Algorithm Theoretical Basis Document (ATBD)

MapBiomas Fire

Collection 2.0

Version 1

General coordinator
Ane A. Alencar
Team
Dhemerson E. Conciani
Diego Pereira Costa
Eduardo Reis Rosa
Eduardo Vélez Martin
Heinrich Hasenack
Luiz Felipe Morais Martenexen
João Paulo Fernandes Márcio Ribeiro
Julia Shimbo
Marcos Rosa
Mariana Dias
Natalia Crusco
Newton Coelho Monteiro
Sarah Moura B. dos Santos
Soltan Galano Duverger
Tasso Azevedo
Valderli J. Piontekowski
Vera L. S. Arruda
Wallace Vieira da Silva
Washington da Franca Rocha

Table of Contents

1.Introduction

1.1. Overview of the MapBiomas Fire

The objective of this document is to describe the theoretical basis, justification and applied to produce the monthly and annual maps of burned areas in Brazil from 1985 to 2022 of the MapBiomas Fire Collection 2.

The MapBiomas Fire project released its collection 1 of annual maps of fire scars covering the period of 1985 to 2020 in August 2021. Some improvements in the methodology and the collection of samples was made for the MapBiomas Fire Collection 2 of mapping of fire scars in Brazil, which is based on mosaics of images from the Landsat satellites with a spatial resolution of 30 meters. The mapping period was expanded from 1985 to 2022, with monthly and annual data of burned areas covering the entire Brazilian territory.

The entire process was carried out collaboratively between MapBiomas institutions and using machine learning algorithms (deep learning) through the Google Earth Engine and Google Cloud Storage platform which offer immense processing capacity in the cloud.

The classification was organized by biomes and regions, collecting burned and unburned area samples for training the algorithm by regions, and using reference maps, such as MODIS Burned Area (MCD64A1 - https://lpdaac .usgs.gov/products/mcd64a1v006/) with 500 m of spatial resolution and INPE fire scars (https://queimadas.dgi.inpe.br/).

The products of the MapBiomas Fire Collection 2 are the following:

- Monthly and annual burned area maps in Brazil from 1985 to 2022;
- Frequency of annual burned areas in Brazil;
- Accumulated burned areas in Brazil;
- Burned areas over Land Use and Land Cover classes of MapBiomas Collection 7.1

The classification algorithms are available on MapBiomas Github (https://github.com/mapbiomas-brazil/fire).

1.2. How we are organized

MapBiomas is a multi-institutional initiative of the Climate Observatory (a network of NGOs working on climate change in Brazil - http://www.observatoriodoclima.eco.br/en/). The co-creators of MapBiomas involve NGOs, universities, and technology companies. For the MapBiomas Fire, IPAM conducted technological and operational development. The geospatial tech company Ecostage is responsible for the backend and dashboard/website/frontend development of MapBiomas. Expert teams in each biome carried out sampling, evaluation and refinement of the mapping, as shown in the box below.

Biome coordination:

• **Amazon** – Amazon Environmental Research Institute (IPAM).

• **Atlantic Forest** – SOS Atlantic Forest Foundation and ArcPlan.

• **Caatinga** – Geodatin.

• **Cerrado** – Amazon Environmental Research Institute (IPAM).

• **Pampa** – GeoKarten.

• **Pantanal** – SOS Pantanal Institute and ArcPlan.

1.3. Historical Perspective: Existent Maps and Mapping Initiatives:

There are few global products that map large-scale burned areas at higher temporal resolution (e.g., twice a day), such as the MODIS (Moderate Resolution Imaging Spectroradiometer) based product MCD64A1 Collection 6, at 500 m pixel resolution provided by the National Aeronautics and Space Administration (NASA). We used the product MCD64A1 Burned Area Product as a reference data of burned areas (Giglio et al., 2016). Additionally, we used the fire hotspots products developed by the National Institute for Space Research (INPE) in Brazil. The INPE fire hotspot product is based on an automatic mapping approach using 1 km x 1 km pixel size and thermal bands of nine satellites, and the AQUA_M-T (Sensor MODIS) as a reference satellite, providing daily data of fire hotspots since 2000, available at http://www.inpe.br/queimad as/bdqueimadas.

