University of São Paulo
"Luiz de Queiroz" College of Agriculture
Center of Nuclear Energy in Agriculture

Dynamics of land use and land cover in the agricultural frontier of the Brazilian Amazon: driving forces of changes and future scenarios

Andrea Santos Garcia

Thesis presented to obtain the degree of Doctor in Science. Area: Applied Ecology

Piracicaba
2020

Andrea Santos Garcia
Bachelor of Science in Biological Sciences

Dynamics of land use and land cover in the agricultural frontier of the Brazilian Amazon: driving forces of changes and future scenarios

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Advisor:
Profª Drª **MARIA VICTORIA RAMOS BALLESTER**

Thesis presented to obtain the degree of Doctor in Science. Area: Applied Ecology

Piracicaba
2020

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Garcia, Andrea Santos

Dynamics of land use and land cover in the agricultural frontier of the Brazilian Amazon: driving forces of changes and future scenarios/ Andrea Santos Garcia - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2020.

143 p.

Tese (Doutorado) - - USP / Escola Superior de Agricultura “Luiz de Queiroz”. Centro de Energia Nuclear na Agricultura.

1. Mudança do uso da terra 2. Desmatamento 3. Distúrbio 4. Ecótono 5. Cerrado 6. Amazônia I. Título

ACKNOWLEDGEMENTS

Throughout the development of this thesis I have received overwhelming support from faculty, peers, family, and friends. I would like to express my deepest appreciation to my supervisor, Dr. Maria Victoria R. Ballester, whose expertise and encouragement were invaluable during the entirety of my program. I would also like to extend my deepest gratitude to my supervisors at the University of Minnesota, Dr. Paul West and Dr. James Gerber, for kindly sharing their knowledge, opportunities, and friendship.

I would like to extend my sincere thanks to the research committee, Dr. Fernanda Reichardt, Dr. Gabriel Lui, and Dr Silvia Molina for their guidance during these years. A special thank you also to Rosangela Montani and Mara Casarin for their unending and very helpful assistance.

I cannot begin to express my thanks to my partner, Patrick Connerton, who has always supported me in this academic journey with insightful discussions, careful reviews, and inspiring coffee breaks.

I gratefully acknowledge the support I have received from my colleagues at the Center for Nuclear Energy in Agriculture: Vivian Vilela, Dr. Rodnei Rizzo, Dr. Henrique Sawakuchi, Mayara Preto, Dr. Erica Nakai, Dr. Daiana Monteiro, Dr. Juliana Vicentine, and Laura Casarin. Special thanks to Vivian Vilela and Dr. Rodnei Rizzo who have made of the field work a much more enjoyable adventure. I am also thankful for the insightful discussions initiated by colleagues at the Institute on the Environment, with special thanks to the Global Landscape Initiative team and members of the grazing discussion group: Peder Engstrom, Lindsey Sloat, Clare Kazanski, Ginger Kowal, Ethan Butler, Kate Brauman, and Deepak Ray.

I'm extremely grateful to my parents, Maria Alice and Sergio Garcia, who have always supported and nurtured me, and to my brothers, Rafael and Luca Garcia. The path that led me here would not have been nearly as gratifying without you all. A special thanks also to Mary and John Connerton, who have provided me with encouragement and support during my year in Minnesota and beyond.

This work would not have been possible without the financial support provided by the Coordination for the Improvement of Higher Education (CAPES), and the São Paulo Research Foundation (FAPESP). The development of this thesis was supported by grant 2015/05103-8 and 2017/12787-6 from FAPESP, and by the Belmont Forum/funded Xingu project ("Integrating land use planning and water governance in Amazonia: towards improved freshwater security in the agricultural frontier of Mato Grosso", grant number 2013/50180-5).

SUMMARY

RESUMO **6**

ABSTRACT **7**

1. INTRODUCTION **9**

1.1. LAND CHANGE SCIENCES 9
1.2. LAND USE AND LAND COVER CHANGE 11
1.3. LAND CHANGE DETECTION AND MODELLING 12
1.4. THE UPPER XINGU RIVER BASIN - A TROPICAL AGRICULTURAL FRONTIER 15
1.5. OBJECTIVES 16

2. ASSESSING LAND USE/COVER DYNAMICS AND EXPLORING DRIVERS IN THE AMAZON'S ARC OF DEFORESTATION THROUGH A HIERARCHICAL, MULTI-SCALE AND MULTI-TEMPORAL CLASSIFICATION APPROACH **21**

ABSTRACT **21**

2.1. INTRODUCTION 21
2.2. STUDY AREA AND DATA 23
2.2.1. Study area *23*
2.2.2. Data acquisition and pre-processing *24*
2.2.2.1. Landsat data 24
2.2.2.2. MODIS and derived indices 25
2.2.3. Ground reference *25*
2.3. METHODS 26
2.3.1. Decision tree and classification rules *27*
2.3.2. Accuracy assessment *32*
2.3.3. LULC change *32*
2.4. RESULTS 33
2.4.1. Classification approach *33*
2.4.2. Accuracy and class separability *33*
2.4.3. LULC change in the UXRB *36*
2.5. DISCUSSION 39
2.5.1. LULC classification and legend *39*
2.5.2. LULC changes in the UXRB *41*
2.6. CONCLUSIONS 42

3. DEFORESTATION AND DEGRADATION IN AN ECOTONE ZONE BETWEEN AMAZON AND CERRADO BIOMES: COMPARING INDICATORS **49**

ABSTRACT **49**

3.1. INTRODUCTION 49
3.2. DATA 50
3.3. MATERIAL AND METHODS 51
3.3.1. Study area *51*
3.3.2. Image processing *51*
3.3.3. Comparisons *53*
3.4. RESULTS 55
3.5. DISCUSSION 62
3.6. CONCLUSION 63

4. MODEL SELECTION FOR DEFORESTATION, DEGRADATION, AND VEGETATION LOSS SUGGESTS DIFFERENT CAUSATIONS AND FUTURE SCENARIOS **67**

ABSTRACT **67**

4.1. INTRODUCTION 67
4.2. MATERIAL AND METHODS 68
4.2.1. Study area *68*
4.2.2. Dataset *70*
4.2.3. Model fitting and selection *73*

4.3. RESULTS 74
4.4. DISCUSSION 76
4.5. CONCLUSION 77
5. GENERAL CONCLUSION 81
APPENDICES 83
APPENDIX A 83
APPENDIX B 126
APPENDIX C 134

RESUMO

Dinâmica do uso e da cobertura do solo na fronteira agrícola da Amazônia brasileira: forçantes de mudanças e futuros cenários

A fronteira agrícola da Amazônia está localizada em zona de ecótono entre a Floresta Amazônica e o Cerrado. A região apresenta alta biodiversidade, e uma paisagem heterogênea, composta por diferentes tipos de fitofisionomias e usos da terra. Nesta região de fronteira, o Alto da Bacia do Rio Xingu (UXRB), com aproximadamente 170.000km² no estado do Mato Grosso, produz 2% da soja mundial e 0,2% do gado consumido no mundo. No entanto, devido a suas paisagens heterogêneas, essa fronteira não é bem representada por modelos gerais que retratam a detecção e a mudança do uso da terra, ou a perda de vegetação nativa. Nosso objetivo nesta pesquisa foi mapear as mudanças no uso da terra na Bacia do Alto Xingu e modelar a perda de vegetação na região. No primeiro capítulo apresentamos uma visão geral dos diferentes conceitos que foram utilizados ao longo desta pesquisa. No segundo capítulo, construímos um esquema de classificação hierárquica com três níveis de informação, melhorando como os mapas de cobertura e uso da terra capturam a heterogeneidade da região. Observamos que a intensificação agrícola ocorreu principalmente na Amazônia, enquanto o Cerrado sofreu uma expansão na área agrícola. Nas últimas décadas, a região vive uma transição de estágio pioneiro de desenvolvimento para uma fronteira agrícola consolidada, com desenvolvimento orientado para commodities. Desta maneira, o aumento na área agrícola está atrelado a mercados internacionais e à relação dólar americano/real brasileiro. No terceiro capítulo, comparamos diferentes fontes de dados para identificar dois processos distintos de perda de árvores: desmatamento e perturbação. Observamos uma diferença impressionante entre os conjuntos de dados construídos para detectar tanto a perturbação quanto o desmatamento. Este padrão está relacionado ao tipo de vegetação dominante e especificidades nos diferentes modelos. No quarto capítulo, analisamos diferentes variáveis espaciais relacionadas às características biofísicas, infra-estrutura, desenvolvimento econômico e composição e configuração da paisagem para selecionar um modelo que modelasse adequadamente o desmatamento futuro, a perturbação florestal e a perda geral de vegetação nativa (incluindo fisionomias não florestais). Nosso trabalho mostra que variáveis influenciam esses processos de diferentes maneiras, levando-nos a concluir que, para lidar com a perda de vegetação, pesquisadores e formuladores de políticas precisam se concentrar em outros processos além do desmatamento da floresta tropical, o qual é o foco tradicional.

Palavras-chave: Mudança do uso da terra; Desmatamento; Distúrbio; Ecótono; Cerrado; Amazônia

ABSTRACT

Dynamics of land use and land cover in the agricultural frontier of the Brazilian Amazon: driving forces of changes and future scenarios

The Amazon`s agriculture frontier is located in the ecotone zone between the Amazon rain forest and the Cerrado (Brazilian savanic formations). The region presents high biodiversity concentrations along with a heterogeneous landscape, comprised of different types of vegetational physiognomies and land uses. In this frontier region, the Upper Xingu River Basin (UXRB), draining approximately 170,000 km² in Mato Grosso state produces 2% of the world's soybeans and 0.2% of the world's consumed cattle meat. However, due to their heterogeneous landscapes, these frontiers are usually poorly represented by general models portraying land use and land cover detection and change, or native vegetation loss. Our goal in this research was to map land use change in the Upper Xingu River Basin and to model vegetation loss in the region. In the first chapter we present an overall overview of different concepts that were applied throughout this research. In the second chapter, we show the results of a hierarchical classification scheme built with three levels of information for improving how land cover and land use maps capture the region's heterogeneity. We observed that agricultural intensification occurred mainly in the Amazon while the Cerrado has undergone an expansion in agricultural area. In the last decades, the region has been experiencing a transition from a pioneer stage of development to a consolidated frontier, with commodity-oriented development. Thus, increases in agricultural areas are tied to both international markets and the American dollar/Brazilian real ratio value. In the third chapter, we compare different data sources to identify two distinct tree loss processes: deforestation and disturbance. We observed an impressive difference between datasets built to detect both disturbance and deforestation. This pattern is related to the dominant vegetation type and specificities in the different models. In the fourth chapter, we analyzed different spatial variables related to biophysical characteristics, infrastructure, economic development, and landscape composition and configuration to select a model which would adequately represent future deforestation, forest disturbance, and general native vegetation loss (including physiognomies other than forest). Our work shows that variables influence these processes in different ways, leading us to conclude that to tackle vegetation loss, both researchers and policy makers have to focus on processes other than the rain forest deforestation, which has been the traditional focus.

Keywords: Land use change; Deforestation; Disturbance; Ecotone; Cerrado; Amazon

1. INTRODUCTION

This thesis is a product derived from a researched project supported by FAPESP and inserted in a Belmont Forum international collaboration research initiative on fresh water security intitled "XINGU Project - Integrating land use planning and water governance in Amazonia: towards improved freshwater security in the agricultural frontier of Mato Grosso" hosted at the Center for Nuclear Energy in Agriculture (CENA/USP) by the Geoprocessing and Environmental Analysis Lab. Additionally, it was developed partly in an internship with the Landscape Initiative team in the Institute on the Environment at the University of Minnesota (USA).

In this chapter presents a briefly introduction of the main subjects of this thesis: land change science, land use and land cover change, land change detection and modelling, tropical agricultural frontier (area of study), and the goals of this goals. In the following three chapters we present research projects developed in order to explore the land use change dynamics the Upper Xingu River Basin, one of the Amazon`s agriculture frontiers. The last chapter presents the general conclusions of this work.

1.1. Land change sciences

Changes in the landscape are observed at a range of spatial and temporal scales – from local to global and from daily to millennial frequencies. Such changes are caused by natural forces (glaciation, earthquakes, flooding, and drought) and by anthropogenic forces (urban expansion, forest conversion into agriculture). Since the 18th century, anthropogenic actions have been driven changes on Earth's surface faster than these changes would naturally occur. Crutzen (2002) have described this period as a new geological era - the Anthropocene. Such rapid changes are associated with the humans increasing capacity to exploit ecosystem good and service (e.g.: food and fibre production, mineral extraction, etc) in order to sustain population growth and new consumption patterns. Currently, 12% of the land surfaces on Earth are used for cropping systems and 30% are used as pasturelands (FAO, 2011). The conversion of natural areas into productive regions, and the increasingly intensified use of these areas, has led to a range of impacts, e.g.: about 25 % of the land managed in the world is already highly degraded; increased biogeochemistry cycles alteration such as the higher global concentrations of carbon dioxide and methane in the atmosphere; nutrient pollution such as by excessive nitrogen application; and freshwater discharge alterations (Coe et al., 2011; Don et al., 2011; FAO, 2011; Mueller et al., 2014).

Land change science (also known as land system science or land use science) is an interdisciplinary area of knowledge which aims to understand how land changes, mainly in response to social-ecological interactions (Müller and Munroe, 2014). Human environmental sciences and geographical information/remote sensing sciences are the main disciplines contributing to this field (Verburg et al., 2015). As land use science aims to describe and explain transition processes, the clarification of pattern and process concepts assists in analysis interpretation and evaluation (Garcia and Ballester, 2016). As demonstrated by Turner and colleagues (2007), land change observations consist of patterns of land cover and land use observed through time or space. Processes are measured by the patterns they produce; thus, these processes are indirectly observed. Different land cover change processes generate distinct spatial and temporal patterns according to their nature – for instance, Garcia and Ballester (2016) observed distinct temporal and spatial patterns produced by deforestation in comparison with forest degradation at a Cerrado region (Brazil).

Land use science terminology is commonly imprecise which is disadvantageous when describing processes. To address this issue, Meyfroidt (2016) reviewed and systematised land use science terminologies, which is

adopted as the terminology used in this thesis. According to Meyfroidt (2016), proximate causes of land use change are those direct changes such as agricultural expansion, urbanization, and reforestation. Underlying (or indirect) causes are fundamental forces that are inherently more diffused and that constitute systemic conditions in human-environment relations, such as regulations. A combinatory cause (or INUS) is a term derived from Boolean logic and represents an insufficient but necessary part of a combination of causes which is itself unnecessary but sufficient for the outcome (when combined with other causes). For example, demand for agricultural products related to population growth is a necessary cause for agriculture expansion but not a sufficient one, since expansion patterns can be correlated to other factors as biophysical conditions (Garcia et al., 2017). Any event, fact, or variable employed in an explanation is a "factor". Causes enrolled in an INUS cause are divided into either "contextual" or "trigger" causes: "contextual" if the cause is a slowly changing factor, explaining changes prevalence; "trigger" if the cause is a rapidly changing factor, explaining the precise location or timing of an event. The processes through which a factor produces its effect is termed a "causal mechanism", and a series of causal mechanisms which links an underlying cause is termed a "causal chain". "Driving force" is a factor that is a typical or hypothetical cause of environmental change, but for which the evidence is not sufficient to firmly establish causal effects. When the factor contributes to statistical explanation of an outcome (or other spatial characteristics such as spatial pattern or structure), it is a "(spatial) determinant factor".

Land cover (LC) is considered the observed biophysical component on the Earth's surface while land use (LU) is defined by the activities undertaken in a certain area, representing a direct link between land cover and the actions of people in their environment (FAO, 2016; Turner et al., 2007). LC change can be the result of natural disturbances or human-induced modifications through the direct or indirect change in LU. Aside from pristine land cover conversion, LU changes also include livestock replacement by crops or silviculture, modification of land management practices and inputs, and the resulting interaction of both agricultural cover and management practices, as for instance land degradation or secondary vegetation growth (Garcia and Ballester, 2016). Identifying causative factors requires an understanding of how decisions are made to drive land use change. LCLU patterns both affect and are affected by processes operating at multiple scales. Decision making from land owners to national policies or global agreements influence LCLU change (Rounsevell et al., 2012). Local decision-making depends mainly on production costs, output prices, taxes, subsidies, credit access and regulations that directly cause land use change (Geist et al., 2006). Government's programs, regulations, and market strategies represent the underlining causes of changes. Additionally, processes of LC and LU change are modelled according to landscape characteristics that determine the spatial distribution of LC and LU change dynamics. For example, in the Cerrado biome, pasturelands are preferably located near streams, in areas that are easily accessible by cattle herds, and croplands areas primarily occur near more fertile soil (Garcia et al., 2017; Garcia and Ballester, 2016).

Land use science has increasing importance in societal and scientific agendas, due to looming sustainability problems world-wide. Land use change and its consequences are at the core of many science agendas such as the Belmont Forum (https://www.belmontforum.org/), the National Ecological Observatory Network (NEON) Grand Challenge framework (http://www.neonscience.org/), and the Nexus Network (http://www.thenexusnetwork.org/). Despite large advances the field has achieved, there are still many challenges to meet. Deforestation associated with agricultural expansion has received much attention throught single and meta-studies which mainly address impacts on biodiversity and biogeochemical cycles. However, less attention has been given to factors driving urbanization, wetland, savanna and grassland conversion, forest degradation, and land use intensification (van Vliet et al., 2016). Additionally, standard approaches and terminology, socio-economic dynamics,

trajectory changes and feedback loops in causal chains are rarely addressed in sysnthesis and outcomes assessment (Meyfroidt, 2016; Müller and Munroe, 2014; van Vliet et al., 2016).

1.2. Land use and land cover change

Classically, land use and land cover (LULC) change is defined as a consequence of the demand for food, water, fiber, and energy (Lambin et al., 2001). More recently, different studies have also described the relationship between LCLU change and other direct and indirect factors such as socioeconomic changes and consumers preference (Lambin and Meyfroidt, 2011), land market (Richards, 2015), and ecosystem services depreciation (Kust et al., 2017). Figure 1 shows how different factors can influence LULC change and possible feedbacks mainly considering rural areas and agriculture frontiers.

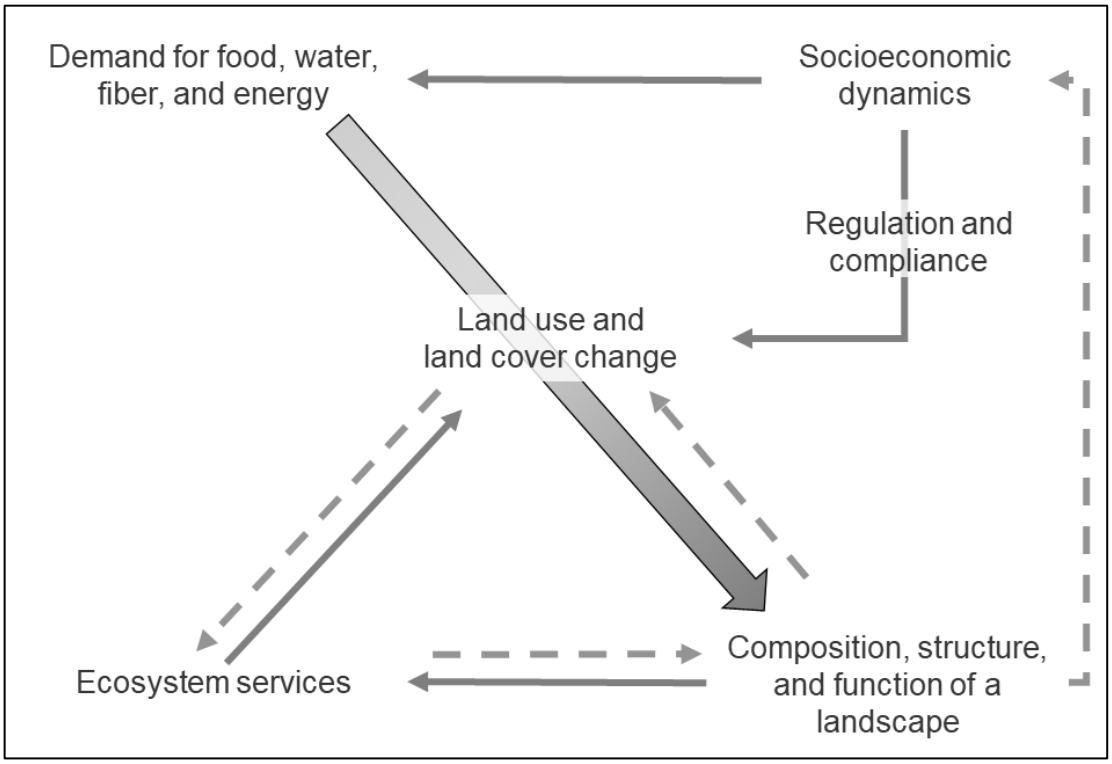


Figure 1. Flow chart of driving forces of land use and land cover change. Dotted arrows represent possible feedbacks.

LCLU also affects a large range of processes shaping ecosystem services and socio-economic dynamics and, thus, plays an important role in meeting demand for food and water (Foley et al., 2005). Deforestation is considered the primary causal mechanism of soil degradation and decreases in soil organic carbon (Chhabra et al., 2006; Don et al., 2011). The conversion of woody cerrado into pasturelands increases soil degradation through superficial erosion, silting streams, and threating water resources (Garcia and Ballester, 2016). While per capita food production increases, agricultural area expansion degrades food security for vulnerable populations in Malawi (Johnson et al., 2013). Such outcomes are a result of lower vitamin A-rich food consumption and water quality depreciation associated with deforestation (Brown, 2016).

Land use change is commonly considered to provide social and economic benefits (Foley et al., 2005). This notion is based on findings of increases in local income; however, this assumption is controversial. Rodrigues and colleagues (2009) found that at the Brazilian Arc of Deforestation, relative standards of living, literacy, and life expectancy increased as deforestation begun but then declined as deforestation frontiers expand. Rather than

increasing income, such pattern may indicate income concentration and no local benefits. Still, vulnerable populations are exploited at short and long-term land use change due its highly dependence on wild food, water and soil quality (Brown, 2016). Working with different scenarios in the US, Polasky and colleagues (2011) estimated that returns to landowners are highest in a scenario with large-scale agricultural expansion. However, such a scnerio also generates the lowest net social benefits across all scenarios considered because of large losses in stored carbon and negative impacts on water quality. Such findings are compatible with several studies in tropical regions in which large agricultural expansion and intensification increases socio-economic benefits for those related to the activity, but at the cost of ecosystem services which mostly affects vulnerable populations.

In recent years changes in land tenure, privatization, and internationalization of agricultural commodity chains make causal factors and drivers of land use change more complex and interconnected (Gasparri and de Waroux, 2015; Verburg et al., 2015). In such recent context, analytical concepts of teleconnection and telecoupling were proposed in order to gain knowledge on drivers of land use change. The former refers to distal environmental and socio-economic drivers of land change while the latter also incorporates feedbacks and multidirectional flows between land systems (Friis et al., 2016; Liu et al., 2013). The assessment of such dynamics is important in guiding land governance. Other emerged challenges related to land governance are to quantify impacts and their potential feedbacks, taking into consideration small and large scales, as well as short- and longer-term time series (Verburg et al., 2015).

With the ending of governmental colonization programs and the lack of state in tropical forest regions, land owners and heads of agribusiness enterprises are assuming an increasingly prominent role in agricultural expansion (Rudel, 2007). Dense supply chains result in more effective lobbying which supports agricultural expansion by encourage infrastructure building, information flow, and flexible environmental regulations (Garrett et al., 2013). Different theories support that on this stage, the pioneer frontier turn into a market-oriented and technified frontier (Meyfroidt et al., 2018). Concurrently, in the Amazon Basin, deforestation was reduced by the implementation of the "Zero Deforestation" campaign and other schemes (Ferreira et al., 2007), as well as by the Brazil's Soy Moratorium (Gibbs et al., 2015). A large expansion of agriculture due to supply chain agglomeration was found in counties with more flexible environmental regulations and outside the moratorium scope than in the counties inside the Amazon biome which are under more protective environmental regulations estabilished by governament and market regulations (Garrett et al., 2013). Nevertheless, governmental regulations such as the Brazilian Forest Code are reported to be ineffective on their own. Conservation initiatives are only secuessful when areas are inspected and when agencies are able to create management plans in junction with landowners (Garcia and Ballester, 2016).

1.3. Land change detection and modelling

Land change detection is the process of identifying differences in the state of a target surface by observing it at different times (Lu et al., 2004). Land use science relies on spatial and temporal data sets, which are extensively derived from remote sensing data. Remote sensing is the process of acquiring data or images of an object distant from the sensor such as aerial photographs, satellite images and radars (Paranhos Filho et al., 2008). In general, aerial photographs and satellite images are representative of spectral intensity, which quantifies the electromagnetic energy reflected from a target surface. All matter above absolute zero (0ºK) emits electromagnetic energy, absorbing a specific portion of the electromagnetic spectrum while reflecting the remaining portions of the spectrum. The result is a distinct spectral signature for each type of land cover. The basic premise in using remotely sensed data for

change detection is that changes in the objects of interest will result in changes in reflectance values or local textures - changes in values of some pixels that constitute an object (Lu et al., 2004).

Currently, Landsat is the longest continuously acquired collection of space-based medium-high resolution remote sensing data. Such characteristics makes Landsat images the most commonly used product in landscape change research (Hansen and Loveland, 2012). Studies based on Landsat images frequently employ a two or three-year anniversary window, with an interval between imagery generation depending on the calendar acquisition dates and the change in interval length (temporal resolution). Anniversary windows are used to minimize discrepancies in reflectance caused by seasonal variation in vegetation and Sun angle. However, for several studies, the availability of satellite sensor data of acceptable quality is a limiting factor which defines acquisition date (Coppin et al., 2004). To detect seasonal change, satellites that return to the same target in shorter intervals (lower temporal resolution) are preferable. In this context, MODIS has been extensively used due to its lower temporal resolution and its higher spectral resolution (a larger wavelength range of the electromagnetic spectrum is sampled). A slight disadvantage is that MODIS presents a moderate resolution with the smallest pixel size being 250 m (https://modis.gsfc.nasa.gov/). Besides Landsat and MODIS, the two most extensively used imagery sets, there are also other sources of imagery available, both freely (SPOT, CBERS, etc) and commercially (Quickbird, IKONOS, etc). Each imagery source has its advantages and disadvantages, thus, each one is recommended for a given application.

There has been much research devoted to evaluating change detection methodologies and the best approach depends on the context and goals (Chen et al., 2012; Coppin et al., 2004; Hansen and Loveland, 2012; Hussain et al., 2013; Lu et al., 2004; Tewkesbury et al., 2015). Change detection can be conducted at different units of analysis (Tewkesbury et al., 2015), been the most common ones pixel-based, object-based, subpixel, kernel and hybrid units. The pixel is the fundamental unit of information in an image and, thus, the fundamental unit of analysis. However, a subpixel analysis is still possible through calculation of fractions of different land covers within a pixel. An object is a group of similarly clustered pixels, which together define a target such as a forest fragment, a building, and so on. Objects are obtained through image segmentation according to the spatial distribution and similarity among pixels (Chen et al., 2012). A kernel (or moving window) is a systematic approach to considering contextual information; it analyses pixels inside of a window of a given size and calculates statistical indices among those pixels (Canty, 2014). Hybrid approaches utilize the combining of more than one approach together (Tewkesbury et al., 2015).

Analysis techniques can be divided into two distinct groups: no-classification (or pre-classification) and post-classification approaches (Canty, 2014). Techniques which do not use classification to detect land change examine differences in two images based on radiance values (Canty, 2014; Lu et al., 2004). Layer arithmetic is a straightforward method which compares the radiance of two images to identify change. Transformation methods use mathematical transformations to highlight variance between images. These methods include, but are not limited to, Fourier analysis, discrete wavelet, Principal Component Analysis (PCA) and Canonical Correlation Analysis (CCA). Another method of no classification change detection is spectral change vector analysis, which detects changes by using chosen thresholds (Canty, 2014). The drawback of most no classification approaches is that they do not offer a complete matrix of change information (Lu et al., 2004). Different from these methods, post-classification change detection defines changes by comparing pixels in a pair of classified images, in which pixels have already been assigned to classes. Post-classification change detection typically reports changes as a summary of the "from-to" changes of categories between two dates. Post-classification change detection is most commonly employed methodology. It minimizes radiometric errors, phenological differences between multi-temporal images, and

provides a complete matrix of change information (Lu et al., 2004). A disadvantage of this method is error propagation, that is errors in any of the input maps are directly translated to the changing results (Tewkesbury et al., 2015).

Image classification is the process of sorting pixels into a finite number of individual classes, or categories, of data based on their pixel values. If a pixel satisfies a certain set of criteria, then the pixel is assigned to the class that corresponds to that criterion. Due to differences among images, classifying multiple images requires operational geoprocessing planning. Hansen and Loveland (2012) offer a review of large area classification methods using Landsat images. In summary, image classification has been traditionally divided according to the minimum area of analysis (pixel-based or object-oriented), and according to the need of providing samples a priori (supervised and unsupervised methods). The LCLU map is composed of a legend of groups obtained throughout a classification scheme adapted to a certain geographical area. As classification schemes are normally built for a specific area of interest, the harmonization and comparability of the resulting data derived from different study regions is an issue in land use science. Such incompatibility is a result of the concepts and methodologies normally used in the geoprocessing flow, which build the legends based only on local biophysical characteristics and sensor characteristics (Di Gregorio and O'Brien, 2012).

Simulation models of land use predict or describe land-use change over space and time and constitute one of the methods available to unravel the dynamics of the land use system. Recent overviews of land-use simulation models show an overwhelming amount of different types of models and applications (van Schrojenstein Lantman et al., 2011). The wide variety of models may be explained by the different scales, research areas and policy questions that land use science addresses and the different disciplines it is originated from (Verburg et al., 2015). Theoretical foundations of land-use simulation models relate to the core principles that are used to explain land-use change and the concepts that are applied to translate these principles into a functioning model of land-use change (Verburg et al., 2004). Thus, a modelling algorithm is a translation of a land use change concept into mathematical rules by use of a computer simulation. Land use models are based on at least one of the following four core principles of land-use change (Lantman et al., 2011): (1) Continuation of historical development - future land use can be predicted by means of historical changes; (2) Suitability of land – based on profit maximization; (3) Result of neighborhood interaction - transition from one use of land to another is dependent on the land use of its surrounding cells; and (4) Result of actor interaction - land-use change is the result of interaction between actors. Given the variety of models, there is no singular superior method for modelling land use change. The choice of which model to use depends research questions and data availability.

Model types present some important distinctions (Verburg et al., 2006). The first distinction is between spatial and non-spatial models. The former aims to achieve spatially explicit representations of land-use change, while the latter focuses on modelling the rate and magnitude of land-use change without specific attention to its spatial distribution. Considering temporal characteristics, dynamic models represent the evolution of land use systems, making them useful for projections of future land-use change. Static models do not account for feedbacks and path dependencies, and are useful to explain the spatial distribution of land use changes as a function of a set of factors. Most land-use change models rely on the inductive approach in which the model specification is based on statistical correlations between land change and a suite of explanatory variables. As with land use change detection, their analytical unit can also differentiate land use change models: an area (a polygon representing either a field, a plot or census track) or a pixel as part of a raster-based representation. Land-use changes are calculated for these spatial objects, directly resulting in maps that show changes in land use patterns. Another group of models use individual

agents as units of simulation. Several characteristics define agents: they are autonomous, they share an environment through agent communication and interaction, and they make decisions that tie behaviour to the environment. Such multi-agent systems give emphasis to agent's decision-making process and to the social organization and landscape in which these individuals are embedded.

1.4. The Upper Xingu River Basin - a tropical agricultural frontier

Agriculture expansion can be linked to the occupation of new areas through deforestation or the occupation of unproductive areas with the potential of expansion/intensification. Expansion on new areas in a consistent pattern characterize agricultural frontier expansion. The Upper Xingu River Basin (hereafter denoted as "UXRB"), is located in Mato Grosso state in Brazil, and in the ecotone between the Amazon rainforest and the neotropical savannas, or Cerrado biome. The main economic activities in the study area during the late 1970's to 1990's were timber and beef. Forests were degraded or cleared due to livestock expansion and logging activities. In the early 2000's, soybean production emerged as another deforestation driver (Brando et al., 2013; Nepstad et al., 2006).

Xingu River is one of the main Amazon River tributaries. Its headwater region encompasses one-third of the Xingu River Basin, draining ~170,000 km2. The UXRB exhibits a wide array of natural physiognomies, ranging from Amazon rainforest to savanna grasslands (Ivanauskas et al., 2008; Velasquez, C.; Alves, H. Q.; Bernasconi, 2010). The Amazon rainforest and Cerrado are the first and second largest biomes in South America. The Amazon rainforest contains enormous biodiversity and one-third of the world's tropical tree species used for timber (Brazil, 2018). The Cerrado is a global biodiversity hotspot and contains the highest land use change rates in Brazil (Brazil, 2018; Klink and Machado, 2005). Due to its position in an ecotone, the basin contains a range of natural vegetational physiognomies, varying from rain forest to savanic formations. The majority of the region is composed of Floresta Estacional Perenifólia – a unique transitional forest with characteristics of both Cerrado and Amazon Biomes (Ivanauskas et al., 2008; Velasquez, C.; Alves, H. Q.; Bernasconi, 2010).

Besides the importance of the region for its natural resources (including biodiversity), the UXRB also encompasses large sociocultural diversity and economic prominence. It is estimated that 16 indigenous ethnic groups and other traditional communities live in the basin (Velasquez, C.; Alves, H. Q.; Bernasconi, 2010). The UXRB exhibits a seasonal tropical climate. The annual mean precipitation is 2000 mm and the annual mean temperature 24 ºC. Seasons are characterized by a dry winter and rainy summer, with a flood season occurring from November to April. The extended and defined rainy season, associated with a flat terrain and being dominated by Oxisol soils (Velasquez, C.; Alves, H. Q.; Bernasconi, 2010), make this region ideal for agribusiness expansion. The main economic activities in the study area are linked to the wood, beef, and soybean industries. Currently, approximately 2% of world's and 9% of Brazil's soybeans are produced in the UXRB, as well as 0.2% of the world's and 13% of Brazil's cattle (FAOSTAT, 2016; IBGE, 2019).

The rapid change in the land use and socioeconomic dynamics in UXRB in the last few decades, and mainly since 2000, make of it an agriculture frontier in development. Combining these intense dynamics with the conditions in the ecotone zone it is located at, the UXRB presents an impressive biophysical, social, and economic diversity. Such characteristics make of this region an ideal study case for land change science which we explore in this thesis.

1.5. OBJECTIVES

To support sustainable development, natural resource management should meet multi-level interests by using reliable information about intervening factors that affect soil, water, and climatic conditions which are directly influenced by patterns of land use. In such context, we aimed to elucidate processes of LCLU change in the Upper Xingu river basin and its relations to the quantitative and qualitative loss of native vegetation, as well as the trends and driving forces of this change; then, design potential future scenarios for the region. We answered the following research questions:

• **Question I:** How has the landscape changed over the last 30 years?

• **Question II:** What is the temporal and spatial distribution of these changes in the landscape, especially regarding the quantitative and qualitative loss of native forests?

• **Question III:** What are the most important factors when predicting future scenarios for the Upper Xingu River Basin and what are these future scenarios?

To answer to these questions, we developed three main research manuscripts presented here in three different chapters. The second chapter of this thesis assessed land use and land cover dynamics in the study area. The third chapter presents the patterns of deforestation and forest degradation as well as compared different indicators which detects both processes. The fourth chapter explore drivers of deforestation, forest degradation, and general vegetation loss (including non-forest vegetation), comparing how variables influence these three different processes.

References

Brando, P.M., Coe, M.T., DeFries, R., Azevedo, A.A., 2013. Ecology, economy and management of an agroindustrial frontier landscape in the southeast Amazon. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 368, 9. https://doi.org/10.1098/rstb.2012.0152

Brasil/MMA, 2018. O Bioma Cerrado [WWW Document]. URL http://www.mma.gov.br/biomas/cerrado (Available)

Brazil/MMA, 2018. Amazônia [WWW Document]. URL http://www.mma.gov.br/biomas/amazônia (Available)

Brown, M.E., 2016. Remote sensing technology and land use analysis in food security assessment. J. Land Use Sci. 11, 623–641. https://doi.org/10.1080/1747423X.2016.1195455

Canty, M.J., 2014. Image Analysis, Classification and Change Detection in Remote Sensing, Third. ed. CRC Press, New York.

Chen, G., Hay, G.J., Carvalho, L.M.T., Wulder, M.A., 2012. Object-based change detection. Int. J. Remote Sens. 33, 4434–4457. https://doi.org/10.1080/01431161.2011.648285

Chhabra, A., Geist, H., Houghton, R.A., Haberl, H., Braimoh, A.K., Vlek, P.L.G., Patz, J., Xu, J., Ramankutty, N., Coomes, O., Lambin, E.F., 2006. Multiple Impacts of Land-Use/Cover Change, in: Lambin, E., Geist, H. (Eds.), Land-Use and Land-Cover Change. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 71–116. https://doi.org/10.1007/3-540-32202-7_4

Coe, M.T., Latrubesse, E.M., Ferreira, M.E., Amsler, M.L., 2011. The effects of deforestation and climate variability on the streamflow of the Araguaia River, Brazil. Biogeochemistry 105, 119–131. https://doi.org/10.1007/s10533-011-9582-2

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., Lambin, E., 2004. Digital change detection methods in ecosystem monitoring: a review. Int. J. Remote Sens. 25, 1565–1596. https://doi.org/10.1080/0143116031000101675

Crutzen, P.J., 2002. Geology of mankind. Nature 415, 23. https://doi.org/10.1038/415023a

Di Gregorio, A., O'Brien, D., 2012. Land-Cover, Overview of Their, Classifications and Interoperability, in: Giri, C.P. (Ed.), Remote Sensing of Land Use and Land Cover: Principles and Applications,. CRC Press, Boca Raton, pp. 37–46.

Don, A., Schumacher, J., Freibauer, A., 2011. Impact of tropical land-use change on soil organic carbon stocks - a meta-analysis. Glob. Chang. Biol. 17, 1658–1670. https://doi.org/10.1111/j.1365-2486.2010.02336.x

FAO, 2016. Land Resources [WWW Document]. L. use. URL http://www.fao.org/nr/land/use/en/ (accessed 5.13.16).

FAO, 2011. The state of the world's land and water resources for food and agriculture (SOLAW) - Managing systems at risk. Food and Agriculture Organization of the United Nations, Rome, and Earthscan, London.

FAOSTAT, 2016. Country Indicators [WWW Document]. URL http://www.fao.org/faostat/en/#home

Ferreira, N.C., Ferreira, L.G., Miziara, F., 2007. Deforestation hotspots in the Brazilian Amazon: Evidence and causes as assessed from remote sensing and census data. Earth Interact. 11, 1–16. https://doi.org/10.1175/EI201.1

Foley, J.A., Defries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J. a, Prentice, I.C., Ramankutty, N., Snyder, P.K., 2005. Global consequences of land use. Science 309, 570–574. https://doi.org/10.1126/science.1111772

Friis, C., Nielsen, J.Ø., Otero, I., Haberl, H., Niewöhner, J., Hostert, P., 2016. From teleconnection to telecoupling: taking stock of an emerging framework in land system science. J. Land Use Sci. 11, 131–153. https://doi.org/10.1080/1747423X.2015.1096423

Garcia, A.S., Ballester, M.V.R., 2016. Land cover and land use changes in a Brazilian Cerrado landscape: drivers, processes, and patterns. J. Land Use Sci. 11, 538–559. https://doi.org/10.1080/1747423X.2016.1182221

Garcia, A.S., Sawakuchi, H.O., Ferreira, M.E., Ballester, M.V.R., 2017. Landscape changes in a neotropical forest-savanna ecotone zone in central Brazil: The role of protected areas in the maintenance of native vegetation. J. Environ. Manage. 187, 16–23. https://doi.org/10.1016/j.jenvman.2016.11.010

Garrett, R.D., Lambin, E.F., Naylor, R.L., 2013. The new economic geography of land use change: Supply chain configurations and land use in the Brazilian Amazon. Land use policy 34, 265–275. https://doi.org/10.1016/j.landusepol.2013.03.011

Gasparri, N.I., de Waroux, Y. le P., 2015. The Coupling of South American Soybean and Cattle Production Frontiers: New Challenges for Conservation Policy and Land Change Science. Conserv. Lett. 8, 290–298. https://doi.org/10.1111/conl.12121

Geist, H., McConnell, W., Lambin, E.F., Moran, E., Alves, D., Rudel, T., 2006. Causes and Trajectories of Land-Use/Cover Change, in: Lambin, E., Geist, H. (Eds.), Land-Use and Land-Cover Change. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 41–70. https://doi.org/10.1007/3-540-32202-7_3

Gibbs, H.K., Rausch, L., Munger, J., Schelly, I., Morton, D.C., Noojipady, P., Soares-Filho, B., Barreto, P., Micol, L., Walker, N.F., 2015. Brazil's Soy Moratorium. Science (80-.). 347, 377–378. https://doi.org/10.1126/science.aaa0181

Hansen, M.C., Loveland, T.R., 2012. A review of large area monitoring of land cover change using Landsat data. Remote Sens. Environ. 122, 66–74. https://doi.org/10.1016/j.rse.2011.08.024

Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D., 2013. Change detection from remotely sensed images: From pixel-based to object-based approaches. ISPRS J. Photogramm. Remote Sens. 80, 91–106. https://doi.org/10.1016/j.isprsjprs.2013.03.006

IBGE, 2019. Sistema IBGE de Recuperação Automática (SIDRA): banco de dados sobre indicadores, população, economia e geociências.

Ivanauskas, N.M., Monteiro, R., Rodrigues, R.R., 2008. Classificação fitogeográfica das florestas do Alto Rio Xingu. Acta Amaz. 38, 387–402. https://doi.org/10.1590/S0044-59672008000300003

Johnson, K.B., Jacob, A., Brown, M.E., 2013. Forest cover associated with improved child health and nutrition: evidence from the Malawi Demographic and Health Survey and satellite data. Glob. Heal. Sci. Pract. 1, 237–248. https://doi.org/10.9745/GHSP-D-13-00055

Klink, C. a., Machado, R.B., 2005. Conservation of the Brazilian Cerrado. Conserv. Biol. 19, 707–713. https://doi.org/10.1111/j.1523-1739.2005.00702.x

Kust, G., Andreeva, O., Cowie, A., 2017. Land Degradation Neutrality: Concept development, practical applications and assessment. J. Environ. Manage. 195, 16–24. https://doi.org/10.1016/j.jenvman.2016.10.043

Lambin, E.F., Meyfroidt, P., 2011. Global land use change, economic globalization, and the looming land scarcity. Proc. Natl. Acad. Sci. 108, 3465–3472. https://doi.org/10.1073/pnas.1100480108

Lambin, E.F., Turner, B.L., Geist, H.J., Agbola, S.B., Angelsen, A., Bruce, J.W., Coomes, O.T., Dirzo, R., Fischer, G., Folke, C., George, P.S., Homewood, K., Imbernon, J., Leemans, R., Li, X., Moran, E.F., Mortimore, M., Ramakrishnan, P.S., Richards, J.F., Skånes, H., Steffen, W., Stone, G.D., Svedin, U., Veldkamp, T. a., Vogel, C., Xu, J., 2001. The causes of land-use and land-cover change: Moving beyond the myths. Glob. Environ. Chang. 11, 261–269. https://doi.org/10.1016/S0959-3780(01)00007-3

Liu, J., Hull, V., Batistella, M., DeFries, R., Dietz, T., Fu, F., Hertel, T.W., Izaurralde, R.C., Lambin, E.F., Li, S., Martinelli, L.A., McConnell, W.J., Moran, E.F., Naylor, R., Ouyang, Z., Polenske, K.R., Reenberg, A., de Miranda Rocha, G., Simmons, C.S., Verburg, P.H., Vitousek, P.M., Zhang, F., Zhu, C., 2013. Framing Sustainability in a Telecoupled World. Ecol. Soc. 18, art26. https://doi.org/10.5751/ES-05873-180226

Lu, D., Mausel, P., Brondízio, E., Moran, E., 2004. Change detection techniques. Int. J. Remote Sens. 25, 2365–2401. https://doi.org/10.1080/0143116031000139863

Meyfroidt, P., 2016. Approaches and terminology for causal analysis in land systems science. J. Land Use Sci. 11, 501–522. https://doi.org/10.1080/1747423X.2015.1117530

Meyfroidt, P., Roy Chowdhury, R., de Bremond, A., Ellis, E.C., Erb, K.H., Filatova, T., Garrett, R.D., Grove, J.M., Heinimann, A., Kuemmerle, T., Kull, C.A., Lambin, E.F., Landon, Y., le Polain de Waroux, Y., Messerli, P., Müller, D., Nielsen, J., Peterson, G.D., Rodriguez García, V., Schlüter, M., Turner, B.L., Verburg, P.H., 2018. Middle-range theories of land system change. Glob. Environ. Chang. 53, 52–67. https://doi.org/10.1016/j.gloenvcha.2018.08.006

Mueller, N.D., West, P.C., Gerber, J.S., MacDonald, G.K., Polasky, S., Foley, J. a, 2014. A tradeoff frontier for global nitrogen use and cereal production. Environ. Res. Lett. 9, 054002. https://doi.org/10.1088/1748-9326/9/5/054002

Müller, D., Munroe, D.K., 2014. Current and future challenges in land-use science. J. Land Use Sci. 9, 133–142. https://doi.org/10.1080/1747423X.2014.883731

Nepstad, D.C., Stickler, C.M., Almeida, O.T., 2006. Globalization of the Amazon soy and beef industries: Opportunities for conservation. Conserv. Biol. https://doi.org/10.1111/j.1523-1739.2006.00510.x

Paranhos Filho, A.C., Lastoria, G., Torres, T.G., 2008. Sensoriamento Remoto Ambiental Aplicado: Introdução às Geotcnologias. UFMS Ed, Campo Grande.

