**Deforestation and Secondary Vegetation – Appendix**

**Collection 6**

**Version 1**

# 1 Overview

This document describes the methodology applied to derive annual maps of deforestation and secondary vegetation regrowth from the annual maps of land-use/land-cover (LULC) provided by MapBiomas Collection 6 (MapBiomas). A time series of this dynamic in vegetation cover was produced for all six Brazilian biomes and spanning 1987-2019, by identifying patterns of classification trajectories at the pixel-level that are consistent with loss/regrowth of natural vegetation.

# 2 Methodology

## 2.1 Input Dataset

The main goal of this methodology is to identify events of natural vegetation loss or the return of secondary vegetation after some period of land-use, irrespective of the specific vegetation/land-use classes involved. Therefore, prior to analysis the 21 classes in the original legend of the MapBiomas dataset were aggregated into three generic classes: Anthropic, Natural and Others (Table 1). The resulting time series (1985-2020) was used as input data for the trajectory analysis algorithm described in the section to come.

***Table 1*** *– Aggregation scheme applied to the MapBiomas Collection 6 annual LULC timeseries to produce the input dataset used for classification trajectory analysis.*

| Aggregated class | Original classes included | Raster Value |
|---|---|---|
| Anthropic | Pasture, Forest Plantation, Perennial Crop, Mosaic of Agriculture and Pasture, Soybean, Sugar Cane, Other annual crops, Urban Infrastructure and Mining | 1 |
| Natural | Forest Formation, Savanna Formation, Mangrove, Wetland, Grassland and Other non forest natural formation | 2 |
| Other | Beach and Dune, Other non vegetated area, River, Lake and Ocean, Aquaculture and Non Observed. | 7 |

## 2.2 Classification Trajectory Analysis

Per pixel classification trajectory analysis was conducted within a moving temporal window while applying persistence criteria to differentiate between noisy class transitions (*e.g.* toggle caused by mixed pixels Xie et al., 2020) from transitions that are consistent with deforestation events and secondary vegetation regrowth events. For a given annual map in the input dataset (with three classes), the algorithm identifies pixels in which there was change in relation to the previous year and then check if the classification was persistent before and after the transition. The period for which a pixel had to present constant classification before and after a class change in order to be mapped as vegetation loss or regrowth was named persistence criteria. Changes in the input map that agreed with the defined criteria were classified in the respective loss/regrowth category. Changes that did not agree were reverted to reflect no change in relation to the map in the previous year. The resulting output has five classes (Primary Vegetation, Secondary Vegetation, Loss of Primary Vegetation, Loss of Secondary Vegetation and Regrowth), in addition to the original three classes in the input data. In the next iterative step, which will produce the map for the next year, the output maps from the previous steps are used as reference for past classification trajectories.

For deforestation, the persistence criteria were defined within a temporal kernel of four years: a pixel was mapped as a deforestation event in year *t* if it persisted as Natural for at least two years prior to conversion to Anthropic (*i.e.* Natural in *t*-1 and *t*-2) and persisted as Anthropic for at least one year after the conversion (*i.e.* Anthropic in *t* and *t*+1).

In contrast with deforestation, the regrowth of secondary vegetation is not a discrete event promptly observable from differences in consecutive annual LULC maps. Rather, it is a gradual process that spans for several years, with its duration controlled by several ecological factors: type and duration of the past land-use regime, abundance of propagules sources in the (*i.e.* natural vegetation fragments) in the landscape, climate and topography, among other variables that can vary widely at the biome scale (Aide et al., 2000; Ferreira et al., 2015; Sobrinho et al., 2016; Uriarte et al., 2010). Therefore, we conducted trajectory analysis considering three distinct persistence criteria (*i.e.* considering three different temporal kernels) regarding regrowth of secondary forest, followed by inspection of such versions by specialists in vegetation dynamics in each of the biomes. The evaluation of each version was based on the knowledge of how vegetation regrowth varies as a function of distinct climatic regime, types of vegetation and past land-use regimes are involved. The three sets of persistence criteria to identify pixels of secondary vegetation in year *t* were: (a) persistent classification as Anthropic for at least two years before the conversion (*i.e.* Anthropic in *t*-1 and *t*-2) and persistence as Natural for at least three years after the transition (*i.e.* Natural in *t*, *t*+1 and *t*+2); (b) persistent classification as Anthropic for at least two years before the conversion (*i.e.* Anthropic in *t*-1 and *t*-2) and persistence as Natural for at least five years after the transition (*i.e.* Natural in the *t* to *t*+5 period); (a) persistent classification as Anthropic for at least two years before the conversion (*i.e.* Anthropic in *t*-1 and *t*-2) and persistence as Natural for at least seven years after the transition (*i.e.* Natural in the *t* to *t*+7 period).