We also used the 30-m resolution Global annual Burned Area Map - GABAM, defined as spatial extent of fires that occurs within a whole year generated via an automated global burned area mapping approach based on all the available Landsat images on GEE platform (Long, 2019).

2. Methodological description

We used all available Landsat imagery (Landsat 5, 7, and 8) and Deep Neural Network (DNN) model to detect and map burned areas within the Brazilian biomes between January 1985 to December 2022. The DNN models use artificial intelligence and machine learning algorithms to perform deep learning classifications of complex phenomena generating higher performance results, including for fire mapping (Langford, 2018).

The images were treated in Google Earth Engine (GEE) to create annual Landsat quality mosaics, used to collect burned and unburned spectral signatures, to serve as training samples for the classification model. The training samples and the annual quality mosaics were exported to a Google Cloud Storage Bucket to be used as input in virtual machines to train the DNN models, process the burn scar mapping and produce a dataset of 38 years of monthly burned area for all of Brazil from 1985 to 2022. The image processing and classification routines used to map the monthly burned areas in the Brazilian territory followed six steps including: (1) definition of the classification regions per biome, (2) construction of annual Landsat quality mosaics, (3) collection of training samples containing spectral signatures of burned and unburned areas on the annual quality mosaics, (4) training and development of the DNN prediction model, (5) use of post-classification routines with masks and spatial filters, and (6) accuracy assessment (validation). Figure 1 presents our methodological approach for detecting and mapping burned areas in the Brazilian biomes. (Figure 1).