Polasky, S., Nelson, E., Pennington, D., Johnson, K.A., 2011. The impact of land-use change on ecosystem services, biodiversity and returns to landowners: A case study in the state of Minnesota. Environ. Resour. Econ. 48, 219–242. https://doi.org/10.1007/s10640-010-9407-0

Richards, P., 2015. What Drives Indirect Land Use Change? How Brazil's Agriculture Sector Influences Frontier Deforestation. Ann. Assoc. Am. Geogr. 105, 1026–1040. https://doi.org/10.1080/00045608.2015.1060924

Rodrigues, A.S.L., Ewers, R.M., Parry, L., Souza, C., Veríssimo, A., Balmford, A., Jr, C.S., Veríssimo, A., Balmford, A., 2009. Boom-and-bust development patterns across the Amazon deforestation frontier. Science 324, 1435–7. https://doi.org/10.1126/science.1174002

Rounsevell, M.D.A., Pedroli, B., Erb, K.-H., Gramberger, M., Busck, A.G., Haberl, H., Kristensen, S., Kuemmerle, T., Lavorel, S., Lindner, M., Lotze-Campen, H., Metzger, M.J., Murray-Rust, D., Popp, A., Pérez-Soba, M., Reenberg, A., Vadineanu, A., Verburg, P.H., Wolfslehner, B., 2012. Challenges for land system science. Land use policy 29, 899–910. https://doi.org/10.1016/j.landusepol.2012.01.007

Rudel, T.K., 2007. Changing agents of deforestation: From state-initiated to enterprise driven processes, 1970–2000. Land use policy 24, 35–41. https://doi.org/10.1016/j.landusepol.2005.11.004

Tewkesbury, A.P., Comber, A.J., Tate, N.J., Lamb, A., Fisher, P.F., 2015. A critical synthesis of remotely sensed optical image change detection techniques. Remote Sens. Environ. 160, 1–14. https://doi.org/10.1016/j.rse.2015.01.006

Turner, B.L., Lambin, E.F., Reenberg, A., 2007. The emergence of land change science for global environmental change and sustainability. Proc. Natl. Acad. Sci. 104, 20666–20671. https://doi.org/10.1073/pnas.0704119104

van Schrojenstein Lantman, J., Verburg, P.H., Bregt, A., Geertman, S., 2011. Core Principles and Concepts in Land-Use Modelling: A Literature Review, in: Koomen, E., Borsboom-van Beurden, J. (Eds.), Land-Use Modelling in Planning Practice. Springer, Dordrecht, Heidelberg, London, New York, pp. 35–57. https://doi.org/10.1007/978-94-007-1822-7_3

van Vliet, J., Magliocca, N.R., Büchner, B., Cook, E., Rey Benayas, J.M., Ellis, E.C., Heinimann, A., Keys, E., Lee, T.M., Liu, J., Mertz, O., Meyfroidt, P., Moritz, M., Poeplau, C., Robinson, B.E., Seppelt, R., Seto, K.C., Verburg, P.H., 2016. Meta-studies in land use science: Current coverage and prospects. Ambio 45, 15–28. https://doi.org/10.1007/s13280-015-0699-8

Velasquez, C.; Alves, H. Q.; Bernasconi, P., 2010. FIQUE POR DENTRO: a Bacia do Rio Xingu em Mato Grosso. nstituto Socioambiental, Instituto Centro de Vida, São Paulo.

Verburg, P.H., Crossman, N., Ellis, E.C., Heinimann, A., Hostert, P., Mertz, O., Nagendra, H., Sikor, T., Erb, K.-H., Golubiewski, N., Grau, R., Grove, M., Konaté, S., Meyfroidt, P., Parker, D.C., Chowdhury, R.R., Shibata, H., Thomson, A., Zhen, L., 2015. Land system science and sustainable development of the earth system: A global land project perspective. Anthropocene 12, 29–41. https://doi.org/10.1016/j.ancene.2015.09.004

Verburg, P.H., Kok, K., Pontius Jr., R.G., Veldkamp, A., 2006. Modeling Land-Use and Land-Cover Change, in: Lambin, E.F., Geist, H. (Eds.), Land-Use and Land-Cover Change, Global Change - The IGBP Series. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 117–135. https://doi.org/10.1007/3-540-32202-7

Verburg, P.H., Schot, P.P., Dijst, M.J., Veldkamp, A., 2004. Land use change modelling: Current practice and research priorities. GeoJournal. https://doi.org/10.1007/s10708-004-4946-y

2. ASSESSING LAND USE/COVER DYNAMICS AND EXPLORING DRIVERS IN THE AMAZON'S ARC OF DEFORESTATION THROUGH A HIERARCHICAL, MULTI-SCALE AND MULTI-TEMPORAL CLASSIFICATION APPROACH

ABSTRACT

Land use and land cover (LULC) are intrinsically tied to ecological and social dynamics. Still, classifying LULC in ecotones, where landscapes are commonly heterogeneous and have a wide range of physiognomies, remains a challenge. Here we present a three-level hierarchical classification approach, using both Landsat and MODIS images, and both pixels and objects as units of information. We applied this multi-temporal and -spatial approach to classify land use in the Upper Xingu River Basin (~170,000 km²), located in the arc of deforestation of the Brazilian Amazon. The first level includes five classes and differentiates managed land from native vegetation with high overall accuracy (93%). The second level has 11 classes (overall accuracy = 86%) and separates main land uses and native vegetation domains. The third level has 16 classes (overall accuracy = 83%) and addresses productivity of both managed and natural systems. We find that this new method presented here is more efficient than existing regional and global land cover products. Applying this approach to assess land cover transitions in the basin from 1985-2015, we find that agricultural production increased, yet manifested itself differently in the northern (Amazon biome) and southern (Cerrado biome) portions of the basin. Analyzing land use change in different levels, we identify that agricultural intensification occurred mainly in the Amazon while the Cerrado has undergone an expansion in agricultural area. The method presented here can be adapted to other regions, improving efficiency and accuracy of classifying land cover in heterogeneous landscapes.

Keywords: Land use change; Amazon; Cerrado; Ecotone; Physiognomies

2.1. Introduction

Land cover refers to the observed biophysical component on the Earth's surface, while land use is defined by the activities undertaken in a certain area (Di Gregorio, 2016; Turner et al., 2007). Land use and land cover (LULC) changes are key components of biological, physical, and socioeconomic processes taking place on the Earth surface (de Chazal and Rounsevell, 2009; Don et al., 2011; Foley et al., 2005; Lathuillière et al., 2017; Macedo et al., 2013). Therefore, LULC is a fundamental input for several models in many research areas and can help inform decision making processes (Bondeau et al., 2007; Döll et al., 2003; FAO, 2011, 1993; Ge et al., 2007). Remotely sensed products are the main data source for LULC mapping, and much research has been devoted to evaluating methodologies with a specific emphasis on classification techniques (Coppin et al., 2004; Gómez et al., 2016; Hansen and Loveland, 2012; Hussain et al., 2013; Tewkesbury et al., 2015; Wang et al., 2017). Despite the large efforts and achievements in improving LULC classification, accurately transforming remotely sensed data into thematic maps remains a challenge when modelling complex landscapes (Gómez et al., 2016), limiting our ability to address patterns and processes of LULC change, especially in tropical regions (Müller et al., 2015; Pendrill and Persson, 2017; Toniol et al., 2017).

Ecotones, agriculture frontiers and biomes with multiple physiognomies present large disagreement in global and continental datasets (Giri et al., 2005; Herold et al., 2008; Pendrill and Persson, 2017). For example, Herold and colleagues (2008) identified that at least three out of four global data sets disagree on LULC classes

attributed to large portions of the Amazon's Arc of Deforestation and neotropical savanna biomes. Factors that contribute to this disagreement are associated with the large biophysical variability observed in diverse landscapes. Inadequately assessing the spatial and temporal variability of these areas precludes the possibility of adequately capturing their LULC classes. Additionally, accuracy of LULC maps generally decreases with increasing class complexity (Lu and Weng, 2007), which is very often important for decision-making. Still, improving LULC monitoring is critical to evaluating impacts on the extent, condition, and productivity of both managed and natural systems. Such dynamics are directly linked to several social or ecosystem services, including food, water and energy security, and biodiversity conservation (Endo et al., 2015; Lal, 2016; Ozturk, 2015). This information is essential when planning land use in regions such as the neotropical savannas (Cerrado) or the tropical rainforest at the Amazon's arc of deforestation. These regions are not only hotspots for biodiversity conservation, but are also undergoing rapid LULC change due to agriculture expansion (Klink and Machado, 2005; Power, 2010). Aside from the importance of these regions in terms of natural resources and agricultural production, spatially explicit LULC information that considers landscape diversity is rarely available.

Classification techniques refer to the process of sorting pixels into a finite number of individual classes, based on their surface reflectance values at a certain time (Lu et al., 2004), or the reflectance behavior of pixels through time (e.g.: MODIS Land Cover products (Friedl et al., 2010)). Use of both high spatial and temporal resolution data is an approach that increases both the amount of data and the modelling complexity, while enabling the study of different native vegetation physiognomies or agriculture productivity (Brown et al., 2007; Ferreira et al., 2003). Due to increases in modelling complexity, current products address either spatial or temporal resolution when classifying LULC. The computational challenge of modelling large satellite imagery data sets can be made more efficient using an object-based approach (Blaschke, 2010). But this approach is not appropriate for landscape formations which are not expected to present a spectral pattern of an object (a group of pixels presenting a similar response). Still, for small scale study areas, approaches that integrate pixel-based and object-based methods out-perform approaches that employ only one of these methods on their own (Aguirre-Gutiérrez et al., 2012; Chen et al., 2018). Recently, the combination of machine (or deep) learning, cloud processing, and big data is shifting classification processing by enabling users to address more complex models of classification (Azzari and Lobell, 2017); mainly when the physical processes resulting in the remotely sensed imagery are not understood, or are not the main focus of the study (Zhu et al., 2017). Such data-driven classifications rely on many redundant explanatory variables (Lebourgeois et al., 2017), and results often are not easily translated into a conceptual model. In this regard, Zhu and colleagues (2017) argue that uniting process-based modeling and machine learning is a promising direction.

Here we present a hierarchical classification approach which applies a multi-temporal/multi-scale and combined object and pixel-based approach in order to contribute to diverse landscape LULC mapping. We demonstrate that through this approach we improve the assessment of LULC change patterns and their causation chains. Moreover, this study is intended to improve classification processes by presenting a workflow which (i) generates a classification allowing the analysis of LULC change in modern agricultural frontiers and ecotone zones, (ii) facilitates the comparison of results with those generated by other studies and/or for other regions, (iii) uses freely available remote sensing data, and (iv) employs a straightforward approach that can be implemented in software commonly used and/or available at no cost to researchers, NGOs, and governmental agencies (R, GRASS, or Orpheum toolbox). We tested this approach by creating a series of multi-temporal LULC maps for the Upper Xingu River Basin, located at the border of the Amazon rainforest and neo-tropical savannas (Cerrado), and describing the changes over three decades. We then explore patterns, trajectories, and causes of LULC change for

this study area. This chapter was published in the journal Remote Sensing Applications: Society and Environment (Garcia et al., 2019).

2.2. Study area and data

2.2.1. Study area

The Upper Xingu River Basin (hereafter denoted as "UXRB"), is located in Mato Grosso state in Brazil, and in the ecotone between the Amazon rainforest and the neotropical savannas, or Cerrado biome. The Xingu River is one of the main Amazon River tributaries. Its headwater region encompasses one-third of the Xingu River Basin, draining ~170,000 km2. The area spreads from about 9.5ºS to 15ºS and from 51ºW to 55.5ºW (Figure 2). The UXRB exhibits a wide array of natural physiognomies, ranging from Amazon rainforest to savanna grasslands (Ivanauskas et al., 2008; Velasquez, C.; Alves, H. Q.; Bernasconi, 2010). The Amazon rainforest and Cerrado are the first and second largest biomes in South America. The Amazon rainforest contains enormous biodiversity and one-third of the world's tropical tree species used for timber (Brazil, 2018). The Cerrado is a global biodiversity hotspot and contains the highest land use change rates in Brazil (Brazil, 2018; Klink and Machado, 2005).

Besides the importance of the region for its natural resources (including biodiversity), the UXRB also encompasses large sociocultural diversity and economic prominence. It is estimated that 16 indigenous ethnic groups and other traditional communities live in the basin (Velasquez, C.; Alves, H. Q.; Bernasconi, 2010). The main economic activities in the study area during the late 1970's to 1990's were timber and beef. Forests were degraded or cleared due to livestock expansion and logging activities. In the early 2000's, soybean production emerged as another deforestation driver (Brando et al., 2013; Nepstad et al., 2006). The defined rainy season, flat terrain, and dominance of good physical structured Oxisol soils (Velasquez, C.; Alves, H. Q.; Bernasconi, 2010), make the UXRB ideal for agricultural expansion and intensification. Currently, approximately 2% of world's and 9% of Brazil's soybeans are produced in the UXRB, as well as 0.2% of the world's and 13% of Brazil's cattle (FAOSTAT, 2016; IBGE, 2016).

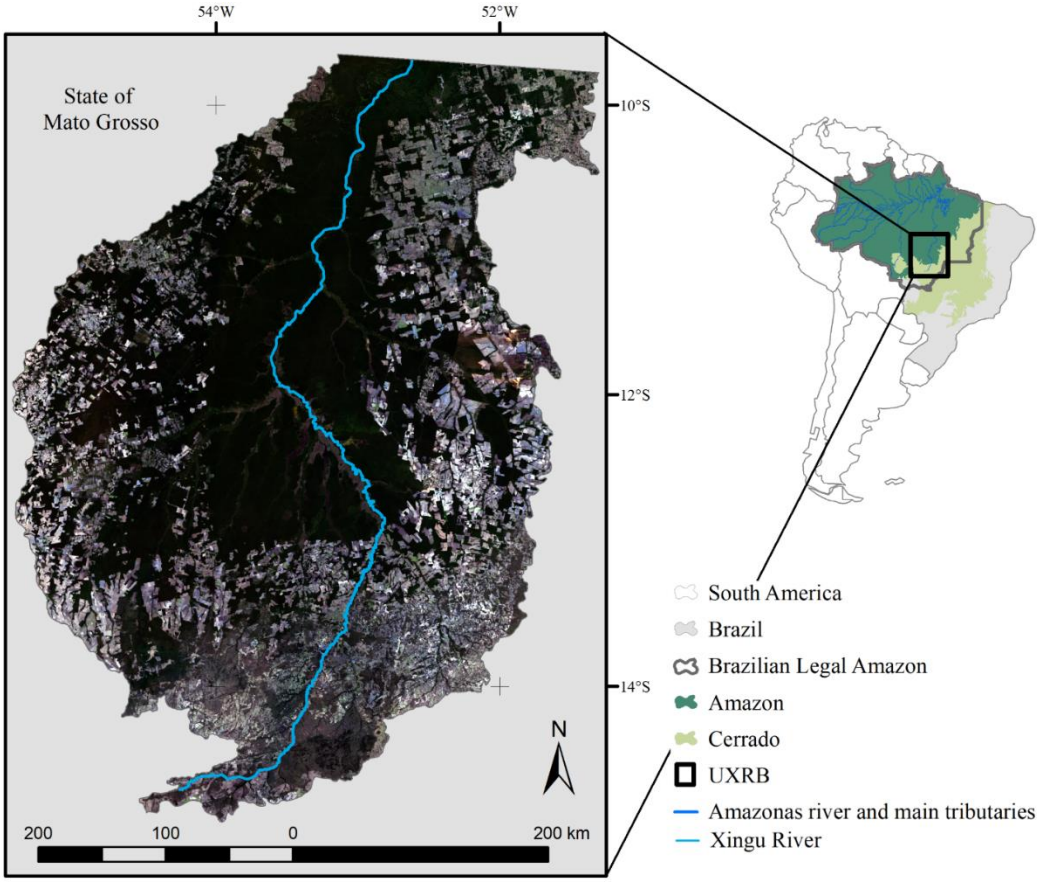


Figure 2. Upper Xingu River Basin (UXRB) is located in the Brazilian Legal Amazon, state of Mato Grosso. It drains an area of ~170,000 km² into the Xingu River, one of the main tributaries of the Amazonas River. The area is in the ecotone between the Amazon and Cerrado biome (also known as neotropical savannas). The image on the left is a mosaic of Landsat 8 images acquired in July and August of 2015. The mosaic is shown in a true-color composition (RGB 432).

2.2.2. Data acquisition and pre-processing

LULC maps for 1985, 1990, 1995, 2000, 2005, 2010, and 2015 of the UXRB were derived from a combination of Landsat (L5, L7 or L8, level 2) and MODIS (Aqua and Terra, collection 6) products. These images were processed based on information gathered through the Brazilian vegetation map (IBGE, 2004), ground reference points, and fine resolution images (Rapideye).

2.2.2.1. Landsat data

Landsat surface reflectance images were obtained from the USGS/Landsat Higher Level Science Data products (http://espa.cr.usgs.gov/). The path/row and acquisition date of all 120 Landsat scenes used in this study are presented in Appendix A - Supplementary Material 1. Landsat Higher Level Science Data products are delivered pre-corrected for geometrically and radiometrically by a homogenous correction chain (Masek et al., 2006; http://espa.cr.usgs.gov/). Although previously corrected, we registered mosaicked and resampled (30 m pixel resolution) images by using autosync workstation (ERDAS, 2014a) in order to guarantee a pixel by pixel overlap between years. As the UXRB presents a flat relief, we implemented a 2D polynomial model of third order based on the green band of each image. We considered 15 m, half pixel resolution, as the maximum acceptable error value (Root Mean Square Error < 0.5 pixel) and resampled the images through cubic convolutions. This process was

particularly important for earlier years, as shown in Appendix A - Supplementary Material 1. A histogram matching procedure was applied when necessary during mosaicking processes based on central path (225 in the WRS-2 system) overlapping areas. Since all raster in the time series share a similar histogram shape, this technique changes each of the pixel's values according to the target histogram but does not change the histogram shape.

No radiometric correction was applied with the exception of Landsat 7 images for 2005 and 2010 which were striped - images collected after 2003, when the Scan Line Corrector (SLC) failed. To correct this error, we applied a gap filling technique, which consists on acquiring a second image (also from Landsat 7 and for same year as the original one) to fill the gaps in the first scene (Scaramuzza et al., 2004). Rather than using these corrected images as primary sources of data to build LULC maps, we used them to visually verify Landsat 5 image consistency. Since Landsat 5 was launched in 1984, and old sensors can present degradation problems in more recent acquisitions. We did not use any image with signs of degradation.

2.2.2.2. MODIS and derived indices

MODIS Enhanced Vegetation Index (EVI), with a 250 m spatial resolution and 8 day temporal resolution were acquired on demand from the University of Natural Resources and Life Sciences, Vienna - BOKU (Vuolo et al., 2012). This database is based on MODIS Level-3, 16-day composite EVI from both Terra and Aqua satellites. The combination of 16-day composites from both satellites allowed us to derive a time series of 8-day temporal resolution. EVI time-series are delivered mosaicked, smoothed, and gap-filled. Such processing steps are based on the "MODIS package" (Mattiuzzi et al., 2016) developed in R (R Development Core Team, 2011), and the Whittaker filter (Vuolo et al., 2012). We tested EVI images against our Landsat mosaic using the previously described process, but no geometric correction was needed. An annual EVI series beginning in August of each analyzed year and ending in August of the following one was built for 2000, 2005, 2010 and 2015. In total, we used 832 MODIS scenes, or ~210 for each year.

2.2.3. Ground reference

We carried out field work to collect ~2000 ground reference points in two independent campaigns. The first campaign, carried out in June 2015, was used to improve systematic comprehension of the study area's LULC composition and to create thresholds for LULC class separability criteria. This dataset is composed of ~400 ground reference points collected on the ground. In addition, ancillary reference data points (~100 points) were collected by comparing Landsat 8 images obtained in 2015 with high spatial resolution RapidEye images acquired in 2012/13 (http://geocatalogo.mma.gov.br/). Hereafter this dataset is called "training data" and its overview is presented in Appendix A - Supplementary Material 2. The field work to develop the training data was mostly concentrated at the contact zone between tropical rainforest and savannas, as well as at areas with more intense land use change.

During a second campaign (October 2016), we acquired 1500 ground reference points throughout the entire basin for accuracy analyses. The points were as distributed in the basin as practical, but concentration around the main roads was inevitable due to access constrains and study area size. In addition, an ancillary set of data points (~500 points) collected by an independent researcher though a comparison of Landsat 8 images of the year 2016 with RapidEye images was also used. This complementary approach was necessary in order to sample large areas which are difficult to assess. This dataset, hereafter called "validation data", was used to verify the quality of the final classification.

2.3. Methods

To derive LULC maps in highly dynamic tropical ecotone zones, we developed an assessment approach through a hierarchical classification model which allows us (i) to assess relevant information for regional landscape evaluation, (ii) to have different levels of information and accuracy for a diverse range of applications, (iii) to harmonize our thematic mapping with LULC schemes produced for other tropical regions or globally, and (iv) to apply the same classification methodology in other study areas. The hierarchical classification model was built in ERDAS Imagine Expert Classifier workstation (ERDAS, 2014b). We applied a hierarchical rule-based approach (a decision tree based on expert knowledge) to integrate variables of different sources and formats. Each branch describes the conditions under which constituent information (variables) gets abstracted into a set of higher-level informational classes. The rules were set based on literature and expert information, and field work observation. This approach focuses on expected reflectance signals based on knowledge of LULC behavior though space and time. It differs from currently popular machine learning algorithms, which fish a signal from data through multiple variable interactions (Maxwell et al., 2018). The development of classifications we applied here advances variable selection based on patterns and process behavior concepts, and can be used in any kind of classification, including machine learning. It is also important to note that learning algorithms demand a large and well distributed field sample to train the algorithms, which still is not always the case for extensive tropical areas. Still, a preliminary classification as we apply here could contribute with sample allocation to develop even more powerful classification processes.

Legend can be defined as the application of a classification scheme in a specific study area. Here we used three levels of information to address LULC legend, each level composed for both land cover classes and land use classes. We considered representative LULC classes for the ecotone zone between Amazon rainforest and Cerrado both in area and in meaning of LULC change. The border between both biomes extends for about 6,300 km and overlaps with the Amazon`s arc of deforestation. The developed LULC legend follows hierarchical concepts presented by FAO's Land Cover Classification System (Di Gregorio and Jansen, 2000). First, we adopted a general level of information on LULC, followed by a differentiation of the main land use and native vegetation domains. The third level of classification is linked to productivity and addresses land use intensity as well different native physiognomies. It is important to point out that we do not differentiate all 11 savanna vegetation physiognomies, which range from forest to grassland fields with no trees (see Oliveira and Marquis (2002) for further description). Instead, we aggregated them into three different classes according to tree density. Previous global (GLC2000 (Bartholomé and Belward, 2005) and MODIS Global land cover (Friedl et al., 2010)), and regional (TerraClass Cerrado (Scaramuzza et al., 2017) and TerraClass Amazonia (Almeida et al., 2016)) efforts produced legends compatible with our intermediate classification level, but no more detailed information on productivity systems or physiognomies are provided.

The spatial-temporal resolution of the final classification products is also of high interest when assessing LULC change. Here we took advantage of Landsat spatial resolution (30 x 30 m) and EVI MODIS temporal resolution (7 days) by combining both products and addressing LULC change in terms of cover and management. EVI MODIS time series presents a 250 m spatial resolution, moreover to overcome this problem, we segmented Landsat images and used the derived objects (larger than 250m) to extract data from EVI MODIS layers by calculating the modal (discrete variables) or mean value (continuous variables) of pixels contained in each object. This approach is referred to as "zonal" in following sections. Level 1 and level 2 products are made available for

each 5-year period from 1985 to 2015. Level 3 was only developed for 2000 - 2015 due to MODIS availability. Nevertheless, the employment of both Landsat and MODIS sensors allows us to rely on the spatial resolution of the former and the temporal resolution of the latter to address mainly the Level 3 of our classification.

2.3.1. Decision tree and classification rules

Our hierarchical classification model consisted of three levels as presented in Table 1. Level 1 (L1) classifies the region of interest into five major LULC types: (1) natural and semi-natural vegetation, (2) cultivated and managed terrestrial areas, (3) burned areas, (4) surface water, and (5) unclassified (see Figure 3 for decision tree overview). This legend is fairly comprehensive and can be harmonized with a broad range of other researches, reducing conversion uncertainties among different products. L1 assesses native vegetation conversion and, thus, supplies information for policies and laws, such as the Brazilian Forest Code (Soares-Filho et al., 2014). To develop this level, first, we applied an unsupervised classification method using ISODATA algorithm to classify Landsat mosaicked images. In this process, each pixel is assigned to a class based on its spectral profile. There is no need for prior knowledge of the number or identification of the different classes present in the imaged area. We set the ISODATA algorithm to cluster pixels in up to 40 classes based on up to 20 iterations, with a minimum size of 0.01% of the study area for each class, a maximum standard deviation of 5, a maximum of 2 merges, and a convergence threshold of 0.95. Finally, we visually aggregated each of the output classes into one of the five legend classes mentioned before.

Table 1. Description of the land cover and land use classes used in this study. Level refers to the hierarchical level in the classification scheme.

Level	Class	Description
1	Natural and semi-natural vegetation	Natural or semi-natural vegetation cover, including forest, savannas, and floodplains
	Cultivated and managed terrestrial areas	Any area under management which modifies the physiognomy of a natural system
1	Burned areas	Either natural or managed areas that have recently being burned
2	Surface water	Area covered by water
3	Unclassified	Any pixel that did not fit in any of the other categories
2	Forest	Forest formations, including rain forest, semi-deciduous, deciduous, riverine, cerradão (cerrado woodland with closed canopy)
	Savanna formation	Wood-grassland ecosystem with open canopy
	Croplands	Cultivated areas
	Pasturelands	Areas with planted grassland for cattle ranching
2	Wetlands	Flooded area, either permanently or seasonally with high proportion of vegetation.
3	Secondary complex	Vegetation in regeneration or disturbed through natural processes or removal
	Bare soils	Exposed bare soil lacking any vegetation
	Urban	Concentrated built-up structures
3	Forest	Forest formations in the Amazon domain such as rain and deciduous forests
	Forest cerrado	Forest formations in the Cerrado domain such riparian forest and woodland
	Woody Cerrado	Woody formation with open canopy and at least 40% of tree cover
	Shrub-grassland cerrado	Grassland formation with no trees to formations with up to 40% of tree cover
	Single crop	Cultivated areas with one harvest per year, rain fed
	Double crop	Cultivated areas with two harvests per year, rain fed
	Irrigated crop	Cultivated areas which are artificially irrigated
	Pasturelands	Areas with regular planted grassland for cattle ranching
	Degraded pasturelands	Areas with planted grassland for cattle ranching with presence of bare ground and poor greenness recovery

Level 2 (L2) encompasses a more specific set of LULC classes, but is broad enough to be adapted or compared to similar regions throughout the tropics. This legend level assesses biophysical characteristics related to LULC of the target landscape. For example, L2 is useful as an input for biogeochemical and hydrological models as it differentiates vegetation types. LULC information and user`s knowledge acquired in field work was necessary to map LULC at the second level of detail (see Figure 3 for decision tree overview). LULC legend for the L2 included: (1) forest (2) savanna formation (3) wetlands (4) secondary complex, (5) croplands, (6) pasturelands, (7) bare soils, (8) urban, (9) burned areas, (10) surface water, and (11) unclassified. L1 unsupervised classification output was disaggregated into forest (closed canopy), savanna formation (open canopy), wetlands, and secondary complex. We disaggregated these classes by overlapping the Brazilian Institute of Geography and Statistics vegetation map as a masking conditional variable, with the classes resulting from the unsupervised classification (described above) being classified as Natural/Semi-natural (L1). Areas where both maps did not match were considered secondary complexes. Urban, burned areas and surface water were imported from L1 without any modification. Forest, savanna, urban, burned areas and surface water formation were inputted in L2 classification based on pixel since we do not expect them to necessarily present object patterns.

We segmented Landsat images to reclassify managed areas. Segmentation was conducted in ArcGIS 10.3 using bands 5, 6, and 4 of mosaicked Landsat 8 images (bands 4, 5, and 3 for years using Landsat 5 or Landsat 7) based on the spectral similarity and spatial proximity among neighbor pixels, with a minimal area of 9 ha. This approach assumes a minimal aggregated number of pixels or a specific shape in crops and pasturelands. We arbitrary chose a minimum area of 9 ha with the goal of adding variance to the analyses but also allowing the algorithm to capture the smallest area as possible. We chose 9 ha so we could evaluate more than one Modis' pixel (pixel resolution equal to

250m, resampled into 30m, and resulting in ~70 pixels), and 100 Landsat pixels (pixel resolution equal to 30m). We classified each object according to L1 classes by defining the major class inside each object and identified Managed Areas which were then reclassified in crops, pasturelands or bare soils. To differentiate these three classes based on MODIS EVI, we calculated three indexes based in relation to the agriculture calendar for Mato Grosso state (Appendix A - Supplementary Material 3). We used the objects created with Landsat images to calculate the mean value for each object (referred to in Figure 3 as zonal approach) and each index. Indexes based on MODIS EVI are as follow:

(Index i) Minimum EVI during soil preparation/sowing season (EVImin):

$$EVI_{min} = \text{MIN}\ (\ EVI_{so(doy\ x)}, EVI_{so(doy\ x+7)}, \dots, EVI_{so(day\ y)}\ )$$

where, EVI_{so} series represents a sequence with 7-day interval from the first day (doy x) of the soil preparation/sowing season to the last one (*doy y*), *doy* corresponds to day of the year in the julian calendar;

(Index ii) Maximum EVI in the growing season (EVI_{max}):

$$EVI_{max} = \text{MAX}\ (\ EVI_{gro(doy\ x)}, EVI_{gro(doy\ x+7)}, \dots, EVI_{gro(doy\ y)}\ )$$

where, EVI_{gro} series represents a sequence with 7-day interval from the first day (*doy x*) of the growing season to the last one (*doy y*);

(Index iii) Differential greenness enhancement index (Rizzi et al., 2009):

$$\text{DGEI} = \frac{EVI_{max} - EVI_{min}}{EVI_{max} + EVI_{min}}.$$

We assumed that crop areas presented both states: bare soil or haystacks due to soil preparation and sowing period (very low EVI values), and also at least one green period during a year due to the growing season (highest EVI values). Moreover, they presented a higher DGEI when compared to pasturelands. As DGEI is the contrast between the peak high and low values of EVI in a season (defined as a few months), it is not sensitive to normal seasonality and precipitation shifts. Still, it can potentially be affected by drastic droughts which negatively impact both crop and pasture growth. Additionally, cropland abandonments can potentially impact this index. Such dynamics prevent land covers from reaching their expected maximum or minimum EVI values. Optimal DGEI thresholds values were identified based on training sampled data. However, since EVI time-series from MODIS are available after 2000, we applied other methods to separate crops from pastures for 1985, 1990 and 1995. For these years, we identified cropland polygons by selecting objects which presented Landsat EVI values lower than 0.3 (bare soil or stover – based on visual selection of sample fields which remained unchanged between 1995 and 2000, and were classified as croplands in 2000) and minimum field size of 40 ha. Cropland area is certainly under estimated for those dates. However, as croplands account for a maximum of 1% of the UXRB basin before 1995 (IBGE, 2016), we trust this procedure did not compromise our results.

The main goal of Level 3 (L3) is to cope with ecological characteristics related to LULC such as regional agriculture greenness or ecological intensification (Arvor et al., 2012). It is also useful, for example, when analyzing socio-economic characteristics associated with crop and cattle ranching intensification or to analyze which natural physiognomies are more likely to be replaced by certain agricultural activity (see Figure 3 for decision tree overview). LULC legend for L3 included: (1) forest, (2) forest cerrado, (3) woody Cerrado, (4) shrub-grassland cerrado, (5) wetland, (6) secondary complex, (7) single crop, (8) double crop, (9) irrigated crop, (10) pasturelands, (11) degraded pasturelands, (12) bare soils, (13) urban, (14) burned areas, (15) surface water, and (16) unclassified. L3 maps include classes that rely on high resolution temporal images, available only from 2000 forwards due to MODIS availability. We identified low productivity pasturelands (or degraded pasturelands) by using a combination of two indices derived from MODIS EVI: annual EVI_{sum} and annual DGEI.

(Index iv) Sum of EVI in a certain year (EVI_{sum}):

$$EVI_{sum} = \sum (EVI_{gro(doy\ x)}, EVI_{gro(doy\ x+7)}, \dots, EVI_{gro(doy\ y)})$$

where EVI_{sum} series represents a sequence with 7-day interval from the first day (*doy x*) of the analyzed period through the last one (*doy y*).

The applied analysis adopts the hysteresis principle, which means that degraded pasturelands present a lower resilience and, thus, they are not able to recover their original state after a disturbance period (Searle et al., 2009; Yengoh et al., 2015). Grazing, fire, and water shortages are forms of disturbance, and the lack of recovery implies a loss of pasture productivity. This loss is liked to biological degradation (decrease in carbon accumulation) but also with some agricultural degradation processes such as loss of productivity in the form of insect attacks, for example (Dias-Filho, 2001). DGEI calculated as shown previously based on the whole analyzed year was used to estimate how recovering photosynthetic capacity varied within a year. The cumulative vegetation index (EVIsum) was used as a proxy for net primary productivity (Ferreira et al., 2013). We considered the combination of low greenness and low annual variation as indicative of pasturelands in degradation. Both DGEI and EVIsum were used to include the temporal EVI variation associated to rainfall (Yengoh et al., 2015). Although this approach does not addresses the complexities of the pasture degradation process, it is a fair indicator of pasture resilience through different seasons (Yengoh et al., 2015). Croplands were reclassified according to a conditional model of double cropping detection. Whenever croplands presented a large number of days with high vegetative response, we consider them as double cropping areas as shown in the following formula:

(v) Count of days with EVI > 0.5 in a certain year (EVI_{count}):

$$EVI_{count} = \sum (EVI_{veg,doy\ x}, \dots, EVI_{veg,doy\ y})$$

$$EVI_{veg,doy\ x} = 1 \rightarrow EVI_{doy} > 0.5$$

$$EVI_{veg,doy\ x} = 0 \rightarrow EVI_{doy} \leq 0.5$$

where EVI_{veg} represents a sequence with 7-day interval from the first day (doy x) of the analyzed period through the last one (doy y). Each EVI_{veg} is assigned the value 1 if the correspondent EVI value at certain day is higher than 0.5, otherwise it is zero. Irrigated areas were manually extracted based on spatial arrangement and higher EVI reflectance in the dry season. The optimal thresholds were calculated based on the training sample data.

We used the Brazilian vegetation map (IBGE, 2004) to first break forest areas into forest (L2-forest inside the rain forest domain) and forest cerrado (L2-forest inside the savanic domain). We also relied on NDVI values calculated from Landsat images in the dry season (NDVIL), and EVImin, and EVIsum values to reclassify forest and savanic formations into forest, forest cerrado, woody Cerrado, and shrub-grassland cerrado. The optimal thresholds were also calculated based on the training sample data - they are shown in Figure 3.

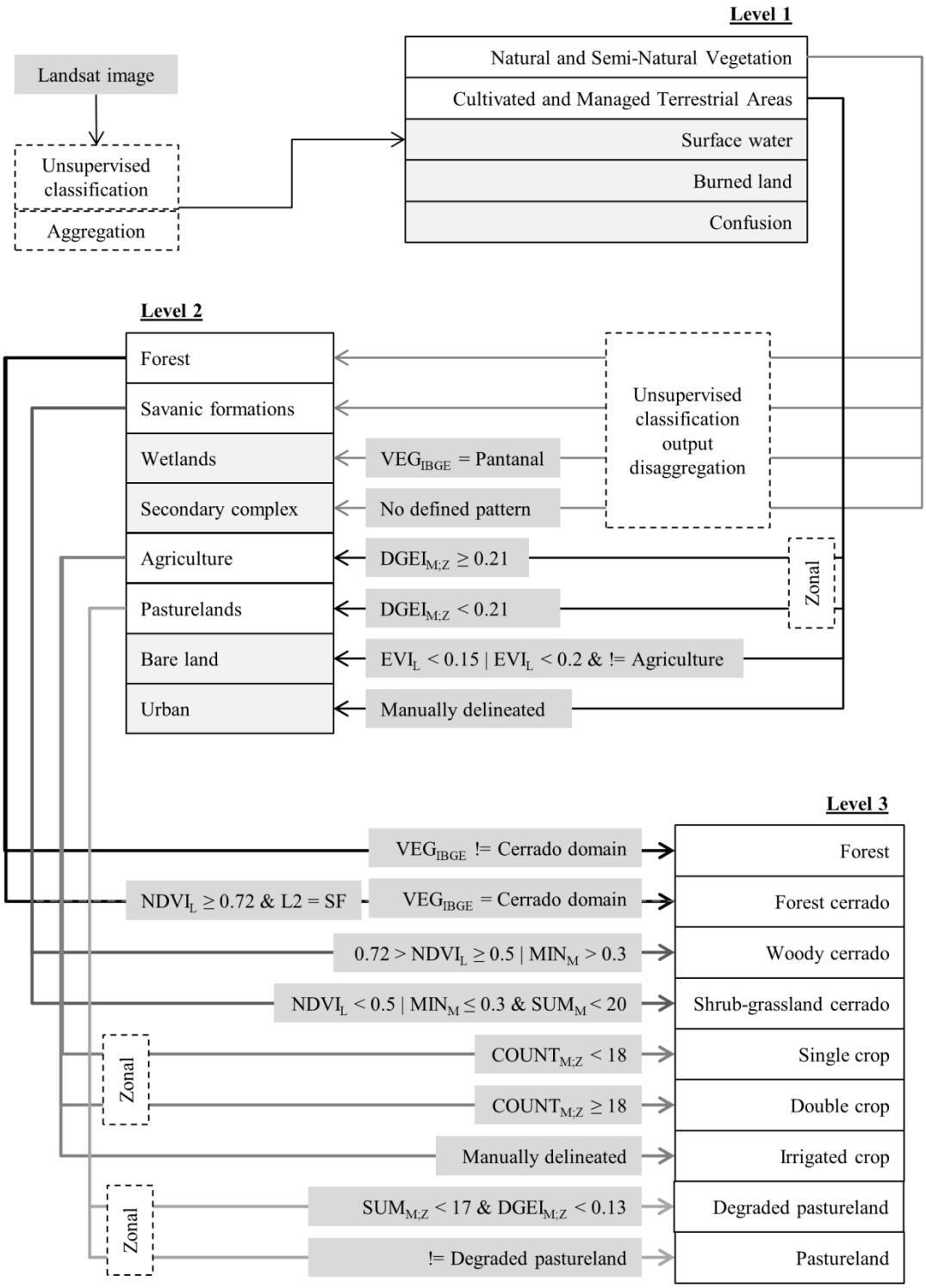


Figure 3. Workflow of the three-level hierarchical classification system developed and applied to derive a 30-year time series of land use and land cover maps of the UXRB, Brazil. Grey boxes indicate the data used in each decision rule designed to obtain a certain class. Solid line boxes represent land use and land cover classes. Land use and land cover classes boxes highlighted in light grey are classes that are carried over to next level of classification without modification. Whenever a class is derived based on an object approach, the arrow defining the respective workflow pass by a "zonal" box delimited by dotted lines. All thresholds presented here were developed for the study area only and can be used as approximate values for similar regions.

2.3.2. Accuracy assessment

We produced a confusion matrix based on (i) reference points collected in 2016 on the ground (total of 1460 ground reference points), and (ii) reference points collected through very high-resolution Rapid Eye images from 2012 and 2013 (total of 489). We transformed the observed sample confusion matrix into an estimated population matrix, as recommended by Pontius and Millones (2011). Then, we calculated overall (dis)agreement, overall quantity disagreement (amount of disagreement due to the quantity of each class), and overall allocation disagreement (amount of disagreement due to the miss-overlap of classes in space). Disagreement and its decomposition were also calculated for classes. Further details on accuracy assessment are available in Appendix A - Supplementary Material 4. As accuracy metrics calculated directly from the observed sample confusion matrix are still very common in the literature, we also present these metrics in Appendix A - Supplementary Material 4.

Once the classification process for the 2015 images returned a satisfactory output, we assumed that the analyst acquired the "know how" of the classification process, and could classify previous years using the same rules. In the process, we visually checked whether different thresholds would improve the classification of sample points with a marked land cover, but no improvement was observed. We did not collected data to estimate accuracy for previous year. But we assumed the 2015 accuracy assessment would also apply for previous year. It's important to point out that burned lands were not considered in such analysis. Burned lands class is present in a short temporal window, which makes it almost impossible to collect ground reference points. Thus, it was omitted from analyses. Still, such class presents a very distinct spectral profile and, then, it would not significantly change the accuracy results. A detailed explanation of reference data collection and analyses is showed in Appendix A - Supplementary Material 4.

In addition to the validation database collected in our field surveys, we used LULC information from Amazon Environmental Research Institute (IPAM, http://ipam.org.br), Socio-Environmental Institute (ISA, https://www.socioambiental.org) and The Brazilian Institute of Geography and Statistics (IBGE, 2016) to evaluate our results in separate analysis. The IPAM database includes LULC ground reference points obtained in 2017 for Querência municipality. ISA database contains maps for multiple years of LULC for Querência, Canarana, São José do Xingu, and Santa Cruz do Xingu municipalities, which in turn represent together about 25% of the UXRB. We randomly sampled these maps and compared each point with the ones from our output maps for the 2010 and 2015 years. ISA's and IPAM's LULC information were not as detailed as our L3 classification for cropping systems. Thus, to analyze L3 we compared IBGE (Brazilian Institute of Geography and Statistics) census data for each municipality in the basin with our maps.

2.3.3. LULC change

Rates of LULC change were calculated based on the continuous rate of change proposed by Puyravaud (2003):

$$LCLUc\ (\%) = \frac{\ln\left(A_2/A_1\right)}{t_2 - t_1}$$

where A indicates the total area occupied by a certain class in time t. The subscripts 1 and 2 represent the earlier and the subsequent year in an analyzed period. LULCc is the rate of change per year expressed in percentage. We analyzed LULC change trajectories in order to address proximate (direct) causes of change. We built transition matrices produced by a pixel-by-pixel identification of the LULC class in an earlier and the subsequent year time. The number of pixels is transformed in area by multiplying the former by the pixel resolution. We explore underlying

causes of LULC change in the UXRB by comparing agricultural expansion to market prices of agricultural products and currency exchange rate (the value ratio between Brazilian Real (R$) and America Dollar (US$)) using the Spearman correlation index. The agricultural products incorporated in the analysis were chosen due their representativeness in the area: (1) cattle herd and beef price, (2) soybean production output and price, (3) maize production output and price. Two sets of prices were used, (1) commodity real prices (corrected by inflation) published by the World Bank (http://www.worldbank.org), and (2) prices received by farmers corrected by inflation (IPCA index for February 2018) and published by the Institute of Applied Economic Research (http://www.ipea.gov.br). Agriculture production was acquired from the Brazilian Institute of Geography and Statistics (https://sidra.ibge.gov.br). All calculations were conducted in R (R Development Core Team, 2011).

2.4. Results

2.4.1. Classification approach

We created a classification scheme based on simple classification routines, remotely sensed information available at no cost, and algorithms that can be implemented in any common and/or free software that offers the possibility of spatial analyses (E.g.: R, GRASS, among others). For this approach, we created a set of rules, classified image mosaics and product accuracy checks based on ground reference points. We expanded the classification scheme to the years 2010, 2005, 2000, 1995, 1990, and 1985 to identify the main changes in LULC. All derived maps are presented in Appendix A - Supplementary material 5. To apply this classification approach in the UXRB, we processed 12 Landsat (1 mosaic) and ~182 MODIS images (~45 mosaics) for each year. Some critical years required more images to double check spectral quality and to replace cloud covered scenes.

2.4.2. Accuracy and class separability

We built a three-level hierarchical model (Figure 3) for image classification with increasing output detail at each level which, thus, presents increasing input and computational demand, and an overall decrease in accuracy (Table 2, Appendix A - Supplementary material 4 for full error matrices). The comparison of our validation database with the produced maps shows that the fraction of correctly classified sample points (overall agreement), considering the proportion of each LULC class in the UXRB, ranged from 84% to 93%. When decomposing the disagreement, all levels presented higher disagreement in quantity than in allocation (Table 2).