Considering that the persistence criteria regarding both loss and regrowth rely on a two years period prior to conversion to check for consistent class changes, the start of the output time series is set at 1987, because years 1985 and 1986 in the input dataset lack two years of information in the past in order to be in the analysis. Similarly, 2020 is the last year in the input dataset and therefore deforestation is mapped until 2019 in the output maps, with 2020 serving only to check for consistent deforestation trajectories. For secondary vegetation regrowth, the final year in the output time series varied according to the respective persistence criteria version: 2018 for the version adopting three years of persistence after class change, 2016 for the five years version and 2014 for the seven years version.

Pixels showing class changes between Natural and Anthropic (or *vice versa*) but not following the defined rules were reclassified in order to correctly represent land-cover/land-use in the next step of the iterative algorithm (*i.e.* when analyzing the next year in the series). For example, when analyzing the 1988 input LULC map, pixels originally classified as Natural in 1987, as Anthropic in 1988 and as Natural again in 1989 were not identified as deforestation in the 1988 output map, because the trajectories do not comply with the persistence criteria for deforestation. Rather, pixels with land-use change trajectories that did not follow the persistence criteria were reclassified to match the classification in the previous year, so that the information available for the next step of the trajectory analysis (1989 in this example) indicates stability until there is a change that follows the persistence criteria.

An overview of the processes through which information in the MapBiomas annual LULC timeseries is used to map vegetation loss or regrowth is given in Figure 1. The seven classes representing vegetation dynamics or stability -- that derive from the trajectory analysis of the original input dataset with three classes – are explained in detail in the next session.

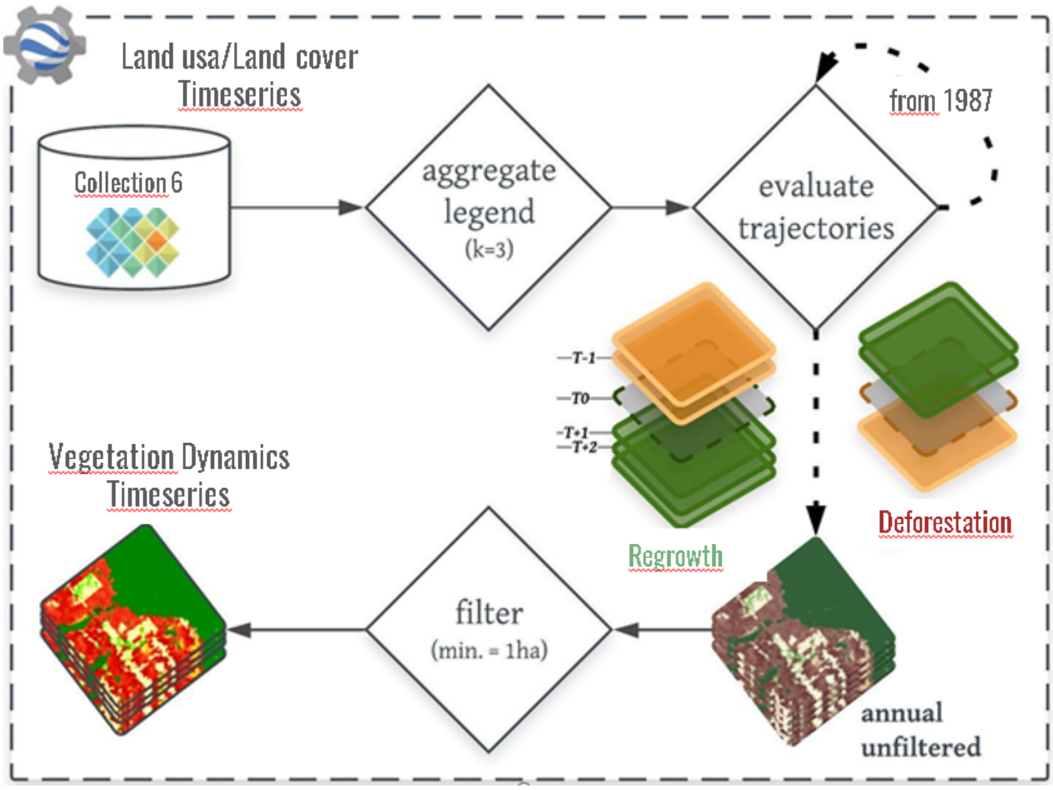


***Figure 1*** *– Overview of the steps needed to map vegetation dynamics using a LULC annual timeseries as input, following the presented methodology. The first step is aggregating 21 LULC classes in the original datasets into three classes. In the second step, pixels in the resulting aggregated annual timeseries have their trajectory analyzed to identify changes that are consistent with the defined persistence criteria. For a pixel to be identified as Regrowth it has to be classified as Natural in the current year of analysis (tile with dashed green border; T0), in (at least) the following two*

