Agriculture and Forest Plantation - Appendix

Collection 6

General coordinator
Bernardo Rudorff

Team
Djonathan Goulart
Kênia Santos
Luciana Oliveira
Marciano Saraiva
Moisés Salgado

1 Overview of the classification method

Mapping 'Agriculture' and 'Forest Plantation' emerged as one of the challenges of the MapBiomas project. The first challenge was in Collection 1, mapping Agriculture and Forest Plantation from 2008 to 2015 in a short period to prove the innovative concept of the project: the production of cheaper, faster, and updated annual maps of coverage and land use for Brazil's territory compared to the methods and practices applied so far. Based on the results from Collection 1, Agrosatelite's team adopted a more appropriate approach for the classification of agriculture. The algorithm developed for the classification of annual and semi-perennial agriculture in MapBiomas Collection 2 (2000 - 2016) incorporated each region's season and off-season periods in Brazil. This algorithm selects the Landsat images available in each scene's specific season period and creates a mosaic from these images. In addition, Collection 2 used the EVI2 and CEI vegetation indexes to train the Random Forest classifier (BREIMAN, 2001).

In Collection 3, the methodology was reformulated. A new approach to obtain metrics was adopted: the use of reducers (minimum, maximum, median, standard deviation, and quality mosaic) applied to the vegetation indexes and spectral bands. A total of 178 bands were created for each annual mosaic. From these bands, we selected those that presented the classifier's best response for each class (more details on the selection of the bands are shown in the topics below). This approach has been used in Collections 4, 5 and 6 for the classes mapped by Random Forest algorithm. Specifically for MapBiomas Collection 5 and 6, the most important methodological change was the use of a normalized Landsat series based on Modis data. The normalization of the images provides a series with similar spectral characteristics, thus allowing the use of samples of only one year for training the model and improving the final quality of the classification. Another improvement in cCollection 6 is the coffee map (as perennial crop) and the use of Deep Learning to map rice and citrus classes.

In MapBiomas Collection 6 for the cross-cutting themes 'Agriculture' and 'Forest Plantation' in the Brazilian territory from 1985 to 2020, some improvements were added, especially the addition of new classes, such as soybean class for all MapBiomas temporal series (from 1985 to 2020), rice (irrigated only), coffee and citrus (São Paulo state only) classes, both as beta version, as well as improvement of 'other perennial crops' maps. The evolution of themes 'Agriculture' and 'Forest Plantation' across MapBiomas collections is illustrated in Figure 1.

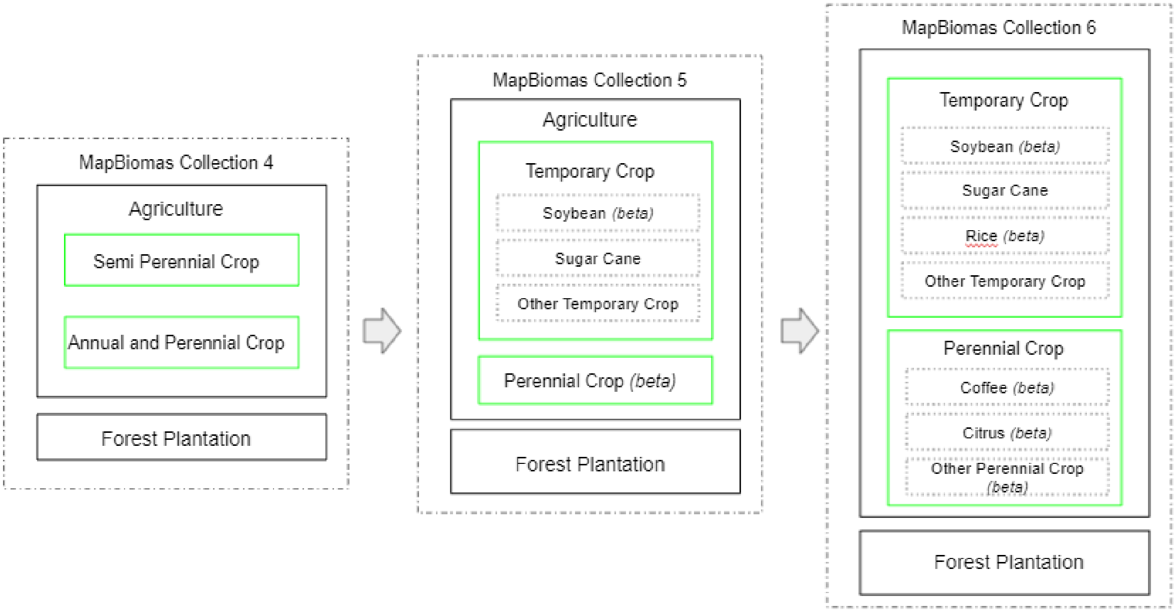


Figure 1. Agriculture and Forest Plantation classes mapped in the MapBiomas Collections 4, 5 and 6.

2 Classification

The MapBiomas-brazil count in GitHub has all the scripts used to classify 'Agriculture' and 'Forest Plantation' classes in MapBiomas Collection 6. The repository link is:

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

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

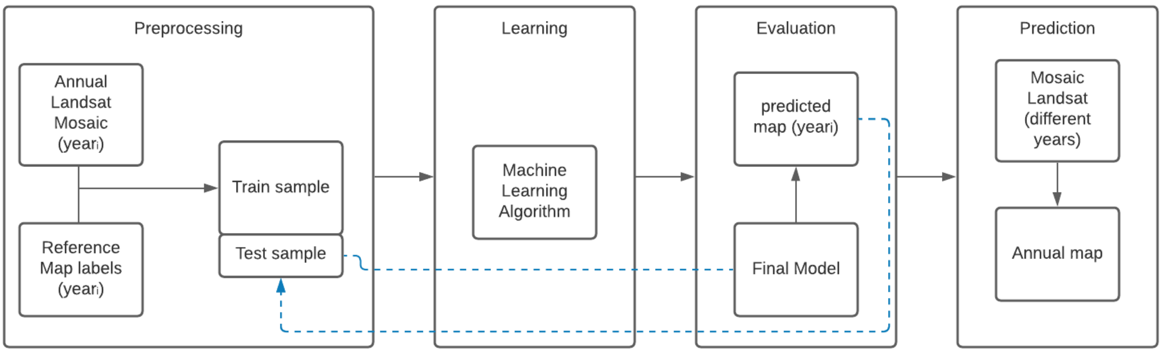


Figure 2. Supervised learning workflow in context of image classification.

The preprocessing step and prediction are the same for both algorithms used in Agriculture and Forest Plantation mapping (i.e. Random Forest and Neural Network). The learning and evaluation steps are specific according to each of the algorithms. The annual rice and citrus 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

2.1.1 Landsat Images availability

The Landsat images availability in Collection 6 period (1985 to 2020) varies among years. Throughout this period, Landsat 5 (1985 to 2012), Landsat 7 (1999 to present), and Landsat 8 (2013 to present) provided the images for the mosaics compositions. Figure 3 shows the variability of available Landsat images for Collection 6 period.

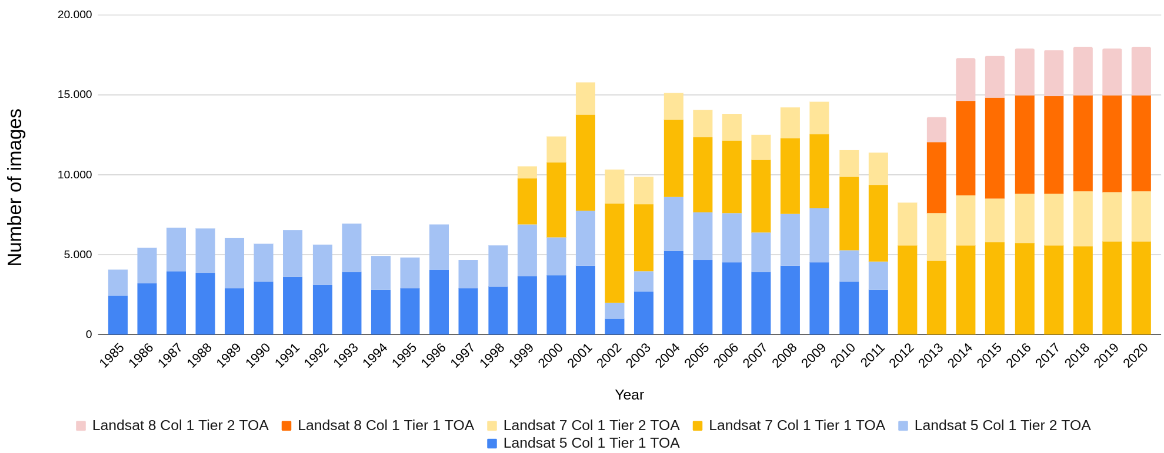


Figure 3. Number of available TOA Landsat images covering the Brazilian territory from 1985 to 2020.

2.1.2 Image selection

To classify the cross-cutting theme of 'Agriculture' and 'Forest Plantation' in Collection 6, in addition to the Landsat images used in previous collections (available on the Google Earth Engine platform), a Landsat normalized time series was created based on the reflectance data from Moderate-Resolution Imaging Spectroradiometer (MODIS). The normalization of reflectance is an important step to guarantee the spectral similarity of the same types of land cover (for more information, see Potapov *et. al.* (2020)). Both collections of Landsat data were used in MapBiomas Collection 6. Table 1 shows the collection used for each class of agriculture and forest plantation classification.

Table 1: Landsat collection used for each class of agriculture and forest plantation classification.

Level 1	Level 2	Level 3	Level 4	Landsat Collection
Farming	Agriculture	Temporary Crop	Soybean	Normalized Landsat Collection
			Sugar Cane	Landsat ToA Collection
			Rice	Landsat ToA Collection
			Other Temporary Crops	Normalized Landsat Collection
		Perennial Crop	Coffee	Normalized Landsat Collection
			Citrus	Normalized Landsat Collection
			Other Perennial Crops	Landsat ToA Collection
	Forest Plantation		Forest Plantation	Landsat ToA Collection

2.1.3 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 theme were taken into account to better distinguish the class of interest from the remaining land cover and land use classes.

The acquisition of Landsat images (TOA and normalized) to compose the mosaics for classification of 'Other Temporary Crops' and 'Soybean' was carried out according to the crop season calendar in six regions in Brazil (Figure 4). The Landsat mosaics used to classify agriculture were built to highlight the seasonal change observed between the season and off-season.

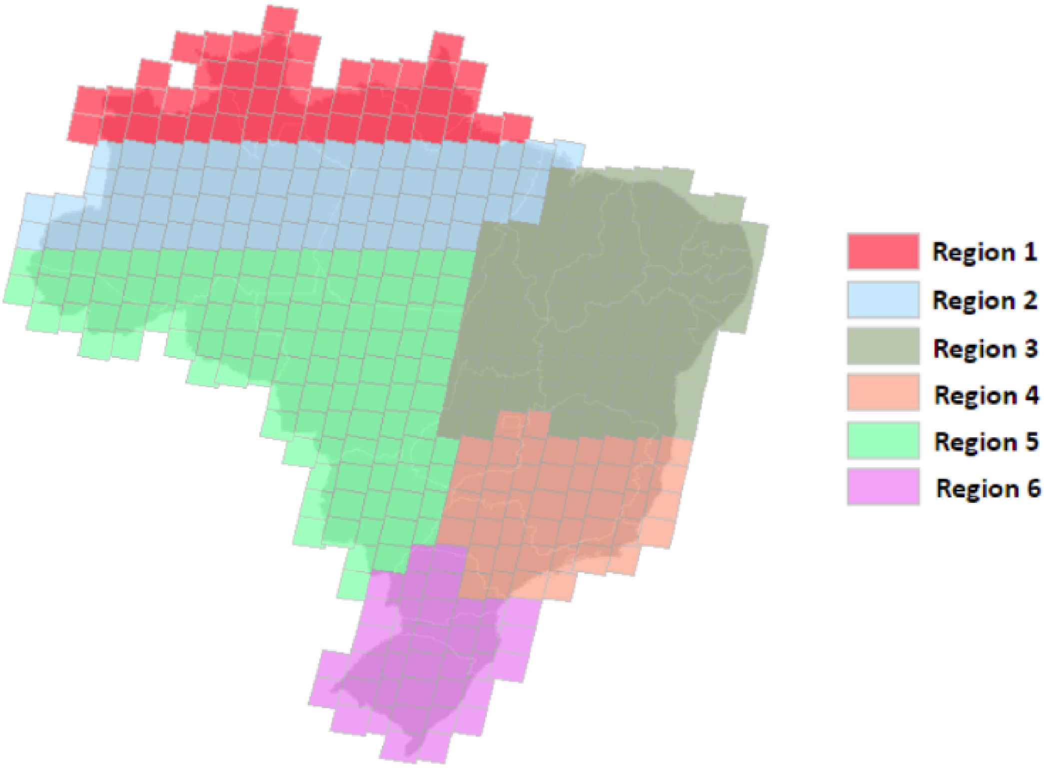


Figure 4. Regional crop calendar differences in Brazil that are considered to build the Landsat mosaics in the classification of the seasonal crops in the MapBiomas Collections.

The off-season also has important characteristics that help to distinguish crops. Season and off-season crop periods used to filter Landsat image collections to classify agriculture in the Collection 6 are shown in Table 2.

Table 2. Periods for the temporal composition of the Landsat mosaics used for the regional crop classification in MapBiomas Collection 6.

Region	season start	season end	off-season start	off-season end
Region 1	04/01/Year	08/31/Year	11/01/Year-1	03/31/Year
Region 2	02/01/Year	06/30/Year	09/01/Year-1	02/15/Year
Region 3	11/15/Year-1	06/15/Year	10/15/Year-1	12/15/Year-1
Region 4	10/15/Year-1	06/15/Year	05/15/Year-1	11/15/Year-1
Region 5	11/01/Year-1	06/30/Year	05/01/Year-1	10/31/Year-1
Region 6	10/01/Year-1	11/15/Year	01/01/Year-1	04/30/Year-1

Table 2 specifies the periods per region for the Landsat images selection in each year to classify seasonal crops. The Fmask (i.e. Function of mask) (Zhu et al., 2014) algorithm was used to detect cloud and cloud shadow in Landsat Images before the composition of the mosaic. Additionally, images from previous years were used to increase the likelihood of cloud-free images acquisition during the seasonal crop to better map the crop fields and, therefore, improve the quality of the classification and final maps.