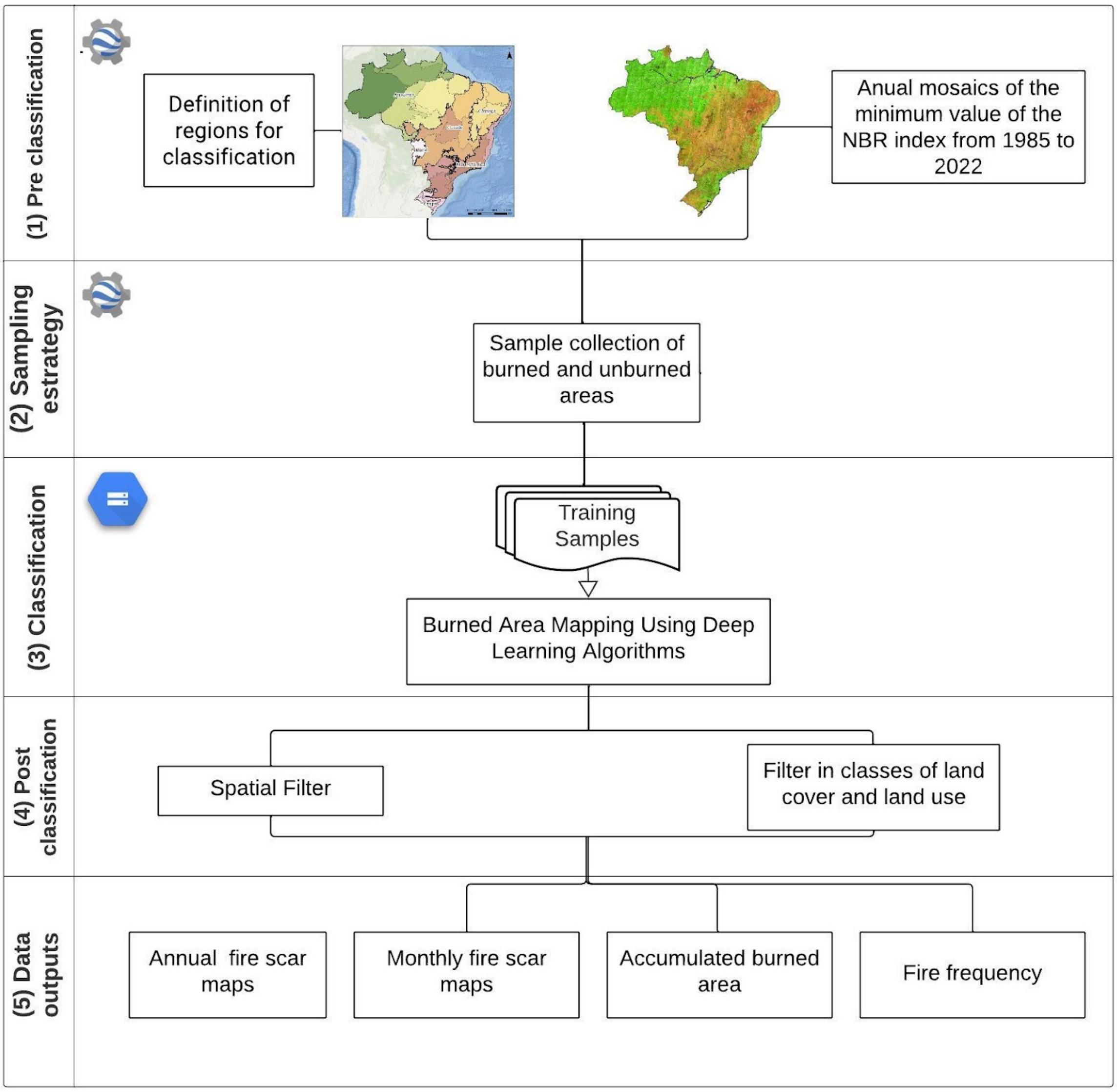

Figure 1. Overview of the method for classifying burned areas in Brazil in MapBiomas Fire Collection 2.0.

2.1. Definition of regions by biome

Considering that the fire regimes and burned area spectral signatures are influenced by climatic conditions and the burned land cover and land use type, we combined edapho-climatic and morphoclimatic data with annual maps of land cover and land use from Map-Biomas Collection 7, to segment each biome into classification regions (Figure 2). This process resulted in 28 classification regions, addressing regional patterns and conferring a more accurate classification of burned areas.