Table 2. Overall agreement, quantity disagreement, allocation disagreement, and number of classes relative to the proposed land use and land cover classification workflow in the Upper Xingu River Basin (Mato Grosso, Brazil) in 3 levels of information. Calculation was based on comparison with 1500 ground reference points collected in 2016. Complete error matrix can be assessed in Appendix A - Supplementary Material 4.

Reference database (2016)			
	Level 1	Level 2	Level 3
Overall agreement (%)	93.21	86.49	83.64
Quantity disagreement (%)	4.43	8.84	10.19
Allocation disagreement (%)	2.35	4.67	6.17
Nº of classes	5	11	16

All classes presented satisfactory agreement when considering their extension in the area - equal to or lower than 7%, independent of the level of classification. The highest disagreement percentages were observed in

cultivated areas in L1 (6%), forest and pasturelands in L2 (7% each), and in forest and pasturelands in L3 (7% each). Forest class error was composed mostly of commission error and quantity disagreement, indicating they were overestimated. Pasturelands in both L1 and L2 presented a balanced error between omission and commission, as well as for quantity and allocation disagreement, indicating both overestimation and spatial dislocation.

When not considering class extension and calculating accuracy directly from observed matrix values, few classes present lower accuracy than expected. In L2, secondary complex and bare land present commission errors of 52% and 54%, respectively. In L3, woody Cerrado (39%), shrub-grassland Cerrado (76%), secondary complex (59%), degraded pasturelands (57%), and bare soil (44%) presented higher omission and commission error percentages. Still, L1 presented great accuracy for all classes. This result was expected, since L1 class separability based on Landsat bands is highly reliable (See Appendix A - Supplementary material 5 – Figure 1 for spectral profile). L2 classification is dependent on Level 1, and requires the use of the spectral indexes DGEI and NDVI to separate croplands from pasturelands, and pasturelands from bare soil (Figure 4). Natural covers could still be separated from Landsat band data (Figure 2 in Appendix A - Supplementary material 5). L3 classification was dependent on previous levels and thus represents the most complex model, requiring multiple indices to separate classes. Classes with higher separability, such as dense forest, and double cropping systems, presented the highest accuracy (Figure 5). Degraded pasturelands were mostly confused with pastureland, indicating that we set a conservative rule when separating those classes.

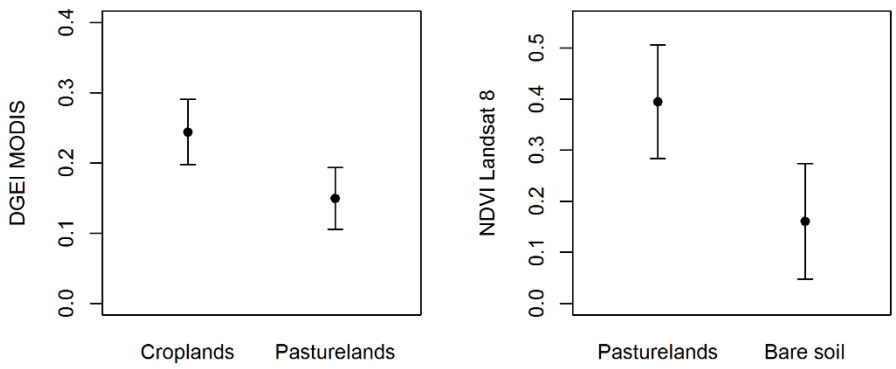


Figure 4. Separability between (a) croplands and pasturelands derived from the differential greenness enhancement index (DGEI) calculated from MODIS; and (b) pasturelands and bare soil derived from NDVI calculated from Landsat. Values were extracted based in ground reference points collected to assess classification accuracy.

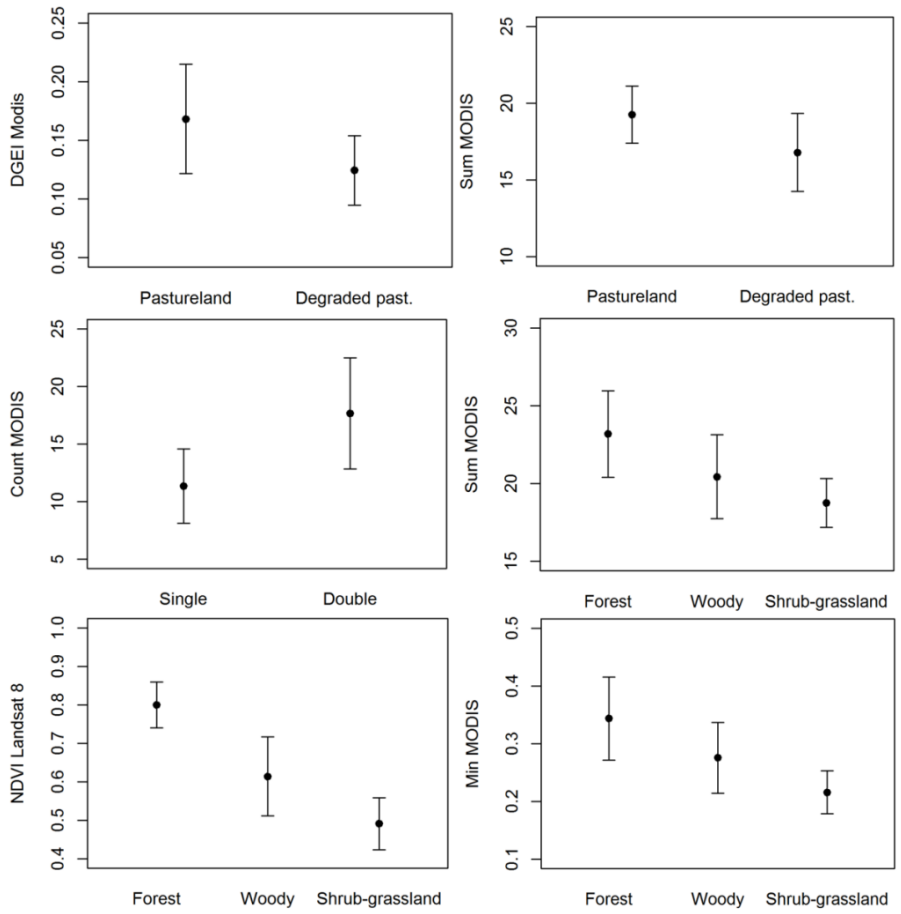


Figure 5. Separability between classes proposed for classification at level 3. Pasturelands and degraded pasturelands are compared derived from the differential greenness enhancement index – DGEI (a) and the sum of EVI values through a year – SUM (b). Single and double cropping systems are compared based on how many days along the year an EVI > 0.5 is present – COUNT(c). Forest, woody and shrub-grassland, all different savanna physiognomies, are compared based on SUM (d), NDVI values (e), and the minimal EVI value through a year (f). Values were extracted based on ground reference points collected to assess classification accuracy.

When we compare our mapping results to those obtained by other institutions using different mapping schemas of remote sensing classification (Table 3), we also achieved high values of overall agreement, indicating a high correspondence among them. However, it is important to highlight that LULC validation analyses based on databases created for other purposes should not be considered as accurate as the analysis produced from our validation dataset. Data from other sources contain their own errors, inherent in their production, in addition to errors which may arise from legend adaptation. Furthermore, IBGE data for planted cropland area also supports our methodology, as we found 99% correspondence between this survey and our classified maps through time (see Figure 6 and the Appendix A - Supplementary material 4 for graphic by municipality).

Table 3. Overall agreement, quantity disagreement, and allocation disagreement calculated when compared to different databases collected in different years. The information corresponds to the LULC classification of the UXRB (Brazil) in 3 levels of information. Databases were developed by *Instituto de Pesquisa Ambiental da Amazônia* (IPAM) and *Instituto Socioambiental* (ISA).

IPAM database (2017)			
	L1	L2	L3
Agreement (%)	79.96	75.38	64.49
Quantity disagreement (%)	13.41	20.41	15.18
Allocation disagreement (%)	6.63	4.21	20.33
ISA database (2010)			
	L1	L2	L3
Agreement (%)	88.97	90.54	-
Quantity disagreement (%)	9.43	5.69	-
Allocation disagreement (%)	1.60	3.77	-
ISA database (2015)			
	L1	L2	L3
Agreement (%)	95.85	94.22	-
Quantity disagreement (%)	2.07	2.99	-
Allocation disagreement (%)	2.08	2.78	-

2.4.1. LULC change in the UXRB

The period with the highest natural vegetation loss rate in the UXRB was observed to be between 2000 and 2005 (- 5% . year^{-1}), while the average rate was about – 2% . year^{-1}. According to L2 mapping, before 1995, most land use change occurred in the savanna formations (9,500 km² of ~17,000 km²), and resulted primarily from pasture expansion (Figure 7; See Appendix A - Supplementary material 7 for transition and area tables). From 1995 to 2010, expansion grew mainly over forested areas, while from 2010 to 2015, savanna formations have again been the main target of agricultural expansion. The observed rates of deforestation between 1995 and 2005 can be spatially divided into two main expansion axes. Until 2000, land use change was most intense at the east and southeast bounds of the UXRB, with large areas of tropical forest converted into pasturelands (~30,000 km²). From 2000 to 2005, when the highest rates of deforestation were observed, forest clearing occurred mainly in the western portion of the basin and represented both crop and pastureland expansion.

While the highest rates of pasture and crop expansion occurred, respectively, between 1990-1995 and 1995-2005, rates dropped to among the lowest between 2005-2010. This last period was also marked by the largest spread of intensification according to L3 mappings - 15% of pasturelands became croplands, and 25% of single crop systems became double cropping systems (Figure 8; See Appendix A - Supplementary material 7 for transition and area tables). Additionally, pastureland decreased in size during this period, as did the rate of improvement of degraded pasturelands. In this period, cropland expansion rate slightly decreased. With the exception of double crop system, cropland expansion rates increased again between 2010-2015. From 2010 to 2015, we observed the largest conversion of pastureland into single or double cropping agricultural systems (~9000 km²), while conversion of native vegetation, mainly Cerrado formations, increased again (~6,500 km² of cerrado and ~2,500 km² of forest were converted into crop systems and pastureland). Additionally, degraded pastureland area also increased during this period.

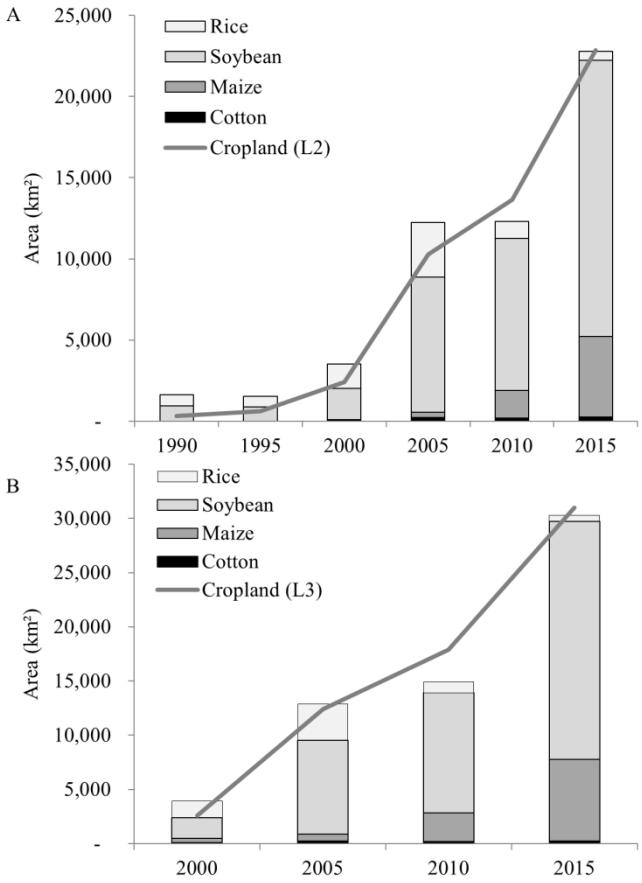


Figure 6. Correlation between the LULC maps produced for UXRB for croplands and the census data made available by IBGE for the top four crops (rice, soybean, maize and cotton). (A) Lines show the evolution of cropland according to L2 classification while bars show planted area data according to IBGE for rice, soybean discounting secondary maize area (double cropping), main crop maize, and cotton. (B) Lines show the evolution of cropland (accounting twice for double cropping areas) according to L3 classification while bars show planted area data according to IBGE for rice, soybean, maize, and cotton. Data is available by municipality in Appendix A - Supplementary material 4 – figure 2.

As expected, the expansion of harvested area and cattle herd is positively correlated with prices and the exchange rate between American Dollar and Brazilian Real (Table 4). For example, the observed peak in the rate of change for cropland (2000-2005) overlaps with the highest prices received by producers in the analyzed period (see Appendix A - Supplementary material 7 for figures). Still, the exchange rate presented a significantly higher correlation to the expansion of agriculture production than commodity prices.

Table 4. Spearman correlation and the corresponding significance at 95% (*) and 99% (**) confidence level between main agriculture products (cattle herd, harvested area of soybean and maize accounting for double crop systems), commodities prices (World Bank), price received by the producers (IPEA), and exchange rate between American Dollar and Brazilian Real (Central Bank of Brazil). One year lagged tests presented very similar results.

	Cattle herd	**Soybeans area**	**Maize area**
Commodity price (US$)	-	0.56*	0.56**
Price received by producers (R$)	0.51*	0.62**	-
Exchange rate (US$:R$)	0.77**	0.84**	0.81**

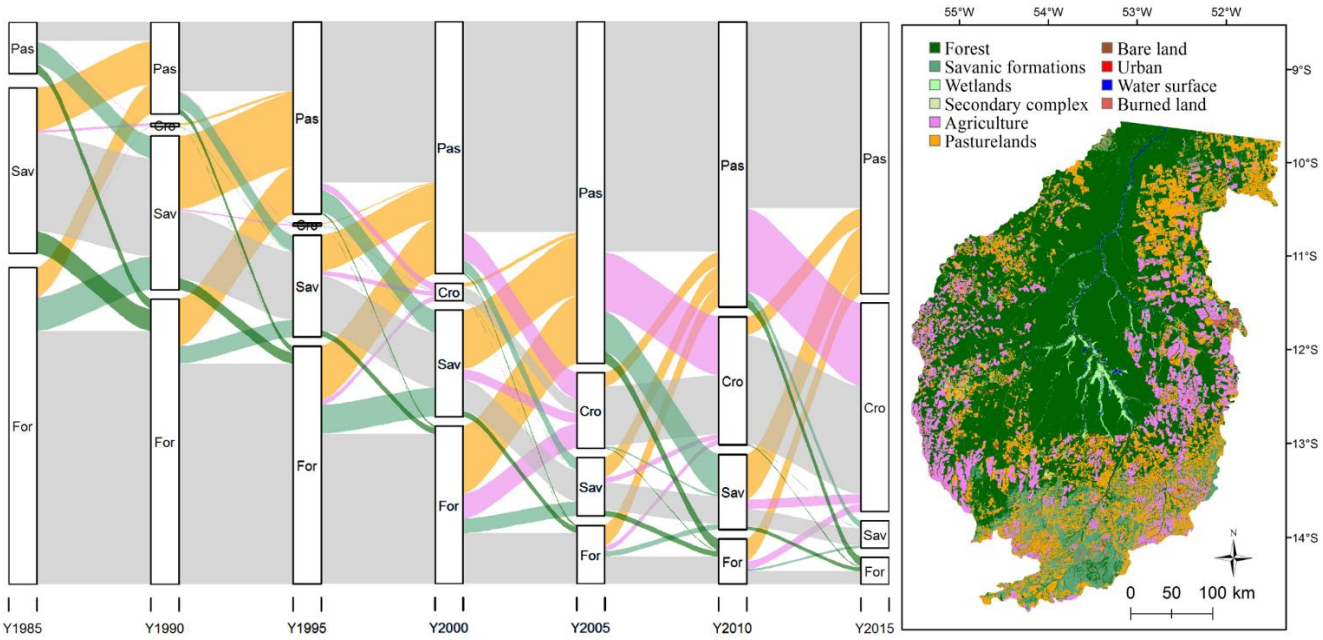


Figure 7. Land use and land cover transition for all pixels that have changed through time and in non-protected areas in the Upper Xingu River Basin from 1985 to 2015, according to level 2 of classification. The vertical boxes represent the proportion of the Upper Xingu River Basin which each land use occupies in a certain year. The flux lines represent the land use and land cover change. The width of each line represents the proportional amount of land being converted into another use, while each color represents the use it was turned into. The color scheme follows the legend of the map which shows the land use and land cover distribution in 2015.

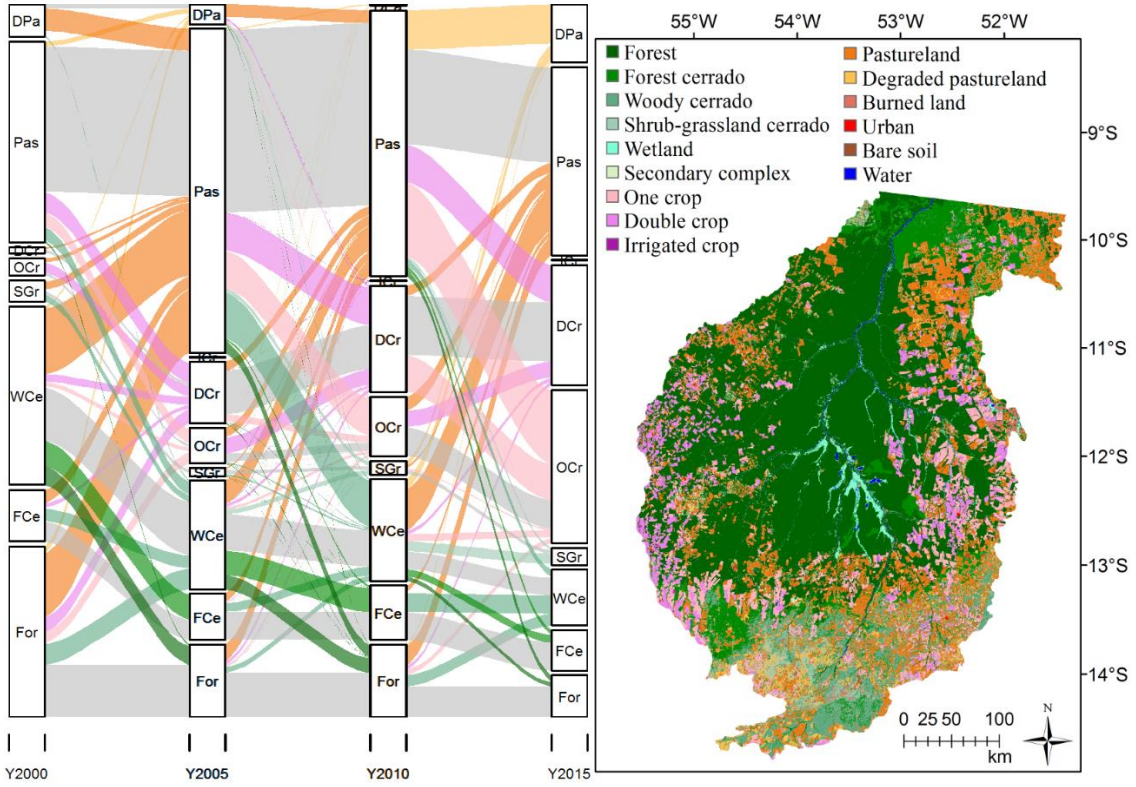


Figure 8. Land use and land cover transition for all pixels that have changed through time and in non-protected areas in the Upper Xingu River Basin from 2000 to 2015, according to level 3 of classification. The vertical boxes represent the proportion of the Upper Xingu River Basin which each land use occupies in a certain year. The flux lines represent the land use and land cover change. The width of each line represents the proportional amount of land being converted into another use, while each color represents the use it was turned into. The color scheme follows the legend of the map which shows the land use and land cover distribution in 2015.

2.5. Discussion

2.5.1. LULC classification and legend

The proposed classification scheme and legend produced spatially explicit information on LULC for the UXRB with more accuracy than we would expect if using global or national products. Besides presenting an overall accuracy between 69-78%, global land cover products achieve poor agreement among themselves in tropical ecotones and diverse landscapes (Herold et al., 2008). By design, regional LULC maps derived from satellite images generally present higher levels of accuracy than global mapping schemas because of their intrinsic variability. Our corresponding product (L2) produced an overall accuracy of 86% for the UXRB. Recently produced regional LULC datasets that encompass the UXRB at least partially, also offer information on the study area. The TerraClass Cerrado has an ~80% overall accuracy (Scaramuzza et al., 2017), while TerraClass Amazonia's accuracy is ~77% (Almeida et al., 2016). Still, these mapping schemes present a simpler legend assessment when compared to our hierarchical approach proposed here at L3, and a poorer performance than both L2 (86% accuracy) and L3 (83% accuracy). Additionally, both datasets together do not provide information for the whole studied area, since no mapping data exists for a large portion of the ecotone zone between the Amazon and Cerrado biomes located inside the TerraClass Amazonia study area.

Also, our accuracy analyses showed similar or higher indexes than others studies applied to similar regions and scale. Sawakushi et al. (2013) applied classification schemes at the Middle Araguaia River Basin, reaching an overall accuracy of 85%, based on 287 ground reference points and delivering a legend which fits in-between our L1 and L2 classifications. Walker (2010) also proposed a hierarchical and multi-level classification approach and applied it to the UXRB. Using a random forest algorithm, and PALSAR- and Landsat-based data, they obtained an overall accuracy from 58% (15 classes) to 92% (2 classes), depending on the classification level. However, all three studies, including ours, achieved poorer results for separating (i) Cerrado land covers, and (ii) pasturelands and degraded pasturelands. Degraded pastureland is an especially important land cover type due to its link to food security and conservation. Even though some studies show an intensification pattern related to improvement in the Brazilian pasturelands since 2005 (Parente and Ferreira, 2018), we believe that degraded pasture is underestimated in our study. The thresholds we implemented linked to the hysteresis principle were likely too conservative.

Few studies have been carried out to improve separability among Cerrado vegetation physiognomies based on remote sensing products. Vegetation indices (VIs) derived from MODIS and calculated indices based on VIs have shown a capacity to separate major physiognomies (Ratana et al., 2005); however, Landsat-simulated VIs presented a better discrimination capability when compared to these methods (Ferreira et al., 2003). In our study at the UXRB, even when employing MODIS and Landsat EVI data together, the overall accuracy for Cerrado physiognomy separability was 61% (L3) when not weighting the full class extension, and even lower values were found for shrub-grassland formations. Since shrub-grassland encompasses only 1% of the landscape, and due to the few samples obtained in field surveys, these results are possibly associated with a poor set of rules in the classification process leading to low accuracy values. Still, the results suggest that further attention should be given to studying the temporal pattern of spectral responses in different Cerrado formations. Toniol and colleagues (2017) evaluated four classifiers for discriminating Cerrado physiognomies in the rainy and dry seasons, and obtained overall accuracies from 26% to 84%, depending on the classifier, season, and metrics used.

MODIS derived VI temporal profiles have also been applied to classify managed land, specifically to extract croplands and to derive crop-specific classes (Brown et al., 2007, 2013). Most of these studies have been

based on modelling annual VI records instead of single indices (Rizzi et al., 2009). Aside from the simplicity of using indices such as DGEI or COUNT, one advantage of such method compared to other approaches is that there is no need to define a time period. In other words, these indices can be applied to other regions that do not necessarily present the same agricultural calendar, allowing their use at regional and global mapping scales. In DGEI calculation, the period of time used can be much broader than the sowing-growth window (as applied here), avoiding seasonal changes.

We achieved a more realistic and representative approach using a segmentation-based approach to turn managed areas into objects, and turning those objects into analysis units (Blaschke, 2010). Through this approach we, avoided problems such as pixel heterogeneity, mixed pixels, and spectral similarity due to crop or pastureland spatial variability (Peña-Barragán et al., 2011; Yu et al., 2016). Additionally, while classification approaches often present a larger allocation disagreement than quantity disagreement, our results shows the opposite (E.g.: Pontius et al (2011)). It suggests that the combined pixel and object approach represents the land configuration better than automatic per-pixel classification methods. Still, the techniques used in this research may not favor mapping of small producers and permanent crops. Such analysis is out of the scope of this research, but it is encouraged for future studies. Indices other than the ones used in this study, such as a clumping index, were not reported to significantly improve separability in neotropical savannas and are highly correlated with VI (Hill et al., 2011). We encourage the application of object-oriented approaches and indices derived from temporal composition in other classification methods, such as machine learning.

The proposed classification approach allowed us to analyze a more complex and realistic scenario of LULC change in the UXRB, and thus address issues related to expansion and intensification, biome specific dynamic changes, and cropland and pastureland roles in native vegetation loss. The high accuracy values attributed to L1 and the applied legend make this level of classification suitable for studies on law and policy enforcement (Goetz et al., 2015). L2 addresses general differences in land use and land cover, offering a reliable product to model ecosystem dynamics such as regional hydrological balance (Dwarakish and Ganasri, 2015). The third level of classification offers further information on regional LULC by addressing vegetative productivity. It allows users to address issues related to habitat suitability and priority areas for conservation by separating native vegetation physiognomies. Such separability issues are even more important and challenging for the Cerrado biome. Although it is recognized as the richest savanna in the world and a hotspot for conservation (Klink and Machado, 2005; Myers et al., 2000), it is still not a major focus of policies and research, nationally or internationally (Nolte et al., 2017). This pattern is related to the lack of monitoring techniques and knowledge of ecological dynamics taking place in the biome. Our proposed methodology offers insights into monitoring techniques and encourages further research. The third level of classification also makes it possible to address food production, land management, and the relationship between productive systems and socioeconomic dynamics. The intensified use of agricultural lands is a key discussion topic related to both land sparing for conservation (Ewers et al., 2009) and socioeconomic development (Martinelli et al., 2017). Still, a clear drawback of the hierarchical approach used in this study is the nested dependency between levels of classification since L3 is dependent on L2, and L2 is dependent on L1. Possible ways of mitigating this problem include assigning a smaller weight to classified maps when using them as source of information in another level of classification, or even avoiding altogether the use of one classified map in the modelling of another level of classification.

2.5.2. LULC changes in the UXRB

Until 2000, LULC change in UXRB was primarily driven by pastureland expansion following patterns observed in Amazon and Cerrado biomes (Ferreira et al., 2012; INPE, 2018). In 1990s and 2000s, the region transitioned from a pioneer landscape to a consolidated frontier, in which agro-production moved from a labor-intensive system to one based on financial capital and integration into market. The consolidation of cattle production in the 90s occurred in all states on the southern frontier of the Amazon (Margulis, 2004). From 2000 to 2015, land use transitions in the UXRB indicated an intensification process, which was manifested through an increase in cultivated area over pastureland, accompanied by a large reduction in deforestation. Intensification processes are a result of environmental regulations, technological changes, economic disincentives for deforestation, and/or market regulations (Gasparri and de Waroux, 2015). By the end of the study period, the UXRB became a frontier which supplies both international and regional markets.

Reductions in deforestation, combined with crop expansion, have been described for the Brazilian biomes in the last decades (Dias et al., 2016; Macedo et al., 2012). Brando (2013) shed light on how the actions of different stakeholders (e.g.: government – zero deforestation act, private sector – soy moratorium, and non-profit organizations through multiple campaigns) helped to reduce deforestation in the UXRB during the 2000s. With available technology, producers were able to increase yields in order to increase profits. However, our analysis shows that intensification is biome-dependent. We observed an increase in agricultural intensification mainly in the Amazon region of the UXRB, between 2000-2005, and 2010-2015. This temporal pattern is correlated with peaks in the exchange rate, and secondarily with commodity prices. Expansion is still the dominant LULC change process in the Cerrado portion of the UXRB, as it is for the Cerrado biome as whole (Lapola et al., 2013). This expansion can be confirmed as we observe a reduction of 20% in savanic formation from 2010 to 2015, a value 10 times higher than the one found for the amazon forest over the same period. The conversion from woody Cerrado to pastureland accounts for the majority of these changes. This process is correlated with an increase in beef prices. But the low number of conservation areas, as well as the lack of market regulation and effective deforestation control programs, are the main differences between protection schemes in the Cerrado and the Amazon biomes (Gibbs et al., 2015; Sparovek et al., 2010). According to intensification theories, without those mechanisms of incentive or control, intensification will not be encouraged (Meyfroidt et al., 2018).

Cash crops primarily replace pasturelands in the UXRB, followed by cerrado formations and Amazonian forest - mainly between 2000-2005 and 2010-2015. Those results are partially in accordance with Macedo (Macedo et al., 2012), which underestimated native vegetation loss at the cerrado region in the UXRB due to cash crop expansion. Rather than intensification and regulation effects, the lower conversion of native vegetation into croplands observed between 2005-2010 is a result of market and currency exchange effects pushing soybean prices down (Richards et al., 2012). The increase in native vegetation loss again from 2010-2015 supports this idea, together with the hypothesis that intensification does not necessarily lead to decreases in the expansion of agricultural frontiers (Barretto et al., 2013; Rudel et al., 2009). Although our goal was not to analyze indirect frontier development through land use change, the displacement of cattle production by cash crops has been observed in the UXRB and elsewhere (Barona et al., 2010; Fehlenberg et al., 2017; Macedo et al., 2012). Aditionally, increased land market values in areas of soybean expansion have been shown to drive migration-based development (Richards, 2015). The increase in deforestation in the Cerrado compared to the Amazon portion of the basin can be considered evidence of a rebound effect. Theories concerning this effect state that technology, policies, and market-drive the intensification of land use as observed in the Amazon portion of the UXRB (Meyfroidt et al., 2018). Land sparing is

observed in a region when land use is restricted by regulations and, as a consequence, production price per area increases due to a required disaggregation of production area. Conversely, expansion is observed elsewhere as a rebound effect when land with limited regulations is available – such as in the Cerrado biome (Soares-Filho et al., 2014); and the production is focused on products for which the prices are elastic based on demand - such as soybean (Heien and Pick, 1991; Meyfroidt et al., 2018).

2.6. Conclusions

Tropical ecotone zones and agricultural frontiers, such as the UXRB, are important for Neotropical biosociodiversity and for agricultural production. However, their inherent complexity poses challenges for LULC change analyses and LULC planning. The combination of multiple classification techniques, as well as the combination of remote sensing and GIS-based information which assesses temporal variability, have been reported as superior approaches compared to traditional techniques. Our proposed classification scheme uses an unsupervised classification approach to initially group pixels into few classes. These classes are then reclassified through decision trees by the integration of multi-sensor and GIS data into an object and pixel context. The multi-data, -temporal and -spatial scale characteristics of the analyses were crucial to maintaining high spatial resolution and the high amount of information throughout the hierarchal classification. Furthermore, the approaches applied here can give insights into continental and global mapping processes such as the use of metrics based on flexible time series with no need to define fixed starting and ending dates. Still, to transfer the methodology to other study areas, thresholds and information should be adapted.

Because of their ability to discriminate between different LULC in complex environments, the LULC products created in this research enabled us to differentiate between different processes that drive agricultural production increases in the UXRB. Beside our observation that the increase in agricultural production in the UXRB being generally correlated with commodity prices and monetary exchange rates, we were also able to map different processes affecting production (expansion, intensification, and degradation). For instance, the area in the basin which overlaps with the Amazon biome experiences intensification, while the Cerrado biome experiences expansion. This result is mainly linked to government and market regulations, which are focused in the Amazon but are lacking in the Cerrado biome.

References

Aguirre-Gutiérrez, J., Seijmonsbergen, A.C., Duivenvoorden, J.F., 2012. Optimizing land cover classification accuracy for change detection, a combined pixel-based and object-based approach in a mountainous area in Mexico. Appl. Geogr. 34, 29–37. https://doi.org/10.1016/j.apgeog.2011.10.010

Almeida, C.A., Coutinho, A.C., Esquerdo, J.C.D.M., Adami, M., Venturieri, A., Diniz, C.G., Dessay, N., Durieux, L., Gomes, A.R., 2016. High spatial resolution land use and land cover mapping of the 1 Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. Acta Amaz. 46, 291–302. https://doi.org/10.1590/1809-4392201505504

Arvor, D., Meirelles, M., Dubreuil, V., Bégué, A., Shimabukuro, Y.E., 2012. Analyzing the agricultural transition in Mato Grosso, Brazil, using satellite-derived indices. Appl. Geogr. 32, 702–713. https://doi.org/10.1016/j.apgeog.2011.08.007

Azzari, G., Lobell, D.B., 2017. Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring. Remote Sens. Environ. 202, 64–74. https://doi.org/10.1016/j.rse.2017.05.025

Barona, E., Ramankutty, N., Hyman, G., Coomes, O.T., 2010. The role of pasture and soybean in deforestation of the Brazilian Amazon. Environ. Res. Lett. 5, 024002. https://doi.org/10.1088/1748-9326/5/2/024002

Barretto, A.G.O.P., Berndes, G., Sparovek, G., Wirsenius, S., 2013. Agricultural intensification in Brazil and its effects on land-use patterns: An analysis of the 1975-2006 period. Glob. Chang. Biol. 19, 1804–1815. https://doi.org/10.1111/gcb.12174

Bartholomé, E., Belward, A.S., 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. Int. J. Remote Sens. 26, 1959–1977. https://doi.org/10.1080/01431160412331291297

Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS J. Photogramm. Remote Sens. https://doi.org/10.1016/j.isprsjprs.2009.06.004

Bondeau, A., Smith, P.C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-campen, H., Müller, C., Reichstein, M., Smith, B., 2007. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Glob. Chang. Biol. 13, 679–706. https://doi.org/10.1111/j.1365-2486.2006.01305.x

Brando, P.M., Coe, M.T., DeFries, R., Azevedo, A.A., 2013. Ecology, economy and management of an agroindustrial frontier landscape in the southeast Amazon. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 368, 9. https://doi.org/10.1098/rstb.2012.0152

Brasil/MMA, 2018. O Bioma Cerrado [WWW Document]. URL http://www.mma.gov.br/biomas/cerrado (Available)

Brazil/MMA, 2018. Amazônia [WWW Document]. URL http://www.mma.gov.br/biomas/amazônia (Available)

Brown, J., Jepson, W., Kastens, J., Wardlow, B., Lomas, J., Price, K., 2007. Multitemporal, Moderate-Spatial-Resolution Remote Sensing of Modern Agricultural Production and Land Modification in the Brazilian Amazon. GIScience Remote Sens. 44, 117–148. https://doi.org/10.2747/1548-1603.44.2.117

Brown, J.C., Kastens, J.H., Coutinho, A.C., Victoria, D. de C., Bishop, C.R., 2013. Classifying multiyear agricultural land use data from Mato Grosso using time-series MODIS vegetation index data. Remote Sens. Environ. 130, 39–50. https://doi.org/10.1016/j.rse.2012.11.009

Chen, Yuehong, Zhou, Y., Ge, Y., An, R., Chen, Yu, 2018. Enhancing Land Cover Mapping through Integration of Pixel-Based and Object-Based Classifications from Remotely Sensed Imagery. Remote Sens. 10, 77. https://doi.org/10.3390/rs10010077

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., Lambin, E., 2004. Digital change detection methods in ecosystem monitoring: a review. Int. J. Remote Sens. 25, 1565–1596. https://doi.org/10.1080/0143116031000101675

de Chazal, J., Rounsevell, M.D.A., 2009. Land-use and climate change within assessments of biodiversity change: A review. Glob. Environ. Chang. 19, 306–315. https://doi.org/10.1016/j.gloenvcha.2008.09.007

Di Gregorio, A., 2016. Land Cover Classification System. FAO, Rome.

Di Gregorio, A., Jansen, L.J.M., 2000. Land Cover Classification System (LCCS): Classification Concepts and User Manual. Fao 53, 179. https://doi.org/10.1017/CBO9781107415324.004

Dias-Filho, M.B., 2001. Degradação de Pastagens: processos, causas e estratégias, 4o. ed, Degradação de Pastagens: processos, causas e estratégias. MBDF, Belém.

Dias, L.C.P., Pimenta, F.M., Santos, A.B., Costa, M.H., Ladle, R.J., 2016. Patterns of land use, extensification, and intensification of Brazilian agriculture. Glob. Chang. Biol. https://doi.org/10.1111/gcb.13314

Döll, P., Kaspar, F., Lehner, B., 2003. A global hydrological model for deriving water availability indicators: Model tuning and validation. J. Hydrol. 270, 105–134. https://doi.org/10.1016/S0022-1694(02)00283-4

Don, A., Schumacher, J., Freibauer, A., 2011. Impact of tropical land-use change on soil organic carbon stocks - a meta-analysis. Glob. Chang. Biol. 17, 1658–1670. https://doi.org/10.1111/j.1365-2486.2010.02336.x

Dwarakish, G.S., Ganasri, B.P., 2015. Impact of land use change on hydrological systems: A review of current modeling approaches. Cogent Geosci. 1. https://doi.org/10.1080/23312041.2015.1115691

ERDAS, 2019. IMAGINE Expert Classifier [WWW Document]. ERDAS IMAGINE 2018 Prod. Descr. URL https://www.hexagongeospatial.com/technical-documents/product-descriptions-2018/erdas-imagine-2018-product-description (accessed 6.30.19).

ERDAS, 2014. ERDAS Imagine [WWW Document]. URL https://www.hexagongeospatial.com/products/power-portfolio/erdas-imagine (accessed 6.1.19).

Ewers, R.M., Scharlemann, J.P.W., Balmford, A., Green, R.E., 2009. Do increases in agricultural yield spare land for nature? Glob. Chang. Biol. 15, 1716–1726. https://doi.org/10.1111/j.1365-2486.2009.01849.x

FAO-Food and Agriculture Organization, 1993. Guidelines for land-use planning, FAO Development Series.

FAO, 2011. The state of the world's land and water resources for food and agriculture (SOLAW) - Managing systems at risk. Food and Agriculture Organization of the United Nations, Rome, and Earthscan, London.

FAOSTAT, 2016. Country Indicators [WWW Document]. URL http://www.fao.org/faostat/en/#home

Fehlenberg, V., Baumann, M., Gasparri, N.I., Piquer-Rodriguez, M., Gavier-Pizarro, G., Kuemmerle, T., 2017. The role of soybean production as an underlying driver of deforestation in the South American Chaco. Glob. Environ. Chang. 45, 24–34. https://doi.org/10.1016/j.gloenvcha.2017.05.001

Ferreira, L.G., Sano, E.E., Fernandez, L.E., Araújo, F.M., 2013. Biophysical characteristics and fire occurrence of cultivated pastures in the Brazilian savanna observed by moderate resolution satellite data. Int. J. Remote Sens. 34, 154–167. https://doi.org/10.1080/01431161.2012.712223

Ferreira, L.G., Yoshioka, H., Huete, A., Sano, E.E., 2003. Seasonal landscape and spectral vegetation index dynamics in the Brazilian Cerrado: An analysis within the Large-Scale Biosphere-Atmosphere Experiment in Amazônia (LBA). Remote Sens. Environ. 87, 534–550. https://doi.org/10.1016/j.rse.2002.09.003

Ferreira, M.E., Jr, L.G.F., Miziara, F., Soares-, B.S., Ferreira, L.G., Miziara, F., Soares-Filho, B.S., 2012. Modeling landscape dynamics in the central Brazilian savanna biome: future scenarios and perspectives for conservation. J. Land Use Sci. 4248, 1–19. https://doi.org/10.1080/1747423X.2012.675363

Foley, J.A., Defries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J. a, Prentice, I.C., Ramankutty, N., Snyder, P.K., 2005. Global consequences of land use. Science 309, 570–574. https://doi.org/10.1126/science.1111772

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168–182. https://doi.org/10.1016/j.rse.2009.08.016

Garcia, A.S., Vilela, V.M. de F.N., Rizzo, R., West, P., Gerber, J.S., Engstrom, P.M., Ballester, M.V.R., 2019. Assessing land use/cover dynamics and exploring drivers in the Amazon's arc of deforestation through a hierarchical, multi-scale and multi-temporal classification approach. Remote Sens. Appl. Soc. Environ. 15, 100233. https://doi.org/10.1016/j.rsase.2019.05.002

Gasparri, N.I., de Waroux, Y. le P., 2015. The Coupling of South American Soybean and Cattle Production Frontiers: New Challenges for Conservation Policy and Land Change Science. Conserv. Lett. 8, 290–298. https://doi.org/10.1111/conl.12121

Ge, J., Qi, J., Lofgren, B.M., Moore, N., Torbick, N., Olson, J.M., 2007. Impacts of land use/cover classification accuracy on regional climate simulations. J. Geophys. Res. 112, D05107. https://doi.org/10.1029/2006JD007404

Gibbs, H.K., Rausch, L., Munger, J., Schelly, I., Morton, D.C., Noojipady, P., Soares-Filho, B., Barreto, P., Micol, L., Walker, N.F., 2015. Brazil's Soy Moratorium. Science (80-.). 347, 377–378. https://doi.org/10.1126/science.aaa0181

Giri, C., Zhu, Z., Reed, B., 2005. A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. Remote Sens. Environ. 94, 123–132. https://doi.org/10.1016/j.rse.2004.09.005

Goetz, S.J., Hansen, M., Houghton, R.A., Walker, W., Laporte, N., Busch, J., 2015. Measurement and monitoring needs, capabilities and potential for addressing reduced emissions from deforestation and forest degradation under REDD+. Environ. Res. Lett. https://doi.org/10.1088/1748-9326/10/12/123001

Gómez, C., White, J.C., Wulder, M.A., 2016. Optical remotely sensed time series data for land cover classification: A review. ISPRS J. Photogramm. Remote Sens. https://doi.org/10.1016/j.isprsjprs.2016.03.008

Hansen, M.C., Loveland, T.R., 2012. A review of large area monitoring of land cover change using Landsat data. Remote Sens. Environ. 122, 66–74. https://doi.org/10.1016/j.rse.2011.08.024

Herold, M., Mayaux, P., Woodcock, C.E., Baccini, A., Schmullius, C., 2008. Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. Remote Sens. Environ. 112, 2538–2556. https://doi.org/10.1016/j.rse.2007.11.013

Hill, M.J., Román, M.O., Schaaf, C.B., Hutley, L., Brannstrom, C., Etter, A., Hanan, N.P., 2011. Characterizing vegetation cover in global savannas with an annual foliage clumping index derived from the MODIS BRDF product. Remote Sens. Environ. 115, 2008–2024. https://doi.org/10.1016/j.rse.2011.04.003

Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D., 2013. Change detection from remotely sensed images: From pixel-based to object-based approaches. ISPRS J. Photogramm. Remote Sens. 80, 91–106. https://doi.org/10.1016/j.isprsjprs.2013.03.006

IBGE, 2019. Sistema IBGE de Recuperação Automática (SIDRA): banco de dados sobre indicadores, população, economia e geociências.

IBGE, 2004. Mapas de Cobertura Vegetal dos Biomas Brasileiros.

INPE - Instituto Nacional de Pesquisas Espaciais, 2019. Monitoramento da Floresta Amazônica Brasileira por Satélite [WWW Document]. URL http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes (accessed 5.10.19).

Ivanauskas, N.M., Monteiro, R., Rodrigues, R.R., 2008. Classificação fitogeográfica das florestas do Alto Rio Xingu. Acta Amaz. 38, 387–402. https://doi.org/10.1590/S0044-59672008000300003

Klink, C. a., Machado, R.B., 2005. Conservation of the Brazilian Cerrado. Conserv. Biol. 19, 707–713. https://doi.org/10.1111/j.1523-1739.2005.00702.x

Lapola, D.M., Martinelli, L.A., Peres, C.A., Ometto, J.P.H.B., Ferreira, M.E., Nobre, C.A., Aguiar, A.P.D., Bustamante, M.M.C., Cardoso, M.F., Costa, M.H., Joly, C.A., Leite, C.C., Moutinho, P., Sampaio, G., Strassburg, B.B.N., Vieira, I.C.G., 2013. Pervasive transition of the Brazilian land-use system. Nat. Clim. Chang. 4, 27–35. https://doi.org/10.1038/nclimate2056

Lathuillière, M.J., Miranda, E.J., Bulle, C., Couto, E.G., Johnson, M.S., 2017. Land occupation and transformation impacts of soybean production in Southern Amazonia, Brazil. J. Clean. Prod. 149, 680–689. https://doi.org/10.1016/j.jclepro.2017.02.120

Lebourgeois, V., Dupuy, S., Vintrou, É., Ameline, M., Butler, S., Bégué, A., 2017. A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated Sentinel-2 time series, VHRS and DEM). Remote Sens. https://doi.org/10.3390/rs9030259

Lu, D., Mausel, P., Brondízio, E., Moran, E., 2004. Change detection techniques. Int. J. Remote Sens. 25, 2365–2401. https://doi.org/10.1080/0143116031000139863

Lu, D., Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. Int. J. Remote Sens. https://doi.org/10.1080/01431160600746456

Macedo, M.N., Coe, M.T., DeFries, R., Uriarte, M., Brando, P.M., Neill, C., Walker, W.S., 2013. Land-use-driven stream warming in southeastern Amazonia. Philos. Trans. R. Soc. B Biol. Sci. 368, 20120153–20120153. https://doi.org/10.1098/rstb.2012.0153

Macedo, M.N., DeFries, R.S., Morton, D.C., Stickler, C.M., Galford, G.L., Shimabukuro, Y.E., 2012. Decoupling of deforestation and soy production in the southern Amazon during the late 2000s. Proc. Natl. Acad. Sci. U. S. A. 109, 1341–1346. https://doi.org/10.1073/pnas.1111374109

Margulis, S., 2004. Causes of deforestation of the Brazilian Amazon. World Bank Work. Pap. 1–77. https://doi.org/10.1596/0-8213-5691-7

Martinelli, L., Batistella, M., Silva, R., Moran, E., 2017. Soy Expansion and Socioeconomic Development in Municipalities of Brazil. Land 6, 62. https://doi.org/10.3390/land6030062

Masek, J.G., Vermote, E.F., Saleous, N., Wolfe, R., Hall, F.G., Huemmrich, F., Gao, F., Kutler, J., T.K. Lim, A., 2006. A Landsat Surface Reflectance Dataset for North America, 1990–2000. Geosci. Remote Sens. Lett. 3, 68–72.

