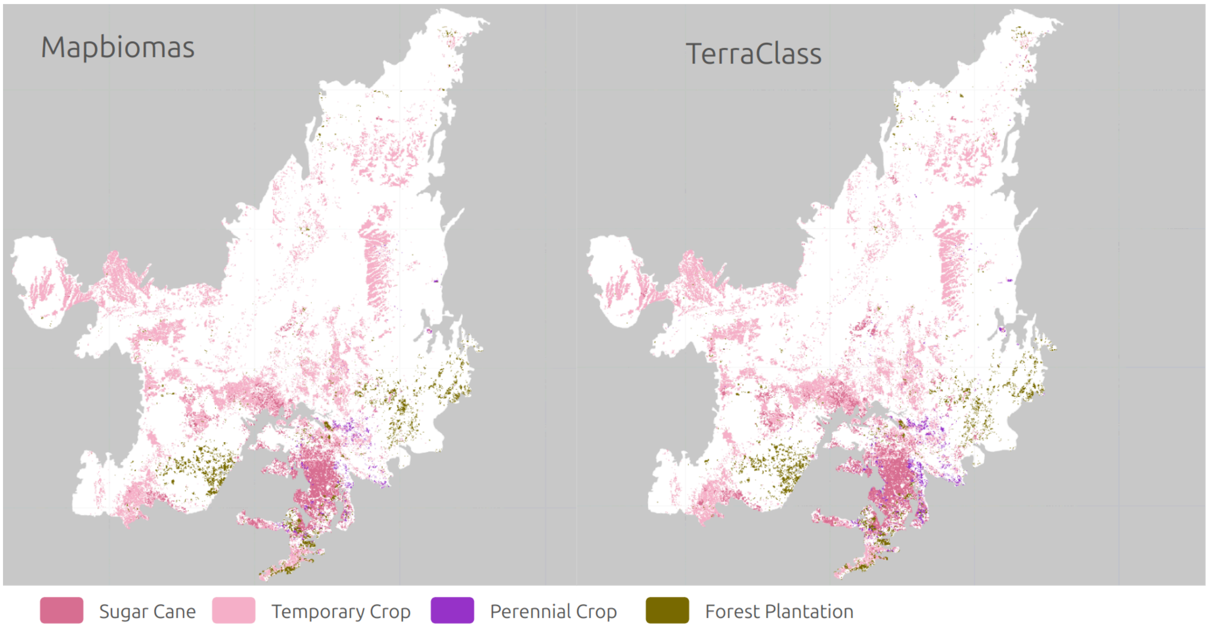


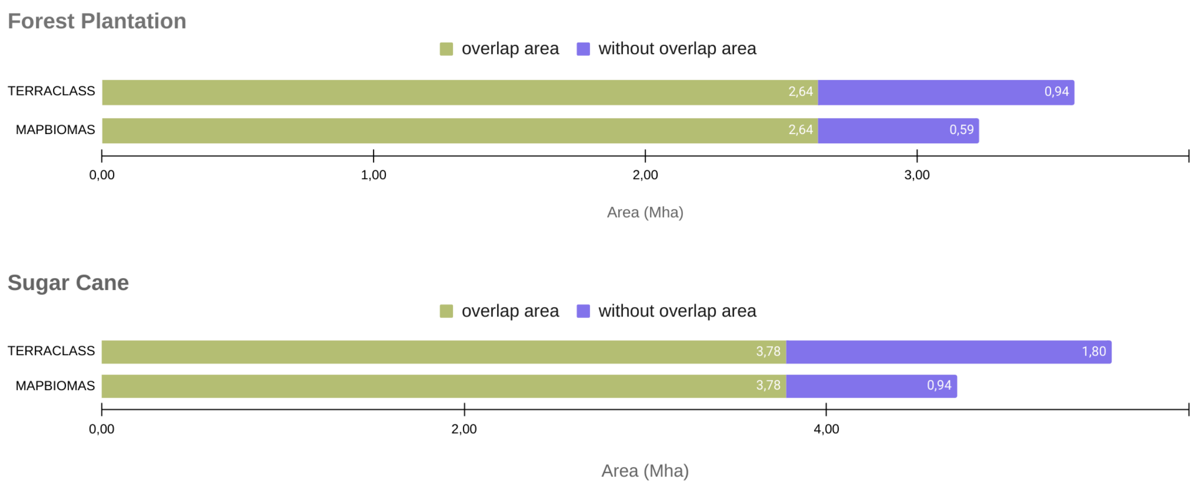

**Figure 32.** Compatible maps of MapBiomas and TerraClass maps to base year 2020.

#### 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 Sugar Cane, Forest Plantation, and Perennial Crop, TerraClass presents higher values than MapBiomas. Otherwise, MapBiomas presented a Temporary Crop area very similar to the TerraClass, as demonstrated in Figure 33.


**Figure 33.** Area per classes from MapBiomas and TerraClass maps to base year 2020.

The overlap area between the maps is shown in Figure 34