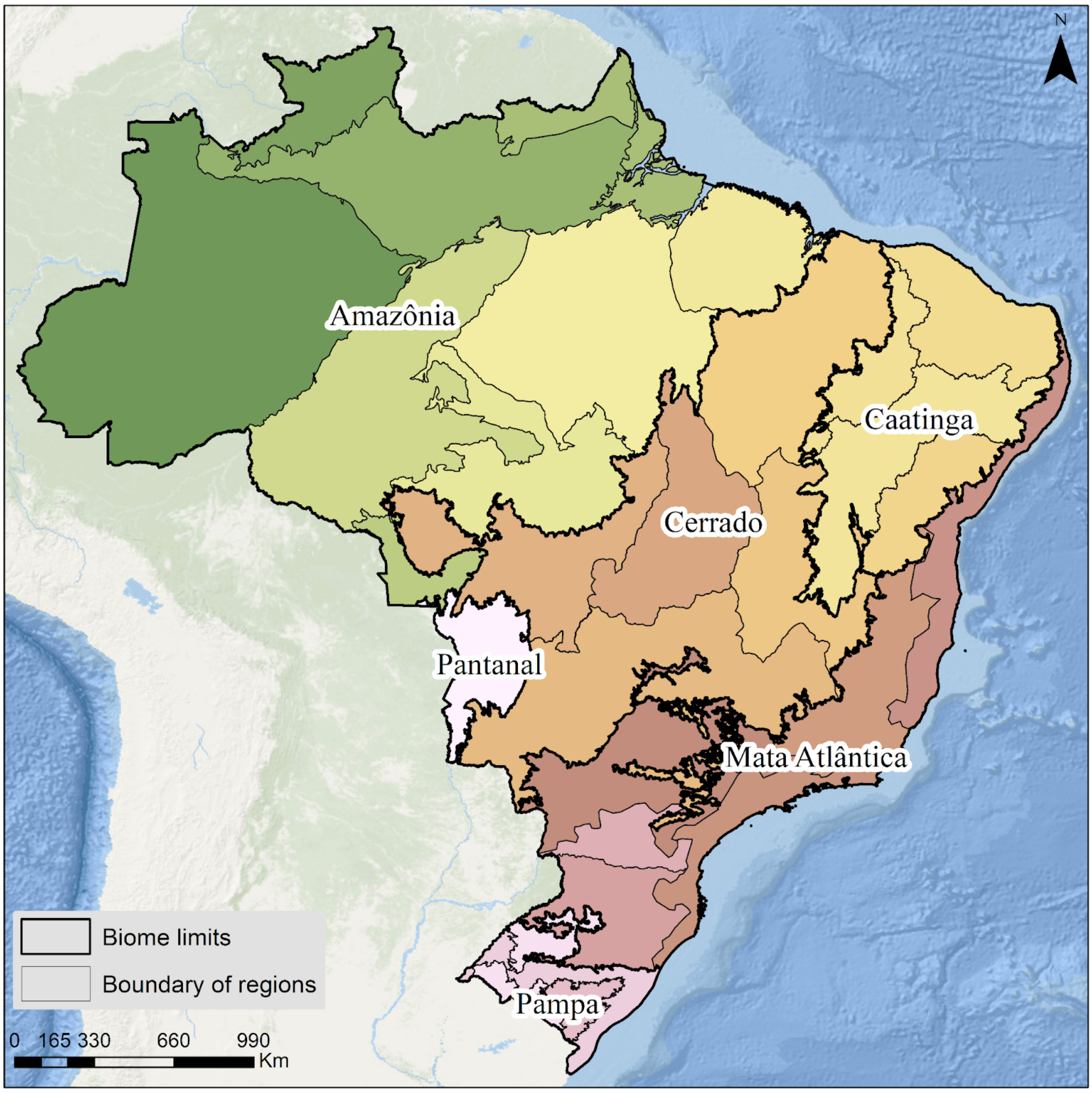

Figure 2. Regions defined for each biome in Brazil to collect training samples and classify burned areas in the MapBiomas Fire Collection 2.

2.2. Annual mosaics

The classification was performed using surface reflectance (SR) USGS Landsat Collection 2 (Tier 1) mosaics (30m × 30 m) constructed for each year from 1985 to 2022. We assessed all the available Landsat 5 (from 1985 to 2011), Landsat 7 (1999 to 2012), and Landsat 8 scenes (2013 to 2022), with a 16-day return interval.

Currently, Landsat Surface Reflectance is accompanied by two Bitwise Quality Assessment bands (QA_PIXEL and QA_RADSAT) that indicate the pixels with radiometric and instrument related problems, including a probability flag. We used the QA_PIXEL band to select and mask the pixels with high confidence levels (67–100%) of 'cloud' and 'shadow'. Then, we used the QA_RADSAT to avoid pixels with radiometric saturation in any surface reflectance band. Finally, we discarded pixels with negative values in the surface reflectance in order to eliminate anomalies and noises in annual quality mosaic composition.

We used a per year statistical approach to summarize this amount of data and optimize the classification without discarding spectral information on a pixel basis. This approach allowed us to create yearly mosaics by performing the composition of all the 16-day images into a single quality mosaic (QM), using the minimum NBR (Normalized Burn Ratio) spectral index (eq. 1 — Key and Benson, 2006) s a per-pixel ordering function, where the pixel with the lowest value of NBR was selected and all the spectral reflectance characteristics, including the scene metadata with the date of that selected pixel, were used to create the annual quality mosaic.