Mattiuzzi, M., Hengl, T., Verbesselt, J., Lobo, A., Klisch, A., Evans, B., Stevens, F., Mosher, S., Maspons, J., Detsch, F., Hufkens, K., 2016. MODIS: MODIS acquisition and processing package.

Maxwell, A.E., Warner, T.A., Fang, F., 2018. Implementation of machine-learning classification in remote sensing: An applied review. Int. J. Remote Sens. https://doi.org/10.1080/01431161.2018.1433343

Meyfroidt, P., Roy Chowdhury, R., de Bremond, A., Ellis, E.C., Erb, K.H., Filatova, T., Garrett, R.D., Grove, J.M., Heinimann, A., Kuemmerle, T., Kull, C.A., Lambin, E.F., Landon, Y., le Polain de Waroux, Y., Messerli, P., Müller, D., Nielsen, J., Peterson, G.D., Rodriguez García, V., Schlüter, M., Turner, B.L., Verburg, P.H., 2018. Middle-range theories of land system change. Glob. Environ. Chang. 53, 52–67. https://doi.org/10.1016/j.gloenvcha.2018.08.006

Müller, H., Rufin, P., Griffiths, P., Barros Siqueira, A.J., Hostert, P., 2015. Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape. Remote Sens. Environ. 156, 490–499. https://doi.org/10.1016/j.rse.2014.10.014

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–8. https://doi.org/10.1038/35002501

Nepstad, D.C., Stickler, C.M., Almeida, O.T., 2006. Globalization of the Amazon soy and beef industries: Opportunities for conservation. Conserv. Biol. https://doi.org/10.1111/j.1523-1739.2006.00510.x

Nolte, C., le Polain de Waroux, Y., Munger, J., Reis, T.N.P., Lambin, E.F., 2017. Conditions influencing the adoption of effective anti-deforestation policies in South America's commodity frontiers. Glob. Environ. Chang. 43, 1–14. https://doi.org/10.1016/j.gloenvcha.2017.01.001

Oliveira, P., Marquis, R., 2002. The cerrados of Brazil, The cerrados …. https://doi.org/10.1663/0013-0001(2003)057[0656:DFABRE]2.0.CO;2

Parente, L., Ferreira, L., 2018. Assessing the spatial and occupation dynamics of the Brazilian pasturelands based on the automated classification of MODIS images from 2000 to 2016. Remote Sens. 10. https://doi.org/10.3390/rs10040606

Peña-Barragán, J.M., Ngugi, M.K., Plant, R.E., Six, J., 2011. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. Remote Sens. Environ. 115, 1301–1316. https://doi.org/10.1016/j.rse.2011.01.009

Pendrill, F., Persson, U.M., 2017. Combining global land cover datasets to quantify agricultural expansion into forests in Latin America: limitations and challenges. PLoS One. https://doi.org/10.1371/journal.pone.0181202

Pontius, R.G., Millones, M., Pontius, Robert, Gilmore, J., Millones, M., Pontius, R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int. J. Remote Sens. 32, 4407–4429. https://doi.org/10.1080/01431161.2011.552923

Power, A.G., 2010. Ecosystem services and agriculture: tradeoffs and synergies. Philos. Trans. R. Soc. B Biol. Sci. 365, 2959–2971. https://doi.org/10.1098/rstb.2010.0143

Puyravaud, J.P., 2003. Standardizing the calculation of the annual rate of deforestation. For. Ecol. Manage. 177, 593–596. https://doi.org/10.1016/S0378-1127(02)00335-3

R Development Core Team, R., 2011. R: A Language and Environment for Statistical Computing. R Found. Stat. Comput., R Foundation for Statistical Computing. https://doi.org/10.1007/978-3-540-74686-7

Ratana, P., Huete, A.R., Ferreira, L., 2005. Analysis of cerrado physiognomies and conversion in the MODIS seasonal-temporal domain. Earth Interact. 9. https://doi.org/10.1175/1087-3562(2005)009<0001:AOCPAC>2.0.CO;2

Richards, P., 2015. What Drives Indirect Land Use Change? How Brazil's Agriculture Sector Influences Frontier Deforestation. Ann. Assoc. Am. Geogr. 105, 1026–1040. https://doi.org/10.1080/00045608.2015.1060924

Richards, P.D., Myers, R.J., Swinton, S.M., Walker, R.T., 2012. Exchange rates, soybean supply response, and deforestation in South America. Glob. Environ. Chang. 22, 454–462. https://doi.org/10.1016/j.gloenvcha.2012.01.004

Rizzi, R., Risso, J., Epiphanio, R.D. V, Rudorff, B.F.T., Formaggio, A.R., Shimabukuro, Y.E., Fernandes, S.L., 2009. Estimativa da área de soja no Mato Grosso por meio de imagens MODIS, in: Simpósio Brasileiro de Sensoriamento Remoto. INPE, São José dos Campos, pp. 387–394.

Rudel, T.K., Schneider, L., Uriarte, M., Turner, B.L., DeFries, R., Lawrence, D., Geoghegan, J., Hecht, S., Ickowitz, A., Lambin, E.F., Birkenholtz, T., Baptista, S., Grau, R., 2009. Agricultural intensification and changes in cultivated areas, 1970-2005. Proc. Natl. Acad. Sci. 106, 20675–20680. https://doi.org/10.1073/pnas.0812540106

Sawakuchi, H.O.., Ballester, M.V.R., Ferreira, M.E., 2013. The Role of Physical and Political Factors on the Conservation of Native Vegetation in the Brazilian Forest-Savanna Ecotone. Open J. For. 03, 49–56. https://doi.org/10.4236/ojf.2013.31008

Scaramuzza, C.A. de M., Sano, E.E., Adami, M., Bolfe, E.L., Coutinho, A.C., César, J., Mora, D., Eduardo, L., Maurano, P., Narvaes, S., José, F., Oliveira, B. de, Rosa, R., Barbosa, E., Valeriano, D. de M., Victoria, D. de C., Bayma, P., Oliveira, G.H. de, Bayma-Silva, G., 2017. LAND-USE AND LAND-COVER MAPPING OF THE BRAZILIAN CERRADO BASED MAINLY ON LANDSAT-8 SATELLITE IMAGES. Rev. Bras. Cartogr. a (2017), Edição Fotogrametia e Sensoriamento Remoto 69, 1041–1051.

Scaramuzza, P., Micijevic, E., Chander, G., 2004. SCL Gap-Filled Products. Phase One Methodology. USGS-United States Geol. Surv.

Searle, K.R., Gordon, I.J., Stokes, C.J., 2009. Hysteretic Responses to Grazing in a Semiarid Rangeland. Rangel. Ecol. Manag. 62, 136–144. https://doi.org/10.2111/08-200.1

Soares-Filho, B., Rajao, R., Macedo, M., Carneiro, A., Costa, W., Coe, M., Rodrigues, H., Alencar, A., 2014. Cracking Brazil's Forest Code. Science (80-.). 344, 363–364. https://doi.org/10.1126/science.1246663

Sparovek, G., Berndes, G., Klug, I.L.F., Barretto, A.G.O.P., 2010. Brazilian agriculture and environmental legislation: Status and future challenges. Environ. Sci. Technol. 44, 6046–6053. https://doi.org/10.1021/es1007824

Tewkesbury, A.P., Comber, A.J., Tate, N.J., Lamb, A., Fisher, P.F., 2015. A critical synthesis of remotely sensed optical image change detection techniques. Remote Sens. Environ. 160, 1–14. https://doi.org/10.1016/j.rse.2015.01.006

Toniol, A.C., Galvão, L.S., Ponzoni, F.J., Sano, E.E., Amore, D. de J., 2017. Potential of hyperspectral metrics and classifiers for mapping Brazilian savannas in the rainy and dry seasons. Remote Sens. Appl. Soc. Environ. 8, 20–29. https://doi.org/10.1016/j.rsase.2017.07.004

Turner, B.L., Lambin, E.F., Reenberg, A., 2007. The emergence of land change science for global environmental change and sustainability. Proc. Natl. Acad. Sci. 104, 20666–20671. https://doi.org/10.1073/pnas.0704119104

Velasquez, C.; Alves, H. Q.; Bernasconi, P., 2010. FIQUE POR DENTRO: a Bacia do Rio Xingu em Mato Grosso. nstituto Socioambiental, Instituto Centro de Vida, São Paulo.

Vuolo, F., Mattiuzzi, M., Klisch, A., Atzberger, C., 2012. Data service platform for MODIS Vegetation Indices time series processing at BOKU Vienna: current status and future perspectives, in: Michel, U., Civco, D.L., Ehlers, M., Schulz, K., Nikolakopoulos, K.G., Habib, S., Messinger, D., Maltese, A. (Eds.), . p. 85380A. https://doi.org/10.1117/12.974857

Walker, W.S., Stickler, C.M., Kellndorfer, J.M., Kirsch, K.M., Nepstad, D.C., 2010. Large-Area Classification and Mapping of Forest and Land Cover in the Brazilian Amazon: A Comparative Analysis of ALOS/PALSAR and Landsat Data Sources. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 3, 594–604. https://doi.org/10.1109/JSTARS.2010.2076398

Wang, X., Dong, X., Liu, H., Wei, H., Fan, W., Lu, N., Xu, Z., Ren, J., Xing, K., 2017. Linking land use change, ecosystem services and human well-being: A case study of the Manas River Basin of Xinjiang, China. Ecosyst. Serv. 27, 113–123. https://doi.org/10.1016/j.ecoser.2017.08.013

Yengoh, G.T., Dent, D., Olsson, L., Tengberg, A.E., Tucker III, C.J., 2015. Use of the Normalized Difference Vegetation Index (NDVI) to Assess Land Degradation at Multiple Scales, SpringerBriefs in Environmental Science. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-24112-8

Yu, W., Zhou, W., Qian, Y., Yan, J., 2016. A new approach for land cover classification and change analysis: Integrating backdating and an object-based method. Remote Sens. Environ. 177, 37–47. https://doi.org/10.1016/j.rse.2016.02.030

Zhu, X.X., Tuia, D., Mou, L., Xia, G.S., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. IEEE Geosci. Remote Sens. Mag. https://doi.org/10.1109/MGRS.2017.2762307

3. DEFORESTATION AND DEGRADATION IN AN ECOTONE ZONE BETWEEN AMAZON AND CERRADO BIOMES: COMPARING INDICATORS

ABSTRACT

Both deforestation and forest degradation are important processes shaping biodiversity and ecosystem service patterns in the tropics. Nevertheless, few datasets and automated models have been built to improve both deforestation and forest degradation, and each of the existent dataset capture different patterns of these processes. In this chapter, we evaluate how well three popular forest loss datasets match, and how they are affected by predominant vegetation type. We calculated quantity and allocation disagreement, as well as the Jaccard's index, for a modified version of the Global Forest Change (GFC) dataset, as well as for the output of the highly automated software CLASlite, and both INPE`s PRODES and DEGRAD programs. We found that the area mapped under deforestation by CLASlite is the largest when compared to the modified GFC or PRODES. The largest degraded area was captured by DEGRAD. Transforming these datasets from raw data into hotspots did not improve the similarity among them. We observed that disagreement was mostly due to misallocation, while the degree of agreement was related to vegetation type. Regions dominated by cerrado forest and semi-deciduous forests presented a higher level of disagreement while evergreen forests presented the lowest. We argue that researchers and decision-makers should pay careful attention when choosing forest loss indicators to represent a process, and clearly justify their choice.

Keywords: Vegetation type; Global forest change; CLASlite; PRODES; DEGRAD

3.1. Introduction

Vegetation loss monitoring through remotely sensed images has been conducted by research institutes and governmental entities since the 1980's. In Brazil, a well-known example is PRODES (Amazon Deforestation Monitoring Project) which estimates deforested areas in the Legal Amazon annually since the 1990's based on Landsat images (KINTISCH, 2007; INPE, 2019). More recently, significant advancements have been made in vegetation loss detection, mainly by including degradation as a process of forest loss along with deforestation, improving the minimum mapped unit, and extending detection beyond pristine dense forest areas. Two products which show the advances in monitoring loss of native vegetation are the Carnegie Landsat Analysis System – Lite (CLASlite), which maps canopy degradation and deforestation (Alves et al., 2009; Asner et al., 2009), and the Tree Cover Loss mapping under the Global Forest Change model (GLAD, 2018; Hansen et al., 2013), which maps the percentage of tree loss. Still, performance comparisons among these different indicators of vegetation loss are not common in the literature and this lack of performance analysis in diverse highly landscapes represent a noticeable research gap. Diverse landscapes are areas that contain a variety of physiognomies such as ecotone zones. Due to its diversity in biophysical characteristics and ecological relevance, ecotone zones can be used to demonstrate differences among indicators of deforestation and degradation.

Choosing a dataset or automated model to map deforestation and degradation is not a simple task, and is made more difficult by the complexity of different available methodologies and many nuances of any given research project. Michaelsen and colleagues (2017) compared four deforestation datasets modelled with Automated Monte Carlo Unmixing (CLASlite), Tasseled Cap (ERDAS), Bhattacharya classification (SPRING), and Spectral Angle Mapper (ENVI). The authors found a difference between two groups of techniques, CLASlite/Tasseled Cap and

Bhattacharya/Spectral Angle Mapper, which was attributed to the limited ability of CLASlite/Tasseled Cap to recognize mixed forests under hydric stress. Hansen and colleagues (2008) compared deforestation mapped by PRODES and MODIS Land Products. They concluded that because PRODES does not consider non-pristine areas nor vegetation types other than rain forest, MODIS captured 120% more deforestation than the official PRODES estimate, even though MODIS spatial resolution is lower than the Landsat image used by PRODES. Global Forest Change, CLASlite and classification methodologies such as random forest were compared to map deforestation in Indonesia at a local scale, with Global Forest Change and CLASlite presenting higher degrees of accuracy for monitoring vegetation cover change (Arjasakusuma et al., 2018).

Inaccurate vegetation loss mapping can also have impacts on models and decision-making processes which use these types of data as an input. For example, most predictive models of land use change in the Amazon present higher degrees of uncertainty for predictions in the arc of deforestation, which is also an ecotone zone between the amazon forest and the Cerrado, the Brazilian savannas (Rosa et al., 2014). Most of these models use deforestation datasets generated by PRODES, which do not cover the entirety of the biome as suggested by Hassen (2008). Consequently, a broader dataset could be beneficial for improving vegetation loss modelling. In this study, we tested whether three widely used datasets would produce different indications of native vegetation loss - including both degradation and clearcutting deforestation. We analyzed a modified version of the Global Forest Change data (Hansen et al., 2013), PRODES and DEGRAD datasets, and the output of CLASlite models (Asner et al., 2009) for both degradation and clearcutting dynamics by comparing the raw data, the calculated density kernel, and significant hotspot occurance. Our hypothesis was that the three indicators would produce different results due their inherent differences in concepts and methods.

3.2. Data

Two major sources of spatial information on global vegetation loss are the CLASlite software and the Global Forest Change yearly model. Both products offer the possibility to easily monitor deforestation and forest degradation. Aside from using different methodologies to generate their products, both products map forest change based on temporal profiles of spectral metrics processed at the same spatial scale using Landsat images and subpixel information. Another source of data on vegetation loss exclusively for Brazilian biomes is the Monitoramento do Desmatamento da Floresta Amazônica Brasileira por Satélite (PRODES) and the Mapeamento da Degradação Florestal na Amazônia Brasileira (DEGRAD). PRODES produces yearly reports on deforestation, while DEGRAD produced yearly reports on vegetation degradation (project discontinued in 2016 and replaced by DETER). Both mainly rely on visual interpretation of Landsat, CBERS, and IRS-2 data. We present a full comparative table in Appendix B - Supplementary Material 1.

CLASlite is a highly automated and stand-alone image processing software for converting raw satellite imagery into deforestation and forest degradation indicators (Asner et al., 2019, 2009). It is built based on the Automated Monte Carlo Unmixing approach, calculating fractional cover of vegetation canopies, dead vegetation, and bare surface. According to differences in estimated fractional covers between time steps, the software classifies pixels as either deforestation, degradation, or no change. The software can be used to analyze Landsat, Modis, and Sentinel images. CLASlite has been used in a variety of studies to map vegetation loss, including peatlands in Indonesia (Carlson et al., 2012), bamboo-dominated forests in the Amazon (Carvalho et al., 2013), rain forests in Madagascar (Allnutt et al., 2013) and in the Brazilian Amazon (Alves et al., 2009; Hasan et al., 2017), different physiognomies in Belize (Chicas et al., 2016), savanna woodlands in Swaziland (Dlamini, 2017), the entire country of

Madagascar (Yesuf et al., 2019), and even to map urban green cover (Kanniah, 2017) and island areas in the Funafuti Atoll - southwestern Pacific Ocean (Hisabayashi et al., 2018).

The Global Forest Change (hereafter referred to as GFC) dataset is an end-product developed by the University of Maryland (HANSEN et al., 2013; GLAD, 2018). It is calculated based on per-band metrics, using bands 3, 4, 5, and 7 of the LANDSAT satellites, starting in 2000, which are processed by learning algorithms in decision trees. GFC is available in platforms such as Google Earth Engine and makes available information on forest loss, forest gain, and year of gain or loss. Although the model methodology is not based on a complete unmixed approach as in CLASlite, the GFC model also applies subpixel detection. GFC also has been used in a variety of studies, including the detection of degraded forests in the state of Mato Grosso, Brazil (Shimabukuro et al., 2017), identifying global forest loss drivers (Curtis et al., 2018), estimating primary loss in rain forests in Brazil, Democratic Republic of the Congo, and Indonesia (Turubanova et al., 2018), and quantifying small scale forest degradation in the(?) Congo (Tyukavina et al., 2018).

3.3. Material and methods

3.3.1. Study area

We tested both datasets in the Upper Xingu River Basin (UXRB), Mato Grosso State, Brazil (Figure 9), which contains around 170,000 km2 and drains into the Amazon River. The study area is located in the ecotone zone in-between the Amazon Rain Forest and Brazilian Savannas - the Cerrado, thus, it is formed by a diverse landscape, from dense forests to grasslands (Garcia et al. 2019). From the 1980s to the year 2000, timber and pasture for cattle raising were the main land use activities. Starting around 2000, grain production became an increasingly prominent activity in the region (FAOSTAT, 2016; IBGE, 2019).

3.3.2. Image processing

In order to model change through CLASlite from 2010 to 2015, we processed images sensed by Landsat 8 and Landsat 5, made available by the USGS at the Landsat Higher Level Science Data products (http://espa.cr.usgs.gov/). On this platform, the USGS delivers images which are geometrically and radiometrically corrected by a homogenous correction chain (Masek et al., 2006). Still, in order to guarantee a pixel-by-pixel overlap between years and datasets, we registered all images to a Landsat 8 image sensed in 2015. As recommended by the CLASlite team (Asner et al., 2019), we mosaicked the images from the same year into one image to compose the whole study area. We matched the images' histogram when necessary, based on central path overlapping areas. This technique changes each of the pixel's values according to the target histogram but does not change the overall histogram shape.

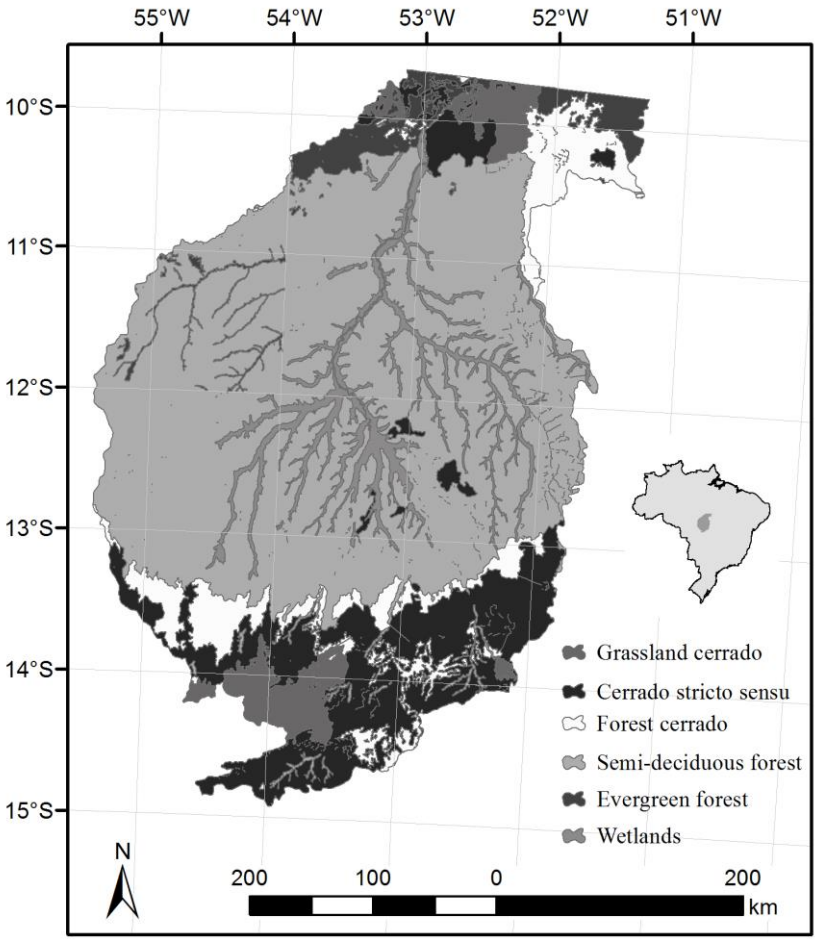


Figure 9. Map showing the Upper Xingu River Basin and the phytophysiognomies found in the basin.

We used the built-in tools in CLASlite to calibrate and transform raw data into surface reflectance values. Then, we implemented the Automated Monte Carlo Unmixing approach (AutoMCU) to model the surface reflectance into subpixel indicators of bare substrate, photosynthetic vegetation, and non-photosynthetic vegetation – all expressed as a percentage (0 - 100%) along with uncertainty estimates. The AutoMCU relies on reference spectra for pure samples of all indicators, which through linear equations calculates the likely percentage of each indicator in a pixel. This approach was first developed for savanna physiognomies, but then was redesigned for tropical forests (Alves et al., 2009; Asner, 1998). Lastly, we compared all mapped years to detect forest change. We used Equation 1 (deforestation) and 2 (degradation) to classify changes as either deforestation or degradation. Additionally, we did not consider any pixels contaminated by shadow or clouds, and removed isolated pixels which were classified as changed. We used land use and land cover maps made available by Garcia and colleagues (2019), along with visual interpretation of Landsat images, to verify CLASlite models output.

$$\text{Deforestation} = (((PV1 - PV2) \geq 20) \text{ OR } ((S1 \leq 5)\ AND\ ((S2 - S1) \geq 15)))\ OR\ ((PV2 < 80)\ AND\ ((NPV2 - NPV1) \geq 20))) \quad \text{Eq (1)}$$

$$\text{Disturbance} = ((((NPV2 - NPV1) \geq 10)\ AND\ ((PV1 - PV2) > 10)) \text{ OR } ((S1 \leq 5)\ AND\ ((S2 - S1) > 10)\ AND\ (S2 \leq 15))) \quad \text{Eq (2)}$$

where, PV is the photosynthetic vegetation fraction, NPV is the non-photosynthetic vegetation fraction, S is the bare substrate fraction, numbers 1 and 2 represent the first and second time series images, respectively.

Both PRODES and DEGRAD datasets were obtained at the TerraBrasilis online platform developed by INPE (http://terrabrasilis.dpi.inpe.br/). We obtained datasets for the Amazon Rain Forest and Cerrado biome in shapefile format, and merged both biomes to build a unique dataset for each process (deforestation | degradation) in each year. Then, we merged the yearly datasets in five-year composition (2010-2015). Differently from CLASlite and GFC, both PRODES and DEGRAD products are identified by visual interpretation of satellite images (mainly Landsat), restricted to analysis of pristine areas, and presenting a minimum mapped area of 6.25 ha.

We obtained GFC from the Google Earth Engine Platform (earthengine.google.com) from 2010 to 2015. The GFC does not differentiate deforestation from degradation. Instead, it retrieves tree loss in a certain year by comparing it with a previous time-step up to the base map produced for the year 2000. Forest cover is considered vegetation 5 m tall or higher (Hansen et al., 2013). We have also used land use and land cover data made available by Garcia and colleagues (2019) to reclassify forest loss as either deforestation or degradation. We used the Level 1 of a classification dataset, which separates (semi) native vegetation from managed areas (overall accuracy of ~ 93%). In other words, when forest loss overlapped with managed areas, we observed that the result of deforestation was a clear-cut area turned into a managed field, and classified these pixels as deforestation. In all other cases these areas were classified as degraded.

Both GFC and CLASlite outputs, as well as the available land use dataset, present a 30 m pixel resolution. PRODES/DEGRAD dataset were transformed into raster and resampled to the same resolution as the other datasets. Every layer used was tested for overlapping with Landsat 8 images from 2015 obtained from USGS/Landsat Higher Level Science Data products (http://espa.cr.usgs.gov/). All layers presented a spatial overlap error value (Root Mean Square Error) lower than a half pixel resolution (15 m).

3.3.3. Comparisons

Often researches do not use raw forest change data to draw conclusions but instead to transform raw data into, for example, hotspot maps (Finer et al., 2018). We compared the datasets in three different popular formats: raw data, kernel density estimation, and hotspot estimation (Figure 10). Thus, we have a spatially explicit dataset for degradation and deforestation, derived from both CLASlite, PRODES/DEGRAD and GFC, and in three different formats. We considered raw data the CLASlite model output, PRODES and DEGRAD products, and the product of the overlapping operation between GFC dataset with land use information - in total six layers. Thus, each layer presents two categories: (1) change, and (2) no change. Kernel density is a surface representing a magnitude-per-unit area. We modeled kernel density by searching for either deforested or degraded pixels in a 3 km radius, allocating the density (occurrences per km2) in a central 30 m2 cell of this search radius. The result is a continuous surface representing the density of change during a time period.

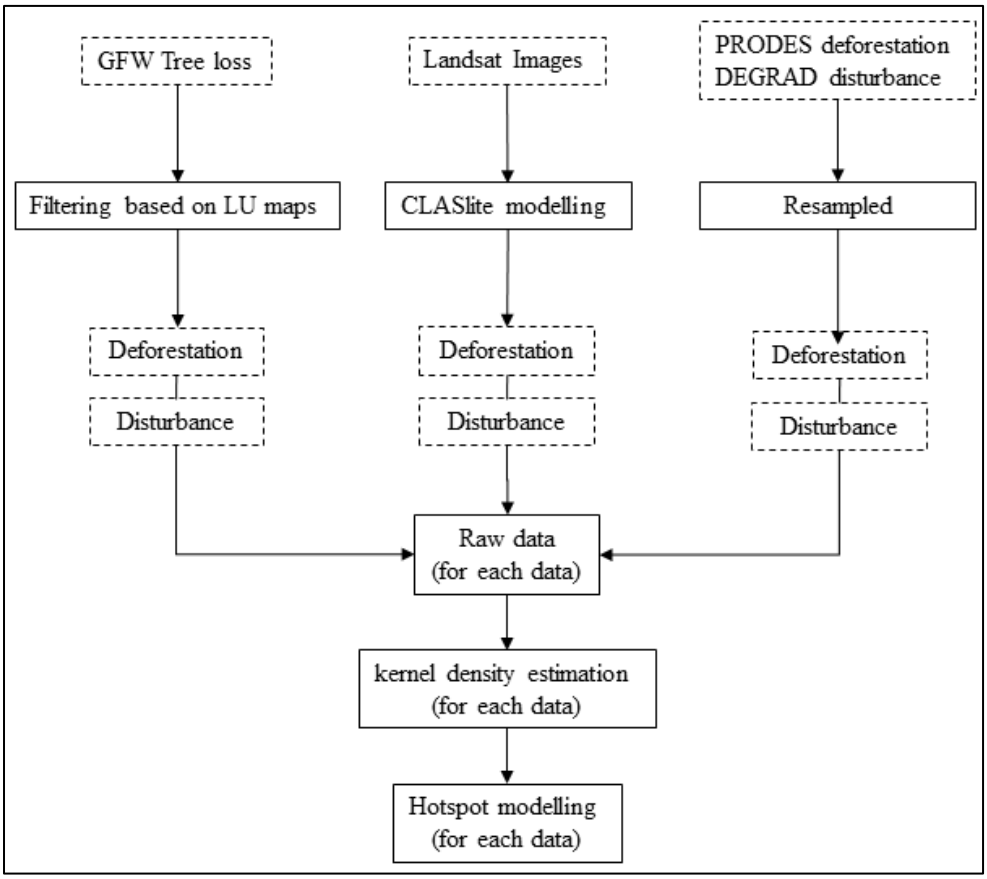


Figure 10. Methodological flow chart. Dashed boxes represent products and solid outlined boxes represent a process/analysis.

We used a fishnet with 1 km2 cell to model hotspots of deforestation and degradation. We counted the number of pixels which were either disturbed or deforested in a radius of 3 km and allocated the value in a 1 km2 cell. Then, we applied Spatial Getis-Ord Gi* statistic (Getis, 1992; Ord and Getis, 1995) together with Bonferroni and False Discovery Rate correction, obtaining a z-score and p-value for each cell in the fishnet. We selected the right tail of the z-score distribution as hotspots for deforestation and degradation, representing confidence levels equal to or larger than 90% (z-score > 1.65). The Gi* statistic is used to analyze local patterns in spatial data, measuring the degree of association between a certain value and its neighboring values, up to a defined distance (Getis & Ord, 1992). We reclassified each analyzed layer according either deforestation or degradation as hotspot (identified with at least 90% of confidence) and no hotspot.

It is important to point out that we choose a 3 km search radius to estimate density kernels and hotspots based on regional farm profiles and tests with larger radii. We assumed that deforestation and degradation, when induced by man, would be the result of a decision made at the property level. Even though property size varies so much that one cannot be consistent using a specific radius, we used this as a starting point. According to information made available by the Rural Environmental Registry – CAR (Roitman et al., 2018), farms in the region present a mean area of 9.2k m2 ± 31km2 and median equal to 2.4 km2. Thus, the search radius representing the variance of regional farm size would be ~1.7 km ± 3.14km. We observed peaks of clustering for radius values of 3 and 5 km, with clustering decreasing for 10 km radius.

We calculated the quantity, exchange and shift components of difference, as well as the overall difference, between three data layers (GFL, PRODES/DEGRAD and CLASlite) representing either deforestation or

degradation. We calculated the difference between layers based on quantity and allocation disagreement as proposed by Pontius and Santacruz (2014). Quantity difference is the amount of difference between two maps regarding proportions of the categories. Exchange is part of allocation disagreement and represents the transition from category i to category j in some pixels and the opposite (from j to i) in an identical number of other pixels. Shift is also part of allocation disagreement when categories are greater than 2, and it consists of the remaining difference in allocation disagreement which is not due to exchange. Overall difference is the sum of the different difference types.

We also used the Jaccard's index to gain insight into the similarity of the raw dataset and hotspot dataset (Bonham-Carter, 1994). Jaccard's index is calculated for binary data as the intersection of two layers for a class of interest (in this case, change) divided by the union of the compared layers. Its value ranges from 0 (complete dissimilarity) to 1 (complete similarity – when both data being compared present full overlapping). We used correlation analysis to compare kernel density layers. Correlation is not a similarity index per se. Still, correlation analysis between two kernel density maps allows for a comparison of the spatial association strength (Bonham-Carter, 1994).

In addition to calculating disagreement, similarity, and correlation for the whole study area, we also calculated these metrics by vegetation physiognomy to compare how vegetation density can contribute to (dis)similarities. Vegetation types present in the UXRB are grassland cerrado (open grassland with few or no shrubs), cerrado stricto sensu (typical cerrado with tree cover between 20 – 70% and up to 8m high), forest cerrado (woody savanna dominated by trees up to 15 m high), semi-deciduous forest (trees up to 25 m high which lose their leaves during the dry season), and evergreen forest (rain forest) (IBGE, 2012; Ribeiro and Walter, 2019).

3.4. Results

CLASlite modelling approach detected the largest area of deforestation, 103% more than the Global Forest Change dataset (GFC) and 76% more than PRODES/INPE (Table 5). When calculating hotspots of deforestation, the difference between datasets decrease - CLASlite detected 11% more than GFC and 13% more than PRODES/INPE. The difference among datasets when analyzing degradation was even larger, with DEGRAD/INPE detecting the largest area. It represented 481% (63% comparing hotspots) more than the area detected by CLASlite and 260% (61% comparing hotspots) more than the area detected by GFC. Nevertheless, INPE`s dataset presented the lowest correlation with other calculated density kernels for both deforestation and degradation (Table 6).

The Jaccard`s index and the disagreement metrics are complementary metrics to understand differences among datasets. Jaccard's index does not consider areas not undergoing any change and, thus, it is a straight-forward measure of overlapping between datasets. The disagreement metrics, on the other hand, break the disagreement into subtypes and are more informative. As expected, the disagreement and Jaccard's index showed a larger disagreement for these datasets which detected a much larger area of either deforestation or degradation. Still, contrary to our expectation, we observed a smaller level of agreement among the datasets when analyzing hotspots than when analyzing the raw datasets. This indicates that hotspot modelling may play a role in extrapolating inconsistencies among datasets (Table 5). We show how datasets overlap for raw data in figure 11 and 12, and for hotspots in figures 13 and 14. In our comparison, disagreement was more due to misallocation of the categories than to quantities of the categories. A marked exception was observed when comparing DEGRAD/INPE with other degradation datasets, which presented a much larger disagreement due to the allocated quantities.

Table 5. Quantity disagreement, allocation disagreement (exchange and shift), overall disagreement, Jaccard's index of agreement calculated for GFC, PRODES/DEGRAD (INPE), and CLASlite outputs, and the size of detected deforestation and degradation processes mapped in the Upper Xingu River Basin, Brazil. Analysis are presented for raw data, showing hotspots mapped with at least 95% of confidence.

DEFORESTATION Raw data							
	Disagreement				**Agreement**	**Area (km²)**	
	Quantity	**Exchange**	**Shift**	**Overall**	**Jaccard**	**Data 1**	**Data 2**
CL - GFW	1.19	1.29	0.00	2.49	0.98	3,986	1,960
CL - INPE	1.02	1.69	0.00	2.71	0.97	3,986	2,264
GFW - INPE	0.18	1.54	0.00	1.72	0.99	1,960	2,264
DEGRADATION Raw data							
	Disagreement				**Agreement**	**Area (km²)**	
	Quantity	**Exchange**	**Shift**	**Overall**	**Jaccard**	**Data 1**	**Data 2**
CL - GFW	0.67	1.81	0.00	2.48	0.98	1,858	2,995
CL - INPE	5.27	1.36	0.00	6.63	0.93	1,858	10,798
GFW - INPE	4.60	1.72	0.00	6.32	0.94	2,995	10,798
DEFORESTATION Hotspot							
	Disagreement				**Agreement**	**Area (km²)**	
	Quantity	**Exchange**	**Shift**	**Overall**	**Jaccard**	**Data 1**	**Data 2**
CL - GFW	0.75	5.46	0.00	6.21	0.94	12,798	11,528
CL - INPE	0.89	6.42	0.00	7.31	0.92	12,798	11,285
GFW - INPE	0.14	7.22	0.00	7.36	0.92	11,528	11,285
DEGRADATION Hotspot							
	Disagreement				**Agreement**	**Area (km²)**	
	Quantity	**Exchange**	**Shift**	**Overall**	**Jaccard**	**Data 1**	**Data 2**
CL - GFW	0.25	6.96	0.00	7.22	0.92	13,776	13,347
CL - INPE	5.15	5.62	0.00	10.76	0.89	13,776	22,504
GFW - INPE	5.40	5.38	0.00	10.78	0.89	13,347	22,504

Table 6. Correlation between density kernels derived from GFC, PRODESDEGRAD (INPE) and CLASlite outputs for the Upper Xingu River Basin, Brazil.

DEFORESTATION Kernel density		
	Correlation	**Significance**
CL - GFW	0.60	$p < 0.05$
CL - INPE	0.52	$p < 0.05$
GFW - INPE	0.52	$p < 0.05$
DEGRADATION Kernel density		
	Correlation	**Significance**
CL - GFW	0.57	$p < 0.05$
CL - INPE	0.48	$p < 0.05$
GFW - INPE	0.49	$p < 0.05$

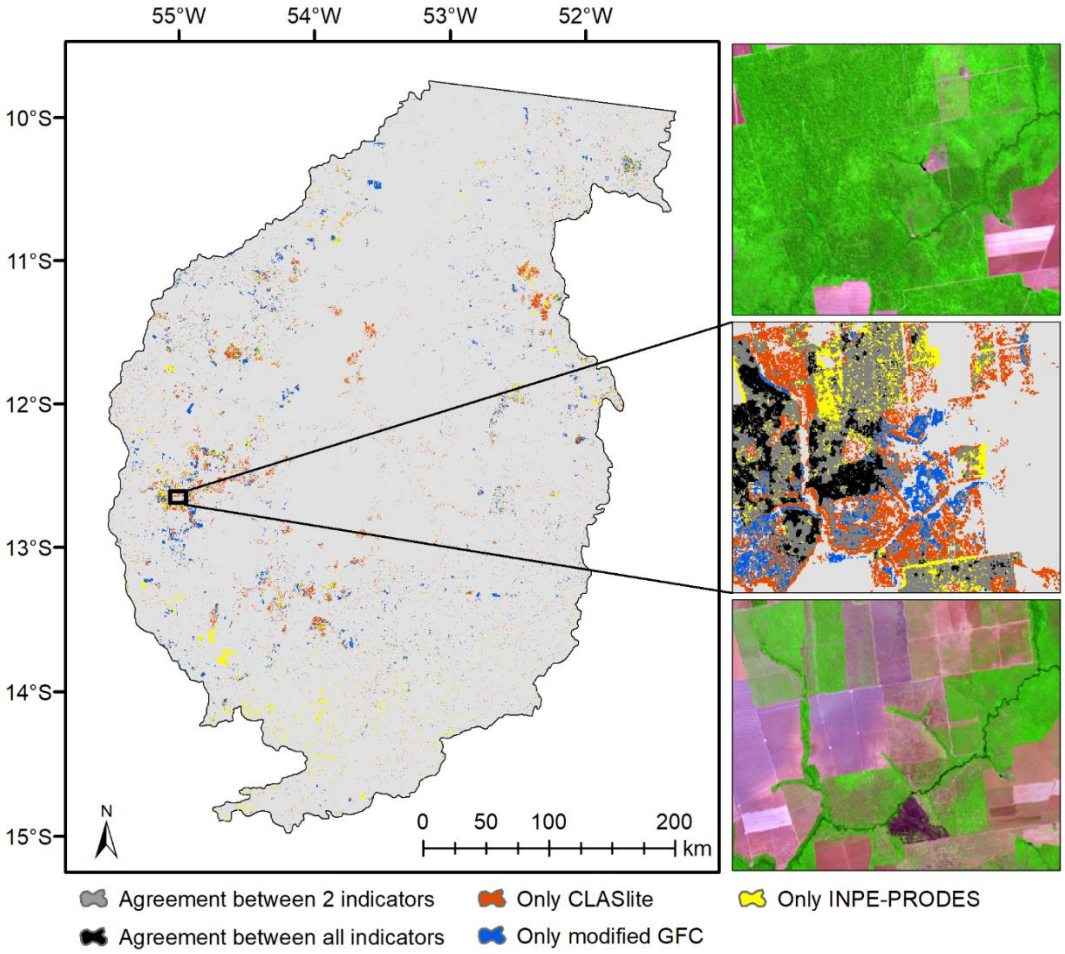


Figure 11. Map showing intersection of GFC, INPE-PRODES, and CLASlite output for deforestation in the Upper Xingu River Basin between 2010 and 2015. Data is shown in raw format. Images from 2010 (Landsat 7 ETM+) and 2015 (Landsat 8 OPI) are shown in false-color composition RGB 543 and RGB 654, respectively.

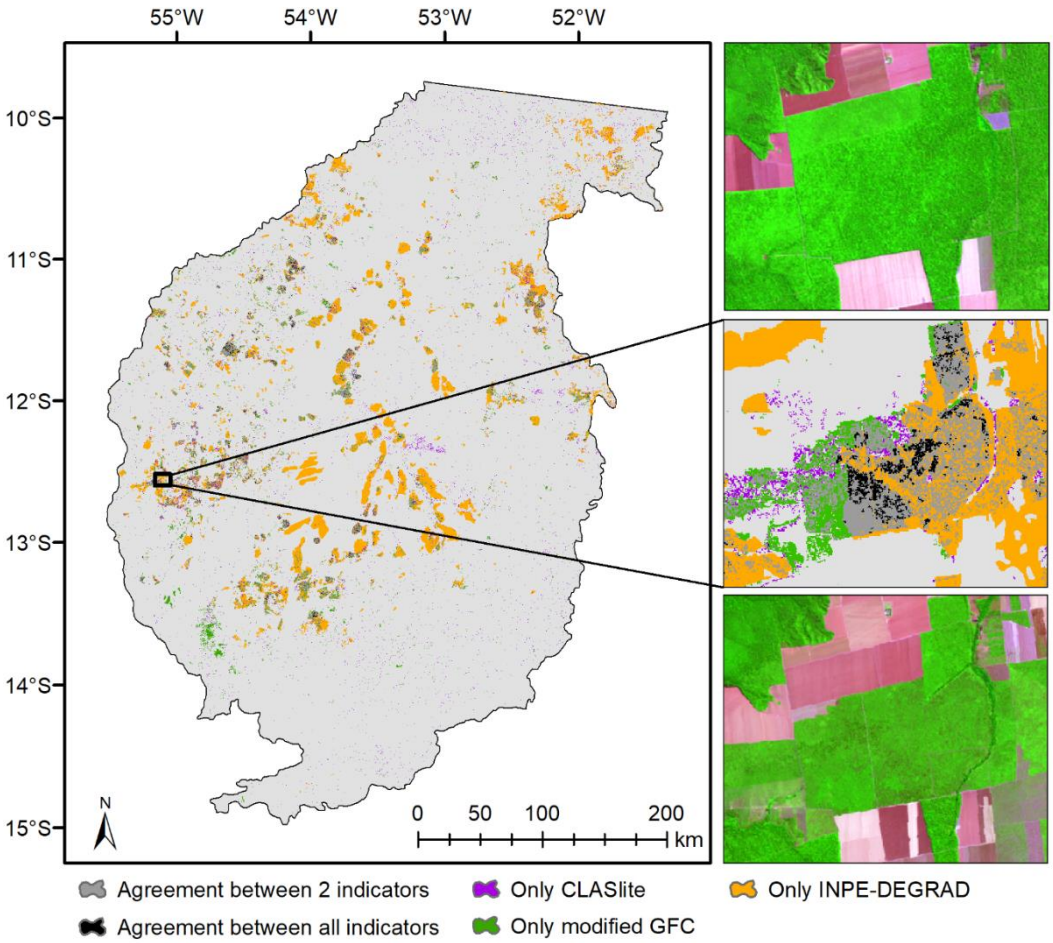


Figure 12. Map showing intersection of GFC, INPE-PRODES, and CLASlite output for degradation in the Upper Xingu River Basin between 2010 and 2015. Data is shown in raw format. Images from 2010 (Landsat 7 ETM+) and 2015 (Landsat 8 OPI) are show in false-color composition RGB 543 and RGB 654, respectively.