2.1.3.1 Soybean

The life cycle of plants is an important factor for monitoring and mapping agricultural crops via Remote Sensing, mainly the short-cycle crops (e.g. two to six months) (FORMAGGIO and SANCHES, 2017), like soybeans crops. The main challenge of this class was the definition of metrics to distinguish this crop from others short-cycle crops. As there are specific months in which soy is usually cultivated, temporal metrics were used to differentiate soy from other short-cycle crops. These metrics were obtained at regular intervals and adapted to the agricultural calendar of each region (Table 3).

Table 3. Period of temporal metrics according to the agricultural calendar of each region.

Period	Region 1	Region 2	Region 3	Region 4	Region 5	Region 6
Month 1	04/01/Year to 05/01/Year	02/01/Yea to 03/01/Year	11/15/Year-1 to 12/15/Year-1	10/15/Year-1 to 11/15/Year-1	11/01/Year-1 to 12/01/Year-1	10/15/Year-1 to 11/15/Year-1
Month 2	05/01/Year to 06/01/Year	03/01/Year to 04/01/Year	12/15/Year-1 to 01/15/Year	11/15/Year-1 to 12/15/Year-1	12/01/Year-1 to 01/01/Year	11/15/Year-1 to 12/15/Year-1
Month 3	06/01/Year to 07/01/Year	04/01/Year to 05/01/Year	01/15/Year to 02/15/Year	12/15/Year-1 to 01/15/Year	01/01/Year to 02/01/Year	12/15/Year-1 to 01/15/Year
Month 4	07/01/Year to 08/01/Year	05/01/Year to 06/01/Year	02/15/Year to 03/15/Year	01/15/Year to 02/15/Year	02/01/Year to 03/01/Year	01/15/Year to 02/15/Year
Month 5	08/01/Year to 09/01/Year	06/01/Year to 07/01/Year	03/15/Year to 04/15/Year	02/15/Year to 03/15/Year	03/01/Year to 04/01/Year	02/15/Year to 03/15/Year
First three observations	04/01/Year to 05/19/Year	02/01/Year to 03/21/Year	11/15/Year-1 to 01/02/Year	10/15/Year-1 to 12/02/Year-1	11/01/Year-1 to 12/19/Year-1	10/15/Year-1 to 12/02/Year-1
Last three observations	07/15/Year to 09/01/Year	05/14/Year to 07/01/Year	02/26/Year to 04/15/Year	01/26/Year to 03/15/Year	02/12/Year to 04/01/Year	01/26/Year to 03/15/Year

The normalized Landsat images of the periods described in Table 3 were used to generate image mosaics for each year. Additional images from the same periods from two previous years were used to provide better results.

2.1.3.2 Sugar cane and Forest Plantation

The classes 'Sugar cane' and 'Forest Plantation' used Landsat mosaics 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 4.

Table 4. Periods used for the selection of mosaic images of sugar cane and forest plantation in Collection 6.

Period	Start	End
season 1	12/01/year-1	01/31/year
season 2	02/01/year	03/31/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
season 3	10/01/year	11/30/year

The Landsat images of the periods described in Table 4 were used to generate image mosaics for each year in addition to images from the same periods from two previous years. Using three years provided better results when compared to the biannual mosaics.

2.1.3.3 Rice

The selection of images was made based on the season period according to the year of mapping carried out in each state (i.e. in Rio Grande do Sul the mapping season was 2019/2020).

Table 5. Periods used for the selection of mosaic images of rice in Collection 6.

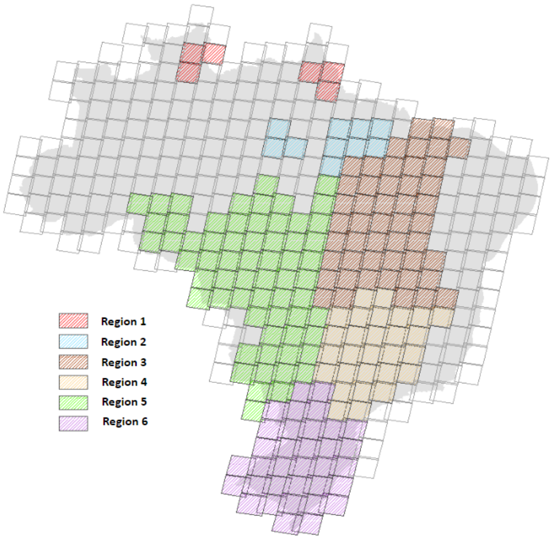
State	season start	season end	off-season start	off-season end
TO	04/01/year-1	07/30/year	08/01/year-1	11/01/year-1
RS	10/01/year-1	04/01/year	01/10/year-1	01/01/year
SC/PR	10/01/year-1	04/30/year	01/01/year	07/30/year

2.1.3.4 Perennial Crop

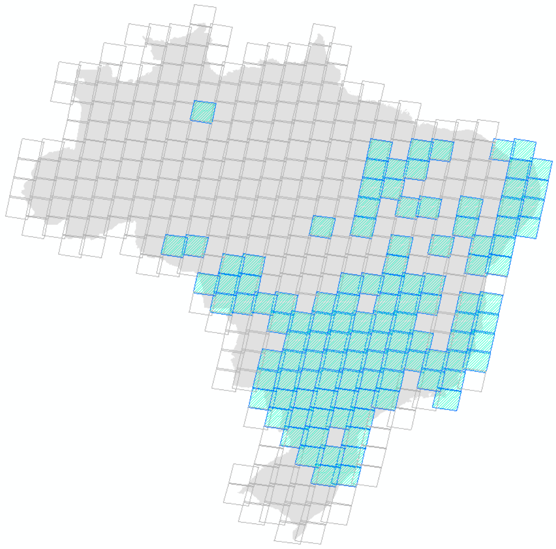
Due to the quantity and complexity of perennial crops existing in Brazil (e.g. coffee, orange, banana, oil palm), from Collection 6 each type of perennial crop will be mapped separately. Thus, an effort was made to train the classifier to specific classes. Therefore, at the first moment, the Perennial Crop class was divided into three subclasses: 'Coffee', ‘Citrus’ and ‘Other Perennial Crop’. The last one doesn’t distinguish between types of crops. For the Coffee and Citrus classification, a median of annual mosaic (i.e. 01-01-year to 12-31-year) was obtained, and the period defined for the classification of the ‘Other Temporary Crop’ class was the same period of the seasonal crop shown in Table 1.

2.1.4 Definition of regions for classification

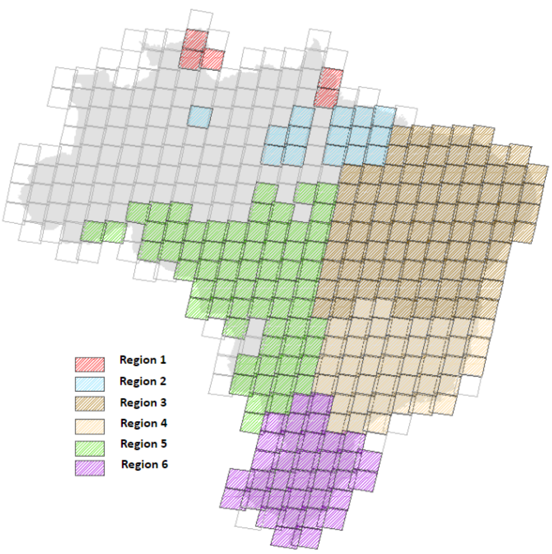
The agriculture and forest plantation are heterogeneously distributed in the Brazilian biomes. Therefore Landsat scenes were selected in regions with the highest occurrence of each class according to the reference maps. Figure 5 illustrates the scenes chosen for each land use class.



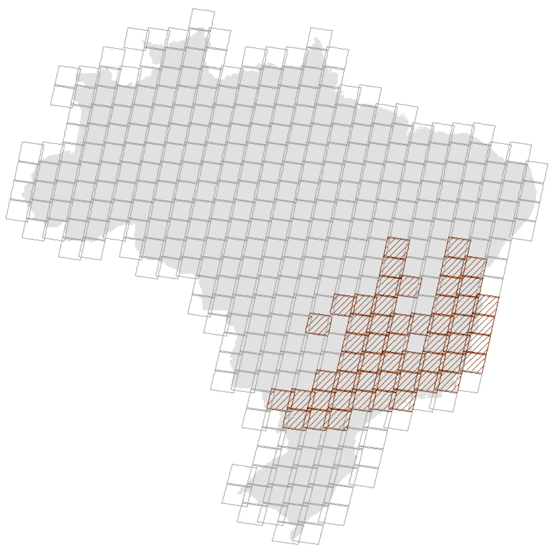
Soybean



Sugar cane



Other Temporary Crop



Coffee

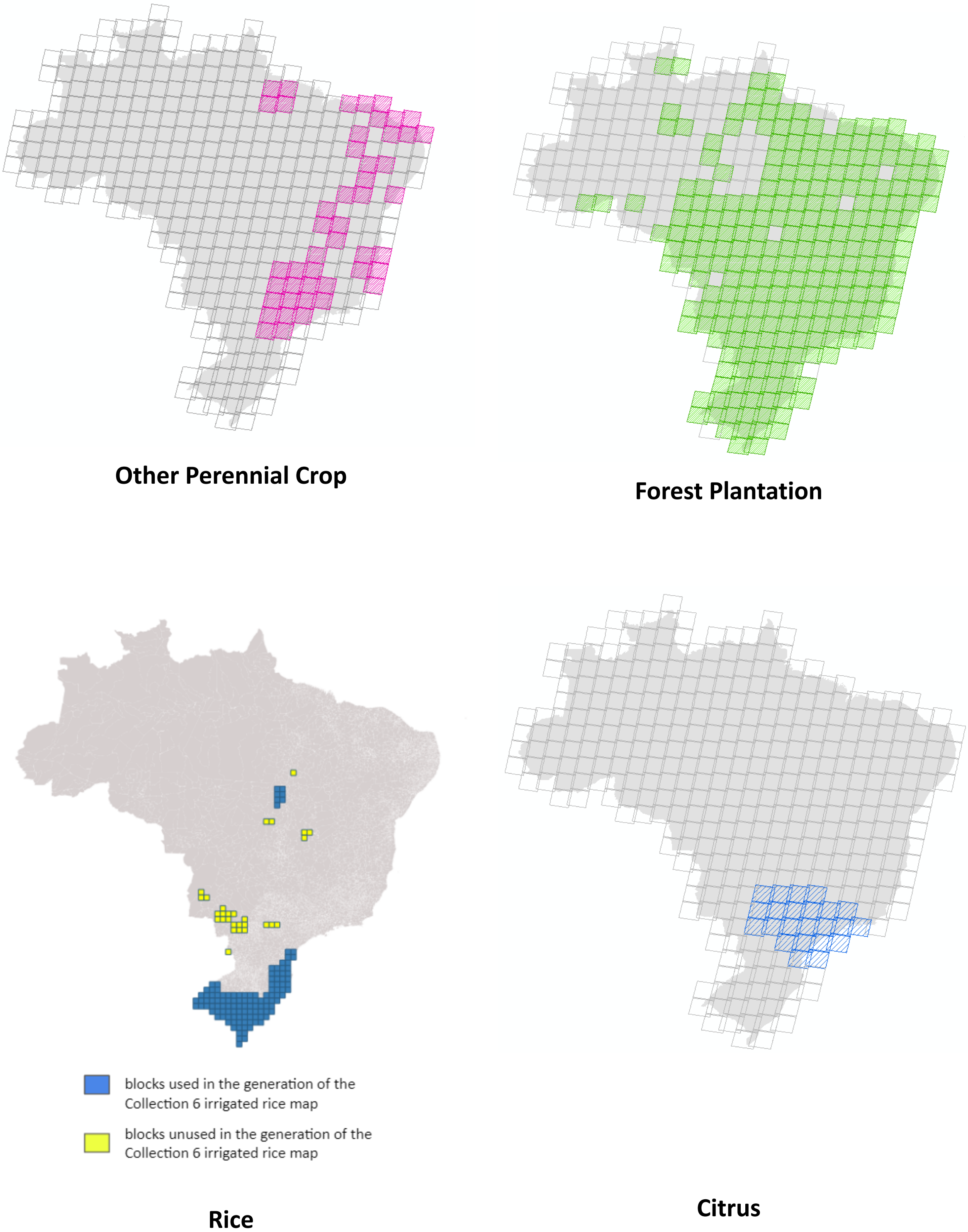


Figure 5. Selected scenes of Landsat series to the classification of maps by land use class.

The rice, coffee and citrus maps (all in beta version) do not cover the total spatial distribution of their crops in Brazil, due in the first moment, the focus was to carry out a proof of concept on the possibility of mapping these crops for all years of the time series.

2.1.5 Feature space

The feature space for ‘Other temporary crops’ in the Collection 6 was composed of Landsat bands, and reducers (minimum, maximum, median, standard deviation and quality mosaic) calculated for each spectral indices presented in Table 6, and each band, resulting in 178 variables.

Table 6. Relation of spectral indices used to classify Agriculture and Forest Plantation classes in Collections 4 to 6.

Index	Expression	Reference
EVI2	2.5 * ((NIR - RED) / (NIR + 2.4*RED + 1))	Jiang et al, 2008
NDVI	(NIR - RED) / (NIR + RED)	Deering, 1978
NDWI	(NIR - SWIR1) / (NIR + SWIR1)	Gao, 1996
CAI	SWIR2 / SWIR1	Nagler et al, 2003
LAI	0.3977 * exp(2.5556 * (NIR - RED)/(NIR + RED))	Chen et al, 2012
CEI	(season_EVI2_max - off-season_EVI2_min) / (season_EVI2_max + off-season_EVI2_min)	Rizzi et al., 2009

The method proposed by Kursa *et al*. (2010) was applied to select all relevant data and reduce the number of variables used in the classification model for ‘other temporary crop’ class. The 50 metrics with greater relevance are shown in Table 7.

Table 7. Metrics used to classify other temporary crops in Collection 6.