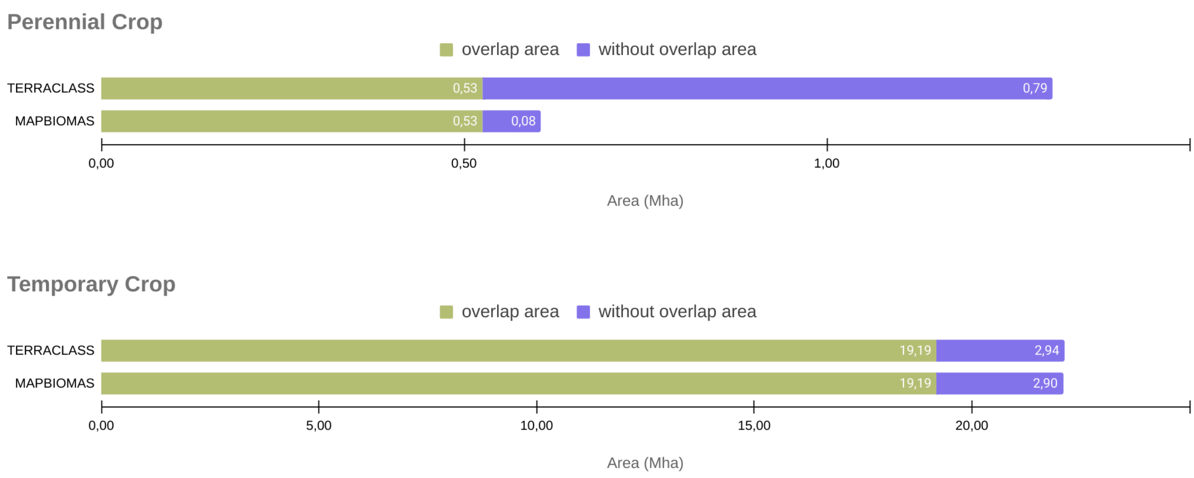



**Figure 34.** Overlap area per classes from MapBiomas and TerraClass project to 2020.

Unlike the pattern found in the Amazon biomes, in the Cerrado, there is a lot of overlap between the mappings from TerraClasse and MapBiomas. The areas obtained for the Temporary Crop class were very similar, with more than 86% of both maps having overlapping areas and the rest being areas that one mapped and the other did not, 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.