$$\lambda QM = [Blue,\ Green,\ Red,\ NIR,\ SWIR1,\ SWIR2] = date\ in\ with\ min\ \left(\frac{\lambda\, NIR - \lambda\, SWIR1}{\lambda\, NIR + \lambda\, SWIR1}\right)\ [\ xi \dots j] \qquad \text{eq. 1}$$

where λ represents the reflectance values of the quality bands that compose the quality mosaic (QM), retrieved from the date in which each pixel reached their minimum (min) NBR value in a given year (x), considering the set of all available scenes, from first (i) to last (j); the λNIR is the Near-Infrared surface reflectance and λSWIR1 is the Short-Wave Infrared surface reflectance used to calculate the NBR spectral index. In other words, we computed the NBR for each pixel with valid observation within a specific year and stacked them into a multi-band image. The pixels with lowest NBR within the multi-band image were selected and their spectral information (Table 1) were used to compose the annual quality mosaic (QM). In addition to the spectral information, we retained the scene metadata information including the date in which each pixel showed its lowest NBR value. The NBR quality mosaic created with the spectral information from the minimum NBR performed well in differentiating burned and unburned land use and cover in the Brazilian biomes (Figure 3).

Table 1. Mosaic bands used as predictors in the classification of burned areas.

Spectral band	**Landsat 5 and 7**		**Landsat 8**	
	Band number	**Band width (µm)**	**Band number**	**Band width (µm)**
Red	3	0.63 - 0.69	4	0.64 - 0.67
NIR	4	0.76 - 0.90	5	0.85 - 0.88
$SWIR_1$	5	1.55 - 1.75	6	1.57 - 1.65
$SWIR_2$	7	2.08 - 2.35	7	2.11 - 2.29

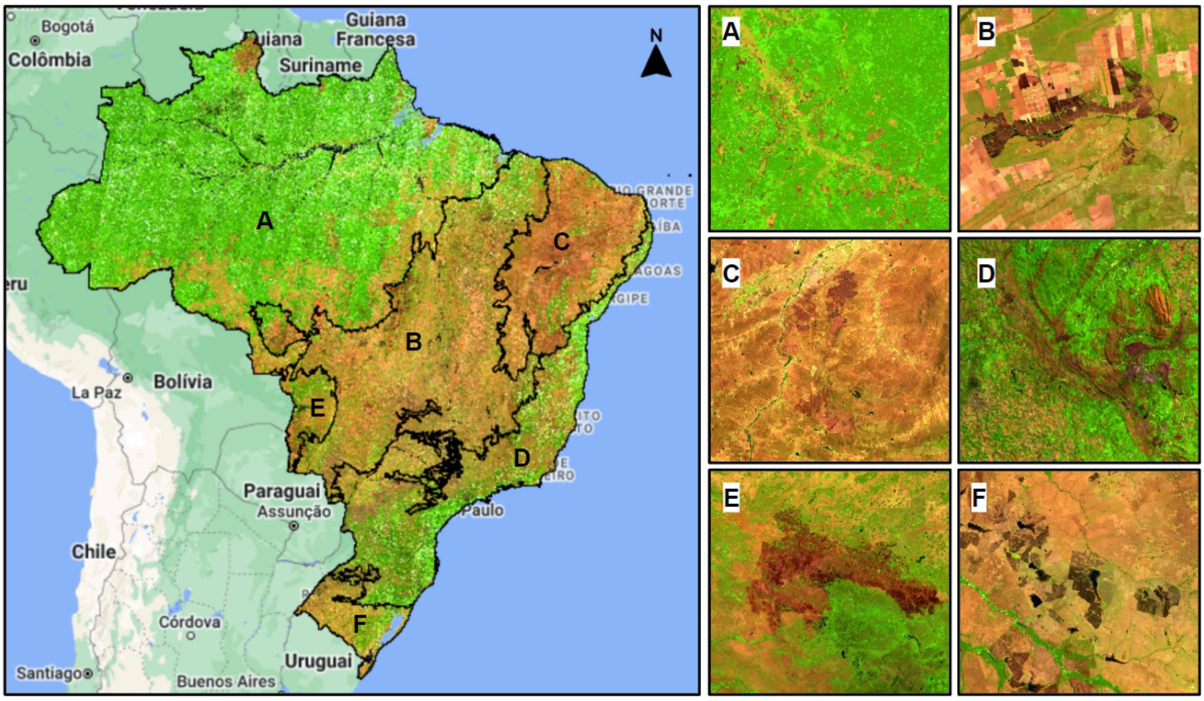

Figure 3. The 2022 quality mosaic (QM) for Brazil (RGB SWIR-1, NIR, RED), created from spectral information retrieved from the minimum NBR pixels in a year, showing examples of the diversity of burn scars by biome: (A) Amazon, (B) Cerrado, (C) Caatinga, (D) Atlantic Forest, (E) Pantanal, and (F) Pampa.