Id	Variable	Description	Statistics	Period
1	season_GREEN_min	Landsat Green band minimum value	minimum	season
2	season_GREEN_median	Landsat Green band median value	median	season
3	season_RED_max	Landsat Red band maximum value	maximum	season
4	season_RED_median	Landsat Red band median value	median	season
5	season_RED_stdDev	Landsat Red band standard deviation value	standard deviation	season
6	season_NIR_qmo	Landsat NIR band selected based on maximum EVI2	maximum	season
7	season_SWIR1_qmo	Landsat SWIR1 band selected based on maximum EVI2	maximum	season
8	season_SWIR1_max	Landsat SWIR1 band maximum value	maximum	season
9	season_SWIR1_min	Landsat SWIR1 band minimum value	minimum	season
10	season_SWIR1_median	Landsat SWIR1 band median value	median	season
11	season_SWIR1_stdDev	Landsat SWIR1 band standard deviation value	standard deviation	season

12	season_SWIR2_max	Landsat SWIR2 band maximum value	maximum	season
13	season_SWIR2_median	Landsat SWIR2 band median value	median	season
14	season_SWIR2_stdDev	Landsat SWIR2 band standard deviation value	standard deviation	season
15	season_TIR1_max	Landsat TIR1 band maximum value	maximum	season
16	season_TIR1_stdDev	Landsat TIR1 band standard deviation value	standard deviation	season
17	season_EVI2_max	Spectral index EVI2 maximum value	maximum	season
18	season_EVI2_min	Spectral index EVI2 minimum value	minimum	season
19	season_EVI2_stdDev	Spectral index EVI2 standard deviation value	standard deviation	season
20	season_NDWI_max	Spectral index NDWI maximum value	maximum	season
21	season_NDWI_min	Spectral index NDWI minimum value	minimum	season
22	season_NDWI_median	Spectral index NDWI median value	median	season
23	season_NDWI_stdDev	Spectral index NDWI standard deviation value	standard deviation	season
24	season_CAI_max	Spectral index CAI maximum value	maximum	season
25	season_CAI_min	Spectral index CAI minimum value	minimum	season
26	season_CAI_median	Spectral index CAI median value	median	season
27	season_CAI_stdDev	Spectral index CAI standard deviation value	standard deviation	season
28	off-season_GREEN_median	Landsat Green band median value	median	off-season
29	off-season_RED_max	Landsat Red band maximum value	maximum	off-season
30	off-season_RED_min	Landsat Red band minimum value	minimum	off-season
31	off-season_RED_median	Landsat Red band median value	median	off-season
32	off-season_SWIR1_max	Landsat SWIR1 band maximum value	maximum	off-season
33	off-season_SWIR1_median	Landsat SWIR1 band median value	median	off-season
34	off-season_SWIR1_stdDev	Landsat SWIR1 band standard deviation value	standard deviation	off-season
35	off-season_SWIR2_max	Landsat SWIR2 band maximum value	maximum	off-season
36	off-season_SWIR2_min	Landsat SWIR2 band minimum value	minimum	off-season
37	off-season_SWIR2_median	Landsat SWIR2 band median value	median	off-season
38	off-season_EVI2_min	Spectral index EVI2 minimum value	minimum	off-season
39	off-season_EVI2_median	Spectral index EVI2 median value	median	off-season
40	off-season_NDWI_qmo	Spectral index NDWI selected based on maximum EVI2	maximum	off-season
41	off-season_NDWI_max	Spectral index NDWI maximum value	maximum	off-season
42	off-season_NDWI_min	Spectral index NDWI minimum value	minimum	off-season
43	off-season_NDWI_median	Spectral index NDWI median value	median	off-season
44	off-season_NDWI_stdDev	Spectral index NDWI standard deviation value	standard deviation	off-season
45	off-season_CAI_qmo	Spectral index CAI selected based on maximum EVI2	maximum	off-season
46	off-season_CAI_max	Spectral index CAI maximum value	maximum	off-season
47	off-season_CAI_min	Spectral index CAI minimum value	minimum	off-season
48	ANNUAL_NIR_cei	Spectral index CEI using NIR band instead EVI2	normalized difference	ANNUAL

49	ANNUAL_EVI2_cei	Spectral index CEI value	normalized difference	ANNUAL
50	ANNUAL_NDWI_cei	Spectral index CEI using NDWI band instead EVI2	normalized difference	ANNUAL

2.1.5.1 Soybean

In addition to the variables shown in Table 7, temporal metrics for soybean classification were added. These time metrics were based on King et al. (2017) and are shown in Table 8.

Table 8. Metrics added to the classification of soybean crops in MapBiomas Collection 6.

Period	Bands and Spectral Indexes	Statistical Metrics
Month 1	RED, NIR, SWIR1, SWIR2, NDVI, NDWI, EVI2	Average
Month 2		
Month 3		
Month 4		
Month 5		
First three observations		Median and Average
Last three observations		

These metrics aimed to obtain spectral characteristics in the phenological period of the crops, which is a relevant characteristic for the distinction between them. A total of 104 layers were used to compose the mosaics for the classification of soybean crops (50 metrics selected for the classification of temporary crops in Collection 6, and 54 metrics added for the classification of soybean).

2.1.5.2 Sugar cane

The metrics used to classify 'Sugar cane' are shown in Table 9.

Table 9. Metrics used to classify sugar cane in MapBiomas Collection 6.

Id	Variable	Description	Statistics	Period
1	season1_GREEN_median	Landsat Green band median value	median	season1
2	season1_RED_median	Landsat Red band median value	median	season1
3	season1_NIR_median	Landsat NIR band median value	median	season1
4	season1_SWIR1_median	Landsat SWIR1 band median value	median	season1
5	season1_SWIR2_median	Landsat SWIR2 band median value	median	season1
6	season1_NDVI_median	Spectral index NDVI median value	median	season1
7	season1_NDWI_median	Spectral index NDWI median value	median	season1

8	season2_GREEN_median	Landsat Green band median value	median	season2
9	season2_RED_median	Landsat Red band median value	median	season2
10	season2_NIR_median	Landsat NIR band median value	median	season2
11	season2_SWIR1_median	Landsat SWIR1 band median value	median	season2
12	season2_SWIR2_median	Landsat SWIR2 band median value	median	season2
13	season2_NDVI_median	Spectral index NDVI median value	median	season2
14	season2_NDWI_median	Spectral index NDWI median value	median	season2
15	off-season1_GREEN_median	Landsat Green band median value	median	off-season1
16	off-season1_RED_median	Landsat Red band median value	median	off-season1
17	off-season1_NIR_median	Landsat NIR band median value	median	off-season1
18	off-season1_SWIR1_median	Landsat SWIR1 band median value	median	off-season1
19	off-season1_SWIR2_median	Landsat SWIR2 band median value	median	off-season1
20	off-season1_NDVI_median	Spectral index NDVI median value	median	off-season1
21	off-season1_NDWI_median	Spectral index NDWI median value	median	off-season1
22	off-season2_GREEN_median	Landsat Green band median value	median	off-season2
23	off-season2_RED_median	Landsat Red band median value	median	off-season2
24	off-season2_NIR_median	Landsat NIR band median value	median	off-season2
25	off-season2_SWIR1_median	Landsat SWIR1 band median value	median	off-season2
26	off-season2_SWIR2_median	Landsat SWIR2 band median value	median	off-season2
27	off-season2_NDVI_median	Spectral index NDVI median value	median	off-season2
28	off-season2_NDWI_median	Spectral index NDWI median value	median	off-season2
29	off-season3_GREEN_median	Landsat Green band median value	median	off-season3
30	off-season3_RED_median	Landsat Red band median value	median	off-season3
31	off-season3_NIR_median	Landsat NIR band median value	median	off-season3
32	off-season3_SWIR1_median	Landsat SWIR1 band median value	median	off-season3
33	off-season3_SWIR2_median	Landsat SWIR2 band median value	median	off-season3
34	off-season3_NDVI_median	Spectral index NDVI median value	median	off-season3
35	off-season3_NDWI_median	Spectral index NDWI median value	median	off-season3
36	season3_GREEN_median	Landsat Green band median value	median	season3
37	season3_RED_median	Landsat Red band median value	median	season3
38	season3_NIR_median	Landsat NIR band median value	median	season3
39	season3_SWIR1_median	Landsat SWIR1 band median value	median	season3
40	season3_SWIR2_median	Landsat SWIR2 band median value	median	season3
41	season3_NDVI_median	Spectral index NDVI median value	median	season3
42	season3_NDWI_median	Spectral index NDWI median value	median	season3

2.1.5.3 Rice

The bands for rice mapping using the U-Net were selected to ensure the greatest highlight between rice crops and other land uses (e.g. other types of agricultural crops). The variables were selected according to the state to be mapped, as shown in table 10.

Table 10: Metrics used to classify rice in MapBiomas Collection 6.

Id	Variable	Description	Regions	Period
1	evi2_off-season	Offseason Enhanced Vegetation Index 2	Tocantins	off-season
2	swir1_off-season	Offseason shortwave infrared 1	Tocantins	off-season
3	swir2_off-season	Offseason shortwave infrared 2	Tocantins Santa Catarina Paraná	off-season
4	cei_evi2	Spectral index CEI using NIR band instead EVI2	Tocantins Santa Catarina Paraná Rio Grande do Sul	Annual
5	cei_ndwi	Spectral index CEI using NIR band instead NDWI	Tocantins Santa Catarina Paraná	Annual
6	swir1_season	season shortwave infrared 1	Rio Grande do Sul	season
7	swir2_season	season shortwave infrared 2	Rio Grande do Sul	season
8	Tir1_season	season temperature band 1	Rio Grande do Sul	season

2.1.5.4 Coffee

The metrics used to classify coffee crops is shown in Table 11.

Table 11. Metrics used to classify coffee in MapBiomas Collection 6.

Id	Variable	Description	Statistics	Period
1	ANNUAL_GREEN_median	Landsat Green band median value	median	ANNUAL
2	ANNUAL_RED_median	Landsat Red band median value	median	ANNUAL
3	ANNUAL_NIR_median	Landsat NIR band median value	median	ANNUAL
4	ANNUAL_SWIR1_median	Landsat SWIR1 band median value	median	ANNUAL
5	ANNUAL_SWIR2_median	Landsat SWIR2 band median value	median	ANNUAL

6	ANNUAL_EVI2_median	Spectral index EVI2 median value	median	ANNUAL
7	ANNUAL_NDWI_median	Spectral index NDWI median value	median	ANNUAL
8	ANNUAL_GREEN_p80	Landsat Green band percentile 80 value	percentile	ANNUAL
9	ANNUAL_RED_p80	Landsat Red band percentile 80 value	percentile	ANNUAL
10	ANNUAL_NIR_p80	Landsat NIR band percentile 80 value	percentile	ANNUAL
11	ANNUAL_SWIR1_p80	Landsat SWIR1 band percentile 80 value	percentile	ANNUAL
12	ANNUAL_SWIR2_p80	Landsat SWIR2 band percentile 80 value	percentile	ANNUAL
13	ANNUAL_EVI2_p80	Spectral index EVI2 percentile 80 value	percentile	ANNUAL
14	ANNUAL_NDWI_p80	Spectral index NDWI percentile 80 value	percentile	ANNUAL
15	ANNUAL_EVI2_max	Spectral index EVI2 max value	maximum	ANNUAL
16	ANNUAL_NDWI_max	Spectral index NDWI max value	maximum	ANNUAL
17	ANNUAL_GREEN_qmo	Landsat Green band selected based on maximum EVI2	quality mosaic	ANNUAL
18	ANNUAL_RED_qmo	Landsat Red band selected based on maximum EVI2	quality mosaic	ANNUAL
19	ANNUAL_NIR_qmo	Landsat NIR band selected based on maximum EVI2	quality mosaic	ANNUAL
20	ANNUAL_SWIR1_qmo	Landsat SWIR1 band selected based on maximum EVI2	quality mosaic	ANNUAL
21	ANNUAL_SWIR2_qmo	Landsat SWIR2 band selected based on maximum EVI2	quality mosaic	ANNUAL
22	ANNUAL_EVI2_qmo	Spectral index EVI2 selected based on maximum EVI2	quality mosaic	ANNUAL
23	ANNUAL_NDWI_qmo	Spectral index NDWI selected based on maximum EVI2	quality mosaic	ANNUAL

2.1.5.5 Citrus

The bands used for training and classification of citrus were annual compositions generated from the median of the five images with less cloud cover in each point orbit. The bands used are shown in Table 12.

Table 12. Metrics used to classify citrus in MapBiomas Collection 6.

Id	Variable	Description	Statistics	Period
1	ANNUAL RED median	Landsat Red band median value	median	ANNUAL
2	ANNUAL_NIR_median	Landsat NIR band median value	median	ANNUAL
3	ANNUAL_SWIR1_median	Landsat SWIR1 band median value	median	ANNUAL

2.1.5.6 Other Perennial Crop

Part of the 'Other Perennial Crop' map came from the separation of that class from the class 'Annual and Perennial Crop' of MapBiomas Collection 4. Therefore, this map was created from two classification processes: 1) in Collection 4, the classifier was trained to classify 'Annual and Perennial Crop' without distinction; 2) in this collection, these maps resulting from the first classification were submitted to a second classification, in which the classifier was trained with new feature spaces to distinguish the pixels of short-cycle crops and long-cycle crops. The resulting map of perennial crops became part of the class 'Other Perennial Crop', while the resulting map of temporary crops wasn't used in this collection (it was processed again using the methodology described before). Figure 6 illustrates the processes performed to generate the class 'Other Perennial Crop'.

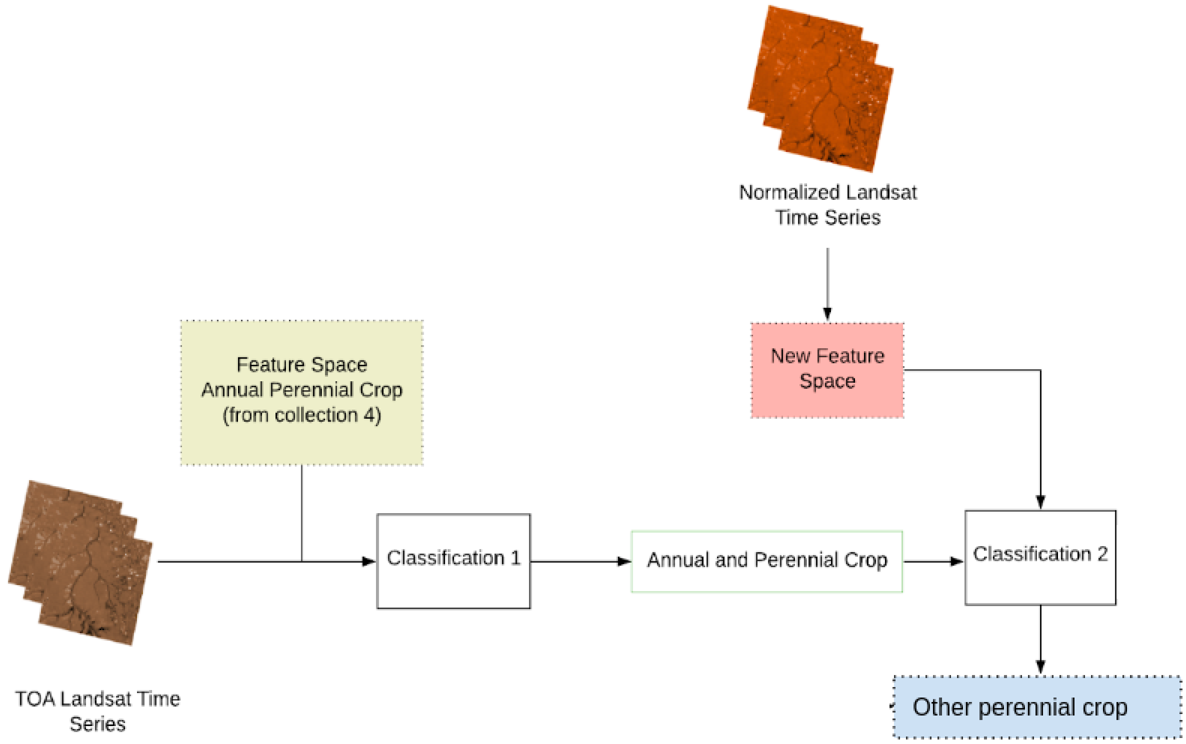


Figure 6. Steps to separate 'Other Perennial Crop' from the previous class 'Annual and Perennial Crop' of the Collection 4.