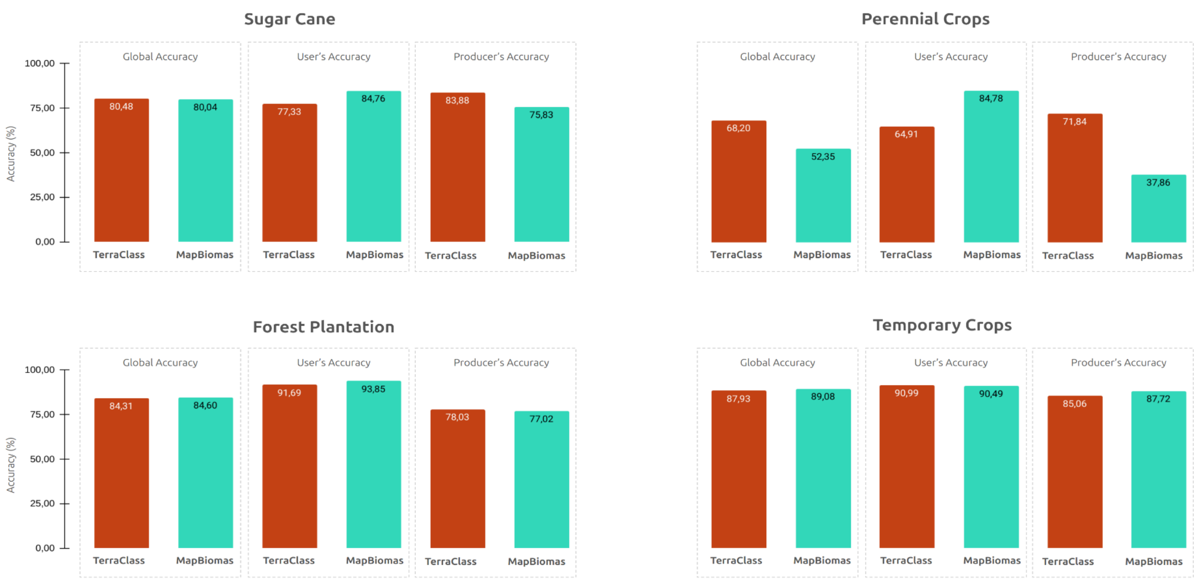



**Figure 35.** Accuracy analysis with Lapig validation points between MapBiomas and TerraClass project to 2020.

The result shows very similar results for all four classes analyzed. The biggest difference is in the Perennial Crop class, where TerraClass showed higher results for producer accuracy and global accuracy, but the values did not exceed 72%, showing that the perennial crop class is a challenge for both mappings. The rest of the classes showed some variations between producer and user accuracy, but overall global accuracy did not differ by more than 1% between the maps analyzed.

### 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, the national inventory of anthropogenic emissions by sources and anthropogenic removals by sinks of all greenhouse gases (GHG) is drawn up periodically. From 1990 to 2016, this inventory was carried out in accordance with the "2006 Intergovernmental Panel on Climate Change (IPCC) Guidelines for National Greenhouse Gas Emissions Inventories''. One of the results obtained from 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 comparison with the results 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 in a similar way to the comparison between MapBiomas and TerraClass data. The Figure 36 therefore shows the compatibility between the classes used in the analysis.