3. Training samples

We created a spectral library with 294,456,236 sampled pixels based on manual delineation of burned areas (95,845,700 sampled pixels; 32%) and unburned areas (198,610,536 sampled pixels; 68%) to be used as training samples. These samples were stratified by Landsat sensors (collected in different years) and each biome. Finally, we divided our spectral library into 28 packs (one for each classification region) and used it as input in the classification step.

2.4. Classification

The classification model used was the Deep Neural Network (DNN), which consists of computational models based on mathematical calculations capable of performing deep learning and visual pattern recognition. The structure we used was the Multi-Layer Perceptron Network (MLPN), that incorporates several layers of interconnected computational units, where each node (neuron) in one layer is connected to a node in the next layer (Hu, Wenk, 2009). The layers are divided into: input, hidden, and output layers. For this DNN model the input layers were the spectral bands RED, NIR, SWIR1 and SWIR2, and the output layers were the classes burned and unburned. The burned area mapping algorithm consisted of two steps: training and prediction.

In the training phase, the following parameters were defined, based on prior tests: learning rate (0.001), batch size (1000), number of interactions (7000), and inputs for classification (Arruda et al. 2021). The classification inputs used in this model were the SR spectral data retrieved from the annual quality mosaics using the training samples of burned and unburned areas.

Based on the spectral library from the burned and unburned training samples, the following spectral bands were used as inputs for the burned area classification model: red (RED—0.65 µm), near infrared (NIR—0.86 µm), and short-wave infrared (SWIR 1–1.6 µm and SWIR 2–2.2 µm). These spectral Landsat bands were chosen based on their sensitivity to fire events among distinct land use and covers.

The training data input was divided into two datasets: 70% of the samples were used for training and 30% for testing, in order to estimate the ability of the DNN algorithm to map burned areas. The classification was performed using the annual Landsat quality mosaics, for each one of the 28 regions, and for each sensor (Landsat 5, Landsat 7, and Landsat 8), resulting in 38 maps of burned areas for all of Brazil (Figure 4).

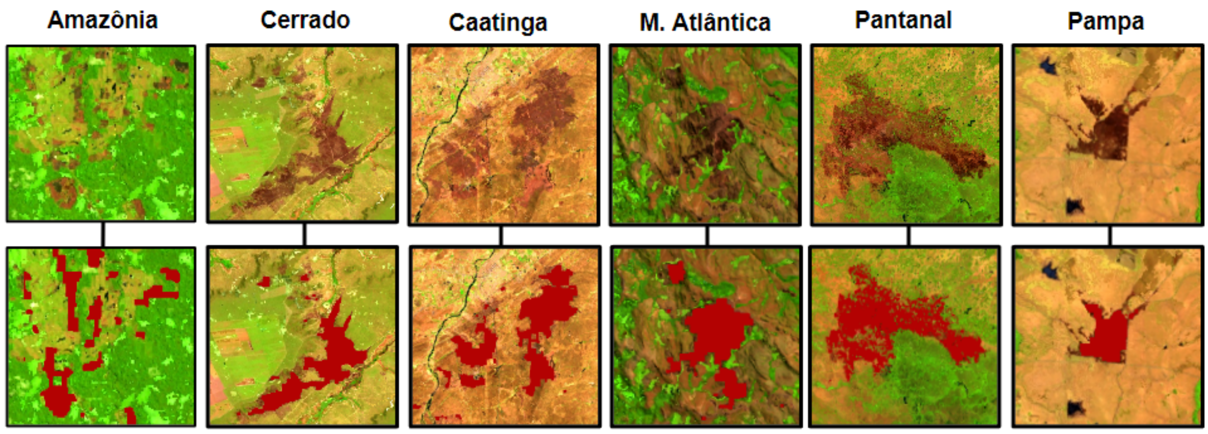

Figure 4. Examples of the burned areas classification for different types of fire, with the Landsat mosaic used for classification, and the area classified as burned in red.