The cycle of temporary crops tends to have greater annual variation in the spectral response than perennial crops, which are more stable over time. Therefore, metrics were selected to highlight this difference between temporary and perennial crops (Table 13).

Table 13. Metrics used to separate perennial and other temporary crops in Collection 6.

Id	Variable	Description	Statistics	Period
1	season_NDWI_stdDev	Spectral index NDWI standard deviation value	standard deviation	season
2	season_NDWI_min	Spectral index NDWI minimum value	minimum	season
3	season_EVI2_stdDev	Spectral index EVI2 standard deviation value	standard deviation	season

4	off-season_EVI2_min	Spectral index EVI2 minimum value	minimum	off-season
5	ANNUAL_EVI2_amplitude	Spectral index EVI2 amplitude value	amplitude	ANNUAL
6	ANNUAL_EVI2_stdDev	Spectral index EVI2 standard deviation value	standard deviation	ANNUAL
7	ANNUAL_EVI2_min	Spectral index EVI2 minimum value	minimum	ANNUAL
8	ANNUAL_EVI2_p10	Spectral index EVI2 10th percentile value	10th percentile	ANNUAL
9	ANNUAL_EVI2_median	Spectral index EVI2 median value	median	ANNUAL
10	ANNUAL_EVI2_mean	Spectral index EVI2 average value	average	ANNUAL
11	ANNUAL_NIR_cei	Landsat NIR band normalized difference value	normalized difference	ANNUAL
12	ANNUAL_EVI2_cei	Spectral index EVI2 normalized difference value	normalized difference	ANNUAL
13	ANNUAL_NDWI_cei	Spectral index NDWI normalized difference value	normalized difference	ANNUAL

2.1.5.7 Forest Plantation

The metrics used to classify forest plantation were the same as in Collection 6 (Table 14).

Table 14. Metrics used to classify forest plantation in Collection 6.

Id	Variable	Description	Statistics	Period
1	season1_GREEN_qmo	Landsat Green band selected based on maximum EVI2	quality mosaic	season1
2	season1_RED_qmo	Landsat Red band selected based on maximum EVI2	quality mosaic	season1
3	season1_NIR_qmo	Landsat Nir band selected based on maximum EVI2	quality mosaic	season1
4	season1_SWIR1_qmo	Landsat SWIR1 band selected based on maximum EVI2	quality mosaic	season1
5	season1_SWIR2_qmo	Landsat SWIR2 band selected based on maximum EVI2	quality mosaic	season1
6	season1_NDVI_qmo	Spectral index NDVI selected based on maximum EVI2	quality mosaic	season1
7	season1_LAI_qmo	Spectral index LAI selected based on maximum EVI2	quality mosaic	season1
8	season2_GREEN_qmo	Landsat Green band selected based on maximum EVI2	quality mosaic	season2
9	season2_RED_qmo	Landsat Red band selected based on maximum EVI2	quality mosaic	season2
10	season2_NIR_qmo	Landsat Nir band selected based on maximum EVI2	quality mosaic	season2
11	season2_SWIR1_qmo	Landsat SWIR1 band selected based on maximum EVI2	quality mosaic	season2
12	season2_SWIR2_qmo	Landsat SWIR2 band selected based on maximum EVI2	quality mosaic	season2
13	season2_NDVI_qmo	Spectral index NDVI selected based on maximum EVI2	quality mosaic	season2
14	season2_LAI_qmo	Spectral index LAI selected based on maximum EVI2	quality mosaic	season2
15	off-season1_GREEN_qmo	Landsat Green band selected based on maximum EVI2	quality mosaic	off-season1
16	off-season1_RED_qmo	Landsat Red band selected based on maximum EVI2	quality mosaic	off-season1
17	off-season1_NIR_qmo	Landsat Nir band selected based on maximum EVI2	quality mosaic	off-season1
18	off-season1_SWIR1_qmo	Landsat SWIR1 band selected based on maximum EVI2	quality mosaic	off-season1
19	off-season1_SWIR2_qmo	Landsat SWIR2 band selected based on maximum EVI2	quality mosaic	off-season1
20	off-season1_NDVI_qmo	Spectral index NDVI selected based on maximum EVI2	quality mosaic	off-season1
21	off-season1_LAI_qmo	Spectral index LAI selected based on maximum EVI2	quality mosaic	off-season1

22	off-season2_GREEN_qmo	Landsat Green band selected based on maximum EVI2	quality mosaic	off-season2
23	off-season2_RED_qmo	Landsat Red band selected based on maximum EVI2	quality mosaic	off-season2
24	off-season2_NIR_qmo	Landsat Nir band selected based on maximum EVI2	quality mosaic	off-season2
25	off-season2_SWIR1_qmo	Landsat SWIR1 band selected based on maximum EVI2	quality mosaic	off-season2
26	off-season2_SWIR2_qmo	Landsat SWIR2 band selected based on maximum EVI2	quality mosaic	off-season2
27	off-season2_NDVI_qmo	Spectral index NDVI selected based on maximum EVI2	quality mosaic	off-season2
28	off-season2_LAI_qmo	Spectral index LAI selected based on maximum EVI2	quality mosaic	off-season2
29	off-season3_GREEN_qmo	Landsat Green band selected based on maximum EVI2	quality mosaic	off-season3
30	off-season3_RED_qmo	Landsat Red band selected based on maximum EVI2	quality mosaic	off-season3
31	off-season3_NIR_qmo	Landsat Nir band selected based on maximum EVI2	quality mosaic	off-season3
32	off-season3_SWIR1_qmo	Landsat SWIR1 band selected based on maximum EVI2	quality mosaic	off-season3
33	off-season3_SWIR2_qmo	Landsat SWIR2 band selected based on maximum EVI2	quality mosaic	off-season3
34	off-season3_NDVI_qmo	Spectral index NDVI selected based on maximum EVI2	quality mosaic	off-season3
35	off-season3_LAI_qmo	Spectral index LAI selected based on maximum EVI2	quality mosaic	off-season3
36	season3_GREEN_qmo	Landsat Green band selected based on maximum EVI2	quality mosaic	season3
37	season3_RED_qmo	Landsat Red band selected based on maximum EVI2	quality mosaic	season3
38	season3_NIR_qmo	Landsat Nir band selected based on maximum EVI2	quality mosaic	season3
39	season3_SWIR1_qmo	Landsat SWIR1 band selected based on maximum EVI2	quality mosaic	season3
40	season3_SWIR2_qmo	Landsat SWIR2 band selected based on maximum EVI2	quality mosaic	season3
41	season3_NDVI_qmo	Spectral index NDVI selected based on maximum EVI2	quality mosaic	season3
42	season3_LAI_qmo	Spectral index LAI selected based on maximum EVI2	quality mosaic	season3

2.1.6 Classification algorithm, training samples and parameters

2.1.6.1 Reference Maps

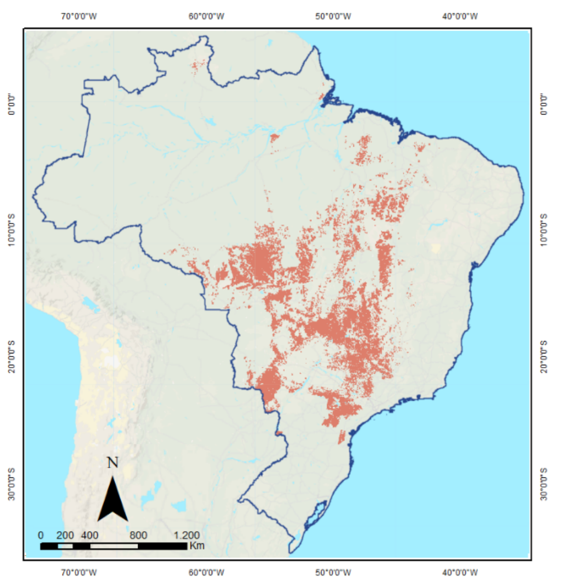
The reference maps used to obtain samples to train the classifier are shown in Table 15.

Table 15. Reference maps used in the Random Forest classification for the classes Agriculture and Forest Plantation in Collection 6.

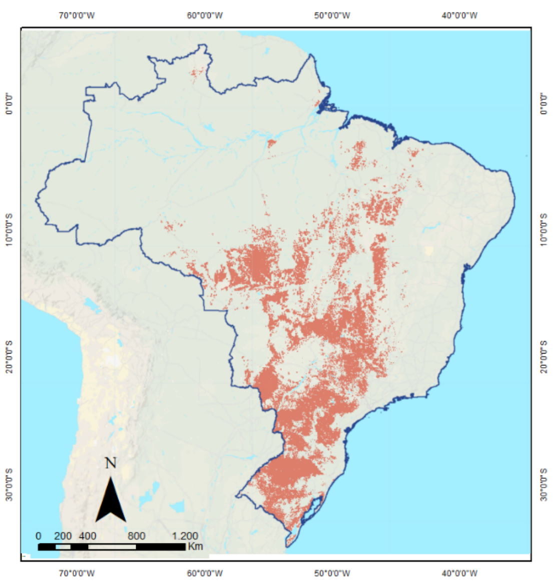
Class	Landsat time series	Number of training samples	Rule	Type	Year of acquisition	Reference
Soybean (2000 - 2020)	Normalized	10,000	Minimum of 20% for the interests class	stable samples	2016	Agrosatélite (2020A) Agrosatélite (2020B) Song (in presso)
Soybean (1985 - 1999)	L5 TOA	10,000	-	stable samples	2000	Agrosatélite (2020A) Agrosatélite (2020B) Song (in press)

Sugar cane	TOA	10,000	-	annual samples	2003 - 2019	Rudorff et al. (2010)
rice	TOA		-	chips	2017-2020	Agência Nacional de Águas (ANA) e Companhia Nacional de Abastecimento (Conab) (2020)
Other Temporary Crop	Normalized	5,000	Minimum of 20% for the interests class	stable samples	2016	Agrosatélite (2020A) Agrosatélite (2020B) Agrosatélite
Coffee	Normalized	10,000	Minimum of 500 samples for coffee and others	stable samples	2015, 2016, 2017, 2018, 2019	Conab (2015, 2016, 2017, 2018, 2019)
Citrus	TOA		-	chips	2020	Agrosatélite
Other Perennial Crop	Normalized	5,000	Minimum of 20% for the interests class	stable samples	2016	Agrosatélite
Forest Plantation	TOA	10,000	-	annual samples	2012 - 2014	Global Forest Watch, Transparent World (2015)

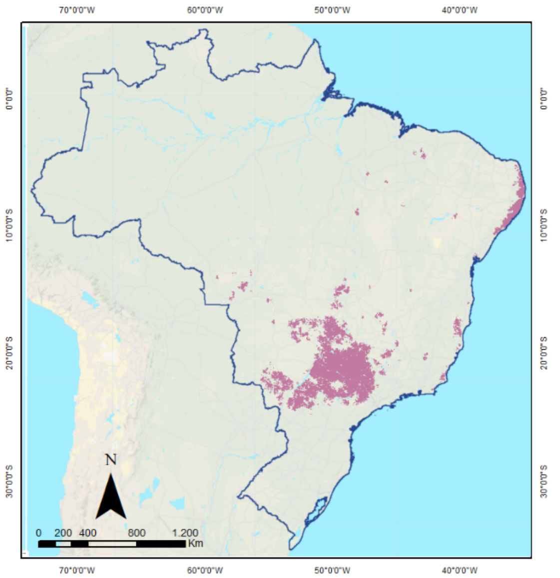
The reference maps used are shown in Figure 7.



Soybean reference map - Agrosatélite



Soybean reference map - Agrosatélite and Songet al. (in press) maps

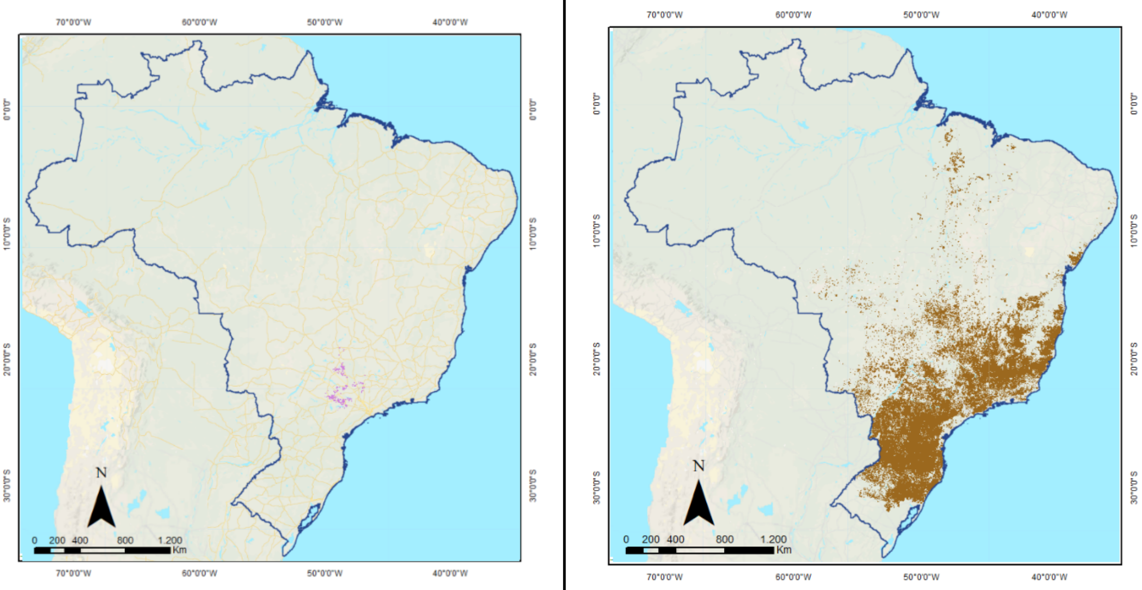


Canasat project (RUDORFF *et al.*, 2010) map of 2018/2019

Rice reference map
Conab/ANA (2020)

Perennial Crop reference map
(Agrosatélite, 2020b)

Coffee reference map
Conab (2015, 2016, 2017, 2018, 2019)



Citrus reference map
(Agrosatélite)

Global Forest Watch
(TRANSPARENT WORLD, 2015)

Figure 7. Reference maps representing the areas with training samples for the classification of 'Agriculture' and 'Forest Plantation' in Brazil in Collection 6.