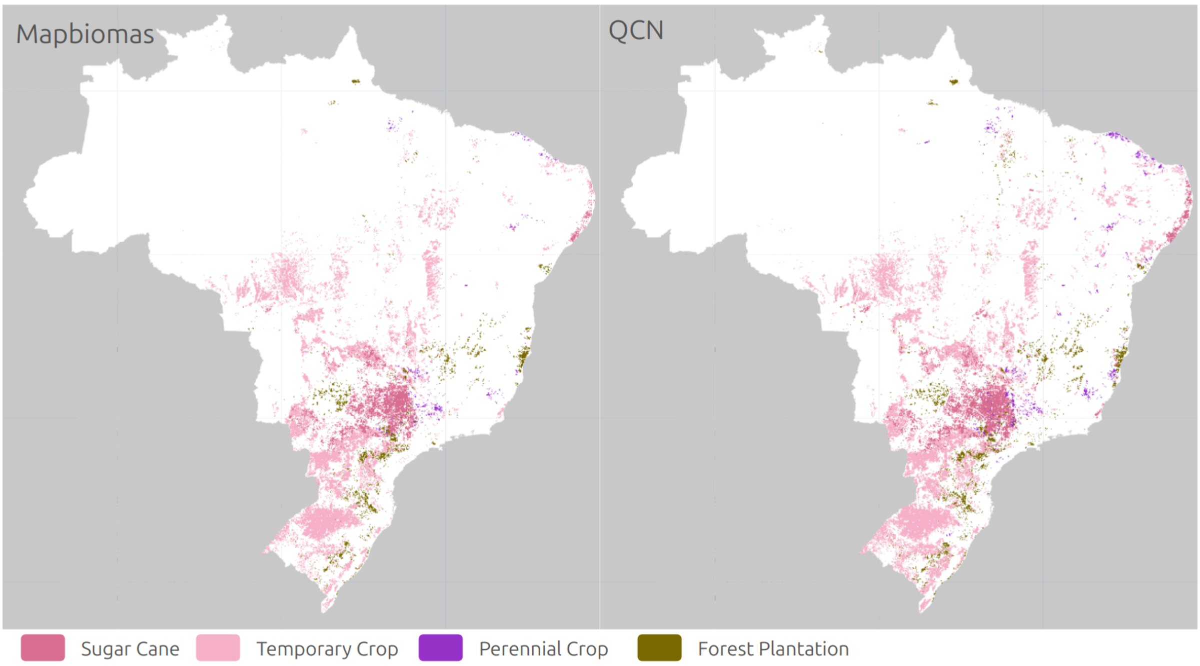



**Figure 36.** Compatible maps of MapBiomas and Fourth National Communication to the United Nations Framework Convention on Climate Change (UNFCC) maps to base year 2016.