*years (green tiles; T+1 and T+2) and also be classified as Anthropic in the two years immediately before the year of analysis (yellow tiles, T -1 and T -2, yellow tiles). For a pixel to be identified as Deforestation (*i.e. *Loss of Primary Vegetation or Loss of Secondary Vegetation) it has to be classified as Anthropic in the current year of analysis (tile with dashed yellow border; T0), in the following years (yellow tile; T+1) and also be classified as Natural (Primary vegetation or Secondary Vegetation) in the two years immediately before the year of analysis (green tiles; T-1 and T-2). The process is carried on iteratively starting by the 1987 map (1985 and 1986 input maps used to check persistence criteria) and the result is an annual timeseries mapping seven classes, which can represent either a type of land cover or a class change event: Primary Vegetation (cover), Secondary Vegetation (cover), Anthropic (cover), Regrowth (change), Loss of Primary Vegetation (change) and Loss of Secondary Vegetation (change). Post processing of the annual timeseries that results from the trajectory analysis consisted of a two-step spatial filter that eliminates small (less than ten 30mx30m pixels), isolated patches of pixels.*

## 2.3 Classification Scheme

The final annual maps produced through trajectory analysis contain seven classes, which can represent either a type of land cover or a class change event: Primary Vegetation (cover), Secondary Vegetation (cover), Anthropic (cover), Regrowth (change), Loss of Primary Vegetation (change) and Loss of Secondary Vegetation (change). The definition of these classes and the persistence rules related to each are shown in Table 2.

***Table 2*** *– Description of the classes mapped in the annual vegetation dynamics timeseries produced by the presented methodology.*

| Class | Description | Rule | Raster Value |
|---|---|---|---|
| Anthropic | Pixels classified as Anthropic in the input dataset and that did not experiment change in the year of analysis. | NA | 1 |
| Regrowth | Areas presenting a historic of Anthropic use followed by change to Natural vegetation specifically in the year of analysis. | Persistent classification as Anthropic for at least two years before the year of analysis and persistent classification as Natural for three/five/seven years after the conversion. | 5 |
| Primary Vegetation | Natural vegetation that persisted from the start of the series (1987) up to the year of analysis. All natural vegetation classes in the input dataset were considered. | Persistent classification as Natural in the input dataset, from the beggining of the series up to the year of analysis. | 2 |
| Secondary Vegetation | Areas presenting a historic of Anthropic use followed by change to Natural vegetation prior to the year of analysis. | Classified as Natural in the input dataset in the year of Analysis and classified as either Regrowth of Secondary Vegetation in the previous iterative step (*i.e.* in map produced for the previous year) | 3 |
| Loss of Primary Vegetation | Areas changing from Primary Vegetation to Anthropic in the year of analysis. | Persistent classification as Primary Vegetation in the maps produced in previous iterative steps (*i.e.* previous years), for at least two years, followed by persistent classification as Anthropic in the input dataset for two years including the year of analysis. | 4 |
| Loss of Secondary Vegetation | Areas changing from Secondary Vegetation to Anthropic in the year of analysis. | Persistent classification as Secondary Vegetation in the maps produced in previous iterative steps (*i.e.* previous years), for at least two years, followed by persistent classification as Anthropic in the input dataset for two years including the year of analysis. | 6 |
| Other | Areas mapped in the input dataset as not Anthropic nor Natural vegetation. Among others, includes the classes Rocky Outcrop, Water Body and Other non-vegetated Area in the MapBiomas dataset. | Classified as Other at any timepoint of the input dataset | 7 |

## 2.4 Post-processing

A two-step spatial filter was applied to all versions of the resulting dataset. In the first step, all pixels that experienced vegetation regrowth throughout the timeseries (*i.e.* classified as Regrowth at least once) were accumulated into a single layer. Patches (*i.e.* connected pixels of the same class) containing less than ten pixels within this mask were eliminated. Such pixels were reclassified according to the mode in the spatial vicinity (considering a three-by-three square kernel). The second step was conducted year-wise instead of considering a temporal accumulated map, using the result of the first step as input; it consisted of reclassifying pixels of Primary Vegetation Loss, Secondary Vegetation Loss or Other that were contained in small patches (less than ten pixels). For each year, eliminated pixels were reclassified according to the mode in the spatial vicinity (three by three square kernel).