2.1.6.2 Random Forest

For the classes mapped by Random Forest algorithm, the process steps is: a) initially, an annual Landsat mosaic is created, according to the period of the year (i.e. season and off-season), specific for each class; b) bands are built with specific metrics for each class; c) simple 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. In order 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 utilized in the training process must be from the same year as the reference map used. The classification using Random Forest Algorithm is illustrated in Figure 8.

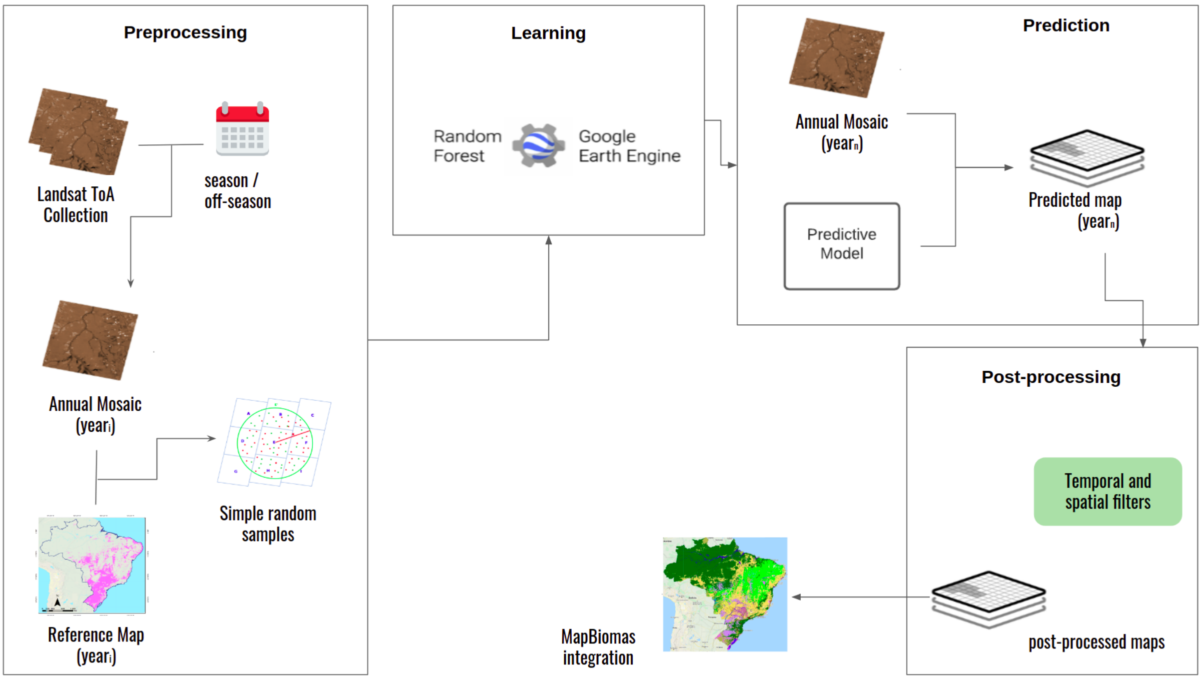


Figure 8. Fluxogram of agriculture and Forest Plantation classification using Random Forest algorithm.

The classes mapped with Random Forest algorithm were: soybean, sugar cane, other temporary crops, coffee, other perennial crops and forest plantation. All classes were trained with 100 trees.

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 were included inside an E' buffer of radius R, in which the center of that radius corresponds to the center of the target scene (E), as shown in Figure 9.

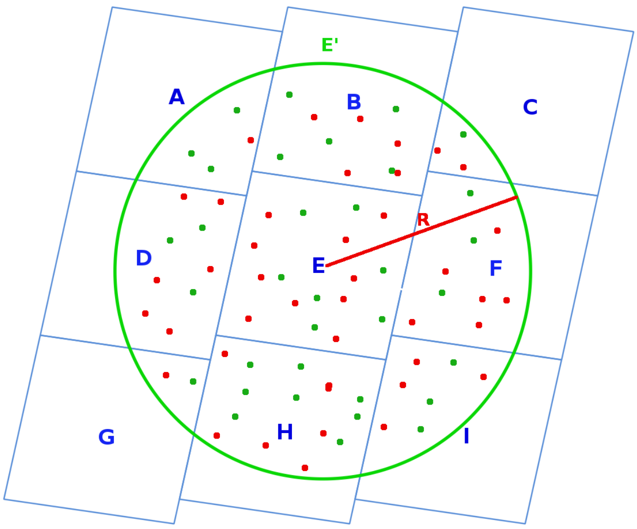


Figure 9. Scheme for sample acquisition for the regionalized training of the Random Forest classifier in Agriculture and Forest Plantation.

Knowing that there are no reference maps available for all classes in all years of the time series (1985 to 2020), stable samples were created. However, these samples were only obtained in classes which used the normalized Landsat series, due to the characteristics of this time series mentioned above. The fluxogram (Figure 10) illustrates the classification process with Normalized Landsat Time Series.

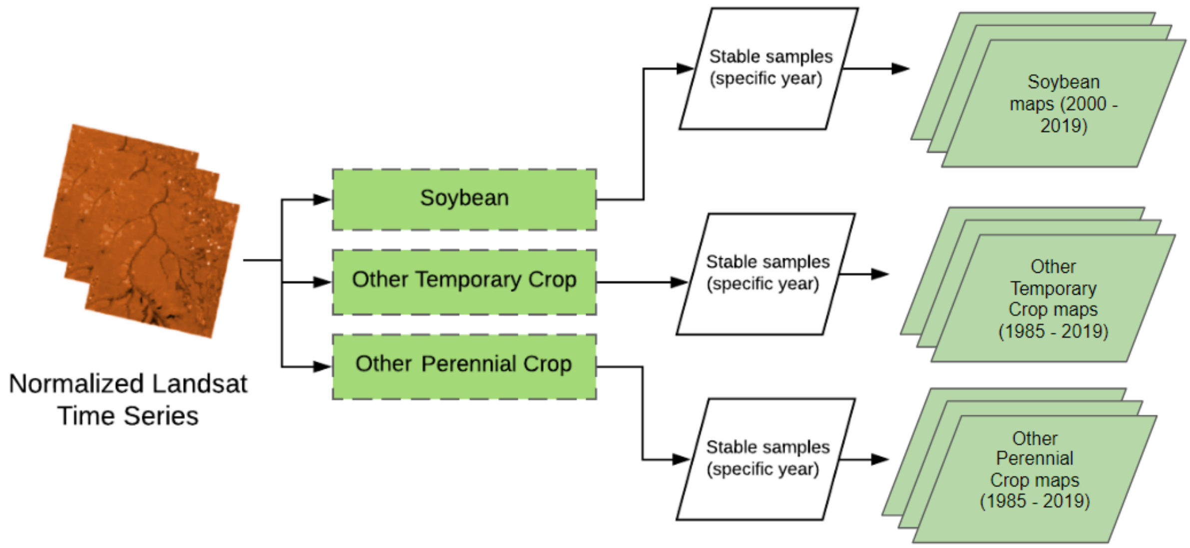


Figure 10. The use of stable samples in classes classified with the Normalized Landsat Time Series.

For the group of classes obtained from the TOA Landsat time series, as a reference map was not available for each year to be classified, annual samples were used on the available reference maps for training and classification only for those years with available reference maps. 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 (as illustrated in Figure 11).

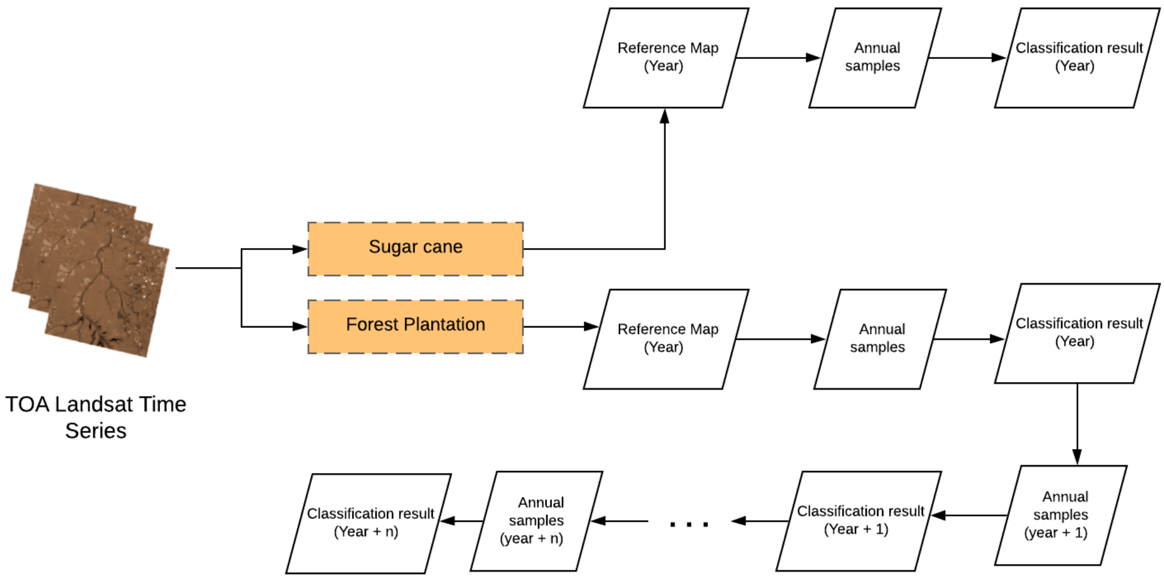
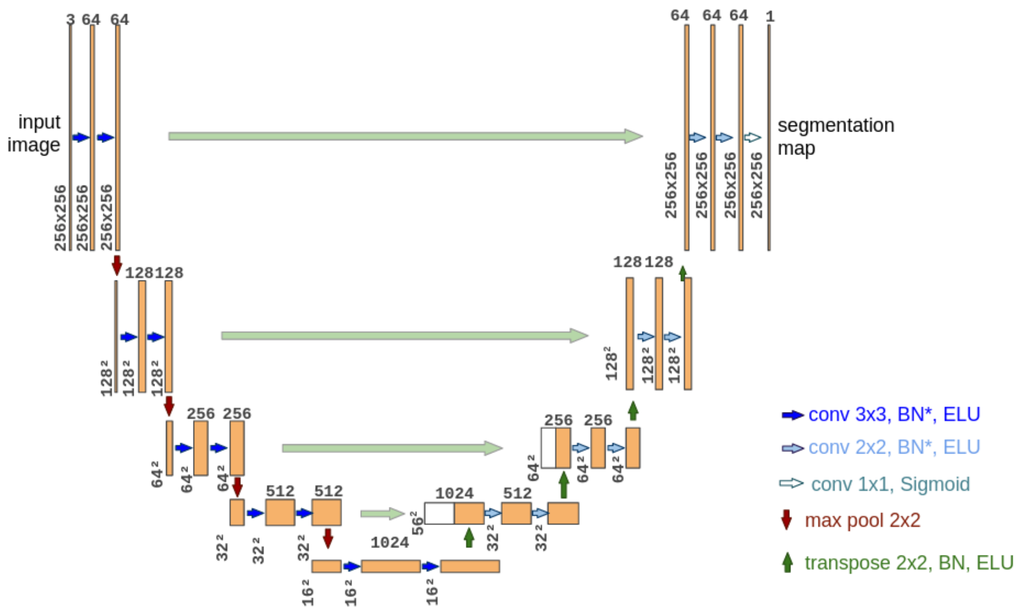


Figure 11. The use of annual samples in classes classified with the TOA Landsat Time Series.

2.1.6.3 Deep Learning

For the mapping of rice and citrus, an adaptation of the U-Net convolutional neural network was used. Unlike machine learning algorithms that classify each pixel without considering the surrounding pixels, this architecture uses the context in which the pixels are. This architecture is illustrated in Figure 12.



*BN = Batch Normalization

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

This architecture was developed in Python, using the TensorFlow 2.0 library. The entire training and mapping process was carried out using the Google Colab platform. To enable the Google Colab platform to have access to satellite images, Google Drive was used to store the images.

To obtain the training and validation sets, each training block was covered to generate chips with 256 x 256 pixels. Then, the chips were divided into 70% for training and 30% for validation for each block. After data separation, the pixel values of each image band were normalized. Normalization scales the numerical values for a given range, making each band have the same weight for the classifier.

2.1.6.4 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 season period according to the year of mapping carried out in each state (i.e. in Rio Grande do Sul the mapping season was 2019/2020). 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 with the states of interest of Collection 6, as illustrated in Figure 13.