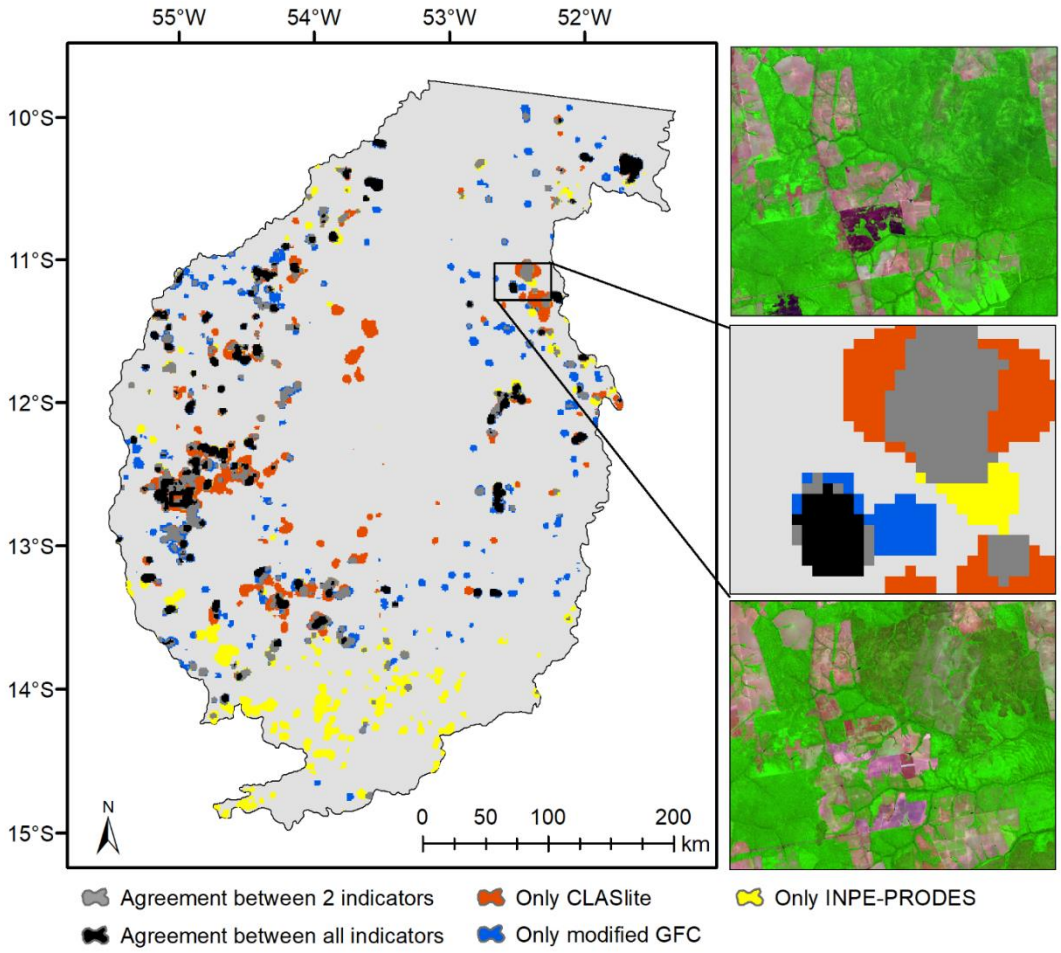


Figure 13. Map showing intersection of GFC, INPE-PRODES, and CLASlite output for deforestation in the Upper Xingu River Basin between 2010 and 2015. Data is shown as hotspot areas with significance ≥ 0.95. Images from 2010 (Landsat 7 ETM+) and 2015 (Landsat 8 OPI) are show in false-color composition RGB 543 and RGB 654, respectively.

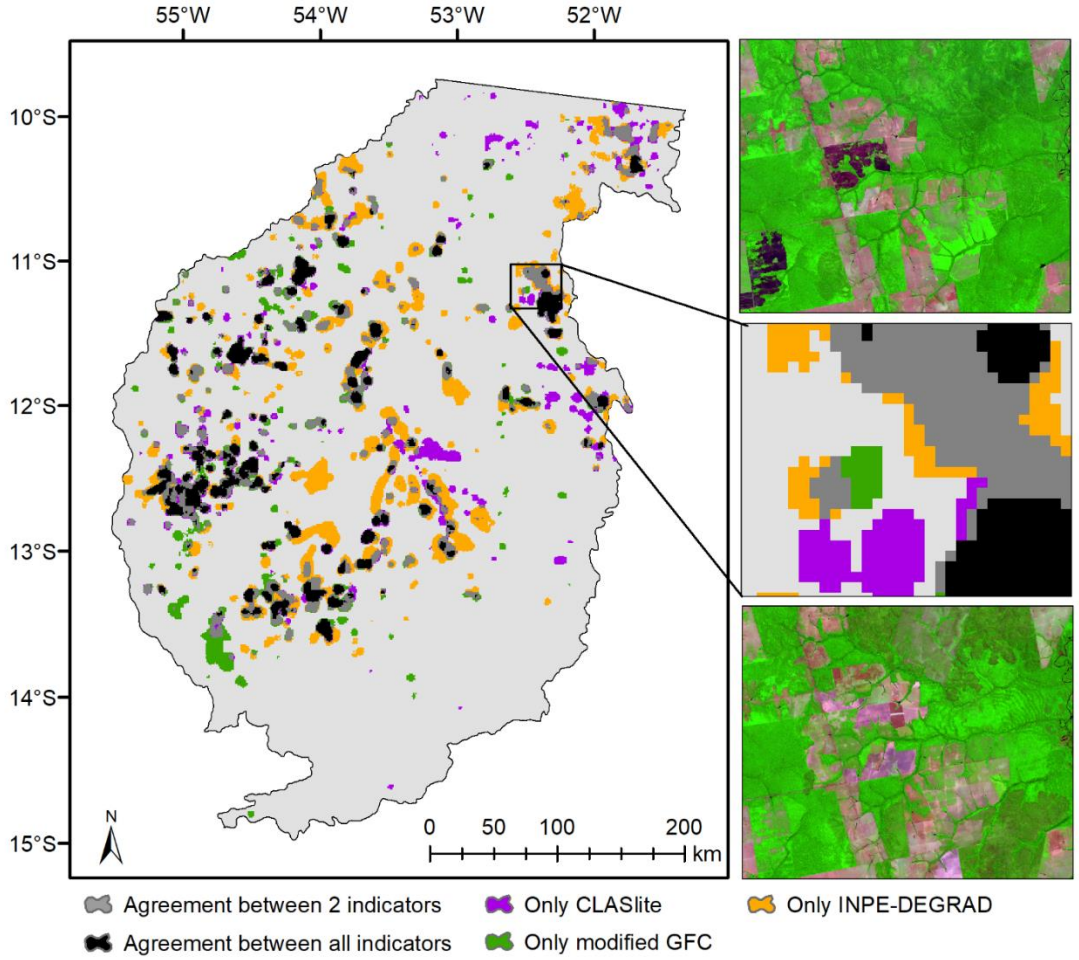


Figure 14. Map showing intersection of GFC, INPE-PRODES, and CLASlite output for degradation in the Upper Xingu River Basin between 2010 and 2015. Data is shown as hotspot areas with significance ≥ 0.95. Images from 2010 (Landsat 7 ETM+) and 2015 (Landsat 8 OPI) are show in false-color composition RGB 543 and RGB 654, respectively.

One important aspect when modelling vegetation change in diverse landscapes is how different models interact with different vegetation types or physiognomies. To shed light on this issue, we calculated similarity and disagreement metrics for each of the vegetation physiognomies in the Upper Xingu River Basin. Figures 15 and 16 show overall disagreement index and the detected area (as a proportion of the vegetation type in which it was detected) for either deforestation or degradation in raw data format. These figures show that, as expected, grassland cerrado and cerrado stricto sensu presented the smallest detected area of both deforestation and degradation. These vegetation types have much smaller number of trees when compared to the other types of vegetation. The trees observed in these vegetation types are scattered around the landscape or can densely clustered as riparian forest formations. When considering more representative vegetation types, forest cerrado and semi-deciduous forest present the largest disagreements. The same pattern was observed in the hotspot analysis, with the exception of the deforestation comparison, in which larger hotspot areas were created for the INPE dataset. All calculated values and the additional plots for hotspot analyses are shown in the supplementary material section (Appendix B).

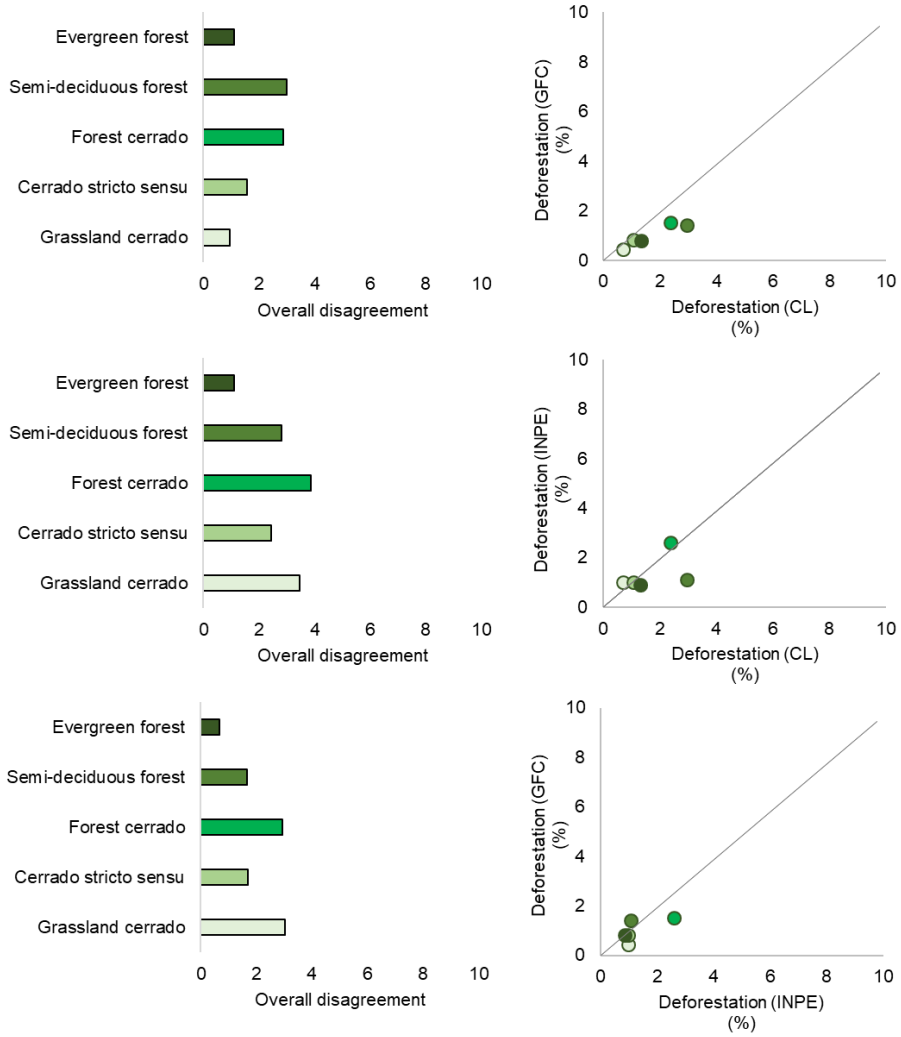


Figure 15. Overall disagreement (left) among indicators and their area as a proportion (right) of the vegetation type in which they are observed. The graphs show a comparison of the Global Forest Change (GFC), CLASlite (CL) and PRODES (INPE) datasets for deforestation detected between 2010 and 2015, and are calculated based on a raw data format.

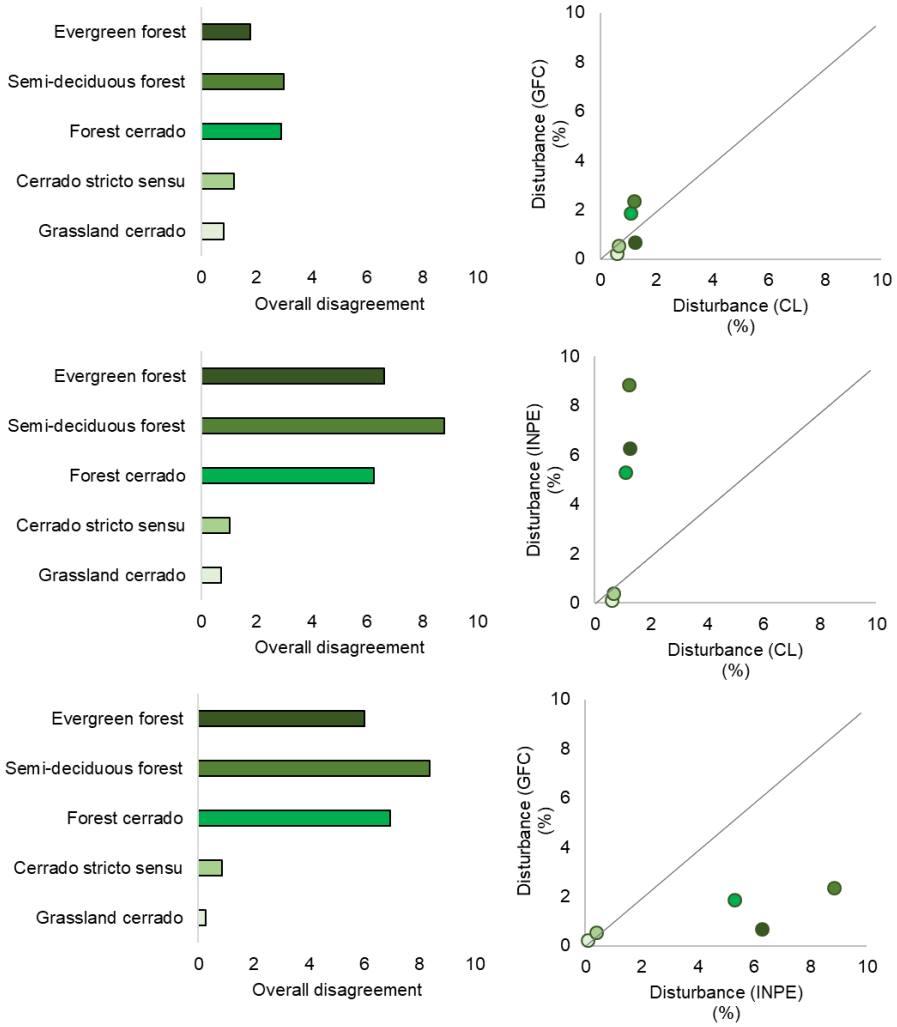


Figure 16. Overall disagreement (left) among indicators and their area as a proportion (right) of the vegetation type in which they are observed. The graphs show a comparison of the Global Forest Change (GFC), CLASlite (CL) and DEGRAD (INPE) datasets for degradation detected between 2010 and 2015, and are calculated based on a raw data format.

3.5. Discussion

Both GFC and CLASlite have been employed to map deforestation and general disturbance (deforestation and degradation considered together) in tropical regions such as Indonesia (Arjasakusuma et al., 2018), Africa (Dlamini, 2017; Lui and Coomes, 2015), and, in conjunction with PRODES/INPE in the Amazonia (Cabral et al., 2018; Hansen et al., 2008; Milodowski et al., 2017). These studies indicate that GFC can overestimate (Arjasakusuma et al., 2018) or underestimate (Dlamini, 2017; Milodowski et al., 2017) deforestation detection (Tropek et al., 2014). But overall, when compared to other methods, GFC delivers a satisfactory result, mainly when considering the cost-benefit of having a ready-to-use dataset (Burivalova et al., 2015). CLASlite outputs have been considered more accurate than GFC, mainly when studying areas with high densities of mature forest (Dlamini, 2017; Lui and Coomes, 2015). Still, both approaches have been regarded as more accurate to detect degradation in general than classification techniques even when they are applied on the local/regional scale (Arjasakusuma et al., 2018; Lui and Coomes, 2015; Milodowski et al., 2017). Still, authors have cited the INPE's PRODES program as very reliable dataset on deforestation, but they have also expressed that PRODES underestimates deforestation due

to its minimum mapped area and the conception of deforestation used in the program. PRODES only maps the deforested "pristine" areas in the Legal Amazon region.

Degradation is a result of any action which degrades tree density, resulting in a thinner canopy. It can be driven by forces such as selective logging or natural dynamics (Alves et al., 2009). As the contrast between degraded areas and their surroundings is not as evident as the contrast in the case of clear-cut deforestation, we expected that degradation detection would be more problematic, presenting a larger disagreement between approaches. Although we observed a larger disagreement for degradation than for deforestation, this pattern as mostly due to the DEGRAD/INPE dataset. When comparing GFC and CLASlite, the disagreement was similar to deforestation. These datasets employ a subpixel approach and Wang and colleagues (2005) have argued that subpixel methods retrieve better results when dealing with degradation than any other method because this approach can capture more than the no green/green tree canopy. Nevertheless, employing multiple approaches on larger scales to investigate degradation may be important to capture patterns in different vegetation types (Wang et al., 2005). The efforts are necessary since degradation can be as wide-ranging as deforestation in certain regions (Alves et al., 2009).

The performance by different remotely sensed products on the detection of deforestation and degradation according to vegetation types or phytophysiognomies is poorly discussed in the literature. The definition of forest is a singular concept when building models to detect deforestation or forest degradation. In the GFC dataset, forest is defined as all vegetation taller than 5 m and typically with > 50% canopy closure (Hansen et al., 2013). The CLASlite dataset does not explicitly use a forest definition, but states that although initially created to map savannas, it is now developed to analyze dense forest and that using it for other physiognomies will not necessarily be well-suited (Asner et al., 2019). In the PRODES and DEGRAD program, INPE also does not use a clear definition of forest, but states that the detection is developed for "pristine" areas.

Here, we observed that the better match between datasets was observed for evergreen forest when comparing the raw data format or hotspots, followed by semi-deciduous and cerrado forest. It is likely that both semi-deciduous and cerrado forests present more noise due to a thinner canopy, thus remotely sensed images capture both canopy and other components. This was also studied in African savannas, where mixtures of bare soil and healthy vegetation tend to create noise in deforestation detection using CLASlite (Dlamini, 2017). Still, as we could observe in the maps shown in figures 11-14, we believe that the GFC, INPE and CLASlite datasets may be capturing different parts of the disturbance effects. Thus, these different datasets are complementary to each other.

3.6. Conclusion

We recommend that attention be given to selecting appropriate datasets when analyzing both deforestation and degradation of native vegetation, and mainly when they will be used as inputs in any modeling approach as it can magnify patterns that are not useful. The analyzed datasets can have great impacts on the amount and specific location of deforestation and degradation mapped, including hotspot modeling. Rather than classifying different datasets and data formats as more or less accurate, we have shown that different datasets describing the same subject do not necessarily overlap. Rather, they may be capturing different spectrums of the same process depending on the landscape they are applied to and the concepts behind each dataset building. Thus, users should choose their data based on the methods applied in the vegetation loss mapping and their goals when manipulating such data, and should also justify their choice accordingly.

REFERENCES

Allnutt, T.F., Asner, G.P., Golden, C.D., Powell, G.V.N., 2013. Mapping recent deforestation and forest disturbance in northeastern Madagascar. Trop. Conserv. Sci. https://doi.org/10.1177/194008291300600101

Alves, D.S., Morton, D.C., Batistella, M., Roberts, D.A., Souza, C., 2009. The Changing Rates and Patterns of Deforestation and Land Use in Brazilian Amazonia, in: Science. pp. 11–23. https://doi.org/10.1029/2009GM000902

Arjasakusuma, S., Kamal, M., Hafizt, M., Forestriko, H.F., 2018. Local-scale accuracy assessment of vegetation cover change maps derived from Global Forest Change data, ClasLite, and supervised classifications: case study at part of Riau Province, Indonesia. Appl. Geomatics. https://doi.org/10.1007/s12518-018-0226-2

Asner, G., Sousan, S., Knapp, D., 2019. CLASlite [WWW Document]. User Guid. URL http://claslite.ciw.edu/en/

Asner, G.P., 1998. Biophysical and biochemical sources of variability in canopy reflectance. Remote Sens. Environ. 64, 234–253. https://doi.org/10.1016/S0034-4257(98)00014-5

Asner, G.P., Knapp, D.E., Balaji, A., Paez-Acosta, G., 2009. Automated mapping of tropical deforestation and forest degradation: CLASlite. J. Appl. Remote Sens. 3, 033543. https://doi.org/10.1117/1.3223675

Bonham-Carter, G.F., 1994. Tools for Map Analysis: Map Pairs, in: Geographic Information Systems for Geoscientists: Modelling with GIS. Elsevier Science, Tarrytown, New York, New York, pp. 221–266.

Burivalova, Z., Bauert, M.R., Hassold, S., Fatroandrianjafinonjasolomiovazo, N.T., Koh, L.P., 2015. Relevance of Global Forest Change Data Set to Local Conservation: Case Study of Forest Degradation in Masoala National Park, Madagascar. Biotropica. https://doi.org/10.1111/btp.12194

Cabral, A.I.R., Saito, C., Pereira, H., Laques, A.E., 2018. Deforestation pattern dynamics in protected areas of the Brazilian Legal Amazon using remote sensing data. Appl. Geogr. https://doi.org/10.1016/j.apgeog.2018.10.003

Carlson, K.M., Curran, L.M., Ratnasari, D., Pittman, A.M., Soares-Filho, B.S., Asner, G.P., Trigg, S.N., Gaveau, D.A., Lawrence, D., Rodrigues, H.O., 2012. Committed carbon emissions, deforestation, and community land conversion from oil palm plantation expansion in West Kalimantan, Indonesia. Proc. Natl. Acad. Sci. https://doi.org/10.1073/pnas.1200452109

Carvalho, A.L. de, Nelson, B.W., Bianchini, M.C., Plagnol, D., Kuplich, T.M., Daly, D.C., 2013. Bamboo-Dominated Forests of the Southwest Amazon: Detection, Spatial Extent, Life Cycle Length and Flowering Waves. PLoS One. https://doi.org/10.1371/journal.pone.0054852

Chicas, S.D., Omine, K., Saqui, P., 2016. CLASlite algorithms and social surveys to asses and identify deforestation and forest degradation in Toledo's protected areas and forest ecosystems, Belize. Appl. Geogr. https://doi.org/10.1016/j.apgeog.2016.08.012

Curtis, P.G., Slay, C.M., Harris, N.L., Tyukavina, A., Hansen, M.C., 2018. Classifying drivers of global forest loss. Science (80-.). https://doi.org/10.1126/science.aau3445

Dlamini, W.M., 2017. Mapping forest and woodland loss in Swaziland: 1990–2015. Remote Sens. Appl. Soc. Environ. https://doi.org/10.1016/j.rsase.2017.01.004

FAOSTAT, 2016. Country Indicators [WWW Document]. URL http://www.fao.org/faostat/en/#home

Finer, B.M., Novoa, S., Weisse, M.J., Petersen, R., Mascaro, J., Souto, T., Stearns, F., Martinez, R.G., 2018. Combating deforestation: From satellite to intervention. Science (80-.). https://doi.org/10.1126/science.aat1203

Garcia, A.S., Vilela, V.M. de F.N., Rizzo, R., West, P., Gerber, J.S., Engstrom, P.M., Ballester, M.V.R., 2019. Assessing land use/cover dynamics and exploring drivers in the Amazon's arc of deforestation through a hierarchical, multi-scale and multi-temporal classification approach. Remote Sens. Appl. Soc. Environ. 15, 100233. https://doi.org/10.1016/j.rsase.2019.05.002

Getis, A., 1992. The analysis of spatial association by use of distance statistics. Geogr. Anal. 24. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

GLAD - Global Land Analysis & Discovery [WWW Document], 2018. . Glob. For. Chang.

Hansen, M.C., Potapov, P. V, Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S. V, Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. Science (80-.). 342, 850–853. https://doi.org/10.1126/science.1244693

Hansen, M.C., Shimabukuro, Y.E., Potapov, P., Pittman, K., 2008. Comparing annual MODIS and PRODES forest cover change data for advancing monitoring of Brazilian forest cover. Remote Sens. Environ. 112, 3784–3793. https://doi.org/10.1016/j.rse.2008.05.012

Hasan, A.F., Laurent, F., Blanc, L., Messner, F., 2017. The use of Landsat time series for identification of forest degradation levels in the eastern Brazilian Amazon (Paragominas), in: 2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp). IEEE, pp. 1–4. https://doi.org/10.1109/Multi-Temp.2017.8035243

Hisabayashi, M., Rogan, J., Elmes, A., 2018. Quantifying shoreline change in Funafuti Atoll, Tuvalu using a time series of Quickbird, Worldview and Landsat data. GIScience Remote Sens. https://doi.org/10.1080/15481603.2017.1367157

IBGE, 2019. Sistema IBGE de Recuperação Automática (SIDRA): banco de dados sobre indicadores, população, economia e geociências.

IBGE, 2012. Manual Técnico da Vegetação Brasileira, Instituto Brasileiro de Geografia e Estatística - IBGE. https://doi.org/ISSN 0101-4234

INPE - Instituto Nacional de Pesquisas Espaciais, 2019. Monitoramento da Floresta Amazônica Brasileira por Satélite [WWW Document]. URL http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes (accessed 5.10.19).

Kanniah, K.D., 2017. Quantifying green cover change for sustainable urban planning: A case of Kuala Lumpur, Malaysia. Urban For. Urban Green. https://doi.org/10.1016/j.ufug.2017.08.016

Kintisch, E., 2007. Improved monitoring of rainforests helps pierce haze of deforestation. Science (80-.). https://doi.org/10.1126/science.316.5824.536

Lui, G. V., Coomes, D.A., 2015. A comparison of novel optical remote sensing-based technologies for forest-cover/change monitoring. Remote Sens. https://doi.org/10.3390/rs70302781

Masek, J.G., Vermote, E.F., Saleous, N., Wolfe, R., Hall, F.G., Huemmrich, F., Gao, F., Kutler, J., T.K. Lim, A., 2006. A Landsat Surface Reflectance Dataset for North America, 1990–2000. Geosci. Remote Sens. Lett. 3, 68–72.

Michaelsen, A.C., Perz, S.G., Huamani Briceño, L., Fernandez Menis, R., Bejar Chura, N., Moreno Santillan, R., Díaz Salinas, J., Brown, I.F., 2017. Effects of drought on deforestation estimates from different classification methodologies: Implications for REDD+ and other payments for environmental services programs. Remote Sens. Appl. Soc. Environ. 5, 36–44. https://doi.org/10.1016/j.rsase.2017.01.003

Milodowski, D.T., Mitchard, E.T.A., Williams, M., 2017. Forest loss maps from regional satellite monitoring systematically underestimate deforestation in two rapidly changing parts of the Amazon. Environ. Res. Lett. https://doi.org/10.1088/1748-9326/aa7e1e

Ord, J.K., Getis, A., 1995. Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. Geogr. Anal. 27, 286–306. https://doi.org/10.1111/j.1538-4632.1995.tb00912.x

Pontius, R.G., Santacruz, A., 2014. Quantity, exchange, and shift components of difference in a square contingency table. Int. J. Remote Sens. 35, 7543–7554. https://doi.org/10.1080/2150704X.2014.969814

Ribeiro, J.F., Walter, B.M.T., 2019. Tipos de Vegetação do Bioma Cerrado [WWW Document]. Embrapa Cerrados. URL http://www.agencia.cnptia.embrapa.br/Agencia16/AG01/arvore/AG01_23_911200585232.html (accessed 5.10.19).

Roitman, I., Cardoso Galli Vieira, L., Baiocchi Jacobson, T.K., da Cunha Bustamante, M.M., Silva Marcondes, N.J., Cury, K., Silva Estevam, L., da Costa Ribeiro, R.J., Ribeiro, V., Stabile, M.C.C., de Miranda Filho, R.J., Avila, M.L., 2018. Rural Environmental Registry: An innovative model for land-use and environmental policies. Land use policy 76, 95–102. https://doi.org/10.1016/j.landusepol.2018.04.037

Rosa, I.M.D.D., Ahmed, S.E., Ewers, R.M., 2014. The transparency, reliability and utility of tropical rainforest land-use and land-cover change models. Glob. Chang. Biol. 20, 1707–1722. https://doi.org/10.1111/gcb.12523

Shimabukuro, Y.E., Arai, E., Anderson, L.O., Aragão, L.E.O. e C. de, Duarte, V., 2017. Mapping Degraded Forest Areas Caused By Fires During the Year 2010 in Mato Grosso State, Brazilian Legal Amazon Using Landsat-5 Tm Fraction Images. Rev. Bras. Cartogr.

Tropek, R., Sedláček, O., Beck, J., Keil, P., Musilová, Z., Šímová, I., Storch, D., 2014. Comment on "High-resolution global maps of 21st-century forest cover change." Science (80-.). https://doi.org/10.1126/science.1248753

Turubanova, S., Potapov, P. V., Tyukavina, A., Hansen, M.C., 2018. Ongoing primary forest loss in Brazil, Democratic Republic of the Congo, and Indonesia. Environ. Res. Lett. https://doi.org/10.1088/1748-9326/aacd1c

Tyukavina, A., Hansen, M.C., Potapov, P., Parker, D., Okpa, C., Stehman, S. V., Kommareddy, I., Turubanova, S., 2018. Congo Basin forest loss dominated by increasing smallholder clearing. Sci. Adv. https://doi.org/10.1126/sciadv.aat2993

Wang, C., Qi, J., Cochrane, M., 2005. Assessment of tropical forest degradation with canopy fractional cover from landsat ETM+ and IKONOS imagery. Earth Interact.

Yesuf, G., Brown, K.A., Walford, N., 2019. Assessing regional-scale variability in deforestation and forest degradation rates in a tropical biodiversity hotspot. Remote Sens. Ecol. Conserv. https://doi.org/10.1002/rse2.110

4. MODEL SELECTION FOR DEFORESTATION, DEGRADATION, AND VEGETATION LOSS SUGGESTS DIFFERENT CAUSATIONS AND FUTURE SCENARIOS

ABSTRACT

Besides deforestation, degradation and non-forest vegetation loss can greatly depreciate carbon stocks, biodiversity, and alter biogeochemistry dynamics. Still, these processes combined are rarely addressed when modelling and predicting vegetation loss. In this study we fitted logistic regressions to model each of these processes using variables related to land use and land cover, land tenure, climatic indicators, and landscape characteristics. We observed that biophysical characteristics, land use, and land tenure presented great contributions in fitting the models. Besides the importance of these variables in all models, deforestation, degradation, and general native vegetation (both forest and non-forest) were influenced differently by each of them. We show that when tackling each of these three processes, decision makers and researchers should address each of the processes differently. Through a policy perspective it means that limiting one of these processes will not necessarily limits another. As an example, in the studied Amazon`s agricultural frontier, deforestation is likely to happen in the Cerrado biome portion rather than Amazon biome, but degradation and general vegetation loss is likely to occur in both biomes. Such pattern is probably a reflection of policies which focus on avoiding deforestation in rain forest in the Amazon biome region, but do not addresses other processes or biomes.

Keywords: Deforestation; Disturbance; Vegetation loss; Agriculture frontier

4.1. Introduction

Causes of deforestation have been successfully addressed by researchers and policy makers (Boucher et al., 2013; Geist, H and Lambin, E, 2002; Gibbs et al., 2015; Hansen et al., 2013; Lawrence and Vandecar, 2015; Soares-Filho et al., 2006; WRI - World Resources Institute, 2014). However, forest degradation has been rarely considered although it can double biodiversity loss from deforestation and plays an important role in the carbon balance dynamics (Barlow et al., 2016; Espírito-Santo et al., 2014). Degradation is the result of disturbances defined as the loss of trees which does not imply in land cover change, and it can be caused by logging, wildfires, blowdowns, and other processes (Asner et al., 2013; Cochrane, 2003; Nelson et al., 1994; Peres et al., 2006). Still, mapping degradation is not as straight forward as mapping deforestation, and the lack of quality, ready-to-use data is likely related to the lack of models explaining its patterns and predicting future scenarios.

While deforestation is the clearcutting of forested areas, turning it into another type of land use, native vegetation loss is the conversion of any type of native vegetation, e.g. conversion of grasslands into a managed land such as croplands or pasturelands. In the neotropics, efforts to understand, model, and predict native vegetation loss, including deforestation, have been mainly concentrated in the Amazon Rain Forest (Laurance et al., 2002; Rosa et al., 2013; Soares-Filho et al., 2006), with fewer instances looking into the Cerrado (Ferreira et al., 2012). Although the Cerrado is a biodiversity hotspot presenting higher rates of deforestation than the Amazon (Strassburg et al., 2017), studies modelling the expansion of either general native vegetation loss, deforestation or degradation are rather few.

Nevertheless, the ecotone zone between the Amazon and the Cerrado biomes overlaps with the Amazon's agricultural frontier which, in turn, is the region under the highest land use change pressure in the basin (Rosa et al., 2013). This region is composed of a diverse landscape of both Cerrado and Rain Forest physionomies (Garcia et al., 2019, 2017) with high biodiversity rates (Kunz et al., 2009; Smith et al., 2001). Agricultural frontiers can present a significant and distinct pattern of land use change compared to other parts of the same biome. As the frontier develops, it moves from a pioneer stage supported by state policies and infrastructure to a consolidated stage, which is dominated by large-scale capitalized agriculture (Meyfroidt et al., 2018). In practice, it means that different patterns of land use change can be expected compared to those modelled for a whole biome. For example, instead of low productivity pastureland expansion into native vegetation, we may expect to observe cropland expansion directly replacing native vegetation (Maeda et al., 2011). Thus, the Amazon's agricultural frontier presents two challenges when modelling change of native vegetation: to cope with both a diverse landscape and with a unique land use dynamic.

In this study, we analyzed which variables are influencing both general native vegetation loss (any vegetation type), clearcutting deforestation (tree loss resulting in a different land use), and forest degradation (tree loss not resulting in a different land use) in the Amazon's agricultural frontier. Our goal was to test whether these processes are similarly influenced by structural, economic and biophysical factors. We hypothesized that each of the three processes will be influenced by different variables. Additionally, we compared our results with models reported in the literature for the Amazon and Cerrado to check whether our model identified the same most important drivers. The next step of this research is to build future scenarios for general vegetation loss, deforestation and degradation in the Upper Xingu River Basin.

4.2. Material and methods

4.2.1. Study area

The Upper Xingu River Basin (hereafter referred to as UXRB), located in Mato Grosso State, in the southern portion of the Brazilian Amazon Basin, encompasses one-third of the whole Xingu River Basin, which totals ~17 million hectares. As shown in Figure 17, the UXRB is in the transition zone between the Amazon and Cerrado biomes. Due to its position in an ecotone, the basin contains a range of natural vegetational physiognomies, varying from rain forest to savanic formations. The majority of the region is composed of Floresta Estacional Perenifólia – a unique transitional forest with characteristics of both Cerrado and Amazon Biomes (Ivanauskas et al., 2008; Velasquez, C.; Alves, H. Q.; Bernasconi, 2010). Figure 17 shows the location of the UXRB in the ecotone zone and the variety of land covers and land uses that occur throughout the basin. The broad forest cover surrounding the river main stream at the centre of the image represents the Xingu Park, one of the largest indigenous land in Brazil. A large indigenous population, composed by about 20000 inhabitants (IBGE, 2010) from a variety of etnic groups (13) lives in the Upper Xingu basin. Despite the long indigenous occupation of the region, landscape changing began in the late 70's, as the process of non-indigenous occupation of the region expanded. Currently, about 25% of the UXRB is located inside protected areas, of which indigenous reservations make up about 24% (Velasquez, C.; Alves, H. Q.; Bernasconi, 2010).

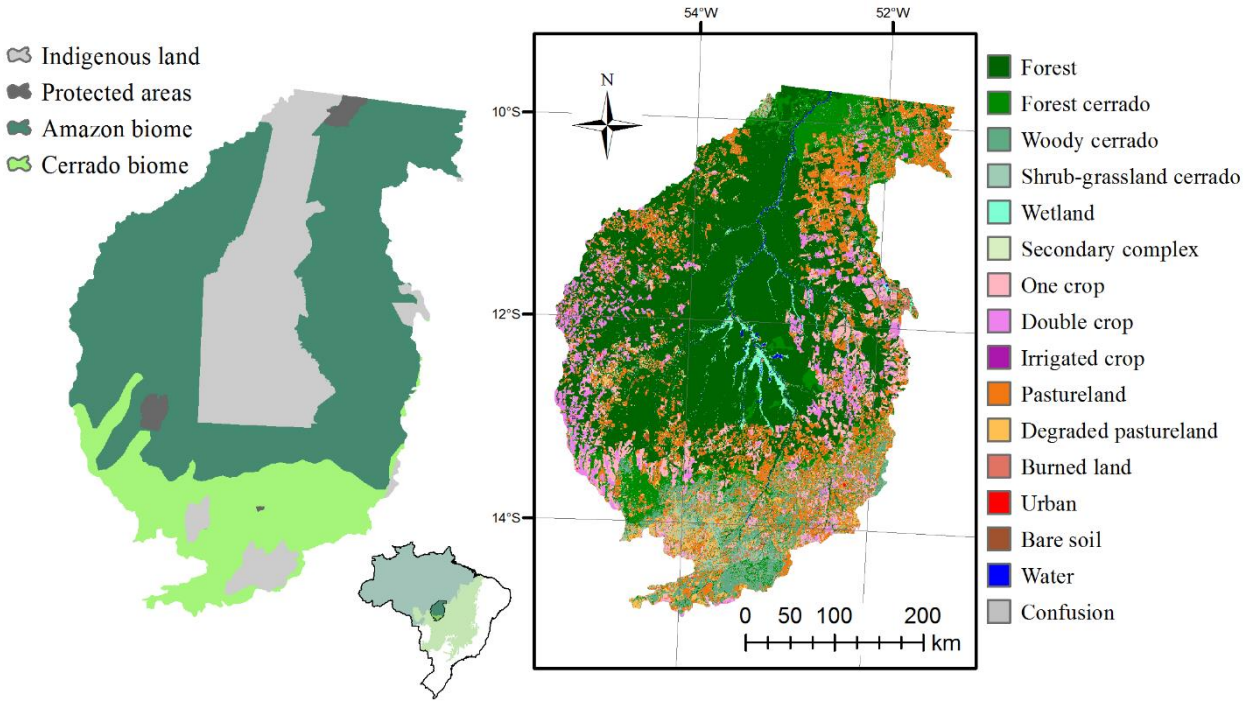


Figure 17. The Upper Xingu River Basin, located in-between the Amazon and Cerrado Biomes. The basin contains a huge variety of land cover and land use types.

The UXRB exhibits a seasonal tropical climate. The annual mean precipitation is 2000 mm and the annual mean temperature 24 ºC. Seasons are characterized by a dry winter and rainy summer, with a flood season occurring from November to April. The extended and defined rainy season, associated with a flat terrain and being dominated by Oxisol soils (Velasquez, C.; Alves, H. Q.; Bernasconi, 2010), make this region ideal for agribusiness expansion. The main economic activities in the study area are linked to the wood, beef, and soybean industries. Until the 2000s, forests were degraded or cleared due to livestock expansion and logging activities. Following this, soybean production emerged as another driver of deforestation (Brando et al., 2013; Garcia et al., 2019; Nepstad et al., 2006). Figure 18 show the evolution of agricultural production in the 25 main municipalities in the UXRB.

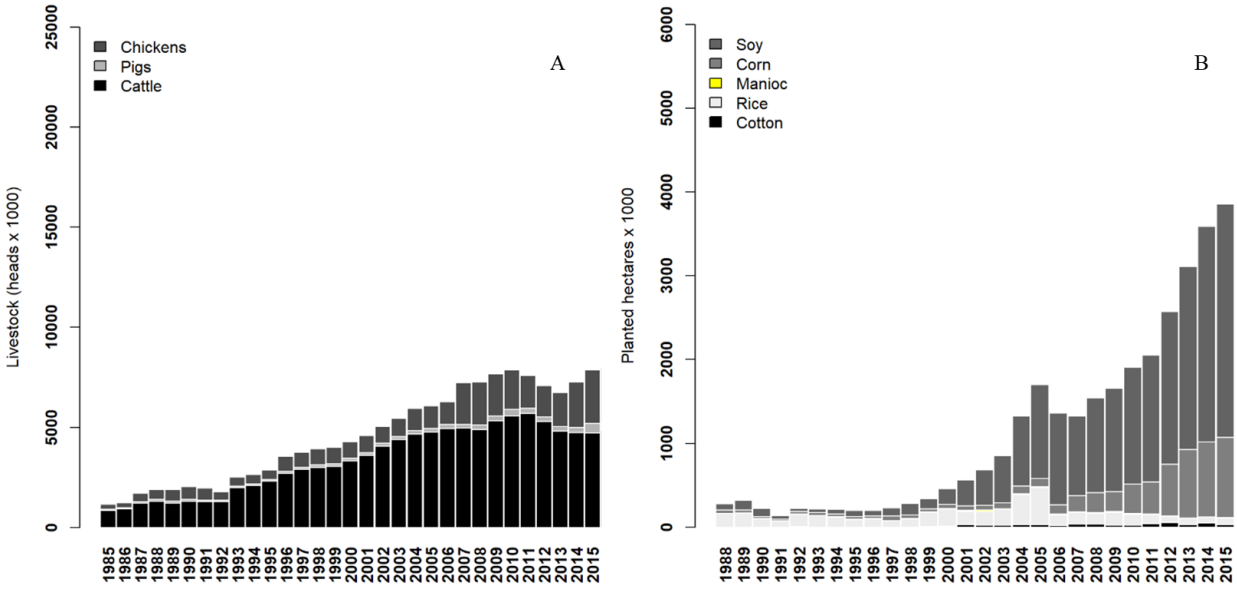


Figure 18. Evolution of agricultural production in the Upper Xingu River Basin. Panel A shows livestock production from 1985 to 2015. Panel B shows planted area in cultivation for the main agricultural products. The data is a compilation for municipalities with at least 20% of their area located inside the UXRB - 25 municipalities in total. The data was obtained from SIDRA/IBGE.

Currently, ~ 9 % of Brazil's soybean harvest and ~13% of Brazil's cattle are produced in the UXRB (FAOSTAT, 2016; IBGE, 2016). Interestingly, wood extraction is not reported as a primary activity according to the Brazilian Institute of Geography and Statistics (IBGE), but is still observed in the basin as discussed in the previous chapter of this thesis. By the end of 2015, the UXRB became a frontier which supplied both international and regional markets with different commodities. A large portion of the study region is in the Cerrado biome and, thus, is not subject to public and market policies toward rain forest protection such as the Brazilian Amazon Moratorium and the Zero Deforestation agenda (Garrett et al., 2013). Accordingly, such areas have exhibited deforestation rates 10-times higher than those in the Amazon Biome, possibly a rebound effect caused by policies developed exclusively for the Amazon (Garcia et al., 2019).

4.2.2. Dataset

We used the land use and land cover maps created by Garcia and colleagues (2019) to obtain general natural vegetation loss data, including forest, shrubland and grasslands loss. Deforestation and degradation patterns were derived from the Global Forest Change (GFC) dataset (Hansen et al., 2013). Trees at least 5 meters tall are classified as forest, regardless of density. Tthis dataset is likely to be overestimating these processes in some areas of the basin, mainly by overestimating the initial area of forest (Tropek et al., 2014). To address this issue, we filtered the dataset using the land use and land cover maps presented (Garcia el a., 2019). We used the first level of information, which presents 93% overall accuracy, to conditionally selected pixels undergoing change in the GFC dataset if they were classified as (semi) natural vegetation in LCLU-L1 in the first time step. To classify tree cover loss as deforestation and degradation, we also conditionally filtered pixels according to the second time step in the change analyses. The pixels which were still considered (semi) natural vegetation after undergoing tree loss were considered degradation, and the ones which became managed areas were considered deforestation. Figure 19 shows the change in natural vegetation cover, and the expansion of deforestation and degradation (disturbance) from 2010 to 2015.

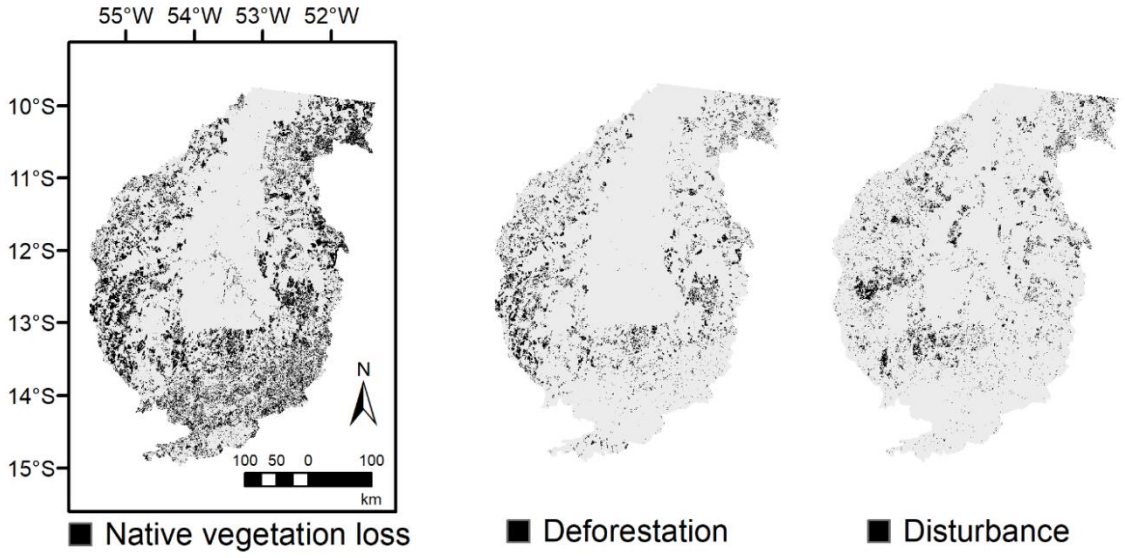


Figure 19. Natural vegetation cover retraction (35,000 km²), deforestation (12,000 km²) and disturbance (11,000 km²) expansion from 2000 to 2015. Black color indicates the change in each of these processes from 2000 to 2015.