Because deep learning methods require powerful computational processing, we conducted our analysis using graphics processing units (GPUs) and specialized hardware components for running parallel arithmetic operations. The computation infrastructure used was core 8vCPU, 32GB RAM with an additional 200 GB disk. Access to GPUs in a virtual machine environment was implemented on the Google Cloud Platform (https://console.cloud.google.com, accessed on 4 April 2022), a suite of cloud computing services provided by Google.

2.5. Post-classification

A spatial filter was applied to remove noise and fill small empty gaps, where burned areas smaller than or equal to 1.4 ha (16 pixels) were removed, and empty gaps (inside and rounded by burned area) smaller than or equal to 5.8 ha (64 pixels) were filled as burned. After evaluating the classification results, post-classification masks were also applied to reduce the commission from the land use, and cover with spectral signatures that are similar to those of recently burned areas, such as water, urban areas, and some crop types. We defined rules per biome to remove pixels that were classified as burned in the distinct land cover and land use classes of the MapBiomas Collection 7.1 as shown in the box below.

• **Amazon:** Water, Urban Area, Mining and Beach, Dune and Sand Spot;

• **Atlantic Forest:** Water, Urban Area, Rice, Mining, Soybean (region 6 and 7); Temporary Crop, Sugar cane and Other Temporary Crops fore region 6;

• **Caatinga:** Water, and Rocky Outcrop;

• **Cerrado:** Water and Urban Area;

• **Pampa:** Rice, Soybeans, Other Temporary Crops, Mosaic of Uses, Urban Area and Water;

• **Pantanal:** Water, Soybean, Cotton and Other Temporary Crops;

To obtain the information of the month in which the fire scar was mapped, post-classification processing was performed to retrieve the date information of the pixel that was burned, from the date of the pixel in which the annual mosaic was built from the minimum NBR.

2.6. Classification evaluation

Evaluations of the classification of burn scars were carried out with Landsat images, with visual inspection, statistics and relationship with land cover and land use data of MapBiomas, in addition to comparison with reference maps (MapBioms Fire Collection 1, MODIS, INPE, GABAM, FIRMS, FireCCi, hotspots) (Figure 5).

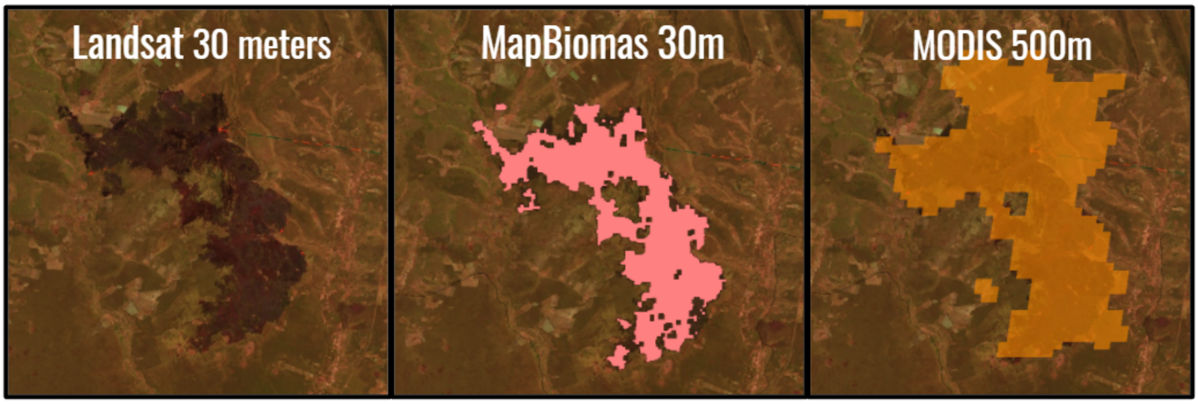

Figure 5: Landsat image with 30 meters resolution, MapBiomas Fire classification and MODIS

(MCD64A1) classification at 500 m resolution.