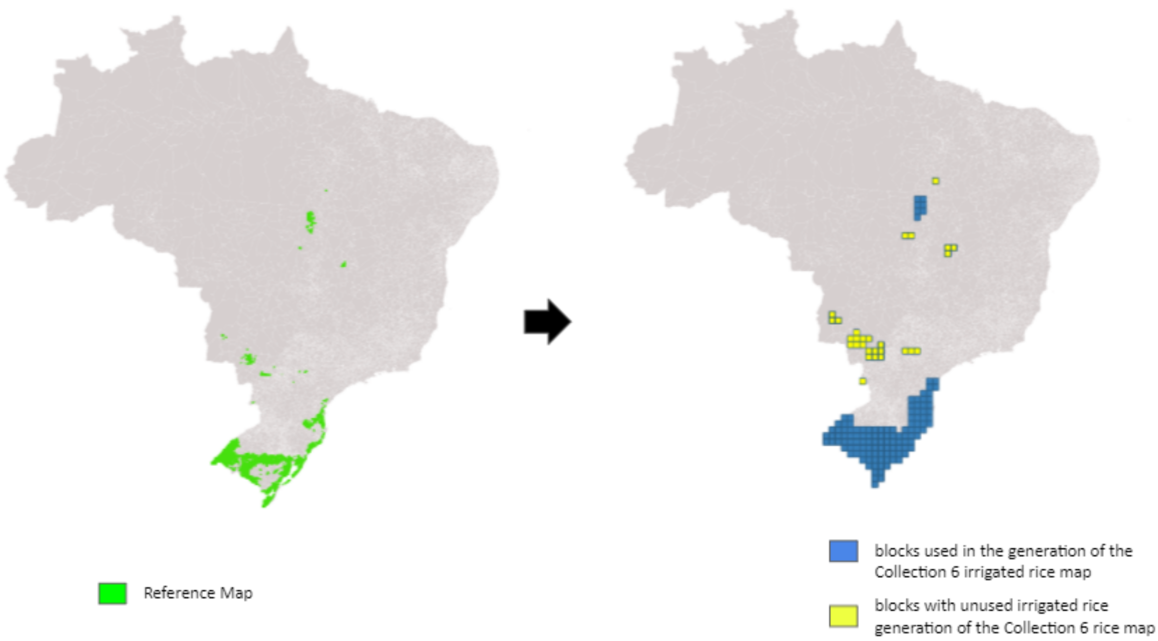
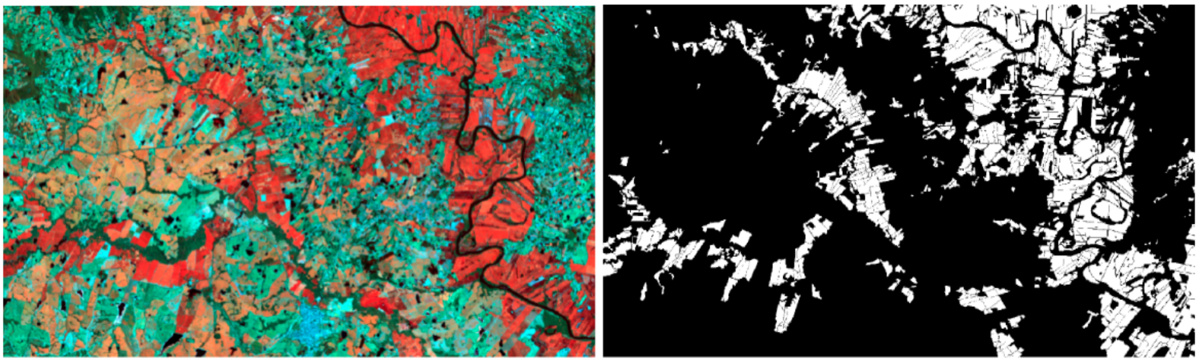


Figure 13. Study area used for the mapping of irrigated rice in the MapBiomas Project.

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 sample to mapping rice



The test data were used for 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 (1985-2020).

2.1.6.4.1 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.

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. This spatial filter removed groups of pixels with 6 or less pixels of the interest class or the “others” class (Figure 15).

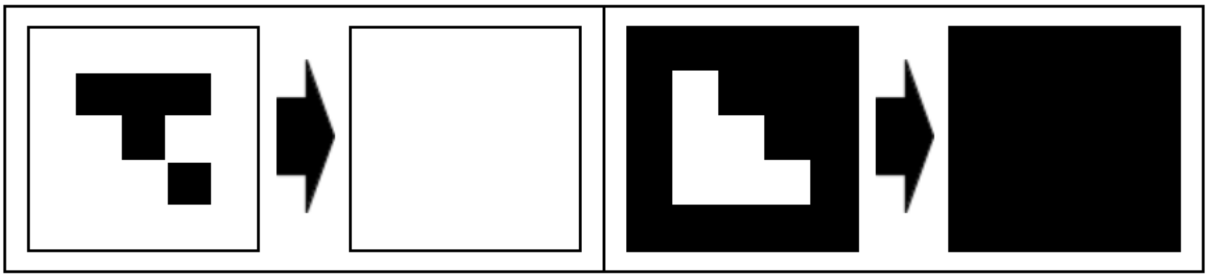


Figure 15. Example of the minimum connected pixels spatial filter. The image on the left shows an exclusion of pixels of the interest class (in black). The image on the right shows an inclusion of pixels of “other classe” (in white) to the interest class.

3.2 Temporal filter

In general, two temporal window filters were applied: using 3 years with 2 years threshold or 5 years with 3 years threshold. The 3-years window excludes the center year when none of the adjacent years are of the interest class, and includes the center year when both adjacent years are of the interest class (Figure 16). The 5-years window excludes the center year when no more than 1 another year is of the interest class, and includes when at least 3 adjacent years are of the interest class (Figure 17).

The 5-year window was applied on Other Temporary Crop, Sugar Cane, Coffee, and Citrus. The 3-years window was applied in all classes (after any other filter to ensure that no isolated year of the interest class remained).

Other specific temporal filters may have been applied in some classes.

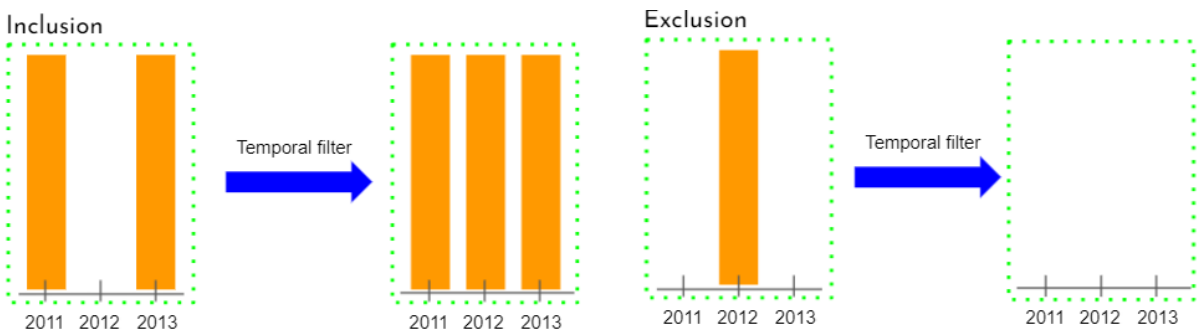


Figure 16. 3-years temporal window filter: The orange bars represent pixels of the mapped class (interest class). The exclusion filter changes a pixel to “others” class when the same pixel was not of the interest class in the adjacent years. The inclusion filter changes a pixel to the interest class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

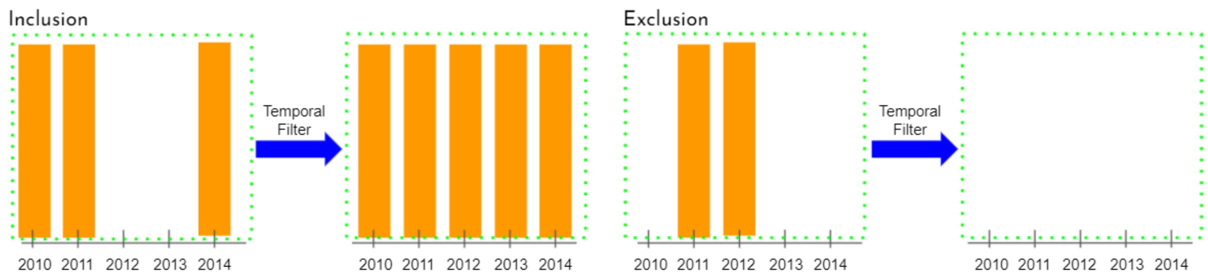


Figure 17. 5-years temporal window filter: The orange bars represent pixels of the mapped class (interest class). The exclusion filter changes a pixel to "others" class when no more than 1 another year is of the interest class. The inclusion filter changes a pixel to the interest class when at least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

Table 16. Temporal filters applied by class.

Class	Temporal Filter	Additional rule
Soybean	3-year window	
sugar cane	5-year window	
Rice	3-year window	inclusion filter only
Other Temporary Crop	5-year window	
Coffee	5-year window	
Citrus	5-year window	
Other Perennial Crop	5-year window	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.

For 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.* 2020), no temporal filter was applied.

For the Forest Plantation class, two temporal window filters were applied: using 5 years with 2 years threshold or 5 years with 3 years threshold. The 5-year with 2 years threshold excludes the center year when no more than 1 another year is of the interest class, and includes when at least 2 adjacent years are of the interest class (Figure 18). The same logic to 5-year with 3 years threshold filter.

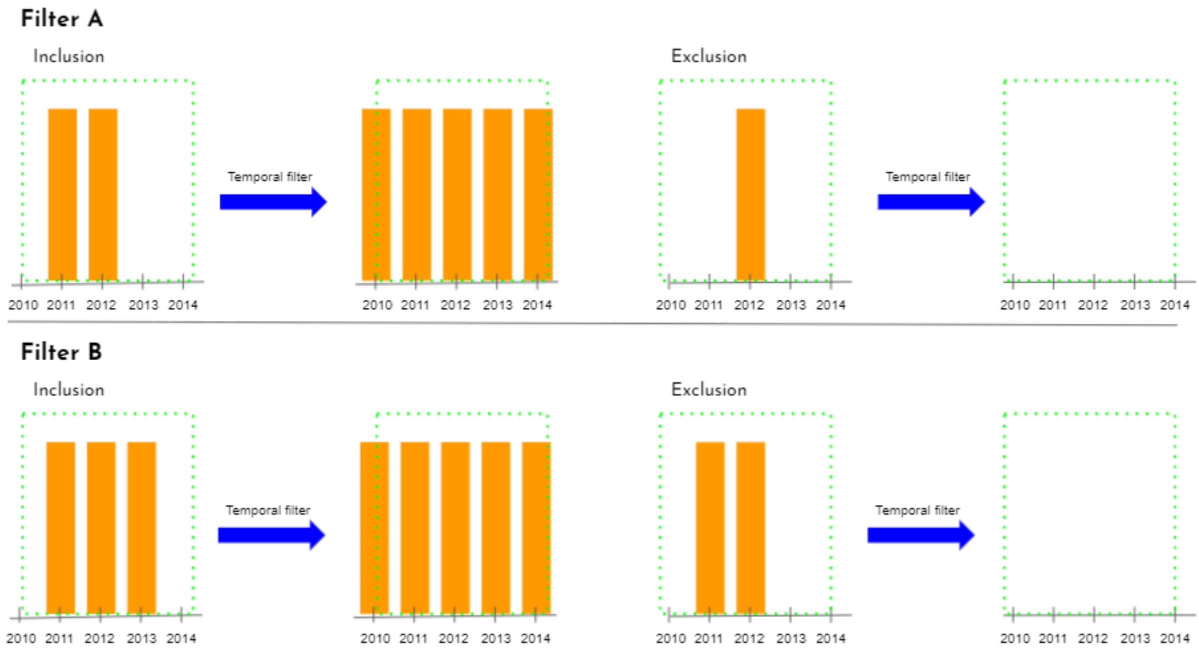


Figure 18. 5-years temporal windows with 2 year (filter A) and 3 years (filter B) thresholds applied in Forest Plantation maps.

Both filters either include or exclude pixels of forest plantation from the maps using a 5-years temporal window (the interest year on the center, two before and two after). The threshold for filter A was 2 years: if in 2 of the 5 years the pixels were forest plantation, the pixel of the center year will be converted to forest plantation (if it wasn't already), otherwise it will be converted to "others" class. Filter B works the same way, but with a threshold of 3 years. Filter A is more inclusive than filter B, pixels remain as forest plantation longer. Filter B is more selective, only converting to forest plantation pixels with higher occurrence of this class.

In addition, another temporal filter was applied to fill longer intervals of non occurrence of forest plantation when it was forest plantation at some year in the past and it became again years after, like the example in Figure 19.

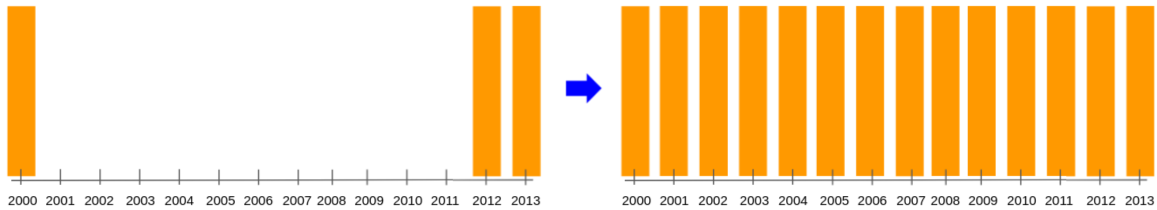


Figura 19. Temporal filter that converted longer intervals into forest plantation when it was in the past and became again years after.

Another consideration was at the end of the series (2017 to 2020). 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 plantations to become identifiable again, to solve this situation pixels from 2016 to 2020 were converted to forest plantation when they were of this class in the 3 years before (2013 to 2015). Figure 20 illustrates this filter.

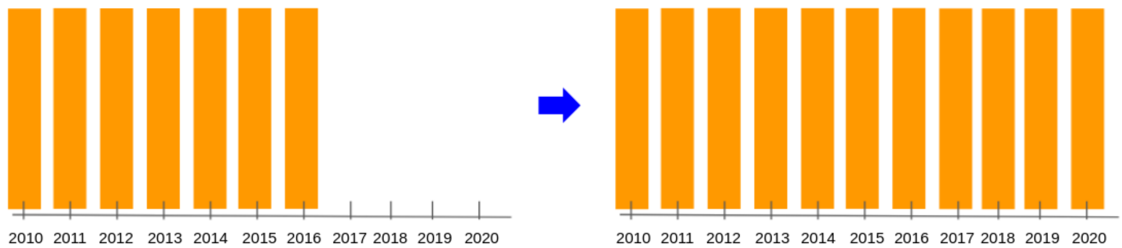


Figure 20. Temporal filter applied in the last years of the forest plantation series.

For the citrus class, another filter applied was a gap-fill filter like the one illustrated in Figure 19. And finally, at the end, another 3-years window temporal filter was applied to all classes to ensure that didn't remain any isolated year of the class of interest or the "others" class, like illustrated on Figure 14.

4 Integration with biomes and themes

After the classification of the Agriculture and Forest Plantation themes, they were integrated to the other land use and land cover classes to compose the MapBiomas Collection 6 final maps. This integration process was based on the overlap order of the classes. The integration process tends to improve the quality of the Agriculture and Forest Plantation maps as it removes some commission errors.