Our dataset also comprised biophysical, economic, infrastructure, and land tenure indicators (Table 7). Biophysical indicators consist of variables representing precipitation, temperature, proximity to rivers, biomes, vegetation types and land covers. Economic variables are gross domestic product (GDP) by municipality, the agriculture portion of municipalities' GDP, cattle production, and grain production. Infrastructure is defined as assess to roads and cities. Land tenure is defined by property ownership, which is categorized as private, settlement (public but used for family farms), or public (either protected areas or indigenous land). The whole dataset was transformed into grids on the same projection (South America Albers Equal Area Conic) and spatial scale (300m).

Table 7. Variables tested in land use change models. This table shows the source, description, unit of measurement, the original data format and resolution. Variables for which the source is listed as LAAG (Laboratorio de Analises Ambientais e Geoprocessamento) were produced by the authors.

Variables	Source	Description	Unit	Data format
Land cover and land use (L1)	LAAG	Land use maps for UXRB at Level 1 of classification	Categorical	Raster (30m)
Land cover and land use (L2)	LAAG	Land use maps for UXRB at Level 2 of classification	Categorical	Raster (30m)
Neighborhood cropland	LAAG	Number of pixels classified as cropland (in L2) in a radius of 3km	Nº of pixels	Raster (30m)
Neighborhood pastureland	LAAG	Number of pixels classified as pastureland (in L2) in a radius of 3km	Nº of pixels	Raster (30m)
Variability of the neighborhood	LAAG	Number of unique LULC classes (in L2) detected in a radius of 3km	Nº of classes	Raster (30m)
Distance from pasturelands	LAAG	Distance from any pixels classified as pastureland (L2)	Meters (m)	Raster (30m)
Distance from croplands	LAAG	Distance from any pixels classified as cropland (L2)	Meters (m)	Raster (30m)
Distance from deforestation	LAAG	Distance from any pixels classified as deforested in the previous period	Meters (m)	Raster (30m)

Distance from disturbance	LAAG	Distance from any pixels classified as disturbed in the previous period	Meters (m)	Raster (30m)
Precipitation Seasonality	WorldClim	Variation in monthly precipitation totals over the course of the year. This index is the ratio of the standard deviation of the monthly total precipitation to the mean monthly total precipitation	Percentage (%)	Raster (10")
Precipitation of driest quarter	WorldClim	Total precipitation that prevails during the driestest month	Millimeters (mm)	Raster (10")
Precipitation of wettest quarter	WorldClim	Total precipitation that prevails during the wettest month	Millimeters (mm)	Raster (10")
Annual precipitation	ANA	Interpolation of precipitation registered in station over 30 years	Millimeters (mm)	Raster (500m)
Slope	Aster GDEM	Slope derived from Aster GDEM	Percentage (%)	Raster (30m)
Soil suitability	EMBRAPA Solos	Soil classified from 1 to 4 according to their suitability to agriculture in mechanized scenario	Discrete	Polygon
GDP	IBGE	Contrast of gross primary production by municipality between 2010 and 2015	R$	Polygon
GDPag	IBGE	Contrast of gross primary production in the agriculture sector by municipality between 2010 and 2015	R$	Polygon
Cattle	IBGE	Contrast of cattle herd by municipality between 2010 and 2015	Heads	Polygon
Grains (corn + soybean)	IBGE	Contrast of main grain production area (corn + soy) by municipality between 2010 and 2015	Planted hectares	Polygon
Farm size	CAR/ IMAFLORA	Private owned proprieties size	Hectares	Polygon
Land tenure	CAR/ IMAFLORA	Either public, private or settlements	Categorical	Polygon
Protected areas	MMA	Protected areas (SNUC). All areas are state units under maximum protection	Categorical	Polygon
Indigenous land	MMA	Indigenous land	Categorical	Polygon
Mines	MMA	Plots with mining (licensed or not)	Categorical	Polygon
Vegetation types	IBGE	Vegetation physiognomies ordered by tree density	Discrete	Polygon
Biomes	MMA	Biomes	Categorical	Polygon
Rivers	ANA	Distance from rivers derived from shapefiles from ANA	Meters	Raster (30m)
Main roads	SEPLAN-MT	Distance from state and federal roads derived from shapefile from SEPLAN	Meters	Raster (30m)
Roads	SEPLAN-MT and CAR	Distance from all roads derived from shapefile from SEPLAN and CAR	Meters	Raster (30m)
Slaughterhouses	Vale et al., 2019	Distance from all slaughterhouses	Meters	Raster (30m)
Silos	CONAB	Distance from all grain silos	Meters	Raster (30m)
Cities	LAAG	Distance from cities	Meters	Raster (30m)

WorldClim (Global Climate Data: www.worldclim.org/bioclim); ANA (National Water Agency of Brazil: metadados.ana.gov.br/geonetwork), Aster GDEM (asterweb.jpl.nasa.gov/gdem.asp), EMBRAPA Solos

(www.embrapa.br/solos/sibcs/bases-de-dados-de-solos), IBGE (Brazilian Institute of Geography and Statistics: sidra.ibge.gov.br/home/ipca15/brasil), CAR/IMAFLORA (Brazil's National Registry of Rural Properties validated by the Instituto de Manejo e Certificação Florestal e Agrícola: atlasagropecuario.imaflora.org/); MMA (The Ministry of Environment: www.mma.gov.br/governanca-ambiental/geoprocessamento), SEPLAN-MT (Secretary of Planing and Coordination of the Mato Grosso state: http://www.seplan.mt.gov.br/); CONAB (The Brazilian National Company of Supply: https://www.conab.gov.br/)

4.2.3. Model fitting and selection

In order to prevent multicollinearity in our models, we performed correlation tests (Pearson) among all pairs of continuous variables and association tests (Tau-Goodman and Kruskal) among all pairs of categorical variables. Correlation among observational variables is common, but it indicates that two or more variables are not independent. It causes false estimation of the relationship between each independent variable with the dependent variable. For categorical variables, we applied chi-squared tests to verify if the distribution of categories was independent. The calculation was performed in R (https://www.r-project.org/), using the package 'raster' (Hijmans et al., 2011). Once correlation was found between a pair of variables, we chose which variable would be included in or eliminated from a model based on its theoretical meaning and performance when fitting it into the models.

We built a different model for the reduction of natural land cover (L1 maps), expansion of deforestation (GFC), and expansion of degradation (GFC) between 2010 and 2015. For each model selection, we fitted both individual variable models and multivariable models. We used the Stochastic Modelling of Land Cover Change (StocModLCC - ROSA et al., 2013) to fit our models and to generate simulations of future patterns. StocModLCC is a probabilistic modelling approach, focused on local-scale and contagious processes. The probability of a pixel to become disturbed or deforested was based in the logistic regression:

$$Pprocess_{x,t} = \frac{1}{(1 + exp^{-k_{x,t}})}$$

where, kx,t represents how variables affect location x at time t, varying from minus infinity to plus infinity. The probability of a pixel to undergo either degradation or deforestation (Pprocess) in location x at time t varies from 0 to 1.

Whenever fitting multiple variables, we refine each model fitting through forward stepwise regression. Through stepwise we chose the model with the lower Akaike Information Criterion (AIC), which represents the smaller number of parameters given a likelihood of the fitted model. Model calibration was evaluated trough pseudo-R2, which compare the likelihood of the model with a null-model (Hosmer and Lemeshow, 2000). We applied cross-validation on 50 % of the samples to allow for selecting models that have higher predictive power. We tested the performance based on the Area Under the Curve of Receiver Operating Characteristics curve (AUC – ROC curve). ROC is a probability curve between false and true positive rates of classification. An AUC equal to 0.5 (half of its potential) means the model has no separability power, while an AUC equal to 1 means the model has full separability power. After fitting testing models to represent each process, the posterior mean and credible interval of each parameter were calculated based on Markov Chain Monte Carlo sampling technique. It allows for the quantification of uncertainty around predictions. The best model was the one with the higher log-likelihood and fewer parameters, but also with all variables showing significant impact on the process being predicted. All details regarding the processing scheme of the StocModLCC are available at Rosa and colleagues (2013, 2014).

4.3. Results

We found notable correlations between few variables such as climate, distance from main roads and pasturelands, main roads and all roads, silos and croplands. We present all variables plotted in maps and the correlation among all pairs of variables in the supplementary material (Appendix C). The final models for natural cover retraction, deforestation expansion, and degradation expansion presented good calibration and discrimination. For natural cover retraction, deforestation expansion, and degradation expansion the AUC values were 0.82, 0.84, and 0.85, respectively (Appendix C). These values indicate that each model had a likelihood value significantly larger than a null-model, and that they can discriminate well between true and false positives.

From all the tested variables, one group significantly contributed to adequately model all three processes (natural cover retraction, deforestation expansion, and degradation expansion). But they did not necessarily have the same influence (Table 8). Among the categorical variables, land cover (forest, savanic formations, managed land), land tenure (private, public-settlements, public-protected areas, public-indigenous land), and biome (Amazon, and Cerrado) contributed to all three model processes. Native vegetation loss was 10-fold more likely in savanic formations than forest, and deforestation almost 2-fold. Contrarily, degradation was 4-fold larger in forest than savanic formations. Protected areas were the most protective type of land. It was 5-fold more protective than private lands for vegetation loss, 293607-fold for deforestation, and 10-fold for degradation. Indigenous land showed to be protective against deforestation only. All processes are more likely to occur in settlements from 1.5 to 2-fold. We observed a diverse response for biome. Native vegetation loss was more likely in the Amazon, but deforestation was more likely in Cerrado, and non-significant difference was observed between biomes for degradation. This can be a reflection of public and market policies which focus on preventing deforestation in the Amazon only and the specifically monitoring of rain forest.

Table 8. Mean coefficient estimation and the 95% confidence intervals for each variable in multiple predictor variables models. The estimates represent the expected change in the odds of P(process) = 1 for a unit increase in the corresponding predictor variable holding the other predictor variables constant. For categorical variables, the coefficient indicates the odds of a pixel classified as such undergo the analyzed process when compared to a pixel classified as a reference. Odds values lower than 1 indicate a protective character as the probabilities tend towards zero.

Variables	**Natural cover retraction**			**Deforestation expansion**			**Disturbance expansion**		
	Odds ratio (95% CI)								
(Intercept)	0.003	(0.002	0.005)	0.002	(0.001	0.004)	0.026	(0.012	0.059)
Neighborhood	1.003	(1.003	1.003)	0.993	(0.992	0.994)	1.001	(1.000	1.002)
Biome Amazon	1.000	(1.000	1.000)	1.000	(1.000	1.000)	ns	(ns	ns)
Biome Cerrado	0.753	(0.732	0.774)	1.401	(1.309	1.498)	ns	(ns	ns)
Land tenure - IL	1.290	(1.247	1.334)	0.266	(0.224	0.314)	1.017	(0.947	1.091)
Land tenure - Private	1.000	(1.000	1.000)	1.000	(1.000	1.000)	1.000	(1.000	1.000)
Land tenure - Settlement	1.433	(1.386	1.482)	1.879	(1.757	2.009)	1.520	(1.406	1.642)
Land tenure - UC	0.195	(0.138	0.267)	0.000	(0.000	0.000)	0.097	(0.042	0.188)
LC - Forest	0.582	(0.567	0.597)	3.881	(3.651	4.128)	7.523	(6.968	8.126)
LC - Managed	1.000	(1.000	1.000)	1.000	(1.000	1.000)	1.000	(1.000	1.000)
LC - Savanic formations	6.025	(5.882	6.167)	6.392	(5.949	6.866)	1.855	(1.660	2.072)
Amount of cropland in 3km	1.000	(1.000	1.000)	1.000	(1.000	1.000)	1.000	(1.000	1.000)
Amount of pastureland in 3km	ns	(ns	ns)	ns	(ns	ns)	1.000	(1.000	1.000)
Distance from cropland	1.000	(1.000	1.000)	ns	(ns	ns)	ns	(ns	ns)

Distance from disturbance	1.000	(1.000	1.000)	0.999	(0.999	0.999)	0.999	(0.999	0.999)
Distance from pastureland	1.000	(1.000	1.000)	1.000	(0.999	1.000)	1.000	(1.000	1.000)
Distance from rivers	1.000	(1.000	1.000)	1.001	(1.000	1.001)	1.000	(1.000	1.000)
Distance from roads	1.000	(1.000	1.000)	1.000	(1.000	1.000)	1.000	(1.000	1.000)
Distance from settlements	1.000	(1.000	1.000)	1.000	(1.000	1.000)	1.000	(1.000	1.000)
Distance from slaughterhouse	1.000	(1.000	1.000)	1.000	(1.000	1.000)	ns	(ns	ns)
Distance main roads	1.000	(1.000	1.000)	ns	(ns	ns)	ns	(ns	ns)
Distance silos	ns	(ns	ns)	ns	(ns	ns)	1.000	(1.000	1.000)
Precipitation mean	1.001	(1.001	1.001)	1.001	(1.001	1.001)	0.998	(0.998	0.998)
Precipitation seasonality	1.016	(1.012	1.020)	ns	(ns	ns)	1.043	(1.035	1.051)
Property size	1.000	(1.000	1.000)	1.000	(1.000	1.000)	1.000	(1.000	1.000)
Slope	0.978	(0.976	0.981)	0.932	(0.924	0.941)	0.982	(0.975	0.988)
Soil suitability	1.102	(1.090	1.114)	0.616	(0.587	0.645)	0.750	(0.728	0.772)
Variability in LU in 3km	1.030	(1.023	1.038)	ns	(ns	ns)	0.953	(0.938	0.968)
Vegetation type	0.972	(0.963	0.980)	1.198	(1.163	1.234)	1.123	(1.101	1.145)
Δ Cattle	ns	(ns	ns)	1.000	(1.000	1.000)	ns	(ns	ns)
Δ GDPag	1.000	(1.000	1.000)	ns	(ns	ns)	1.000	(1.000	1.000)
Δ Grain	1.000	(1.000	1.000)	ns	(ns	ns)	ns	(ns	ns)

While deforestation and degradation were both more likely to take place in more suitable soil for agriculture, native vegetation loss was more likely in less suitable ones. Following this pattern, a larger variability in local (3 km radius) land cover and land use promoted native vegetation loss, while prevented degradation. Additionally, all three process were more likely to be observed in lands with low slopes, but larger precipitation seasonality. The odds related to distance variables seems not impactful enough to influence the model. Still, it is important to note that the reported odds are related to the scale of each variable, and in the case of distance it is one meter. In a landscape the odds became impactful. All processes were more likely near to pasturelands and disturbed areas, but away from rivers. Distance from pasturelands and distance from deforested pixels were 80 % correlated. Distance from roads slightly contributed to all models. Native vegetation loss and deforestation were more promoted near roads, but not degradation. Interestingly, process in the neighborhood were positively related to natural vegetation loss and degradation, but not with deforestation.

The combined probability for all processes is shown in Figure 20, ranging from 0 (no probable future loss) to 1 (high probability of loss). A large portion of the basin presents a medium probability of vegetation loss. Indigenous land presents both areas of very low and very high probability. High probabilities are linked to general vegetation loss due to natural dynamics in wetlands in the Xingu National Park as well as fires in cerrado vegetation inside indigenous land on the southern and eastern portion of the basin. Nevertheless, spots with high probabilities are observed in cerrado vegetation in the Amazon biome (northern portion of the basin), and in a middle latitude between 12ºS and 13ºS which are hotspot areas of land use change into more intensified agriculture productions.

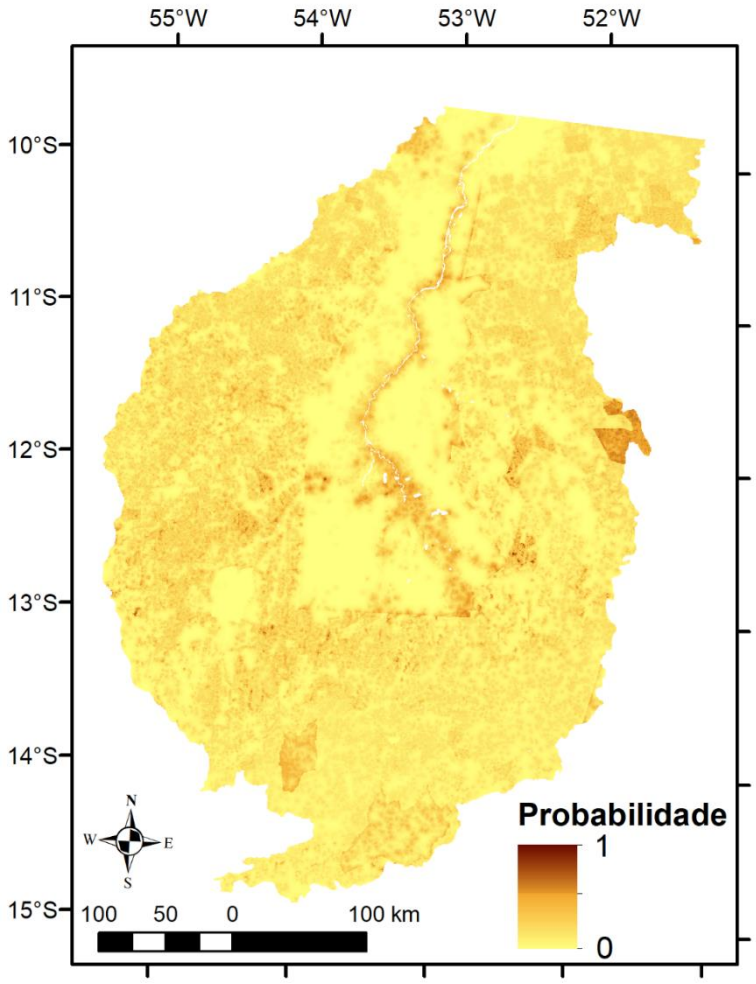


Figure 20. Spatial distribution of the probability of all three processes combined indicating regions with larger probability (brown color) of vegetation loss in the Upper Xingu River Basin in the future.

4.4. Discussion

We expected to select different variables when fitting each of the models for general natural vegetation loss, deforestation, and degradation. We observed that all three models present similarities and distinctions among them, and it is not possible to say that one process is more similar to another. Therefore, such pattern is a reflection of the different implicitly definitions in each process.

The variables we tested in this research were also used to model deforestation by other authors and in the same two biomes. Land cover, slope, soil fertility, distance from river and roads, precipitation, and human attractiveness (population size of municipalities ranked by their respective Human Development Index) were shown to be related to deforestation in the Cerrado biome (Ferreira et al., 2012). Unfortunately, variable importance rank is not reported. Land cover, distance from roads, rivers, and settlements, as well as topography, protected areas, soil fertility, dry season length, and agricultural development (increase or decrease in planted area, cattle head or GDP) were all used to model deforestation in the Amazon basin (Rosa et al., 2013). The authors showed that roads, neighborhood, and protected areas were the most important variables to predict deforestation. Protected areas also showed high importance on our models, but roads or neighborhood do not show great importance. Proximity to roads is observed to be a very strong factor in promoting deforestation in the Amazon basin (Geist, H and Lambin, E, 2002). As the Upper Xingu River Basin is a consolidated agricultural frontier, accessing roads is no longer an

impactful issue when expanding agricultural production (Perz et al., 2009). Thus, access to roads became a weaker factor when modelling deforestation or degradation. Although the impact of road quality may contrast with that of road availability, we have not tested this due to the lack of data on road quality. Additionally, factor the colonization process is different in this region and roads network was not developed following the usual pattern found in Amazonia. The fishbone pattern resulting from road opening by governmental initiatives is less common and medium to large farms are predominant. Therefore, deforestation is less concentrated along roads. Our study shows the same protective effect of publicly owned areas (protected areas and indigenous lands) as has also been show in other works (Rosa et al., 2013; Soares-Filho et al., 2006). In our assessment, protected areas were shown to be more effective than indigenous land.

In a review of land use modelling for tropical basins, Rosa and others (2014) demonstrated that distance from roads, land covers, soil factors, and urban locations were the most commonly used variables in modelling deforestation in South America, Africa and Asia. Land cover and soil factors were important variables in our models, but not distance from settlements or cities. Datasets on land use change and deforestation are made available for both the Cerrado and Amazon biomes (MMA-EMBRAPA-INPE, 2015; ACHARD; HANSEN, 2018; INPE, 2019). Still, more studies are concentrated in the Amazon due to its elevated public profile and high quality annual datasets describing vegetation loss (MISSING-VALUE, 2018; I. M. D. D. Rosa et al., 2014). Similarly, deforestation is more commonly addressed than degradation or general native vegetation loss, due to its detection being more straight-forward than that of degradation. As we expected, degradation was influenced by different factors than deforestation or general native vegetation loss. Degradation represents the loss of trees, but do not imply in a change in land cover. This process can be caused by factors such as logging, fires, blowdown, and others (Asner et al., 2013; Cochrane, 2003; Nelson et al., 1994; Peres et al., 2006). Land use, land tenure, precipitation seasonality, slope, and soil suitability for agriculture showed to be strong variables when modelling degradation. Lower slopes, larger precipitation seasonality, soil suitability for agriculture, land cover and land tenure were the most important characteristics of degradation expansion.

4.5. Conclusion

We have shown that different conditions influence native vegetation loss, deforestation, and degradation patterns. Only four variables significantly influenced the three processes: rural settlement and savanic formations increased the probability of the three processes to occur, while protected areas and slope decreased the probability. Other variables influenced the probability of these processes differently. For example, the presence of indigenous land and low suitability of the land for agriculture increased the probability only of general vegetation loss. Vegetation density increased the probability of deforestation and degradation. Thus, we show that to cope with losses for either carbon or biodiversity though these different processes, stake holders should look into others factors than those influencing only deforestation. Additionally, we have shown that loss in agricultural frontiers may not be positively influenced by the same factors that play an important role in relatively more preserved regions such as proximity to roads. Still, conservation units and indigenous land present a protective pattern as in more conserved regions.

REFERENCES

Asner, G.P., Keller, M., Lentini, M., Merry, F., Souza, C., 2013. Selective Logging and Its Relation to Deforestation, in: Amazonia and Global Change. American Geophysical Union (AGU), San Francisco. https://doi.org/10.1029/2008GM000722

Barlow, J., Lennox, G.D., Ferreira, J., Berenguer, E., Lees, A.C., Nally, R. Mac, Thomson, J.R., Ferraz, S.F.D.B., Louzada, J., Oliveira, V.H.F., Parry, L., Ribeiro De Castro Solar, R., Vieira, I.C.G., Aragaõ, L.E.O.C., Begotti, R.A., Braga, R.F., Cardoso, T.M., Jr, R.C.D.O., Souza, C.M., Moura, N.G., Nunes, S.S., Siqueira, J.V., Pardini, R., Silveira, J.M., Vaz-De-Mello, F.Z., Veiga, R.C.S., Venturieri, A., Gardner, T.A., 2016. Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. Nature. https://doi.org/10.1038/nature18326

Boucher, D., Roquemore, S., Fitzhugh, E., 2013. Brazil's success in reducing deforestation. Trop. Conserv. Sci. https://doi.org/10.1177/194008291300600308

Brando, P.M., Coe, M.T., DeFries, R., Azevedo, A.A., 2013. Ecology, economy and management of an agroindustrial frontier landscape in the southeast Amazon. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 368, 9. https://doi.org/10.1098/rstb.2012.0152

Cochrane, M.A., 2003. Fire science for rainforests. Nature. https://doi.org/10.1038/nature01437

Espírito-Santo, F.D.B., Gloor, M., Keller, M., Malhi, Y., Saatchi, S., Nelson, B., Junior, R.C.O., Pereira, C., Lloyd, J., Frolking, S., Palace, M., Shimabukuro, Y.E., Duarte, V., Mendoza, A.M., López-González, G., Baker, T.R., Feldpausch, T.R., Brienen, R.J.W., Asner, G.P., Boyd, D.S., Phillips, O.L., 2014. Size and frequency of natural forest disturbances and the Amazon forest carbon balance. Nat. Commun. https://doi.org/10.1038/ncomms4434

Ferreira, M.E., Jr, L.G.F., Miziara, F., Soares-, B.S., Ferreira, L.G., Miziara, F., Soares-Filho, B.S., 2012. Modeling landscape dynamics in the central Brazilian savanna biome: future scenarios and perspectives for conservation. J. Land Use Sci. 4248, 1–19. https://doi.org/10.1080/1747423X.2012.675363

Garcia, A.S., Sawakuchi, H.O., Ferreira, M.E., Ballester, M.V.R., 2017. Landscape changes in a neotropical forest-savanna ecotone zone in central Brazil: The role of protected areas in the maintenance of native vegetation. J. Environ. Manage. 187, 16–23. https://doi.org/10.1016/j.jenvman.2016.11.010

Garcia, A.S., Vilela, V.M. de F.N., Rizzo, R., West, P., Gerber, J.S., Engstrom, P.M., Ballester, M.V.R., 2019. Assessing land use/cover dynamics and exploring drivers in the Amazon's arc of deforestation through a hierarchical, multi-scale and multi-temporal classification approach. Remote Sens. Appl. Soc. Environ. 15, 100233. https://doi.org/10.1016/j.rsase.2019.05.002

Garrett, R.D., Lambin, E.F., Naylor, R.L., 2013. The new economic geography of land use change: Supply chain configurations and land use in the Brazilian Amazon. Land use policy 34, 265–275. https://doi.org/10.1016/j.landusepol.2013.03.011

Geist, H, J., Lambin, E, F., 2002. Proximate Causes and Underlying Driving Forces of Tropical Deforestation. Bioscience 52, 143–150.

Gibbs, H.K., Rausch, L., Munger, J., Schelly, I., Morton, D.C., Noojipady, P., Soares-Filho, B., Barreto, P., Micol, L., Walker, N.F., 2015. Brazil's Soy Moratorium. Science (80-.). 347, 377–378. https://doi.org/10.1126/science.aaa0181

Hansen, M.C., Potapov, P. V, Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S. V, Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. Science (80-.). 342, 850–853. https://doi.org/10.1126/science.1244693

Hijmans, R.J., van Etten, J., Mattiuzzi, M., Sumner, M., Greenberg, J.A., Lamigueiro, O.P., Bevan, A., Racine, E.B., Shortridge, A., 2011. Raster: raster: Geographic data analysis and modeling. R Packag. version 2–0.

Hosmer, D.W., Lemeshow, S., 2000. Applied Logistic Regression, Biometrics. John Wiley & Sons, Inc., Hoboken, NJ, USA. https://doi.org/10.1002/0471722146

INPE - Instituto Nacional de Pesquisas Espaciais, 2019. PRODES Cerrado [WWW Document]. URL http://www.obt.inpe.br/cerrado/ (accessed 5.27.19).

Ivanauskas, N.M., Monteiro, R., Rodrigues, R.R., 2008. Classificação fitogeográfica das florestas do Alto Rio Xingu. Acta Amaz. 38, 387–402. https://doi.org/10.1590/S0044-59672008000300003

Kunz, S.H., Ivanauskas, N.M., Martins, S.V., Silva, E., Stefanello, D., 2009. Análise da similaridade florística entre florestas do Alto Rio Xingu, da Bacia Amazônica e do Planalto Central. Rev. Bras. Botânica 32, 725–736. https://doi.org/10.1590/S0100-84042009000400011

Laurance, W.F., Albernaz, A.K.M., Schroth, G., Fearnside, P.M., Bergen, S., Venticinque, E.M., Da Costa, C., 2002. Predictors of deforestation in the Brazilian Amazon. J. Biogeogr. 29, 737–748. https://doi.org/10.1046/j.1365-2699.2002.00721.x

Lawrence, D., Vandecar, K., 2015. Effects of tropical deforestation on climate and agriculture. Nat. Clim. Chang. https://doi.org/10.1038/nclimate2430

Maeda, E.E., de Almeida, C.M., de Carvalho Ximenes A., A., Formaggio, A.R., Shimabukuro, Y.E., Pellikka, P., 2011. Dynamic modeling of forest conversion: Simulation of past and future scenarios of rural activities expansion in the fringes of the Xingu National Park, Brazilian Amazon. Int. J. Appl. Earth Obs. Geoinf. https://doi.org/10.1016/j.jag.2010.09.008

Meyfroidt, P., Roy Chowdhury, R., de Bremond, A., Ellis, E.C., Erb, K.H., Filatova, T., Garrett, R.D., Grove, J.M., Heinimann, A., Kuemmerle, T., Kull, C.A., Lambin, E.F., Landon, Y., le Polain de Waroux, Y., Messerli, P., Müller, D., Nielsen, J., Peterson, G.D., Rodriguez García, V., Schlüter, M., Turner, B.L., Verburg, P.H., 2018. Middle-range theories of land system change. Glob. Environ. Chang. 53, 52–67. https://doi.org/10.1016/j.gloenvcha.2018.08.006

MMA-Embrapa-INPE, 2015. Mapeamento do Uso e Cobertura do Cerrado: Projeto TerraClass Cerrado 2013. Mapeamento do Uso e Cober. do Cerrado Proj. TerraClass Cerrado 2013.

Nelson, B.W., Kapos, V., Adams, J.B., Oliveira, W.J., Braun, O.P.G., 1994. Forest Disturbance by Large Blowdowns in the Brazilian Amazon. Ecology 75, 853–858. https://doi.org/10.2307/1941742

Nepstad, D.C., Stickler, C.M., Almeida, O.T., 2006. Globalization of the Amazon soy and beef industries: Opportunities for conservation. Conserv. Biol. https://doi.org/10.1111/j.1523-1739.2006.00510.x

Peres, C.A., Barlow, J., Laurance, W.F., 2006. Detecting anthropogenic disturbance in tropical forests. Trends Ecol. Evol. https://doi.org/10.1016/j.tree.2006.03.007

Perz, S., Messina, J.P., Reis, E., Walker, R., Walsh, S.J., 2009. Scenarios of future Amazonian landscapes: Econometric and dynamic simulation models, in: Amazonia and Global Change. American Geophysical Union (AGU), San Francisco, pp. 83–100. https://doi.org/10.1029/2008GM000736

Rosa, I.M.D., Purves, D., Carreiras, J.M.B., Ewers, R.M., 2014. Modelling land cover change in the Brazilian Amazon: temporal changes in drivers and calibration issues. Reg. Environ. Chang. 15, 123–137. https://doi.org/10.1007/s10113-014-0614-z

Rosa, I.M.D., Ahmed, S.E., Ewers, R.M., 2014. The transparency, reliability and utility of tropical rainforest land-use and land-cover change models. Glob. Chang. Biol. 20, 1707–1722. https://doi.org/10.1111/gcb.12523

Rosa, I.M.D., Purves, D., Souza, C., Ewers, R.M., 2013. Predictive Modelling of Contagious Deforestation in the Brazilian Amazon. PLoS One 8, e77231. https://doi.org/10.1371/journal.pone.0077231

Shimabukuro, Y.E., Santos, J.R. dos, Formaggio, A.R., Duarte, V., Rudorff, B.F.T., 2018. The Brazilian Amazon Monitoring Program: PRODES and DETER Projects, in: Achard, F., Hansen, M.C. (Eds.), Global Forest Monitoring from Earth Observation. CRC Press, pp. 166–183. https://doi.org/10.1201/b13040-13

Smith, T.B., Kark, S., Schneider, C.J., Wayne, R.K., Moritz, C., 2001. Biodiversity hotspots and beyond: the need for preserving environmental transitions. Trends Ecol. Evol. 16, 431. https://doi.org/10.1016/S0169-5347(01)02201-7

Soares-Filho, B.S., Nepstad, D.C., Curran, L.M., Cerqueira, G.C., Garcia, R.A., Ramos, C.A., Voll, E., McDonald, A., Lefebvre, P., Schlesinger, P., 2006. Modelling conservation in the Amazon basin. Nature 440, 520–523. https://doi.org/10.1038/nature04389

Strassburg, B.B.N., Brooks, T., Feltran-Barbieri, R., Iribarrem, A., Crouzeilles, R., Loyola, R., Latawiec, A.E., Oliveira Filho, F.J.B., De Scaramuzza, C.A.M., Scarano, F.R., Soares-Filho, B., Balmford, A., 2017. Moment of truth for the Cerrado hotspot. Nat. Ecol. Evol. https://doi.org/10.1038/s41559-017-0099

Tropek, R., Sedláček, O., Beck, J., Keil, P., Musilová, Z., Šímová, I., Storch, D., 2014. Comment on "High-resolution global maps of 21st-century forest cover change." Science (80-.). https://doi.org/10.1126/science.1248753

Vale, P., Gibbs, H., Vale, R., Munger, J., Brandão, A., Christie, M., Florence, E., 2019. Mapping the cattle industry in Brazil's most dynamic cattle-ranching state: Slaughterhouses in Mato grosso, 1967-2016. PLoS One 14. https://doi.org/10.1371/journal.pone.0215286

Velasquez, C.; Alves, H. Q.; Bernasconi, P., 2010. FIQUE POR DENTRO: a Bacia do Rio Xingu em Mato Grosso. nstituto Socioambiental, Instituto Centro de Vida, São Paulo.

WRI - World Resources Institute, 2014. Global Forest Watch Open Data Portal [WWW Document]. URL www.globalforestwatch.org (accessed 5.27.19).

5. GENERAL CONCLUSION

The most rapidly changing landscape in the Brazilian Legal Amazon is located in the ecotone zone between the Cerrado and the Amazon Rain Forest. This tropical ecotone zone and agricultural frontier is highly valued for its biosociodiversity, natural resources (e.g.: water and timber), and agricultural production. In this study, we intended to investigate the temporal and spatial dynamics of land use change in one of the Amazon`s Agricultural Frontier by focusing on the Upper Xingu River Basin, and by answering the following questions: (i) How has the landscape changed over the last 30 years?; (ii) What is the temporal and spatial distribution of these changes in the landscape, especially regarding the quantitative and qualitative loss of native forests?; (iii) What are the most important factors when predicting future scenarios for the Upper Xingu River Basin and what are these future scenarios?

The Upper Xingu River Basin landscape complexities, due to its natural heterogeneity and highly dynamic land use change, impose challenges for land cover and land use mapping. Different datasets showed dissimilar profiles for the region regarding both land use and cover as well as change dynamics such as deforestation. Using a combination of multiple remote sensing products and techniques, as well as GIS-based information proved to be essential when modelling land cover and land use in the studied region. In particular, utilizing flexible time series and spatial units of analysis were found to be important characteristics in our mapping approach. The LULC maps created in this thesis allowed us to differentiate among processes that drive agriculture expansion in the region. This expansion occurs in different ways - currently areas of the basin that overlaps the Amazon biome experience prominent intensification, while the Cerrado biome experiences comparatively greater expansion. In general, expansion is correlated with prices and currency exchange rates. Expansion/intensification preferences are linked to government and market regulations. These regulations are focused on the Amazon, generating intensification in the region, but these same regulations are scarce in the Cerrado biome. The increase in agriculture production in the Upper Xingu River Basin was correlated with commodity prices and monetary exchange rates. However, this increase was not spatially homogeneous. The area in the basin which overlaps with the Amazon biome experiences intensification, while the Cerrado biome experiences expansion. This pattern reflects market demands and public policies, which focus on preventing loss of rain forest in the Amazon Biome.

Similarly, the influence of such policies was observed when modelling native vegetation loss. Deforestation was more likely to be observed in the Cerrado biome, while loss of native vegetation (including other than forest) and forest degradation was not. Only few variables significantly influenced deforestation, forest degradation, and vegetation loss in the same way. Other variables had different influences according to the process, indicating that not the three processes can be tackled in the same way. Although disturbance also causes tremendous impacts to biodiversity, carbon stocks, and other ecosystem functions, discussion about the role of policies in preventing disturbance is not substantial.

Nevertheless, we observed that transition areas such as the UXRB, with diverse phytophysiognomies, present particular challenges for remote sensing. Still, different data sources offer different perspectives on loss of natural areas and their processes. These differences are linked to each workgroup's framework and their objectives, but it is important to note that the divergences are not evenly distributed in the landscape. It is recommended that users of this data be cautious in choosing which data to use, taking into account the framework under which the data to be used was generated, the process of interest, and the vegetation types to be studied.

Based on our results and the discussion in the literature, we conclude that general models and framework interpretations of land use change should not be applied in areas as diverse as the Upper Xingu River Basin. Additionally, policies towards specific regions and dynamics such as deforestation in the Amazon Rain Forest will impact both positively and negatively other dynamics. As we have shown, these policies can promote agriculture intensification in target areas, but a leakage in non-target areas and processes is also likely to be observed.

APPENDICES

Appendix A

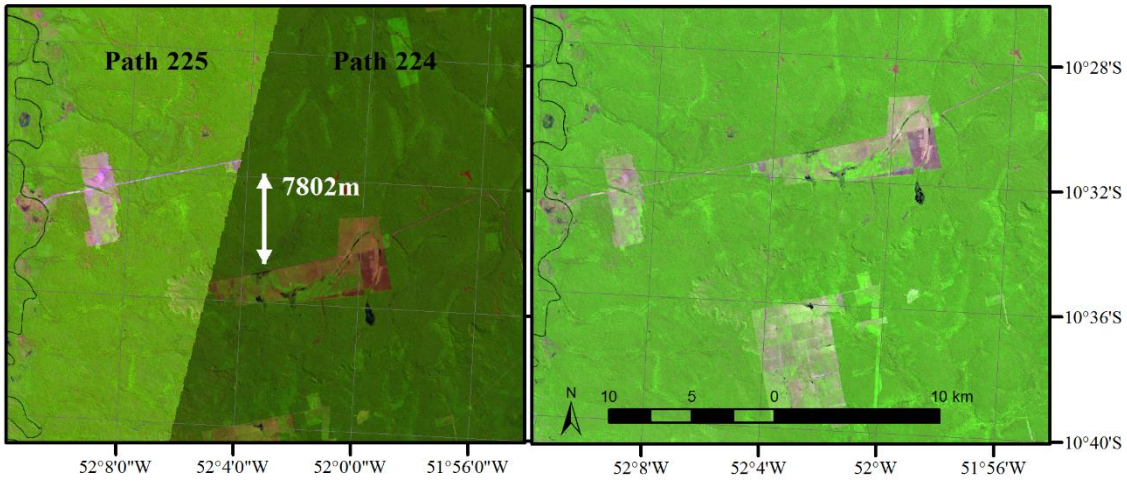
Supplementary material to support Chapter 2: Assessing land use/cover dynamics and exploring drivers in the Amazon's Arc of Deforestation through a hierarchical, multi-scale and multi-temporal classification approach

Supplementary material 1. Here we present the list of Landsat images used in the classification process and the geometric distortion found on Landsat images.

SM 1 - Table 1. List of the 108 Landsat images used to compose the mosaics and to map LCLU. First column indicates the image position (paths and rows). First line and second lines indicate satellite and image acquisition year, respectively. Specified dates represents the dates each image (path/row) was acquired for a given year. All images we obtained at the dry season in order to avoid clouds.

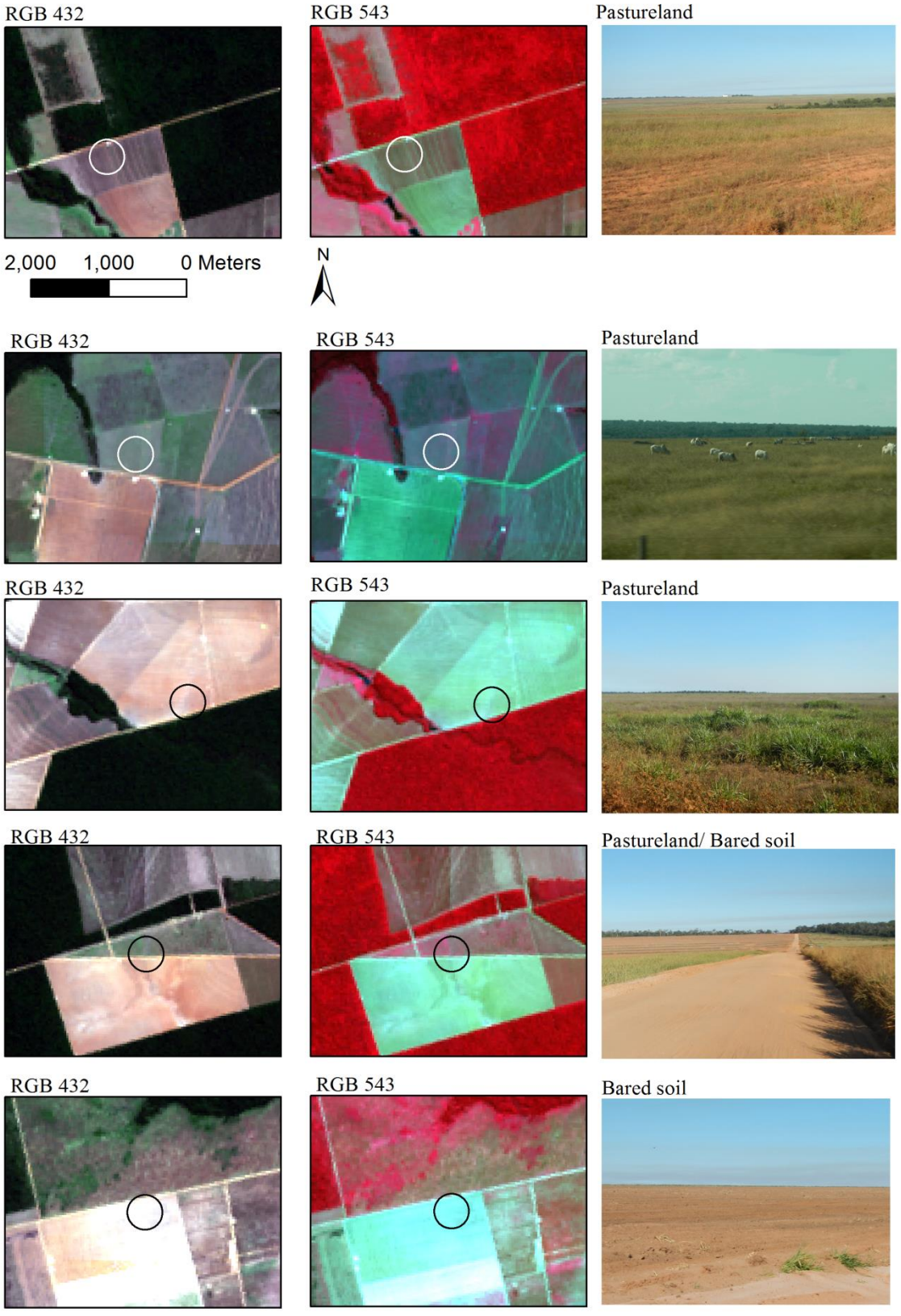
			Landsat 5 TM		
Path/row	**1985**	**1990**	**1995**	**2005**	**2010**
224/67	3/Jun	1/Jun	2/Aug	12/Jul	26/Jul
224/68	3/Jun	1/Jun	2/Aug	12/Jul	26/Jul
224/69	3/Jun	1/Jun	1/Jul	12/Jul	26/Jul
224/70	3/Jun	1/Jun	1/Jul	12/Jul	26/Jul
225/67	10/Jun	8/Jun	8/Jul	3/Jul	15/Jun
225/68	10/Jun	8/Jun	8/Jul	15/Jun	15/Jun
225/69	10/Jun	8/Jun	6/Jun	3/Jul	15/Jun
225/70	10/Jun	8/Jun	6/Jun	3/Jul	15/Jun
226/67	17/Jun	15/Jun	13/Jun	10/Jul	24/Jul
226/68	17/Jun	2/Aug	13/Jun	10/Jul	24/Jul
226/69	20/Ago	2/Aug	13/Jun	10/Jul	24/Jul
226/70	17/Jun	2/Aug	13/Jun	10/Jul	24/Jul
	Landsat 7 ETM+			**Landsat 8 OLI**	
Path/row	**2000**	**2005***	**2010***	**2014**	**2015**
224/67	7/Aug	6/Sep	3/Aug	5/Jul	9/Aug
224/68	7/Aug	20/Jul	3/Aug	5/Jul	9/Aug
224/69	7/Aug	20/Jul	3/Aug	5/Jul	9/Aug
224/70	7/Aug	20/Jul	3/Aug	5/Jul	9/Aug
225/67	29/Jul	27/Jul	7/Jun	12/Jul	31/Jul
225/68	29/Jul	27/Jul	7/Jun	12/Jul	31/Jul
225/69	29/Jul	27/Jul	7/Jun	12/Jul	31/Jul
225/70	29/Jul	27/Jul	7/Jun	12/Jul	31/Jul
226/67	18/Jun	18/Jul	1/Aug	3/Jul	7/Aug
226/68	5/Aug	18/Jul	1/Aug	3/Jul	7/Aug
226/69	5/Aug	18/Jul	1/Aug	3/Jul	7/Aug
226/70	5/Aug	18/Jul	1/Aug	3/Jul	7/Aug

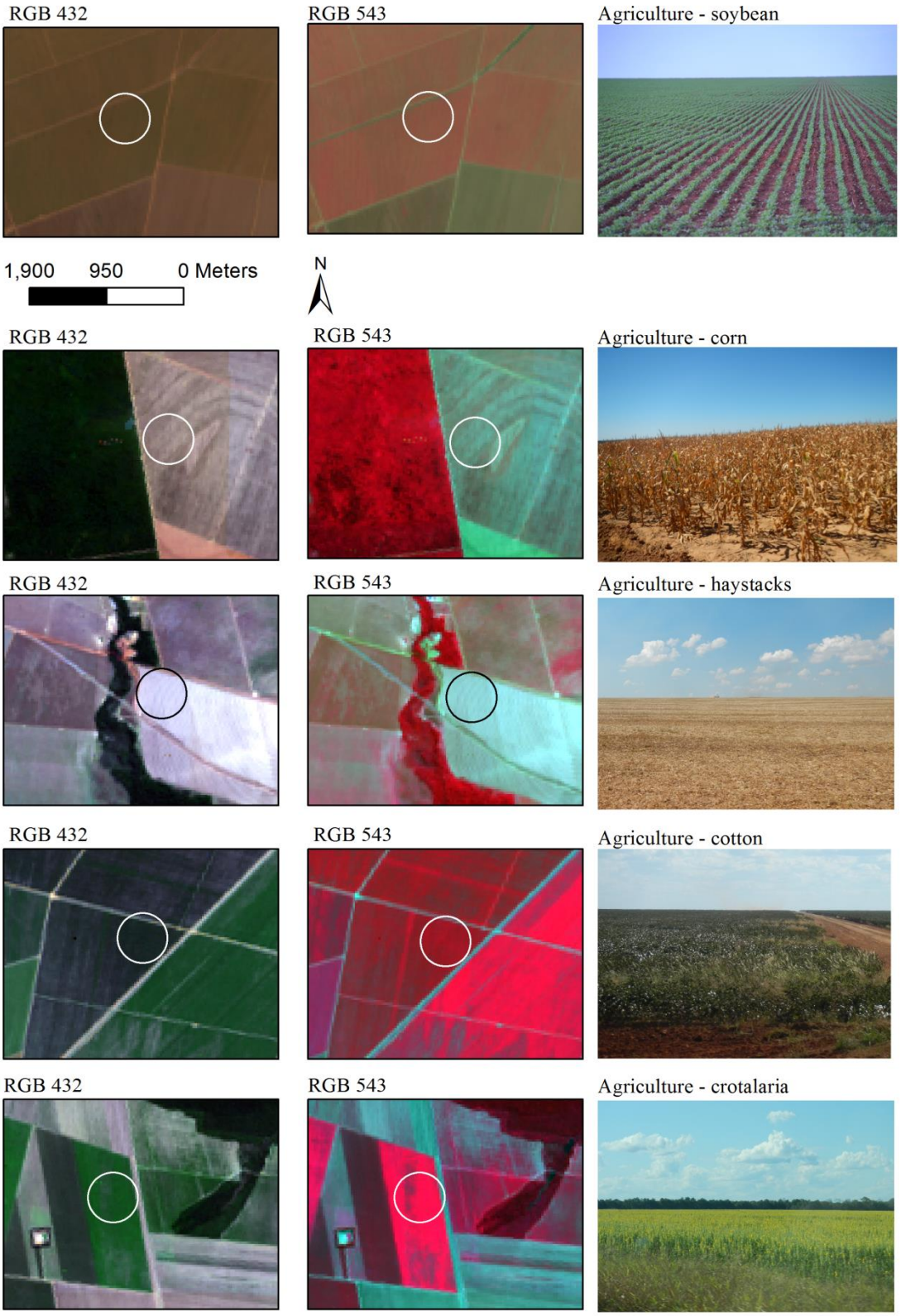
*SLC-off Products.