# 3 Concluding remarks

The methodology presented here conceptualizes categories of vegetation dynamics based on per-pixel LULC classification trajectories, which demands some premises to be adopted. For example, any natural vegetation mapped in the beginning of the input time series is regarded as Primary Vegetation until it experiments change, even though in reality some of those areas of natural vegetation cover in Brazil had already been used prior to 1985. Additionally, the mapping of Secondary Vegetation following the presented methodology is unable to inform about the quality of the developing vegetation and therefore can represent contrasting ecological processes, such as regeneration, restoration or biological invasion (*eg*. Damasceno et al., 2018; Fernandes et al., 2016; Pinheiro & Durigan, 2009).

Even though the quality of the produced maps is tightly linked to the accuracy of the input dataset (MapBiomas), a validation protocol is being produced to allow per biome quality assessment of the vegetation dynamics classification. The main goal is to reduce uncertainties and eliminate bias when estimating area and accuracy metrics for vegetation dynamic classes that are not prevalent in the territory.

# 4 References

Aide, T. M., Zimmerman, J. K., Pascarella, J. B., Rivera, L., & Marcano-Vega, H. (2000). Forest Regeneration in a Chronosequence of Tropical Abandoned Pastures: Implications for Restoration Ecology. *Restoration Ecology*, *8*(4), 328–338. https://doi.org/10.1046/j.1526-100x.2000.80048.x

Damasceno, G., Souza, L., Pivello, V. R., Gorgone-Barbosa, E., Giroldo, P. Z., & Fidelis, A. (2018). Impact of invasive grasses on Cerrado under natural regeneration. *Biological Invasions*, *20*(12), 3621–3629. https://doi.org/10.1007/s10530-018-1800-6

Fernandes, G. W., Coelho, M. S., Machado, R. B., Ferreira, M. E., Aguiar, L. M. de S., Dirzo, R., Scariot, A., & Lopes, C. R. (2016). Afforestation of savannas: an impending ecological disaster. *Natureza & Conservação*, *14*(2), 146–151. https://doi.org/10.1016/j.ncon.2016.08.002

Ferreira, M. C., Walter, B. M. T., & Vieira, D. L. M. (2015). Topsoil translocation for Brazilian savanna restoration: propagation of herbs, shrubs, and trees. *Restoration Ecology*, *23*(6), 723–728. https://doi.org/10.1111/rec.12252

Pinheiro, E. D. S., & Durigan, G. (2009). Dinâmica espaço-temporal (1962-2006) das fitofisionomias em unidade de conservação do Cerrado no sudeste do Brasil. *Revista Brasileira de Botânica*, *32*(3), 441–454. https://doi.org/10.1590/S0100-84042009000300005

Rosa, M. R., Brancalion, P. H. S., Crouzeilles, R., Tambosi, L. R., Piffer, P. R., Lenti, F. E. B., Hirota, M., Santiami, E., & Metzger, J. P. (2021). Hidden destruction of older forests threatens Brazil's Atlantic Forest and challenges restoration programs. *Science Advances*, *7*(4). https://doi.org/10.1126/sciadv.abc4547

Sobrinho, M. S., Tabarelli, M., Machado, I. C., Sfair, J. C., Bruna, E. M., & Lopes, A. V. (2016). Land use, fallow period and the recovery of a Caatinga forest. *Biotropica*, *48*(5), 586–597. https://doi.org/10.1111/btp.12334

Uriarte, M., Schneider, L., & Rudel, T. K. (2010). Land Transitions in the Tropics: Going Beyond the Case Studies. *Biotropica*, *42*(1), 1–2. https://doi.org/10.1111/j.1744-7429.2009.00580.x

Xie, Z., Pontius, R. G., Huang, J., & Nitivattananon, V. (2020). Enhanced intensity analysis to quantify categorical change and to identify suspicious land transitions: A case study of Nanchang, China. *Remote Sensing*, *12*(20), 1–16. https://doi.org/10.3390/rs12203323