5 Validation strategies

The independent validation points provided by the LAPIG of the Goias Federal University (UFG) were used to calculate the global accuracy of the mapping and the accuracy for each land use class. The following section also presents some comparisons between the Random Forest classification results and the reference maps.

5.1 Accuracy analysis

The map accuracy analysis was produced using independent validation points provided by the LAPIG of the Goias Federal University (UFG). We used all points that at least two interpreters considered the same class, resulting in over 12,000 validation points. Accuracy analysis from LAPIG points was performed in the following classes: 'Forest Plantation', 'Perennial Crop' and 'Temporary Crop'. The 'Temporary Crop' class contains the classes 'Soy', 'Sugarcane', 'Rice' and 'Other Temporary Crops', and the 'Perennial Crop' contains the classes 'Coffee', 'Citrus (SP only)' and 'Other Perennial Crops'. LAPIG points used for the accuracy analyzes are shown in Figure 21.

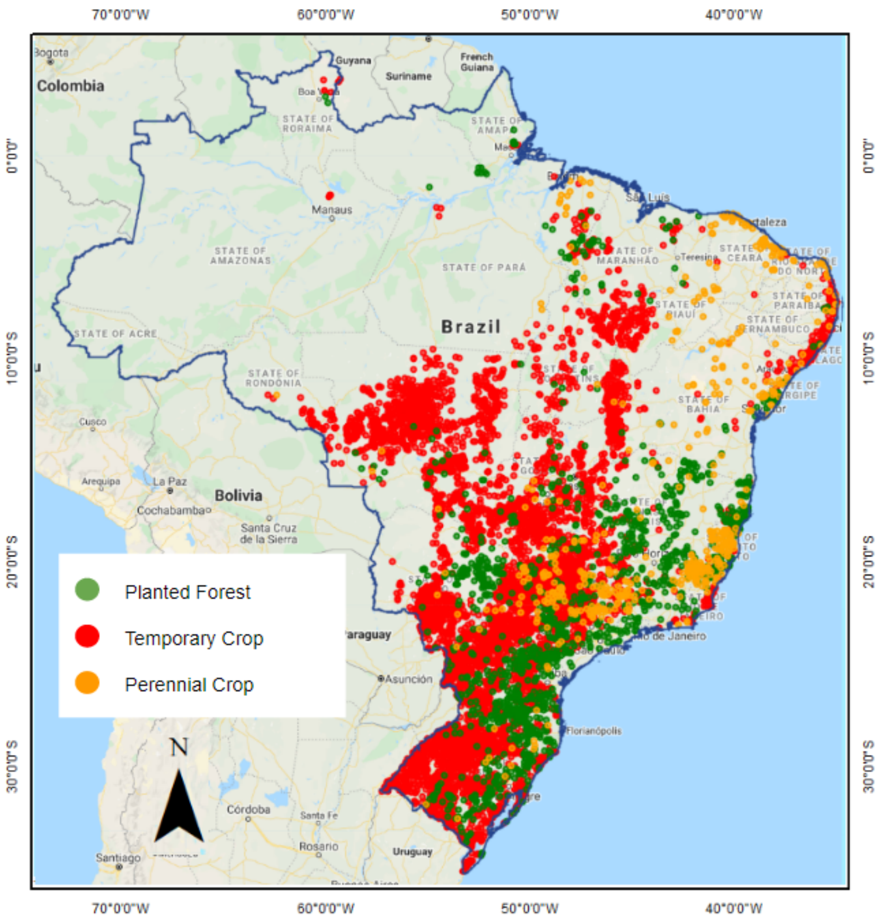


Figure 21. LAPIG points that were used for the accuracy analyzes of 'Temporary Crops', 'Perennial Crops' and 'Forest Plantation' classes.

The result of the accuracy analysis of the 'Temporary Crop' class (Figure 22) showed that the accuracy of this class increases over the years, reaching higher values in the final years of the series. Throughout the series, the maps accuracies were above 60%, reaching the highest values after 2008 (*i.e*. producer and user accuracy above 80%).

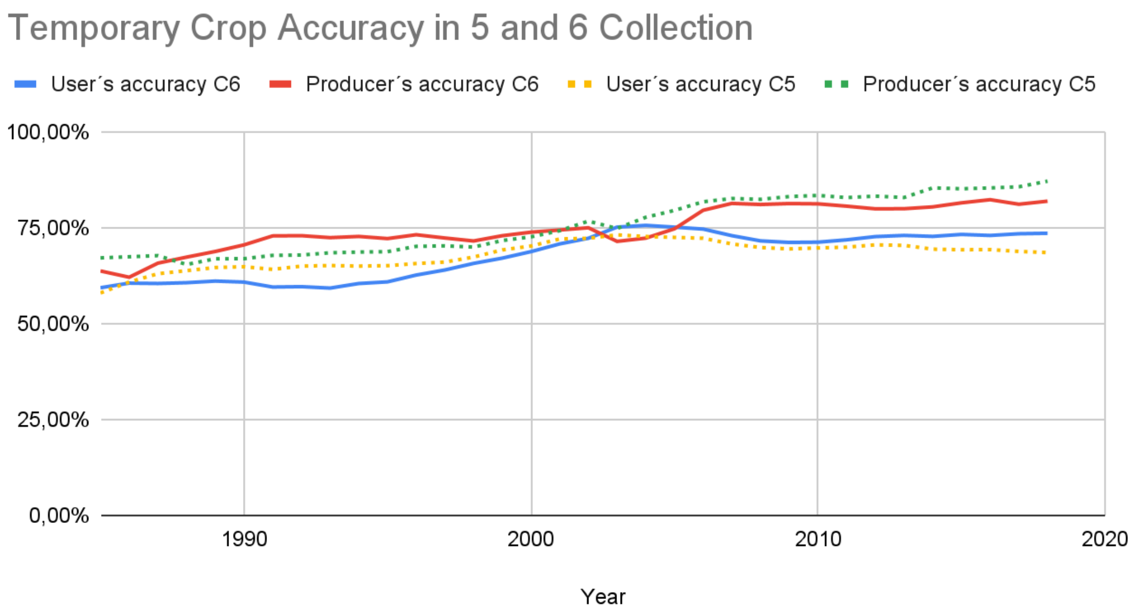


Figure 22. Producer and user's accuracy of the temporary crop class in Collections 5 and 6.

The accuracy of Perennial Crop (including Coffee, Citrus and Other Perennial Crop) can be seen in Figure 23. The producer's accuracy increases throughout the series, reaching its maximum value of 19% in the last year with available validation samples (2018). Although still low, it shows an improvement compared to Collection 5 as a result of the new approach of mapping individual crops (like Coffee and Citrus). The user's accuracy shows lower values from 1985 to 1993 (not more than 20%), with an increase from 1993 to 1996, when it reached 57%. From then, it varied between 49% and 58% until 2013, when it reached 60% and remained above that in the last years, with a maximum of 66% in 2015. When compared to Collection 5, user's accuracy shows a great improvement after 2006, but lower values from 1988 to 1998.

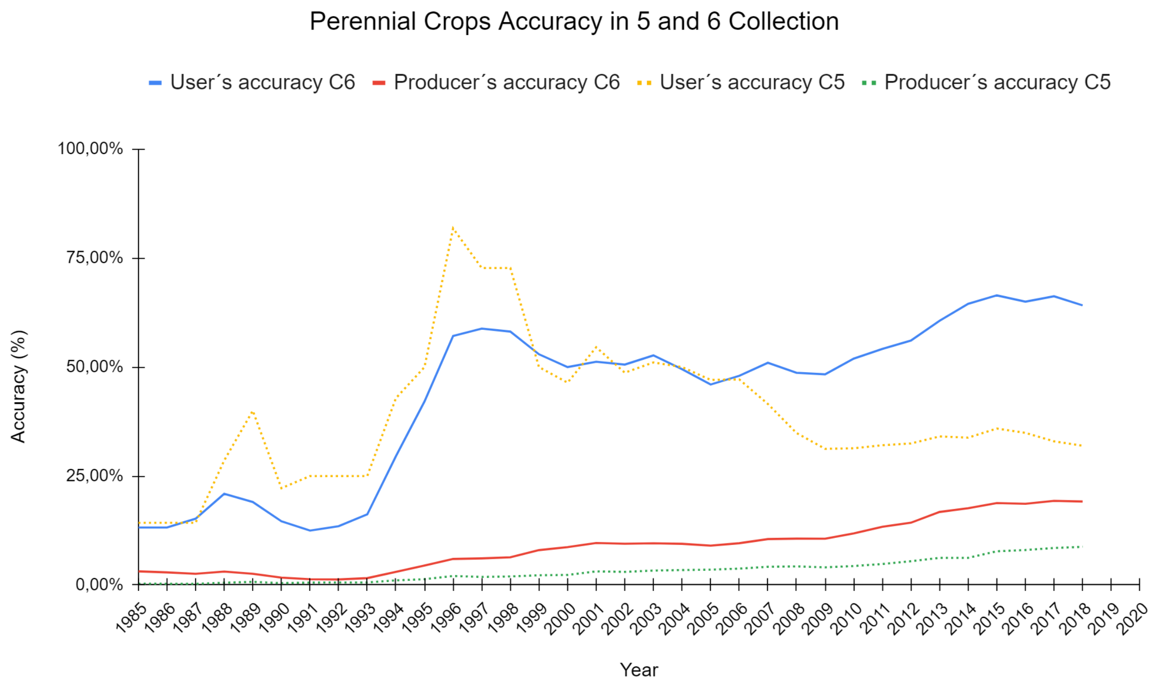


Figure 23. Producer' and user's accuracy of the perennial crop class in Collections 5 and 6.

In MapBiomas Collection 6, the main objective to improve the forest plantation map was: 1) increase the Producer's accuracy in specific regions and b) reduce confusion with other classes, especially in regions of land use changes from forestry to other classes. The Figure 24 shows the Collection 6 accuracy results in comparison with Collection 5.

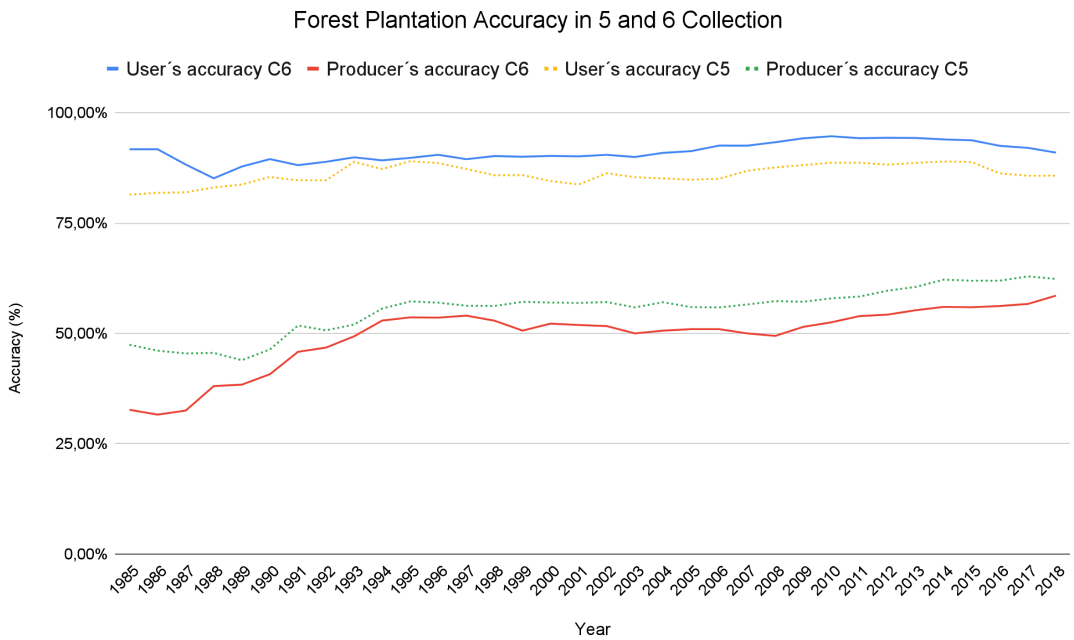


Figure 24. Producer and user's accuracy of the forest plantation class in Collections 5 and 6.

5.2 Comparison with reference data

In addition to the comparison with reference maps and validation points, a comparison between the ‘Agriculture’ and ‘Forest Plantation’ maps of MapBiomas Collection 6 with data from the Systematic Survey of Agricultural Production (LSPA - Levantamento Sistemático da Produção Agrícola), carried out by the Brazilian Institute of Geography and Statistics (IBGE), was also made, and these are considered official data for estimating agricultural area in the country.

5.2.1 Comparison of Temporary Crop area

The graph in Figure 25 shows the comparison between the area of the class ‘Temporary Crops’ (which includes sugar cane, soybean and ‘Other Temporary Crops’) with the areas estimated by LSPA - IBGE.

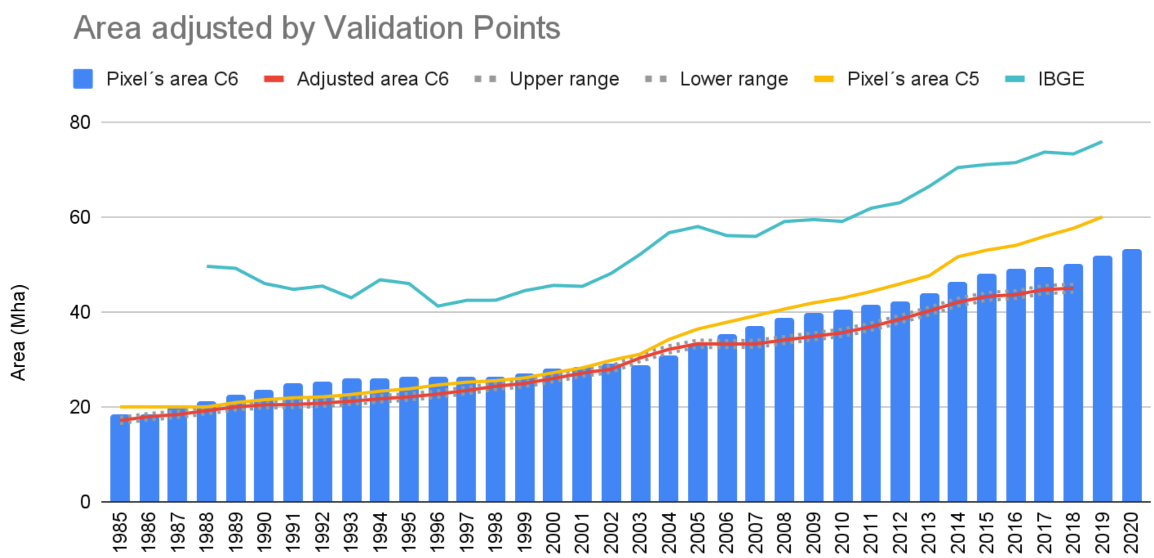


Figure 25. Comparison between MapBiomas Temporary Crop area and LSPA-IBGE Temporary crop Area.