SM 1 - Figure 1. Landsat image acquired with USGS/Landsat Higher Level Science Data products (http://espa.cr.usgs.gov/) from year 1985. Left panel shows path 225 and 224, which are spatially unallied. Right panel shows both images after corrections.

Supplementary material 2. Land use covers in the Upper Xingu River Basin. Photos are taken at the signed point (red triangles) in each image by pointing the camera to the circles near triangles. First column shows Landsat 8 (OLI sensor) images in true colour composition, second column shows a false colour composition using near infrared to enhance vegetation. Both photos and satellite images were acquired in June 2015.




RGB 432
RGB 543
Agriculture - soybean
1,900 950 0 Meters
N
RGB 432
RGB 543
Agriculture - corn
RGB 432
RGB 543
Agriculture - haystacks
RGB 432
RGB 543
Agriculture - cotton
RGB 432
RGB 543
Agriculture - crotalaria

Supplementary material 3. Main agriculture crops in Mato Grosso and their temporal cultivation stages according to agriculture calendar and temporal profiles produced by Arvor et al. (2012), Gusso et al. (2014), and Zhu et al. (2016). Note that the main crop is primarily soybean.

Months	Climate conditions	Crop stages	Soybean	Corn	Cotton
Oct	Rainfall frequency picks up	Soil preparation - sowing	Planting[1]	Planting[1]	
Nov		Growth	Planting[1]	Planting[1]	
Dec					Planting[1]
Jan	Peak of the rainy season	Sowing / Harvest	Early harvest[1] / Planting[2]	Planting[2]	Planting[2]
Feb		Growth / Harvest	Harvest[1] / Planting[2]	Harvest[1] / Planting[2]	Harvest[1] / Planting[2]
Mar		Growth / Harvest	Harvest[1]	Harvest[1] / Planting[2]	
Apr	Scattered rain	Growth / Harvest	Harvest[2]	Harvest[1]	
May	Dry season starts	Harvest	Harvest[2]	Harvest[2]	Harvest[2]
Jun		Harvest	Soybean-Free	Harvest[2]	Harvest[2]
Jul			Soybean-Free	Harvest[2]	Harvest[2]
Aug			Soybean-Free		
Sep	Dry season is ending	Soil preparation - sowing	Soybean-Free / Early planting[1]	Early planting	

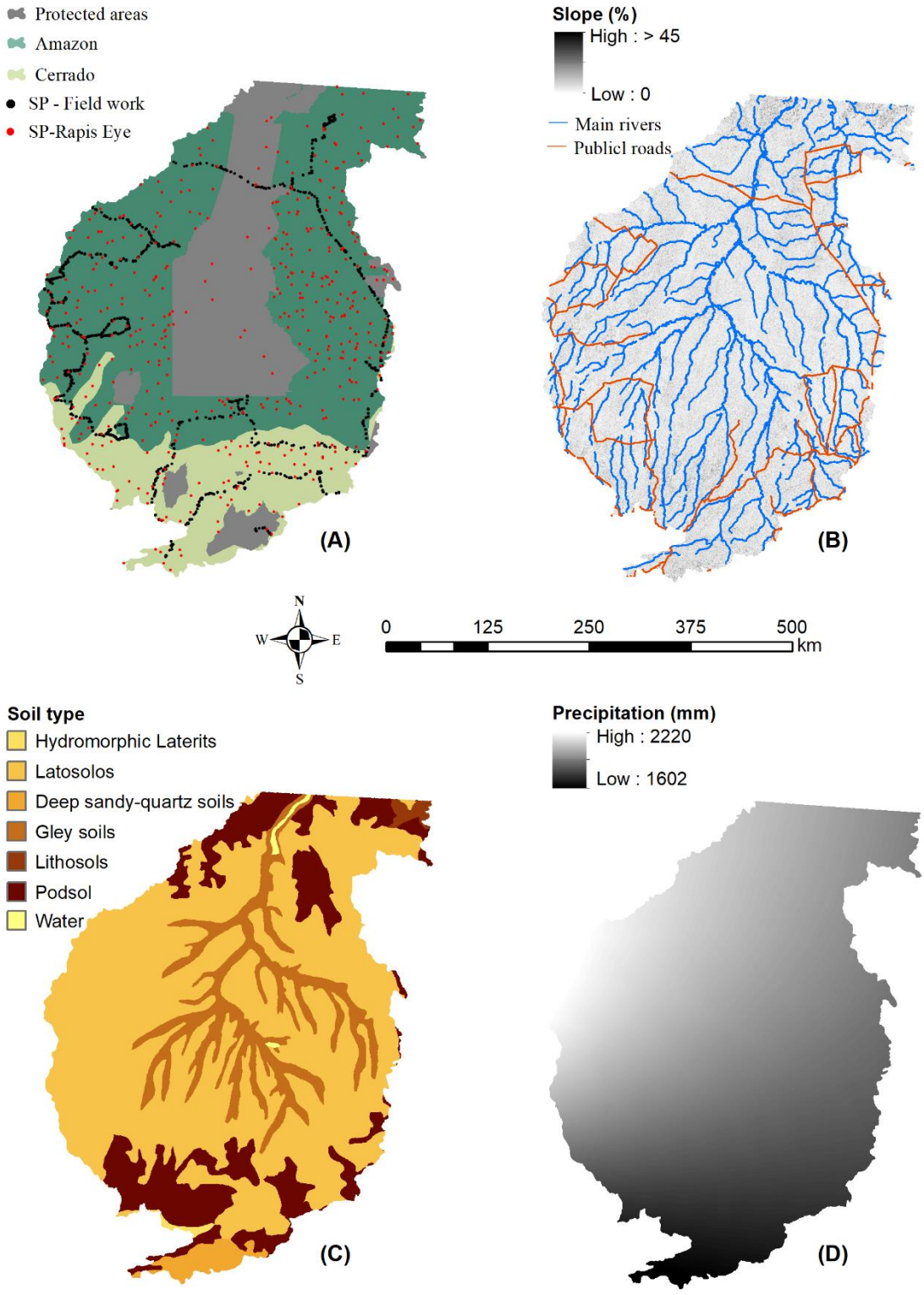
[1]Main crop; [2]Second crop ("safrinha").

Supplementary material 4. This section presents in further detail methods and results for accuracy assessment.

We assessed the accuracy of the modelled land cover maps comparing them with reference data collected in field work and obtained from very high-resolution Rapid Eye images collected in 2012 and 2013. Reference data collected from Rapid Eye images were only used for classes 'water', 'wetlands' and 'urban' because these classes are assumed to be more stable through time. Thus, land use changes would not disrupt comparisons between modelled and reference. To offer more insights into the quality of our land cover and land use maps, we also built confusion matrices by comparing our maps with (1) reference data collected to build land cover maps by IPAM for the municipality of Querência, (2) land cover maps produced by ISA for the municipalities of Canarana, Querência, Santa Cruz do Xingu, e São José do Xingu, (3) comparing cropland area reported in municipal census. Such comparisons are considered secondary as the legend developed for these products had to be adapted to match the legend of the maps we produced, and therefore, adding extra error.

To build the reference dataset based on field work, we performed a systematic and non-equidistributed sampling of information in the basin in 2016. As shown in Figure 1, the largest biophysical variation in the area is consistent with a latitudinal pattern; then, we designed our fieldwork in order to sample the largest amount of points following main roads from north to south, in the east and west portion of the basin. We drove 3500 km inside the UXRB and collected 1460 points of reference. The routine consisted in follow main roads and collect data every ~2 km for every land cover immediately around the stop point. We registered all information by date, geographic location (GPS Garmin Montana 600, WAAS-enabled, interpolating the minimum of 5 satellites), land cover and land use class, brief description, and photographic record.

An ideal reference data should be randomized and equidistributed (stratified or not), which means that every pixel has the same (or almost the same) chance of integrate the sample population (Olofsson et al., 2013; Stehman, 2001). Still, it is recognized that depending on the study area, such sample design is not feasible to implement (Congalton, 2008). According to Stehman (2001), the main process of making sampling design practical is to allocate efforts where collecting data is more likely. Although different strategies can be applied, the increase complexity a study area imposes difficulties in the calculation of the probability of inclusion of pixel in the sample. The UXRB, located in the north portion of the state of Mato Grosso and in the border with Pará, is a difficult area to be assessed in order to collect information for random located or equidistributed points. The access to the region is limited due to (i) violent conflicts related to land tenure, logging, poaching, mining, illegal drug trade, among other issues connected or not with environmental regulation, (ii) the absence of the State as a regulatory institution, resulting in violence and impunity in daily bases, (iii) private landowners rarely allow access to their lands for different reasons, (iv) and the authors had no federal authorization to go inside conservation units and/or indigenous reserves. With the access restriction in the UXRB, each unit (pixel) in the area does not have an equal chance of being included in the sample and, therefore, we did not calculate statistical measures within the confusion matrix and area as demonstrated by Olofsson and colleagues (2014). Systematic and non-equidistributed sampling presents no randomness and, thus, it is statistically weak. Still, theoretically, the overall number of samples can mitigates the lack of randomness by introducing more variability, even though such mitigation could deprecate within each classes (Congalton, 2008; Stehman, 2001).



SM 4 - Figure 1. Distribution of (A) protected areas (both indigenous land and conservation units - MMA), Amazon and Cerrado biomes (IBGE) and reference sample points (SR) collected in field work and collected in very high-resolution image (Rapid Eye); (B) rivers (ANA), public roads (DNIT) and slope (INPE); (C) soil types (IBGE); (D) Annual precipitation (ANA).

Here, we used the method presented by Pontius and Millones (2011), which assumes that sampling intensity may differ between classes. As recommended by the authors, we calculated disagreement and decomposed such metric in quantity disagreement and allocation disagreement – overall and per class. Refer to Pontius and Millones (2011) to calculation of estimated population matrix and metrics. From Table 1 to 10, we show all results.

SM 4 - Table 1. Accuracy metrics for the Upper Xingu River Basin, Brazil. Metrics calculated from estimated population matrix. Reference was composed of field work points collected in 2016, and Rapid Eye images collected from 2012/13. The matrix represents the accuracy results for Level 1 of classification.

			Disagreement	
	Agreement	Disagreement	Quantity	Allocation
Overall	0.93	0.07	0.04	0.02
			Disagreement	
	Omission	Commission	Quantity	Allocation
Natural and Semi-Natural Vegetation	0.03	0.01	0.02	0.02
Cultivated and Managed Terrestrial Areas	0.01	0.06	0.04	0.02
Surface water	0.03	0	0.03	0

SM 4 - Table 2. Accuracy metrics for the Upper Xingu River Basin, Brazil. Metrics calculated from estimated population matrix. Reference was composed of field work points collected in 2016, and Rapid Eye images collected from 2012/13. The matrix represents the accuracy results for Level 2 of classification.

			Disagreement	
	Agreement	Disagreement	Quantity	Allocation
Overall	0.87	0.13	0.09	0.05
			Disagreement	
	Omission	Commission	Quantity	Allocation
Forest	0	0.06	0.06	0.01
Savanic formations	0.03	0.01	0.02	0.02
Wetlands	0.03	0	0.03	0
Secondary complex	0	0	0	0
Agriculture	0.03	0.01	0.02	0.03
Pasturelands	0.02	0.05	0.03	0.04
Bare land	0.01	0	0.01	0
Urban	0	0	0	0
Water	0.01	0	0.01	0

SM 4 - Table 3. Accuracy metrics for the Upper Xingu River Basin, Brazil. Metrics calculated from estimated population matrix. Reference was composed of field work points collected in 2016, and Rapid Eye images collected from 2012/13. The matrix represents the accuracy results for Level 3 of classification.

			Disagreement	
	Agreement	Disagreement	Quantity	Allocation
Overall	0.84	0.16	0.10	0.06
			Disagreement	
	Omission	Commission	Quantity	Allocation
Forest	0	0.07	0.07	0
Forest cerrado	0.03	0.01	0.02	0.01
Woody cerrado	0.01	0.01	0	0.02
Shrub-grassland cerrado	0.01	0	0.01	0
Wetland	0.03	0	0.03	0
Secondary complex	0.01	0	0	0
Single crop	0.01	0.01	0	0.02
Double crop	0.01	0.01	0	0.01
Irrigated crop	0	0	0	0
Pastureland	0.02	0.05	0.03	0.04
Degraded pastureland	0.02	0	0.02	0.01
Urban	0	0	0	0
Bare soil	0.01	0	0.01	0
Water	0.01	0	0.01	0

SM 4 - Table 4. Accuracy metrics for the Upper Xingu River Basin, Brazil. Metrics calculated from estimated population matrix. Maps produced by IPAM for the municipality of Querencia in 2017 was used as reference. The matrix represents the accuracy results for Level 1 of classification.

			Disagreement	
	Agreement	Disagreement	Quantity	Allocation
Overall	0.80	0.20	0.13	0.07
			Disagreement	
	Omission	Commission	Quantity	Allocation
Natural and Semi-Natural Vegetation	0.05	0.14	0.09	0.10
Cultivated and Managed Terrestrial Areas	0.02	0.06	0.04	0.04
Surface water	0.13	0	0.13	0

SM 4 - Table 5. Accuracy metrics for the Upper Xingu River Basin, Brazil. Metrics calculated from estimated population matrix. Maps produced by IPAM for the municipality of Querencia in 2017 was used as reference. The matrix represents the accuracy results for Level 2 of classification.

			Disagreement	
	Agreement	Disagreement	Quantity	Allocation
Overall	0.75	0.25	0.20	0.04
			Disagreement	
	Omission	Commission	Quantity	Allocation
Forest	0.01	0.17	0.16	0.02
Wetlands	0.02	0.01	0.01	0.01
Secondary complex	0.15	0	0.15	0
Agriculture	0.01	0.05	0.03	0.02
Pasturelands	0.02	0.03	0.01	0.03
Urban	0	0	0	0
Water	0.04	0	0.04	0

SM 4 - Table 6. Accuracy metrics for the Upper Xingu River Basin, Brazil. Metrics calculated from estimated population matrix. Maps produced by IPAM for the municipality of Querencia in 2017 was used as reference. The matrix represents the accuracy results for Level 3 of classification.

			Disagreement	
	Agreement	Disagreement	Quantity	Allocation
Overall	0.64	0.35	0.15	0.20
			Disagreement	
	Omission	Commission	Quantity	Allocation
Wetland	0.02	0.04	0.01	0.05
Secondary complex	0.08	0	0.08	0
Single crop	0.07	0.18	0.10	0.15
Double crop	0.02	0.06	0.04	0.05
Irrigated crop	0	0	0	0
Pastureland	0.08	0.08	0	0.16
Urban	0.01	0	0.01	0
Water	0.06	0	0.06	0

SM 4 - Table 7. Accuracy metrics for the Upper Xingu River Basin, Brazil. Metrics calculated from estimated population matrix. Maps produced by ISA for the municipality of Canarana, Querencia, Santa Cruz do Xingu, and São José do Xingu in 2010 was used as reference. The matrix represents the accuracy results for Level 1 of classification.

			Disagreement	
	Agreement	Disagreement	Quantity	Allocation
Overall	0.89	0.11	0.09	0.02
			Disagreement	
	Omission	Commission	Quantity	Allocation
Natural and Semi-Natural Vegetation	0.01	0.1	0.09	0.02
Cultivated and Managed Terrestrial Areas	0.1	0.01	0.09	0.02
Surface water	0	0	0	0

SM 4 - Table 8. Accuracy metrics for the Upper Xingu River Basin, Brazil. Metrics calculated from estimated population matrix. Maps produced by ISA for the municipality of Canarana, Querencia, Santa Cruz do Xingu, and São José do Xingu in 2010 was used as reference. The matrix represents the accuracy results for Level 2 of classification.

			Disagreement	
	Agreement	Disagreement	Quantity	Allocation
Overall	0.90	0.09	0.06	0.04
			Disagreement	
	Omission	Commission	Quantity	Allocation
Forest	0.01	0.05	0.04	0.02
Savanic formations	0	0.02	0.02	0
Wetlands	0.02	0	0.02	0
Agriculture	0.02	0	0.02	0
Pasturelands	0.04	0.02	0.02	0.04
Water	0	0	0	0

SM 4 - Table 9. Accuracy metrics for the Upper Xingu River Basin, Brazil. Metrics calculated from estimated population matrix. Maps produced by IPAM for the municipalities of Canarana and Querencia in 2015 was used as reference. was used as reference. The matrix represents the accuracy results for Level 1 of classification.

			Disagreement	
	Agreement	Disagreement	Quantity	Allocation
Overall	0.96	0.04	0.03	0.02
			Disagreement	
	Omission	Commission	Quantity	Allocation
Natural and Semi-Natural Vegetation	0.01	0.03	0.02	0.02
Cultivated and Managed Terrestrial Areas	0.03	0.01	0.02	0.02
Surface water	0	0	0	0

SM 4 - Table 10. Accuracy metrics for the Upper Xingu River Basin, Brazil. Metrics calculated from estimated population matrix. Maps produced by IPAM for the municipalities of Canarana and Querencia in 2015 was used as reference. The matrix represents the accuracy results for Level 2 of classification.

			Disagreement	
	Agreement	Disagreement	Quantity	Allocation
Overall	0.94	0.06	0.03	0.03
			Disagreement	
	Omission	Commission	Quantity	Allocation
Forest	0.01	0.01	0.01	0.02
Wetlands	0	0	0	0
Agriculture	0.04	0.01	0.03	0.02
Pasturelands	0.01	0.03	0.02	0.02
Water	0	0	0	0

As other metrics are still used by a large amount of researches, we also present the observed confusion matrix for each level of classification along with overall accuracy, overall accuracy' two-tailed binomial test and its upper and lower level, user's accuracy, producer's accuracy, balanced accuracy, and kappa coefficient of agreement.

SM 4 - Table 11. Accuracy measurements employed in this study.

Accuracy measurement	Computation		
Overall accuracy: proportion of points mapped correctly	$Oac = \frac{n_s}{n_t}$	Oac	Overall accuracy
		n_s	Number of success = $\sum x_{ii}$, where x_{ii} is the success for each class
		n_t	Total number of samples
Standard deviation of overall accuracy	$S = \sqrt{\frac{Oac \times (1 - Oac)}{n_t}}$	S	Standard deviation
		Oac	Overall accuracy
		n_t	Total number of samples
User's accuracy: proportion of points which were mapped in a certain class that is actually in that class based on ground reference	$U = \frac{x_{ii}}{x_{i+}}$	U	User's accuracy
		x_{ii}	Number of successes when comparing modelled and reference
		x_{i+}	Marginal sum of a modelled class (row)
Producer's accuracy: proportion of reference points in a certain class which is mapped as the same certain class	$P = \frac{x_{ii}}{x_{+i}}$	P	Producer's accuracy
		x_{ii}	Number of successes when comparing modelled and reference
		x_{+j}	Marginal sum of a reference class (column)
Balanced accuracy: average of sensitivity (rate of true positives or producer's accuracy) and specificity (rate of true negatives)	$B = \frac{P + \frac{n_t - x_{+i} - x_{i+}}{\sum x_{+i} - x_{+i}}}{2}$	B	Balanced accuracy
		P	Producer's accuracy
		n_t	Total number of samples
		x_{+j}	Marginal sum of a reference class (column)
		x_{i+}	Marginal sum of a modelled class (row)

SM 4 - Table 12. Full error matrix comparing modelled land cover and land use information with reference points collected in Upper Xingu River Basin, Brazil, in 2016, and Rapid Eye images collected from 2012/13. The matrix represents the accuracy results for Level 1 of classification.

		Reference					
		Natural	Managed	Water	Total	User's accuracy	Balanced accuracy
Model	Natural and Semi-Natural Vegetation	644	57	7	708	91%	95%
	Cultivated and Managed Terrestrial Areas	30	858	29	917	94%	94%
	Surface water	1	0	323	324	100%	95%
	Total	675	915	359	**1825**		
	Producer's accuracy	95%	94%	90%			**1949**
	Overall accuracy		94%	(92%; 95%)			
	Average user's accuracy		95%				
	Average producer's accuracy		96%				
	Accuracy p-value		0.00				

SM 4 - Table 13. Full error matrix comparing modelled land cover and land use information with reference points collected in Upper Xingu River Basin Brazil, in 2016, and Rapid Eye images collected from 2012/13. The matrix represents the accuracy results for Level 2 of classification.

		Reference											
		For	Sav	Wet	Sec	Agr	Pas	Bar	Urb	Wat	Total	User's accuracy	Balanced accuracy
Model	Forest	352	16	13	0	1	4	1	0	2	389	90%	96%
	Savanic formations	8	130	12	7	2	12	2	0	0	173	75%	89%
	Wetlands	5	5	99	0	0	0	0	0	3	112	88%	85%
	Secondary complex	2	1	4	14	0	4	4	0	0	29	48%	80%
	Agriculture	1	1	2	1	382	25	2	0	5	419	91%	89%
	Pasturelands	2	10	7	1	65	271	13	2	7	378	72%	88%
	Bare land	1	0	1	0	23	11	46	0	19	101	46%	82%
	Urban	0	0	0	0	0	0	0	32	0	32	100%	97%
	Water	0	0	1	0	0	0	0	0	316	317	100%	95%
	Total	371	163	139	23	473	327	68	34	352	**1642**		
	Producer's accuracy	95%	80%	71%	61%	81%	83%	68%	94%	90%			**1950**
	Overall accuracy			84%	(83%; 86%)								
	Average user's accuracy			79%									
	Average producer's accuracy			89%									
	Accuracy p-value			0.00									

SM 4 - Table 14. Full error matrix comparing modelled land cover and land use information with reference points collected in Upper Xingu River Basin, Brazil, in 2016, and Rapid Eye images collected from 2012/13. The matrix represents the accuracy results for Level 3 of classification.

		Reference																
		For	Fce	Wce	SGc	Wet	Sec	SCr	DCr	Icr	Pas	Dpa	Urb	Bar	Wat	Total	User's accuracy	Balanced accuracy
Model	Forest	279	11	2	0	10	2	0	0	0	3	0	0	1	3	312	90%	97%
	Forest cerrado	0	49	13	0	3	0	0	0	0	0	0	0	0	0	66	75%	83%
	Woody cerrado	3	7	82	16	10	4	0	0	0	5	5	0	2	0	135	61%	81%
	Shrub-grassland cerrado	0	4	22	10	2	1	0	0	0	2	0	0	0	0	42	24%	65%
	Wetland	3	1	3	0	104	1	0	0	0	0	0	0	0	3	116	90%	86%
	Secondary complex	3	0	1	0	4	12	0	0	0	4	1	0	4	0	30	41%	77%
	Single crop	0	0	0	0	1	0	76	11	0	23	1	0	2	4	119	64%	85%
	Double crop	0	1	1	0	1	1	17	71	3	1	0	0	0	1	98	73%	89%
	Irrigated crop	0	0	0	0	0	0	0	0	8	0	0	0	0	0	9	100%	86%
	Pastureland	0	0	0	4	5	1	7	4	0	177	27	0	15	4	245	73%	84%
	Degraded pastureland	0	0	2	2	0	0	1	0	0	22	25	2	1	3	59	43%	69%
	Urban	0	0	0	0	0	0	0	0	0	0	0	32	0	0	33	100%	97%
	Bare soil	1	0	0	0	1	0	3	4	0	5	3	0	43	17	78	56%	81%
	Water	0	0	0	0	1	0	0	0	0	0	0	0	0	323	325	100%	95%
	Total	290	74	127	32	143	23	105	91	12	243	62	35	69	359	**1291**		
	Producer's accuracy	97%	67%	65%	31%	73%	55%	73%	79%	73%	73%	40%	94%	63%	90%			**1664**
	Overall accuracy		78%	(76%; 80%)														
	Average user's accuracy		71%															
	Average producer's accuracy		70%															
	Accuracy p-value		0.00															

SM 4 - Table 15. Full error matrix comparing modelled land cover and land use information with land use maps produced by IPAM for the municipality of Querencia in 2017. The matrix represents the accuracy results for Level 1 of classification.

		Reference					
		Natural	Managed	Water	Total	User's accuracy	Balanced accuracy
Model	Natural and Semi-Natural Vegetation	55	2	14	71	77%	82%
	Cultivated and Managed Terrestrial Areas	15	96	3	114	84%	90%
	Surface water	1	0	11	12	92%	69%
	Total	71	98	28	**162**		
	Producer's accuracy	77%	98%	39%			**197**
	Overall accuracy		82%	(76%; 87%)			
	Average user's accuracy		84%				
	Average producer's accuracy		72%				
	Accuracy p-value		0.00				

SM 4 - Table 16. Full error matrix comparing modelled land cover and land use information with land use maps produced by IPAM for the municipality of Querencia in 2017. The matrix represents the accuracy results for Level 2 of classification.

		Reference									
		Forest	Wetlands	Secondary	Agriculture	Pasturelands	Urban	Water	Total	User's accuracy	Balanced accuracy
Model	Forest	39	1	9	0	0	0	2	51	76%	90%
	Wetlands	0	6	0	0	0	0	10	16	38%	77%
	Secondary	0	0	1	0	0	0	0	1	100%	53%
	Agriculture	1	1	6	41	4	0	0	53	77%	88%
	Pasturelands	4	2	2	8	36	1	3	56	64%	88%
	Urban	0	0	0	0	0	4	0	4	100%	90%
	Water	0	0	1	0	0	0	11	12	92%	71%
	Total	44	10	19	49	40	5	26	**138**		
	Producer's accuracy	89%	60%	5%	84%	90%	80%	42%			**193**
	Overall accuracy		72%	(65%; 78%)							
	Average user's accuracy		78%								
	Average producer's accuracy		64%								
	Accuracy p-value		0.00								

SM 4 - Table 17. Full error matrix comparing modelled land cover and land use information with land use maps produced by IPAM for the municipality of Querencia in 2017. The matrix represents the accuracy results for Level 3 of classification.

		Reference										
		Wetland	Secondary complex	Single crop	Double crop	Irrigated crop	Pastureland	Urban	Water	Total	User's accuracy	Balanced accuracy
Model	Wetland	6	2	0	0	0	0	0	8	16	38%	79%
	Secondary complex	0	2	0	0	0	0	0	0	2	100%	58%
	Single crop	1	3	24	2	0	7	0	2	39	62%	78%
	Double crop	0	3	4	12	0	0	0	0	19	63%	90%
	Irrigated crop	0	0	0	0	1	0	0	0	1	100%	100%
	Pastureland	2	2	6	0	0	33	1	1	45	73%	85%
	Urban	0	0	0	0	0	0	4	0	4	100%	90%
	Water	0	1	0	0	0	0	0	11	12	92%	75%
	Total	9	13	34	14	1	40	5	22	**93**		
	Producer's accuracy	67%	15%	71%	86%	100%	83%	80%	50%			**138**
	Overall accuracy		67%	(59%; 75%)								
	Average user's accuracy		78%									
	Average producer's accuracy		69%									
	Accuracy p-value		0.00									

SM 4 - Table 18. Full error matrix comparing modelled land cover and land use information with land use maps produced by ISA for the municipalities of Canarana, Querencia, Santa Cruz do Xingu, and São José do Xingu in 2010. The matrix represents the accuracy results for Level 1 of classification.

		Reference					
		Natural	Managed	Water	Total	User's accuracy	Balanced accuracy
Model	Natural and Semi-Natural Vegetation	1250	173	6	1429	87%	88%
	Cultivated and Managed Terrestrial Areas	31	683	0	714	96%	89%
	Surface water	6	0	6	12	50%	75%
	Total	1287	856	12	**1939**		
	Producer's accuracy	97%	80%	50%			**2155**
	Overall accuracy		90%	(89%; 91%)			
	Average user's accuracy		78%				
	Average producer's accuracy		76%				
	Accuracy p-value		0.00				

SM 4 - Table 19. Full error matrix comparing modelled land cover and land use information with land use maps produced by ISA for the municipalities of Canarana, Querencia, Santa Cruz do Xingu, and São José do Xingu in 2010. The matrix represents the accuracy results for Level 2 of classification.

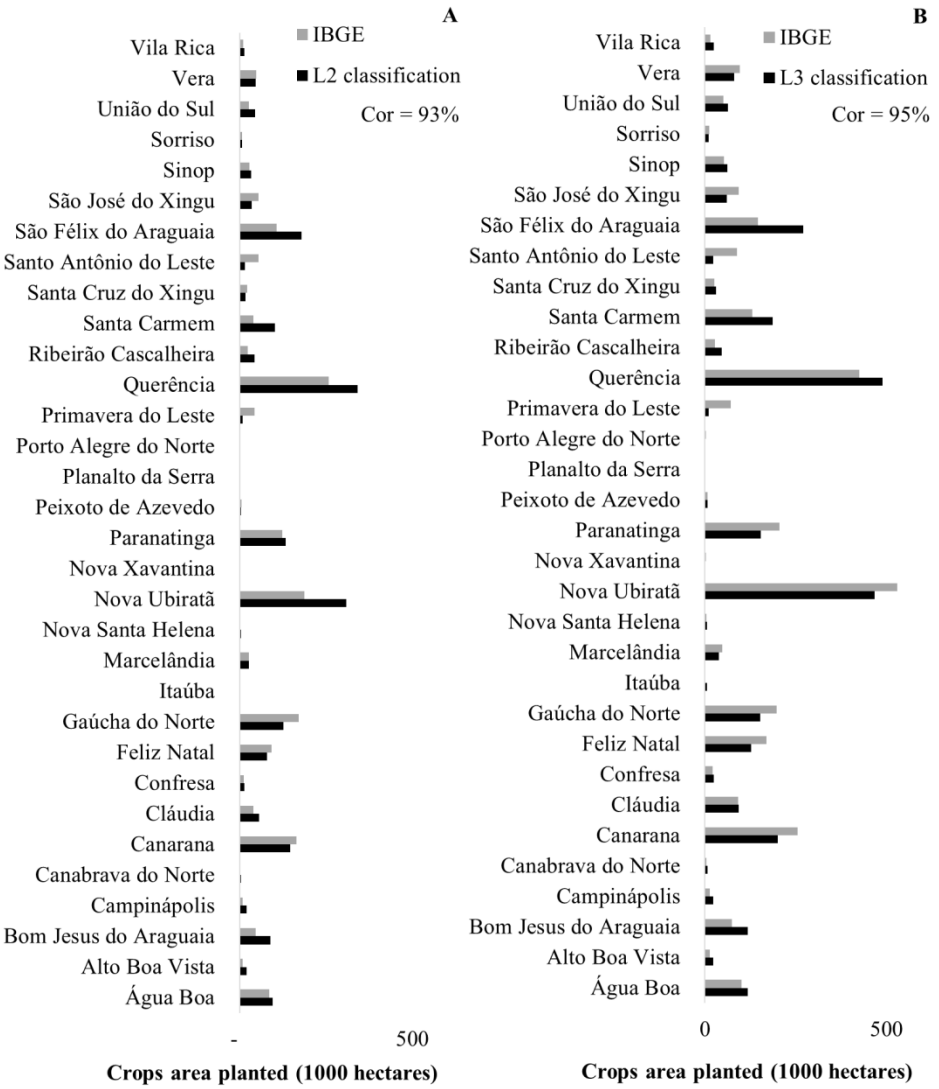
		Reference								
		Forest	Savanic	Wetlands	Agriculture	Pasturelands	Water	Total	User's accuracy	Balanced accuracy
Model	Forest	792	1	19	7	22	2	843	94%	92%
	Savanic formations	54	1	0	9	103	2	169	1%	70%
	Wetlands	15	0	4	0	1	1	21	19%	58%
	Agriculture	0	0	0	33	17	0	50	66%	66%
	Pasturelands	13	0	0	49	444	0	506	88%	85%
	Water	0	0	0	0	0	4	4	100%	72%
	Total	874	2	23	98	587	9	**1278**		
	Producer's accuracy	91%	50%	17%	34%	76%	44%			**1593**
	Overall accuracy		80%	(78%; 82%)						
	Average user's accuracy		61%							
	Average producer's accuracy		52%							
	Accuracy p-value		0.00							

SM 4 - Table 20. Full error matrix comparing modelled land cover and land use information with land use maps produced by ISA for the municipalities of Canarana and Querencia in 2015. The matrix represents the accuracy results for Level 1 of classification.

		Reference					
		Natural	Managed	Water	Total	User's accuracy	Balanced accuracy
Model	Natural and Semi-Natural Vegetation	575	24	2	601	96%	95%
	Cultivated and Managed Terrestrial Areas	16	421	0	437	96%	96%
	Surface water	6	0	2	8	25%	75%
	Total	597	445	4	**998**		
	Producer's accuracy	96%	95%	50%			**1046**
	Overall accuracy		95%	(94%; 97%)			
	Average user's accuracy		72%				
	Average producer's accuracy		80%				
	Accuracy p-value		0.00				

SM 4 - Table 21. Full error matrix comparing modelled land cover and land use information with land use maps produced by ISA for the municipalities of Canarana and Querencia in 2015. The matrix represents the accuracy results for Level 2 of classification.

		Reference							
		Forest	Wetlands	Agriculture	Pasturelands	Water	Total	User's accuracy	Balanced accuracy
Model	Forest	512	1	4	3	1	521	98%	95%
	Wetlands	17	0	0	0	0	17	0%	49%
	Agriculture	5	0	181	11	0	197	92%	85%
	Pasturelands	11	0	69	160	0	240	67%	91%
	Water	6	0	0	0	2	8	25%	83%
	Total	551	1	254	174	3	**855**		
	Producer's accuracy	93%	0%	71%	92%	67%			**983**
	Overall accuracy		87%	(85%; 89%)					
	Average user's accuracy		56%						
	Average producer's accuracy		65%						
	Accuracy p-value		0.00						



SM 4 - Figure 2. Correlation between the LCLU maps produced for Upper Xingu River Basin for croplands and the census data made available by IBGE for soybean, maize and cotton – those crops made up the great majority of cropland in the studied basin. (A) L2 cropland area and IBGE planted area data for soybean discounting maize (secondary crop or "safrinha"), main crop maize, and cotton. (B) L3 cropland defined by single crop area and twice the double crop area and IBGE planted area data for soybean, maize, and cotton.

REFERENCES

Congalton, R., 2008. Sample Design Considerations, in: Assessing the Accuracy of Remotely Sensed Data, Mapping Science. CRC Press, New york, pp. 63–83. https://doi.org/10.1201/9781420055139

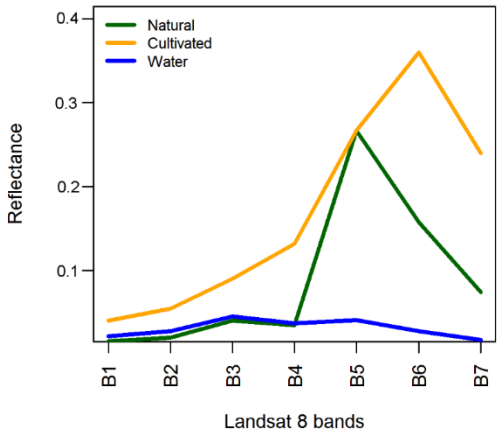
Olofsson, P., Foody, G.M., Herold, M., Stehman, S. V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2014.02.015

Olofsson, P., Foody, G.M., Stehman, S. V., Woodcock, C.E., 2013. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sens. Environ. 129, 122–131. https://doi.org/10.1016/j.rse.2012.10.031

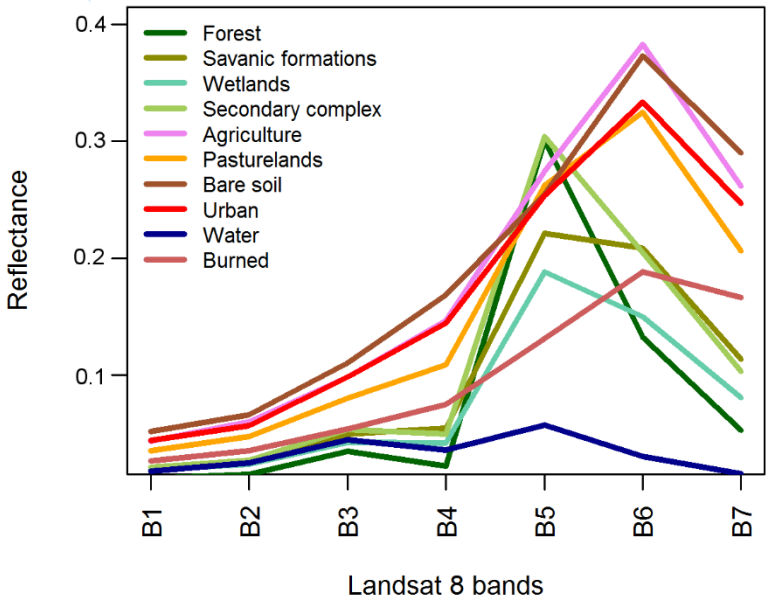
Pontius, R.G., Millones, M., Pontius, Robert, Gilmore, J., Millones, M., Pontius, R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int. J. Remote Sens. 32, 4407–4429. https://doi.org/10.1080/01431161.2011.552923

Stehman, S. V, 2001. Statistical Rigor and Practical Utility in Thematic Map Accuracy Assessment. ISPRS J. Photogramm. Remote Sens. 67, 727–734.

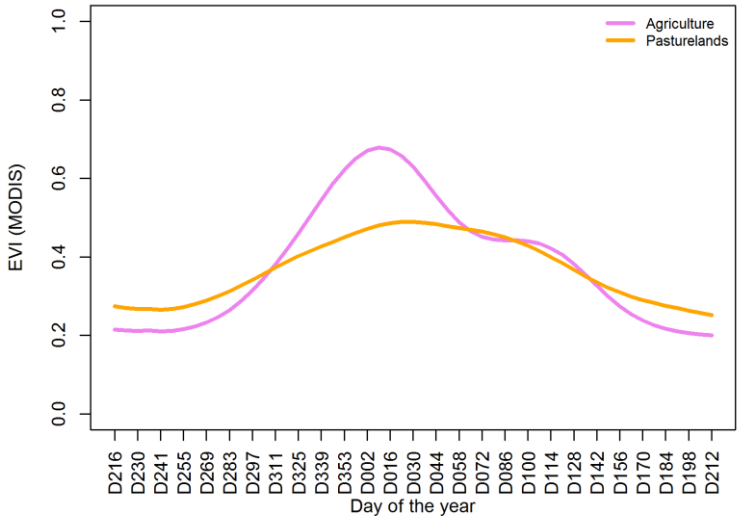
Supplementary material 5. This section presents spectral profiles of land cover and land use classes proposed for Upper Xingu River Basin according to Landsat bands and EVI MODIS yearly variation. See each figure's legend for details.



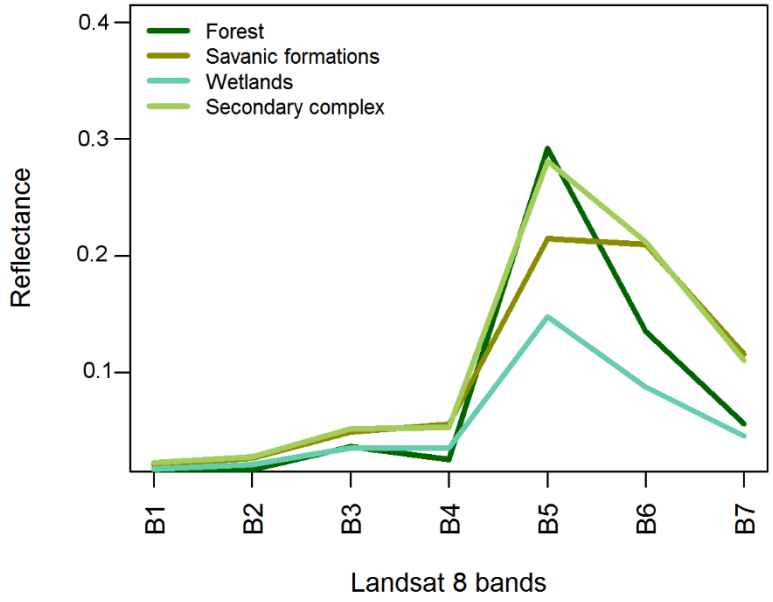
SM 5 - Figure 1. Spectral profile based on Landsat 8 bands. The classes shown in the graph were used to classify the land cover and land use in the Upper Xingu River Basin at Level 1 according to the methods presented in the manuscript. Profile was drwan based on the calculated mean value of each Landsat 8 band sampled with ground true points colected for accuracy assessesment.



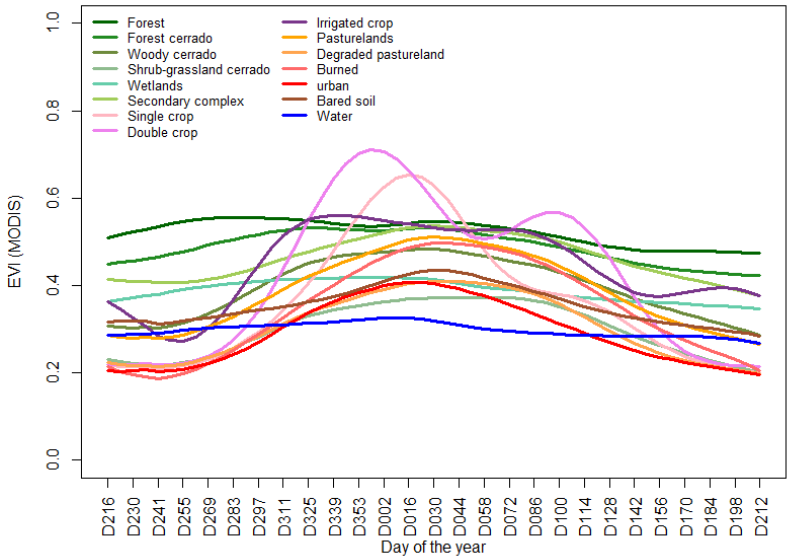
SM 5 - Figure 2. Spectral profile based on Landsat bands. The classes shown in the graph were used to classify the land cover and land use in the Upper Xingu River Basin at Level 2 according to the methods presented in the manuscript. Profile was drwan based on the calculated mean value of each Landsat 8 band sampled with ground true points colected for accuracy assessesment.