#### 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 Table 37.

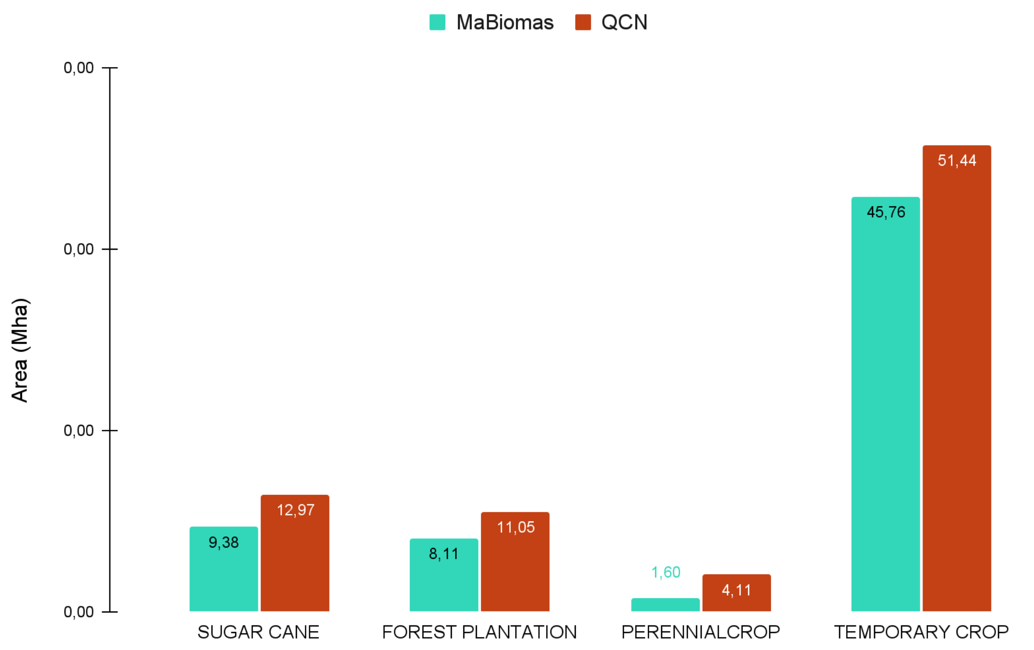
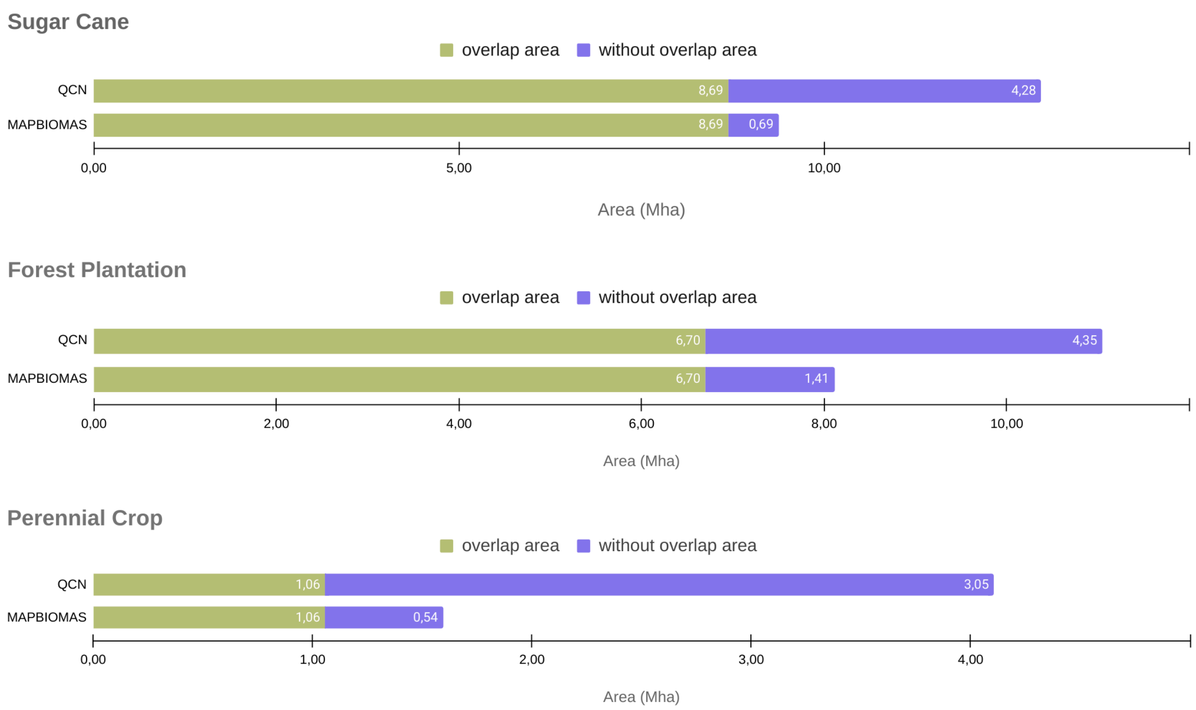



**Figure 37.** Area per classes from MapBiomas and QCN maps to base year 2016.



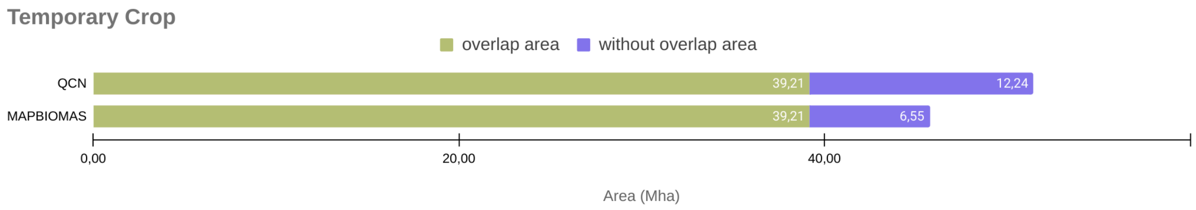



**Figure 38.** Overlap area per classes from MapBiomas and QCN project to 2016.

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, unlike Mapbiomas, which uses pixel-by-pixel classification, which contributes to a more pixelated or noisy map.