2.8. Fire Statistics

2.8.1.Annual Occurrence data by Land Use and Land Coverage

Annual burned area data from 1985 to 2022, with an image with 38 bands where each band corresponds to a year and the pixel has the value of the Land Use and Land Coverage class (Collection 7.1) of the pixel that burned, in the respective year.

2.8.2 Monthly Occurrence data by Use and Coverage

The Monthly Occurrence data by Land Use and Land Coverage also cover the period from 1985 to 2022 and are separated by year, where each band corresponds to the monthly scars for that year. The pixel has the value of the month and the LULC code (Collection 7.1) of the burned area.

2.8.3. Cumulative burned area por uso e Cobertura

The accumulated burned area data available in the MapBiomas platform was built from the increment of the burned area for each year; that is, the same pixel is only counted as fire once, regardless of whether there was more than one fire occurrence.

The accumulated burned area data by land cover and land use type was done based on crossing the occurrence of fire with the land cover and land classes of MapBiomas Collection 7.1, considering the last year of the period. The data represents the total number of times that the same pixel had a fire event, for different periods and with the respective Use and Coverage classes, for the last year.

2.8.4 Fire frequency data

The burning frequency data was produced from the grouping of the burned area in each year, running on a single map with 38 classes for the entire period (1985-2022), where class 1 represents the pixels that burned once, class 2 the pixels that burned 2 times and so on.

The frequency data by land cover type was done by crossing the land cover and land use map of MapBiomas Collection 7.1 of the last year of the period.

3.References

Alencar, Ane A C, Vera L S Arruda, Wallace Vieira, Dhemerson E Conciani, Diego Pereira Costa, Natalia Crusco, Soltan Galano Duverger, et al. 2022. “Long-Term Landsat-Based Monthly Burned Area Dataset for the Brazilian Biomes Using Deep Learning.” Remote Sensing 14 (2510): 29. https://doi.org/doi.org/10.3390/rs14112510.

Arruda, V. L. S.; Piontekowski, V. J.; Alencar, A.; Pereira, R. S.; Matricardi, E. A. T. An alternative approach for mapping burn scars using Landsat imagery, Google Earth Engine, and Deep Learning in the Brazilian Savanna. Remote Sensing Applications: Society and Environment, [s. l.], v. 22, p. 100472, 2021.

Giglio, L., Schroeder, W., Justice, C.O., 2016. The collection 6 MODIS active fire detection algorithm and fire products. Remote Sens. Environ. 178, 31–41. https:// doi.org/10.1016/j.rse.2016.02.054.

Hu, X.; Weng, Q. Remote Sensing of Environment Estimating impervious surfaces from medium spatial resolution imagery using the self-organizing map and multi-layer perceptron neural networks. Remote Sens. Environ. 2009, 113, 2089–2102.

Key, C., Benson, N., 2006. Landscape assessment: remote sensing measure of severity: the normalized burn Ratio. In: FIREMON: Fire Effects Monitoring and Inventory System. USDA Forest Service, Rocky Mountain Research Station, General Technical Report, RMRS-GTR-164-CD:LA1-LA51, Fort Collins, CO: U.S.

Langford, Z.; Kumar, J.; Hoffman, F. Wildfire mapping in interior alaska using deep neural networks on imbalanced datasets. In Proceedings of the 2018 IEEE International Conference on Data Mining Workshops (ICDMW), Singapore, 17–20 November 2018

Long, Tengfei, Zhaoming Zhang, Guojin He, Weili Jiao, Chao Tang, Bingfang Wu, Xiaomei Zhang, Guizhou Wang, and Ranyu Yin. 2019. "30 m Resolution Global Annual

Burned Area Mapping Based on Landsat Images and Google Earth Engine" Remote Sensing 11, no. 5: 489. https://doi.org/10.3390/rs11050489