5.2.2 Comparison of Perennial Crop area

Mapping perennial agriculture has been a challenge throughout the MapBiomas collections. The approach of mapping each type of crop separately used in Collection 6 has enabled the improvement of the perennial crop map, however, the mapped area still underestimates the official area provided by IBGE. It is noteworthy that the area provided by IBGE comprises all perennial crops in Brazil, and the MapBiomas area refers only to citrus, coffee and some concentrations of perennial crops spread throughout the territory. Figure 26 shows the comparison between the MapBiomas Collection 6 perennial area and PAM perennial crop area in 2019.

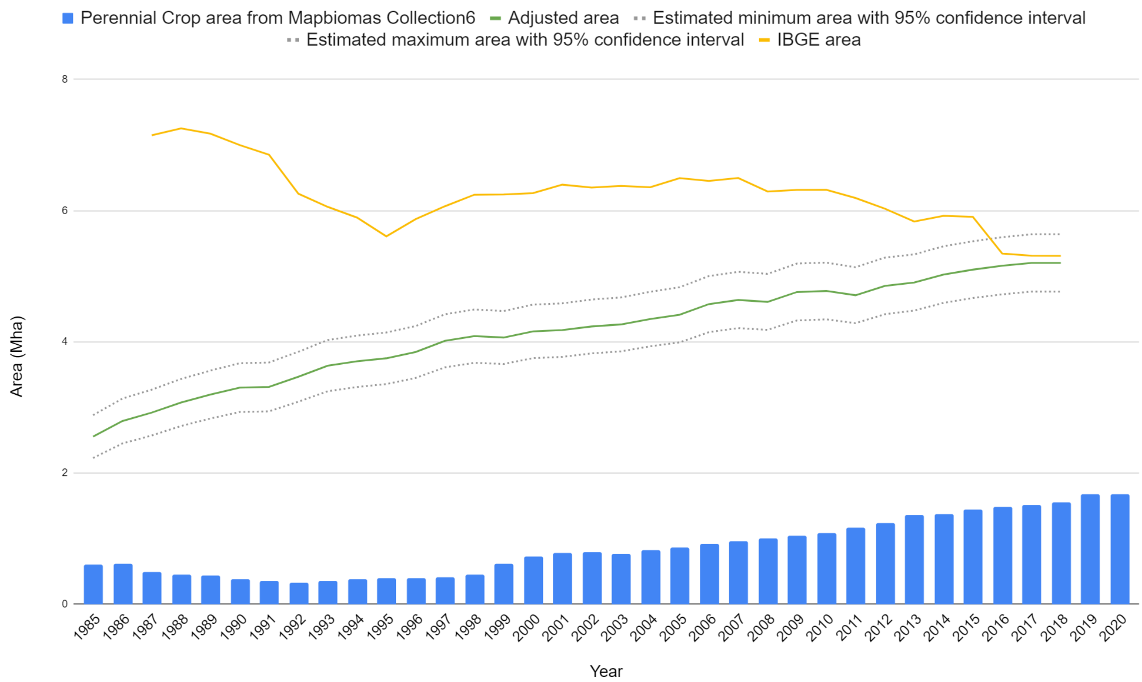


Figure 26. Comparison between MapBiomas Perennial Crop area and LSPA-IBGE Perennial crop Area.

Figure 27 shows the comparison of the coffee area by municipality obtained by the Mapbiomas map and by the one informed by PAM 2019 (IBGE).

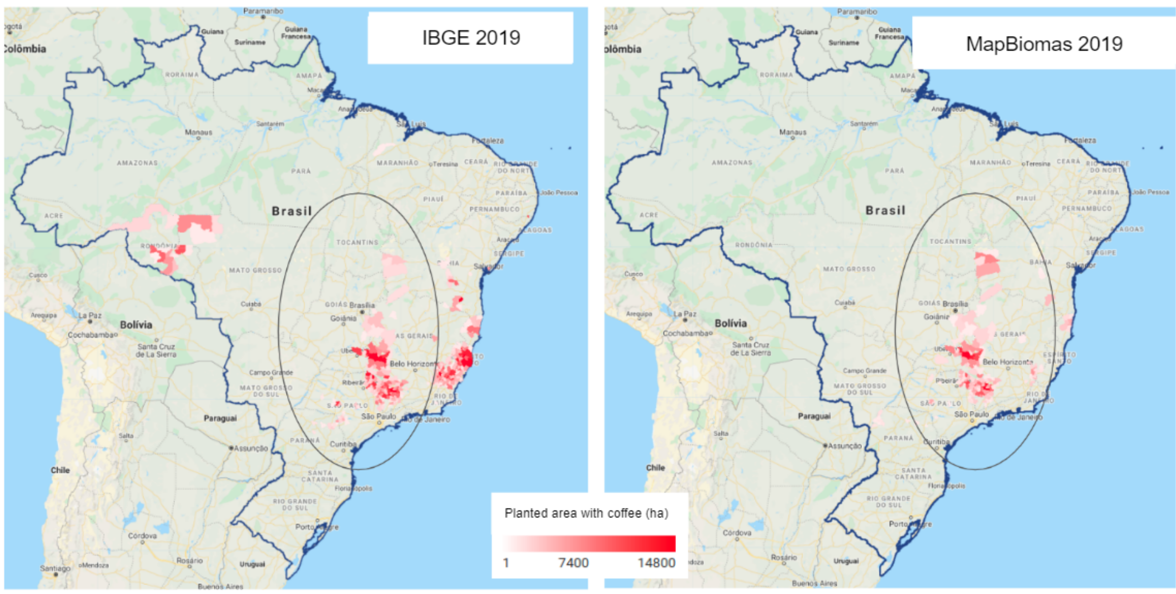


Figure 27. Comparison between MapBiomas coffee map area and Conab coffee map in 2019.

5.2.3 Comparison of Forest Plantation area

Forest Plantation areas obtained from MapBiomas Collection 6 annual maps were also compared with areas from official sources. In order to estimate Brazil's forestry area, a comparison was made between MapBiomas, IBGE's Production of Vegetable Extraction and Forest Plantation (PEVS-IBGE) and Ibá - *Indústria brasileira de árvores* (Brazilian Tree Industry). The results are shows in Figure 28.

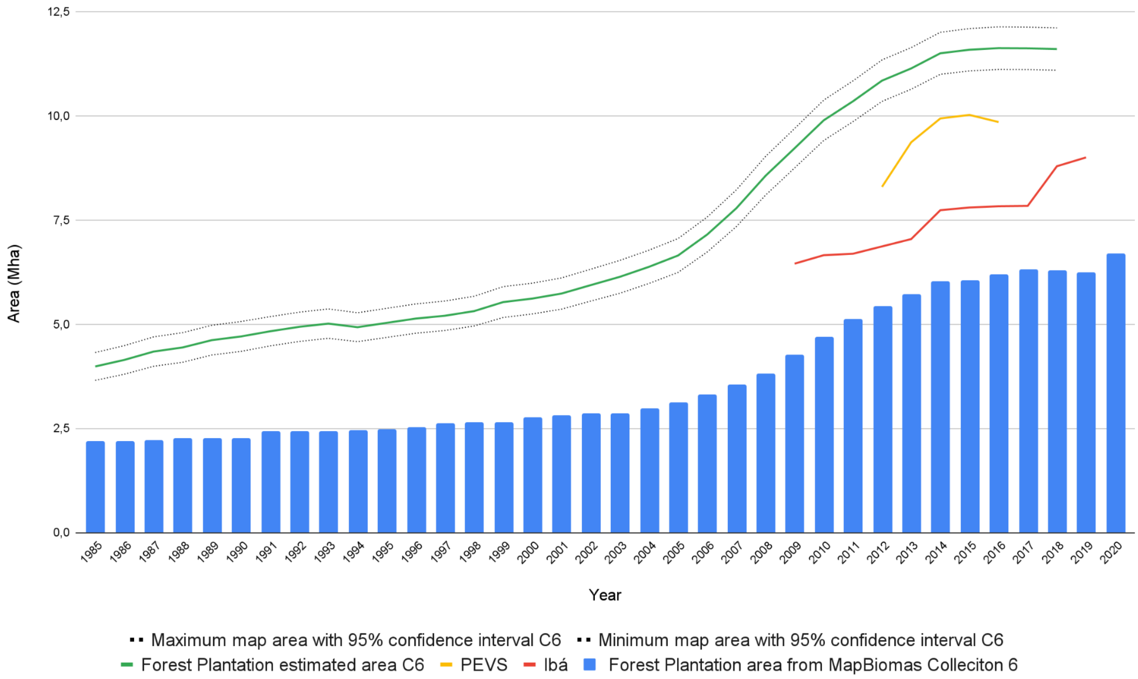


Figure 28. Comparison between MapBiomas, IBGE's Production of Vegetable Extraction and Forest Plantation (PEVS-IBGE) and Ibá - Indústria brasileira de árvores (Brazilian Tree Industry).

6 References

AGROSATÉLITE. Agrosatélite Applied Geotechnology Ltd. Scripts used for Agriculture and Forest Plantation classification on Collection 6 of Mapbiomas project. **GitHub repository**, 2021. Available at: <https://github.com/mapbiomas-brazil/agriculture/tree/mapbiomas60> and <https://github.com/mapbiomas-brazil/forest-plantation/tree/mapbiomas60>.

AGROSATÉLITE. Agrosatélite Applied Geotechnology Ltd. Geospatial Analysis of Soy in the Cerrado Biome: Expansion Dynamics | Soy Agricultural Aptitude | Evaluation System for Financial Compensation: 2001 to 2019. - Florianópolis, 2020a 60 p. : il ISBN: 978-65-991465-1-0. Available at: <https://agrosatelite.com.br/wp-content/uploads/2020/06/relatorio_ccm.pdf>.

AGROSATÉLITE. Agrosatélite Applied Geotechnology Ltd. Soy Moratorium|crop season 2018/2019|. Monitoring by satellite images of soybean plantations in non-compliance with the soy moratorium. - Florianópolis, 2020b. 56p. Available at: <https://agrosatelite.com.br/cases/#moratoria>.

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

CONAB. Companhia Nacional de Abastecimento. Calendário de plantio e colheita de grãos no Brasil. 2019. 75p. ISSN: 2318-6852. Available at: <https://www.conab.gov.br/institucional/publicacoes/outras-publicacoes/item/7694-calendario-agricola-plantio-e-colheita>.

DEERING, Donald Wayne. Rangeland reflectance characteristics measured by aircraft and spacecraft sensors. **Ph. D Dissertation. Texas A&M University**, 1978.

FORMAGGIO, Antonio Roberto, Ieda Del'Arco Sanches | Sensoriamento Remoto em agricultura | São Paulo: Oficina de Textos, 2017. ISBN 978-85-7975-277-3.

GAO, Bo-Cai. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. **Remote Sensing of Environment**, v. 58, n. 3, p. 257-266, 1996.

IBGE, Instituto Brasileiro de Geografia e Estatística. Produção Agrícola Municipal - PAM. Disponível em: <https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9117-producao-agricola-municipal-culturas-temporarias-e-permanentes.html?=&t=o-que-e>

IBÁ, Indústria brasileira de árvores, Disponível em: <https://www.iba.org/publicacoes/relatorios>.

JIANG, Zhangyan et al. Development of a two-band enhanced vegetation index without a blue band. **Remote sensing of Environment**, v. 112, n. 10, p. 3833-3845, 2008.

KURSA, Miron B. et al. Feature selection with the Boruta package. **J Stat Softw**, v. 36, n. 11, p. 1-13, 2010.

KING, L. A., Adusei, B., Stehman, S. V., Potapov, P. V., Song, X. P., Krylov, A., Di Bella, C., Loveland, T. R., Johnson, D. M., & Hansen, M. C. (2017). A multi-resolution approach to national-scale cultivated area estimation of soybean. **Remote Sensing of Environment**, 195, 13–29. https://doi.org/10.1016/j.rse.2017.03.047.

LSPA. Levantamento Sistemático da Produção Agrícola. 2019. 75p. ISSN: 2318-6852. Available at: <https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9201-levantamento-sistematico-da-producao-agricola.html?=&t=o-que-e>.

NAGLER, Pamela L. et al. Cellulose absorption index (CAI) to quantify mixed soil–plant litter scenes. **Remote Sensing of Environment**, v. 87, n. 2-3, p. 310-325, 2003.

PEVS, Produção da Extração Vegetal e da Silvicultura, PEVS. Disponível em:

<https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9105-producao-da-extracao-vegetal-e-da-silvicultura.html?=&t=downloads>

POTAPOV, P., Hansen, M. C., Kommareddy, I., Kommareddy, A., Turubanova, S., Pickens, A., ... & Ying, Q. (2020). Landsat Analysis Ready Data for Global Land Cover and Land Cover Change Mapping. **Remote Sensing**, 12(3), 426.

RIZZI, Rodrigo et al. Estimativa da área de soja no Mato Grosso por meio de imagens MODIS. **Simpósio Brasileiro de Sensoriamento Remoto**, v. 14, p. 387-394, 2009.

SONG, X.-P., HANSEN, M.C., POTAPOV, P.V., ADUSEI, B., PICKERING, J., ADAMI, M., LIMA, A., ZALLES, V., STHEHMAN, S.V., DI BELLA, C.M., CECILIA, C.M., COPATI, E.J., FERNANDES, L.B., HERNANDEZ-SERNA, A., JANTZ, S.M., PICKENS, A.H., TURUBANOVA, S., & TYUKAVINA, A. (In review). Massive soybean expansion in South America since 2000 and implications for conservation. **Nature Sustainability.**

RUDORFF, Bernardo Friedrich Theodor et al. Studies on the rapid expansion of sugarcane for ethanol production in São Paulo State (Brazil) using Landsat data. **Remote sensing**, v. 2, n. 4, p. 1057-1076, 2010.

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

UNIVERSITY OF MARYLAND. Global Land Analysis Discovery - Commodity Crop Mapping and Monitoring in South America. 2000-2019. Avaiable at:< https://glad.umd.edu/projects/commodity-crop-mapping-and-monitoring-south-america>.