For Sugar Cane, the QCN project mapped 12.97 millions of hectares and MapBiomas mapped 9.38 millions of hectares in 2016. This represents a difference of 3.59 millions of hectares more mapped by QCN in Brazil. The Forest Plantation class shows a similar trend, with 11.05 millions hectares mapped by the QCN and 8.11 millions of hectares by the MapBiomas project. The Perennial Crop class showed 4.11 millions of hectares mapped by QCN and 1.6 million of 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.

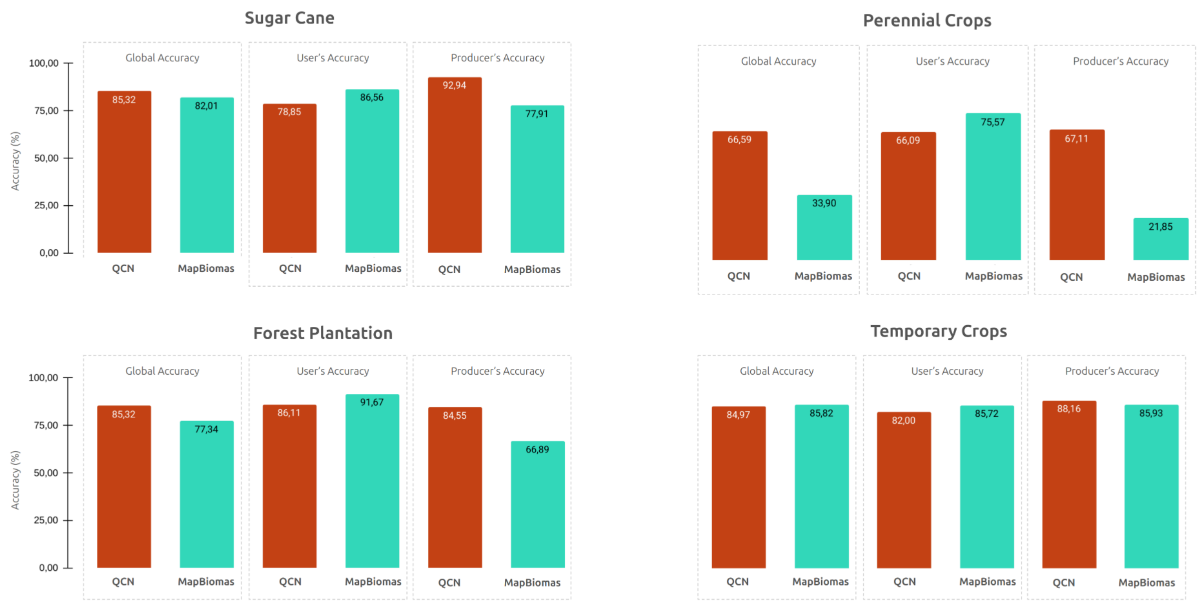



**Figure 39.** Accuracy analysis with Lapig validation points between MapBiomas and QCN project to 2016.

For all the classes analyzed, QCN maps showed higher producer accuracy results than MapBiomas, which corroborates the result of larger mapped areas shown in Figure X. 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 Mha show an overlapping area with the result of the QCN mapping, which corresponds to 86% of the temporary crops map.

These results indicate that MapBiomas and the other land use and land cover mapping sources present very similar results, which shows that the methodology adopted by the project presents promising results for land use and land cover mapping.

# 6 References


Teixeira Junior, Paulo Domingos Teixeira; et al. OBTENÇÃO DE CALENDÁRIO AGRÍCOLA BASEADO EM SÉRIE TEMPORAL DE ÍNDICE DE VEGETAÇÃO. In: ANAIS DO XX SIMPóSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 2023, 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> Acesso em: 15 Jul. 2024.

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

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

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>.

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

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

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