SM 5 - Figure 3. Spectral profile based on year EVI MODIS composition. The classes shown in the graph were used to classify the land cover and land use in the Upper Xingu River Basin at Level 2 according to the methods presented in the manuscript. Profile was drwan based on the calculated mean value of EVI band (MODIS Aqua and Terra) sampled with ground true points colected for accuracy assessesment.

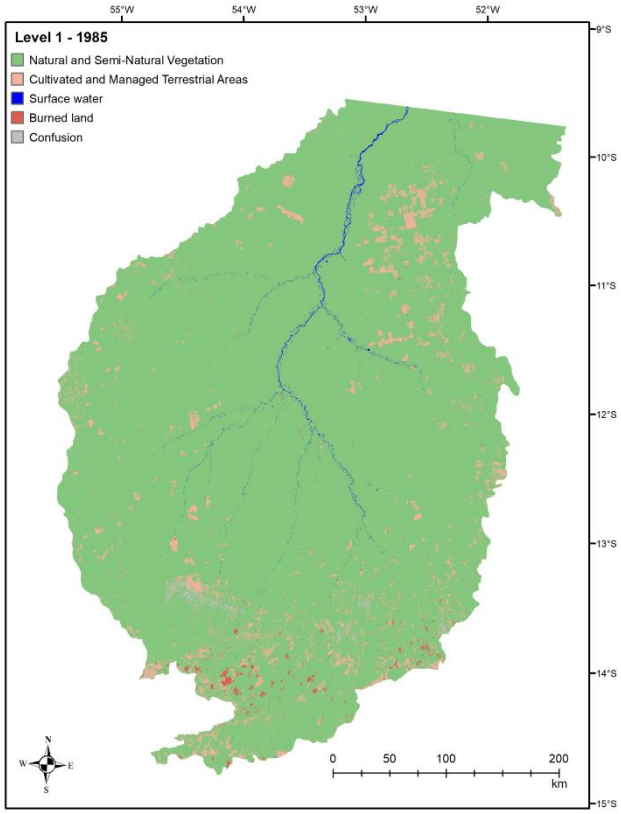


SM 5 - Figure 4. Spectral profile based on Landsat bands. The classes shown in the graph were used to classify the land cover and land use in the Upper Xingu River Basin at Level 3 according to the methods presented in the manuscript. Profile was drwan based on the calculated mean value of each Landsat 8 band sampled with ground true points colected for accuracy assessesment.

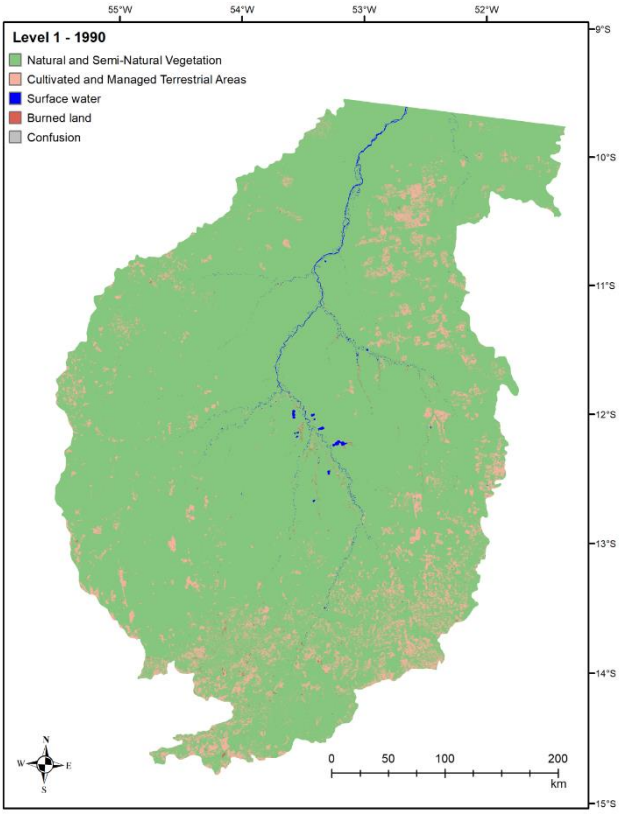


SM 5 - Figure 5. Spectral profile based on year EVI MODIS composition. The classes shown in the graph were used to classify the land cover and land use in the Upper Xingu River Basin at Level 3 according to the methods presented in the manuscript. Profile was drwan based on the calculated mean value of EVI band (MODIS Aqua and Terra) sampled with ground true points colected for accuracy assessesment.

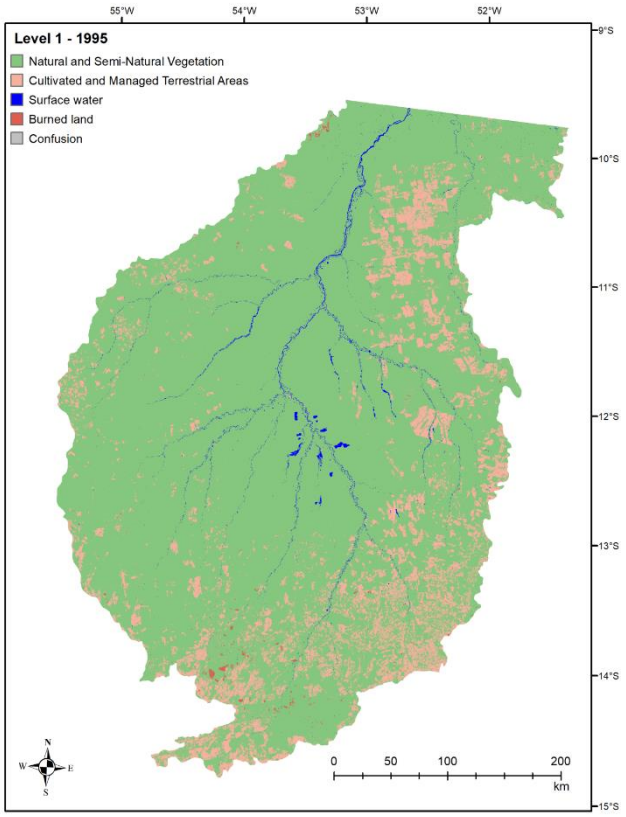
Supplementary material 6. Land cover and land use maps of Upper Xingu River Basin in three different levels and seven different years: L1 (1985, 1990, 1995, 2000, 2005, 2010, 2015), L2 (1985, 1990, 1995, 2000, 2005, 2010, 2015), and L3 (2000, 2005, 2010, 2015).



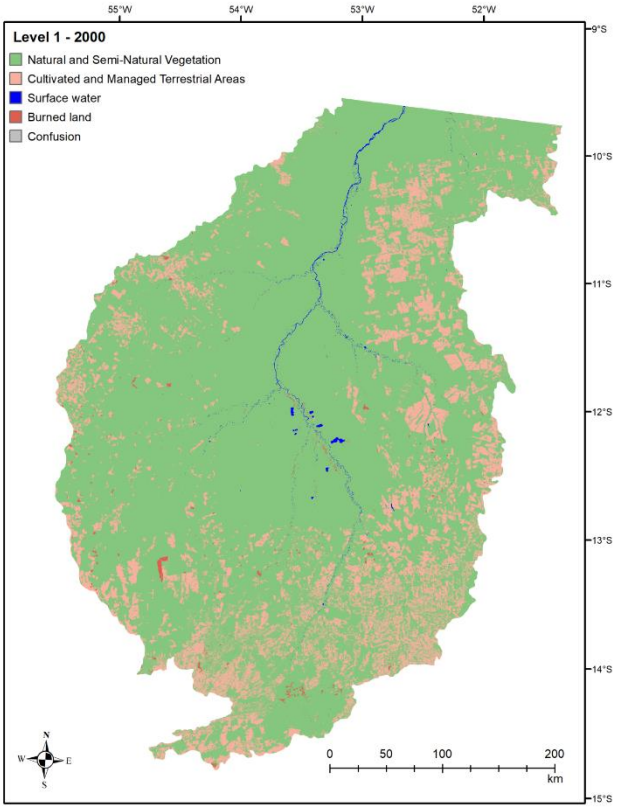
SM 6 - Figure 1. Level 1 of land cover and land use classification scheme for the Upper Xingu River Basin in 1985.



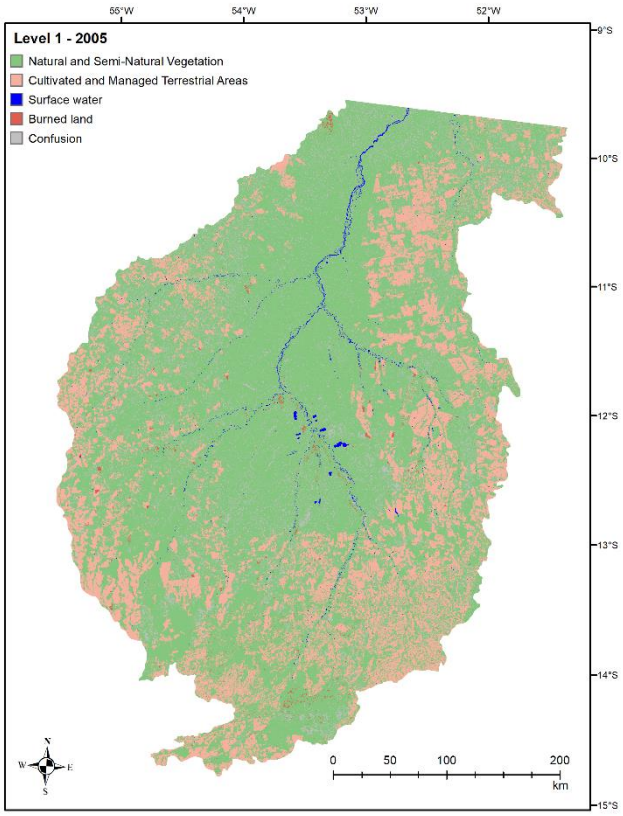
SM 6 - Figure 2. Level 1 of land cover and land use classification scheme for the Upper Xingu River Basin in 1990.



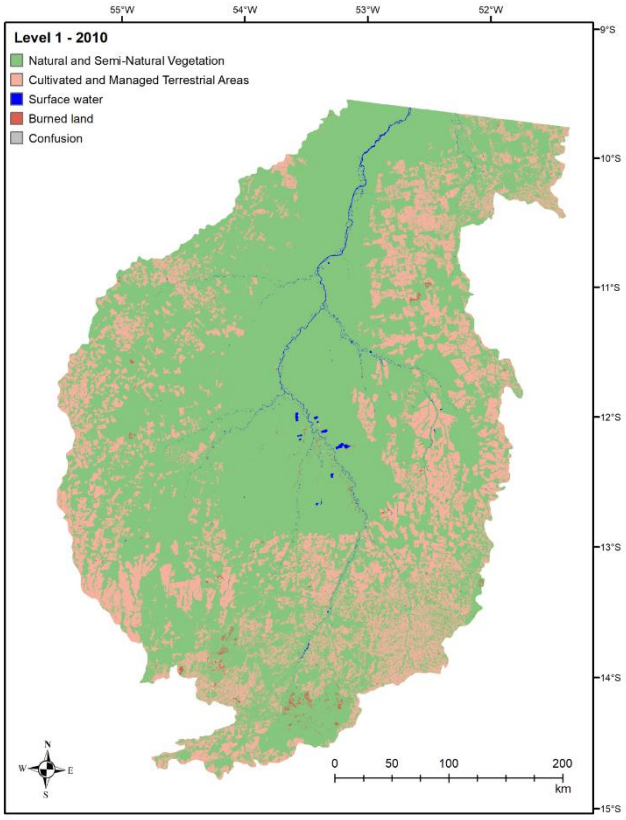
SM 6 - Figure 3. Level 1 of land cover and land use classification scheme for the Upper Xingu River Basin in 1995.



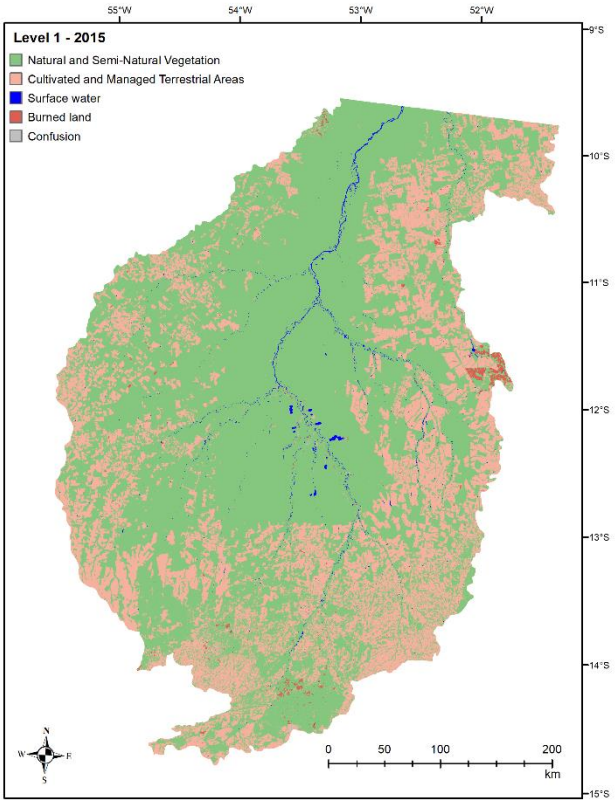
SM 6 - Figure 4. Level 1 of land cover and land use classification scheme for the Upper Xingu River Basin in 2000.



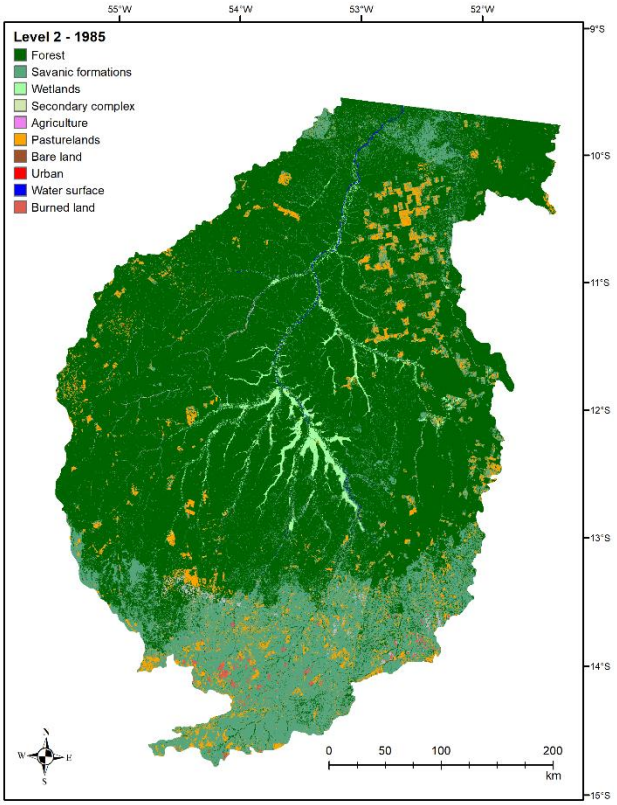
SM 6 - Figure 5. Level 1 of land cover and land use classification scheme for the Upper Xingu River Basin in 2005.



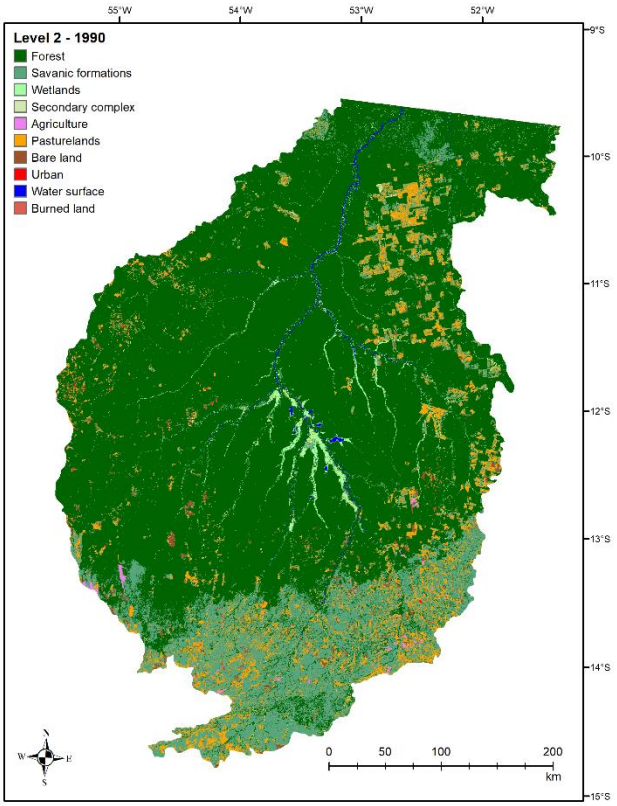
SM 6 - Figure 6. Level 1 of land cover and land use classification scheme for the Upper Xingu River Basin in 2010.



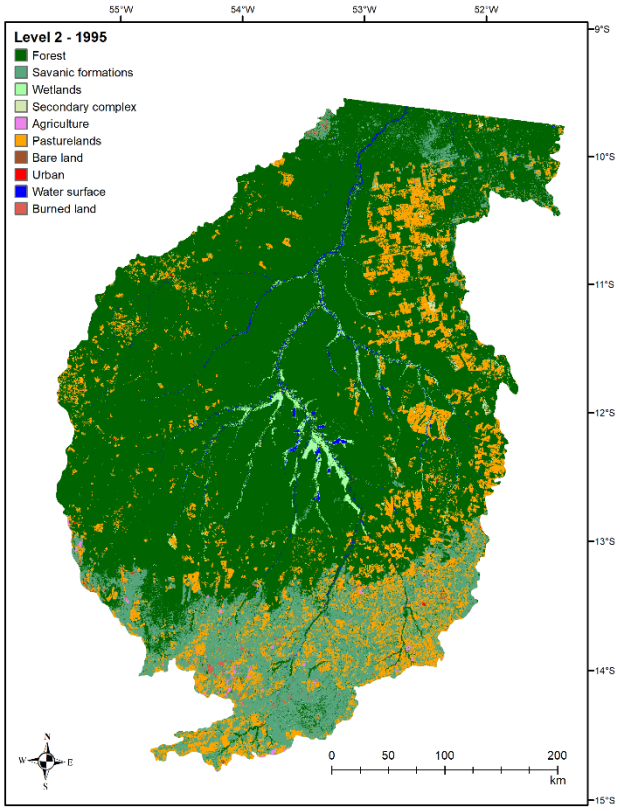
SM 6 - Figure 7. Level 1 of land cover and land use classification scheme for the Upper Xingu River Basin in 2015.



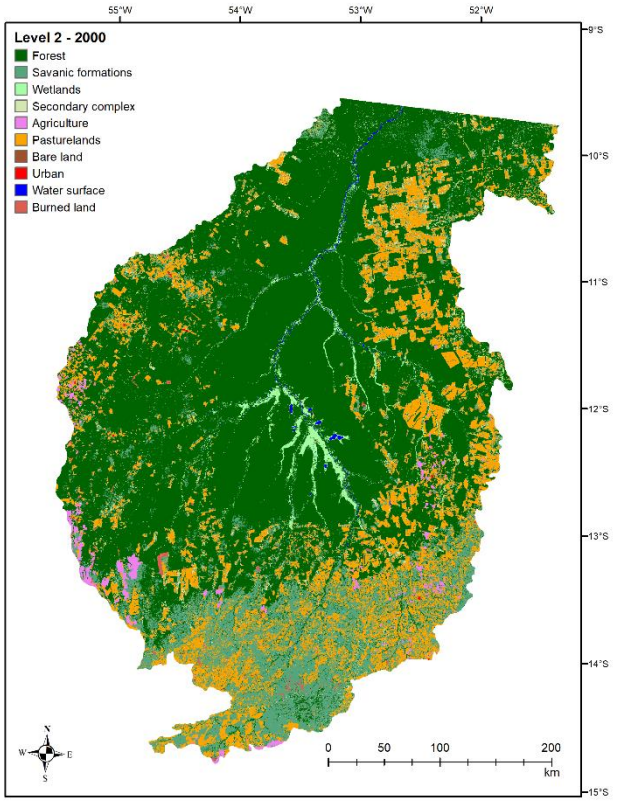
SM 6 - Figure 8. Level 2 of land cover and land use classification scheme for the Upper Xingu River Basin in 1985.



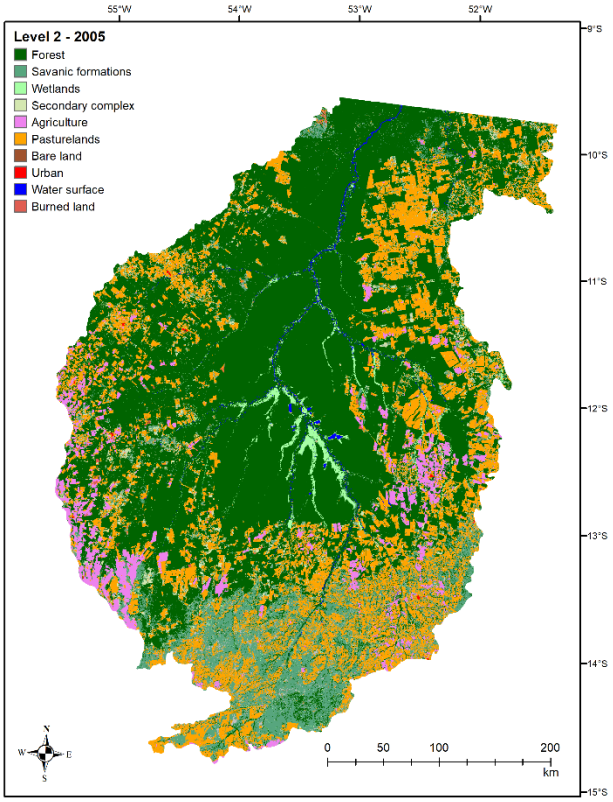
SM 6 - Figure 9. Level 2 of land cover and land use classification scheme for the Upper Xingu River Basin in 1990.



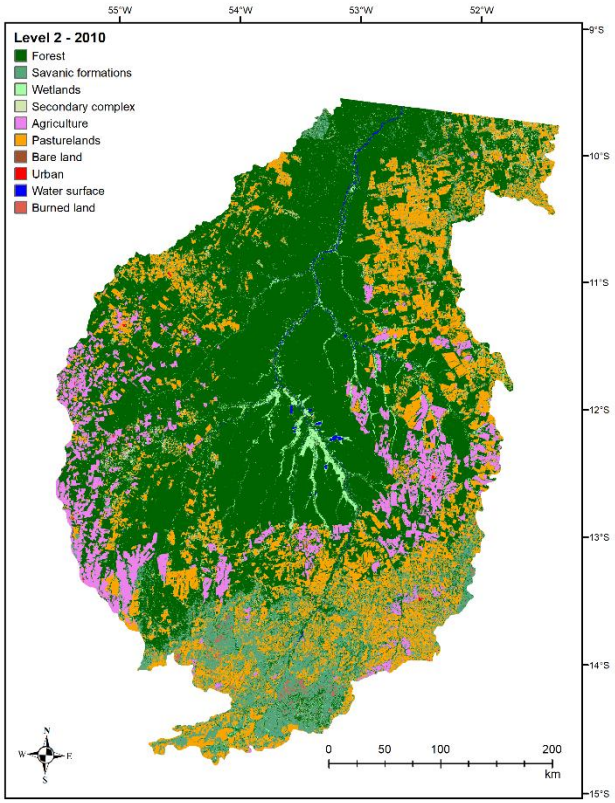
SM 6 - Figure 10. Level 2 of land cover and land use classification scheme for the Upper Xingu River Basin in 1995.



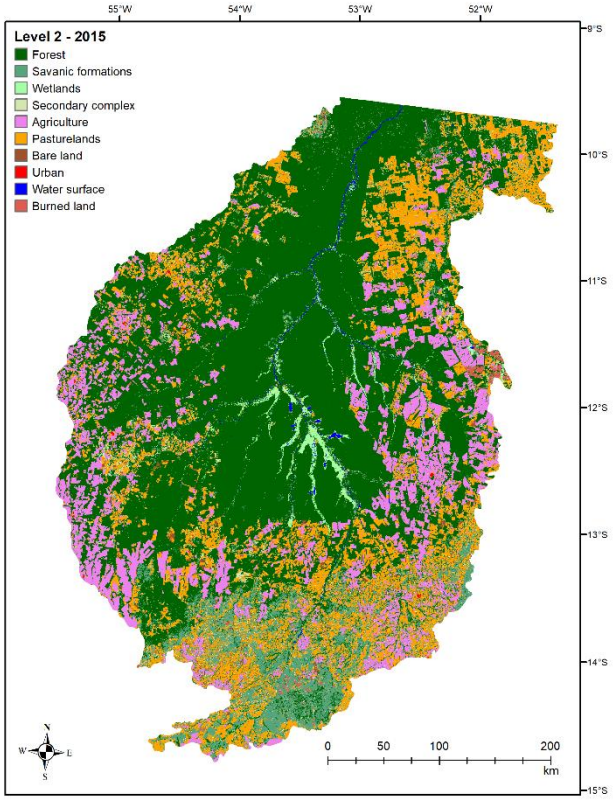
SM 6 - Figure 11. Level 2 of land cover and land use classification scheme for the Upper Xingu River Basin in 2000.



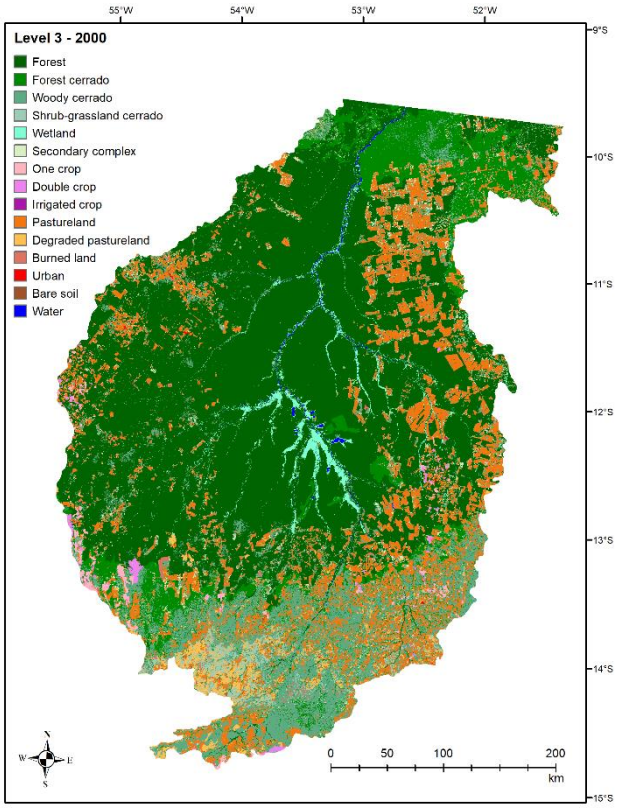
SM 6 - Figure 12. Level 2 of land cover and land use classification scheme for the Upper Xingu River Basin in 2005.



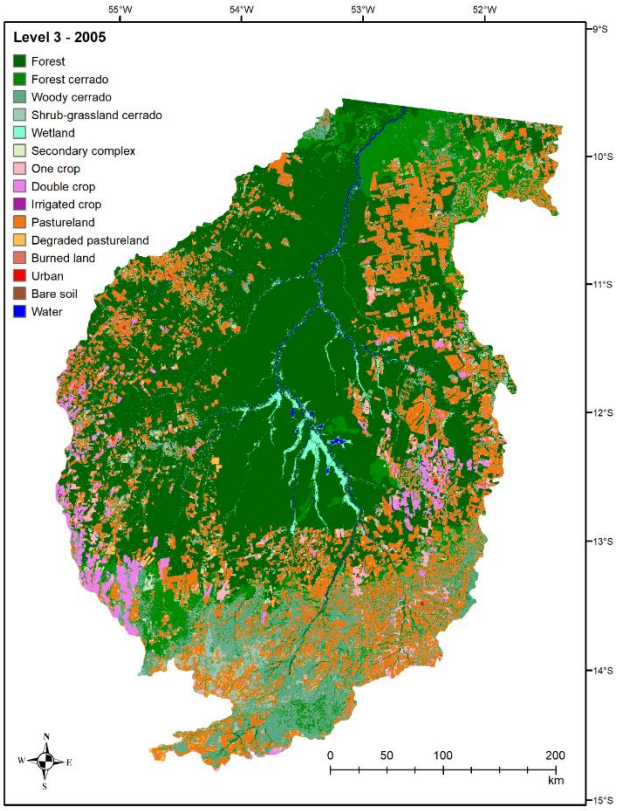
SM 6 - Figure 13. Level 2 of land cover and land use classification scheme for the Upper Xingu River Basin in 2010.



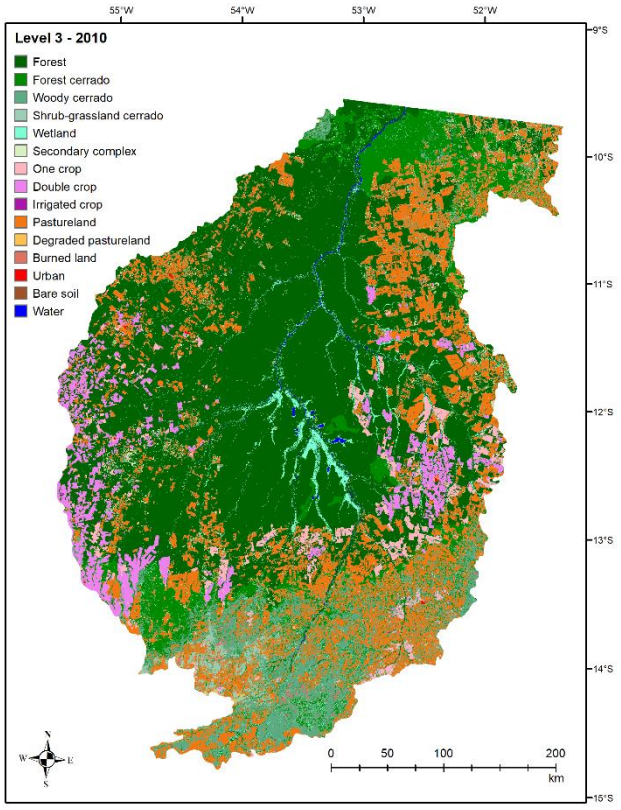
SM 6 - Figure 14. Level 2 of land cover and land use classification scheme for the Upper Xingu River Basin in 2015.



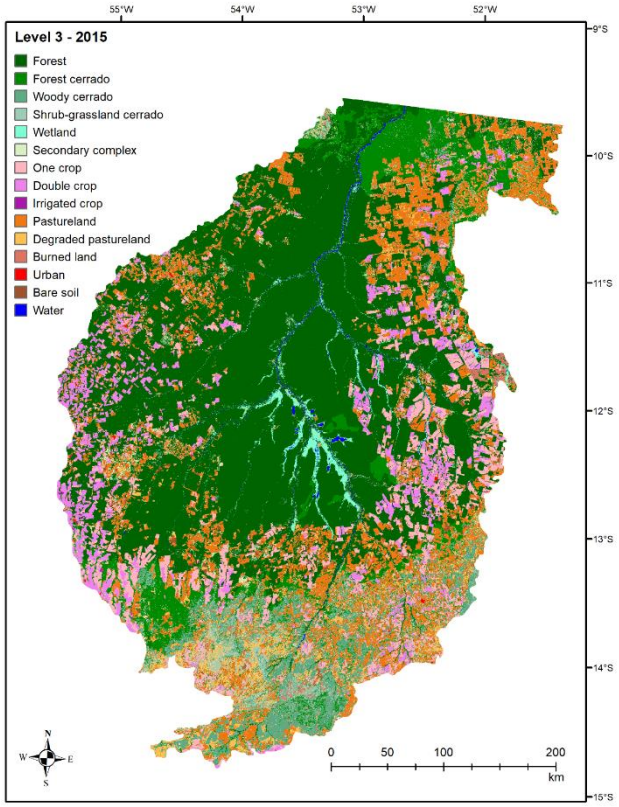
SM 6 - Figure 15. Level 3 of land cover and land use classification scheme for the Upper Xingu River Basin in 2000.



SM 6 - Figure 16. Level 3 of land cover and land use classification scheme for the Upper Xingu River Basin in 2005.



SM 6 - Figure 17. Level 3 of land cover and land use classification scheme for the Upper Xingu River Basin in 2010.



SM 6 - Figure 18. Level 3 of land cover and land use classification scheme for the Upper Xingu River Basin in 2015.

Supplementary material 7. Land cover and land use transition tables and rates of change.

SM7 - Table 1. Land cover change matrix according to Level 1 of classification for the Upper Xingu River Basin, MT, Brazil. Indexes equal to the percentage of a class (rows) that changes to another class in the next 5 years period (columns). Grey shadings represent no change. Land cover classes representing less than 1% of the basin area are not shown.

		1990			
		Natural	**Managed**	**TOTAL (km²)**	**TOTAL (%)**
1985	**Natural**	94	5	160,228	93
	Managed	49	51	7,423	4
		1995			
		Natural	**Managed**	**TOTAL (km²)**	**TOTAL (%)**
1990	**Natural**	93	7	155,314	91
	Managed	27	73	12,622	7
		2000			
		Natural	**Managed**	**TOTAL (km²)**	**TOTAL (%)**
1995	**Natural**	92	8	147,539	84
	Managed	15	85	20,500	12
		2005			
		Natural	**Managed**	**TOTAL (km²)**	**TOTAL (%)**
2000	**Natural**	88	11	123,883	78
	Managed	13	86	28,466	17
		2010			
		Natural	**Managed**	**TOTAL (km²)**	**TOTAL (%)**
2005	**Natural**	92	8	113,020	69
	Managed	13	87	39,073	23
		2015			
		Natural	**Managed**	**TOTAL (km²)**	**TOTAL (%)**
2010	**Natural**	91	8	124,110	68
	Managed	6	93	43,720	31
TOTAL 2015 (km²)		145	143		
TOTAL 2015 (%)		65	0.1		

SM 7 - Table 2. Land cover change matrix according to Level 2 of classification for the Upper Xingu River Basin, MT, Brazil. Indexes equal to the percentage of a class (rows) that changes to another class in the next 5 years period (columns). Grey shadings represent no change. Land cover classes representing less than 1% of the basin area are not shown.

		1990						
		For	**Sav**	**Wet**	**Crop**	**Past**	**TOTAL (km²)**	**TOTAL (%)**
1985	**Forest (For)**	93	4	1	0	2	122,148	72
	Savanna formations (Sav)	30	54	1	1	12	34,818	21
	Wetlands (Wet)	33	7	46	0	3	2,932	2
	Croplands (Crop)	0	0	0	0	0	-	0
	Pasturelands (Past)	15	32	1	0	42	7,430	4

		1995						
		For	**Sav**	**Wet**	**Crop**	**Past**	**TOTAL (km²)**	**TOTAL (%)**
1990	**Forest (For)**	91	4	1	0	3	126,195	75
	Savanna formations (Sav)	10	60	1	1	26	26,646	16
	Wetlands (Wet)	36	3	44	0	3	2,761	2
	Croplands (Crop)	0	15	0	4	72	319	0
	Pasturelands (Past)	8	17	1	1	71	10,482	6

		2000						
		For	**Sav**	**Wet**	**Crop**	**Past**	**TOTAL (km²)**	**TOTAL (%)**
1995	**Forest (For)**	87	5	1	0	4	120,153	71
	Savanna formations (Sav)	10	68	0	2	17	22,813	13
	Wetlands (Wet)	32	4	59	0	5	2,802	2
	Croplands (Crop)	0	11	0	41	45	282	0
	Pasturelands (Past)	1	11	0	3	81	20,732	12

		2005						
		For	**Sav**	**Wet**	**Crop**	**Past**	**TOTAL (km²)**	**TOTAL (%)**
2000	**Forest (For)**	87	3	1	2	6	107,717	64
	Savanna formations (Sav)	16	51	1	4	24	23,509	14
	Wetlands (Wet)	25	1	62	0	3	3,108	2
	Croplands (Crop)	0	2	0	75	23	1,712	1
	Pasturelands (Past)	1	5	0	9	82	25,861	15

		2010						
		For	**Sav**	**Wet**	**Crop**	**Past**	**TOTAL (km²)**	**TOTAL (%)**
2005	**Forest (For)**	93	1	2	0	2	99,141	59
	Savanna formations (Sav)	19	62	1	3	11	16,748	10
	Wetlands (Wet)	28	6	60	1	3	2,905	2
	Croplands (Crop)	1	2	0	77	19	6,924	4
	Pasturelands (Past)	4	12	0	15	67	34,914	21

		2015						
		For	**Sav**	**Wet**	**Crop**	**Past**	**TOTAL (km²)**	**TOTAL (%)**
2010	**Forest (For)**	92	3	1	1	2	98,820	58

Savanna formations (Sav)	10	52	0	5	28	17,257	10
Wetlands (Wet)	45	6	41	0	3	4,111	2
Croplands (Crop)	1	0	0	82	14	12,453	7
Pasturelands (Past)	4	2	0	25	63	31,034	18
TOTAL 2015 (km²)	96,479	12,747	2,742	20,115	29,275		
TOTAL 2015 (%)	57	8	2	12	17		

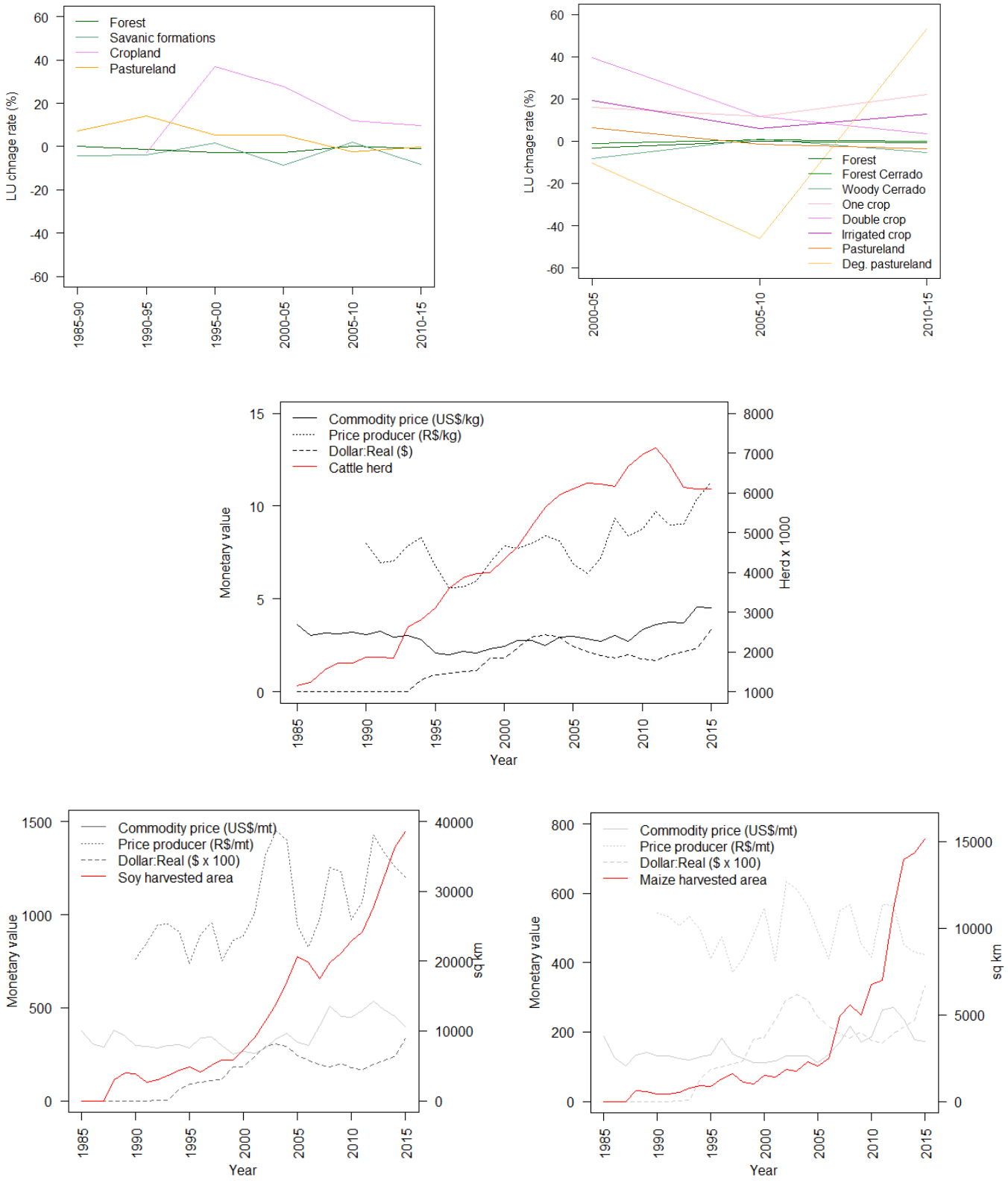
SM 7 - Table 3. Land cover change matrix according to Level 3 of classification for the Upper Xingu River Basin, MT, Brazil. Indexes equal to the percentage of a class (rows) that changes to another class in the next 5 years period (columns). Grey shadings represent no change. Land cover classes representing less than 1% of the basin area are not shown.

		2005										
		For	**Fce**	**Wcer**	**Wet**	**Sec**	**Scr**	**Dcr**	**Pas**	**Dpas**	**TOTAL (km²)**	**TOTAL (%)**
2000	**Forest (For)**	88	0	2	1	2	1	1	5	1	92,097	54.5
	Forest cerrado (Fce)	0	81	7	1	3	0	2	6	0	15,619	9.2
	Woody cerrado (Wcer)	7	10	49	1	5	1	2	22	1	21,288	12.6
	Wetlands (Wet)	23	2	1	62	0	0	0	3	0	3,108	1.8
	Sing crop (Scr)	0	0	2	0	0	14	60	24	0	1,088	0.6
	Double crop (Dcr)	0	0	2	0	0	11	66	20	0	613	0.4
	Pasturelands (Pas)	1	0	5	0	2	4	6	81	1	23,445	13.9
	Degraded pasturelands (Dpas)	0	0	6	1	0	4	4	65	14	2,416	1.4

		2010										
		For	**Fce**	**Wcer**	**Wet**	**Sec**	**Scr**	**Dcr**	**Pas**	**Dpas**	**TOTAL (km²)**	**TOTAL (%)**
2005	**Forest (For)**	94	0	1	2	1	0	0	2	0	83,855	49.6
	Forest cerrado (Fce)	0	84	6	3	2	0	0	4	0	15,286	9.0
	Woody cerrado (Wcer)	7	13	59	1	2	1	1	11	0	15,646	9.3
	Wetlands (Wet)	26	3	5	60	0	0	0	3	0	2,905	1.7
	Sing crop (Scr)	1	0	3	0	0	22	45	27	0	2,437	1.4
	Double crop (Dcr)	1	0	2	0	0	10	72	14	0	4,457	2.6
	Pasturelands (Pas)	3	1	11	0	2	8	8	67	0	33,376	19.7
	Degraded pasturelands (Dpas)	4	0	7	0	2	7	9	62	2	1,538	0.9

		2015										
		For	**Fce**	**Wcer**	**Wet**	**Sec**	**Scr**	**Dcr**	**Pas**	**Dpas**	**TOTAL (km²)**	**TOTAL (%)**
2010	**Forest (For)**	93	0	1	1	1	0	0	2	0	83,478	49.4
	Forest cerrado (Fce)	0	83	10	1	1	0	0	3	0	15,073	8.9
	Woody cerrado (Wcer)	3	5	46	0	2	3	2	22	6	15,102	8.9
	Wetlands (Wet)	36	9	6	42	0	0	0	2	1	4,073	2.4

Sing crop (Scr)	1	0	0	0	0	46	26	22	1	4,389	2.6
Double crop (Dcr)	1	1	0	0	0	32	55	9	0	8,017	4.7
Pasturelands (Pas)	2	1	2	0	1	17	9	54	9	30,762	18.2
Degraded pasturelands (Dpas)	0	0	1	1	0	15	4	31	29	162	0.1
TOTAL 2015 (km²)	81,144	14,312	10,565	2,742	2,575	11,084	8,956	24,628	4,647		
TOTAL 2015 (%)	48	8	6	2	2	7	5	15	3		



SM7 – Figure 1. Top plots present the change rate of different LULC in UXRB. Intermediate and bottom plots show the price and rate exchange fluctuations along with agriculture production in the UXRB. Worth noting, soybean and maize harvested area accounts for main and secondaries harvest.

Appendix B

Supplementary material to support Chapter 3: Deforestation and degradation in an ecotone zone between Amazon and Cerrado biomes: comparing indicators

Supplementary material 1. General description of each dataset.

SM1 - Table 1. Description and comparison among datasets on deforestation and degradation.

Dataset	Reference	Type	Methods	Product/ Output	MMA/ Temporal resolution	Year released/ Availability
Global Forest Change (GFC)	Hansen et al., 2013	End-product	Calculated based on per-band metrics (reflectance value, mean, percentiles and slope of linear regressions considering reflectance), of bands 3, 4, 5, and 7 of the LANDSAT satellites which are processed by learning algorithms in decision trees	Raster in which each of the cells indicates the annual percent vegetation cover with height > 5m, and annual loss or annual gain.	Landsat cell (900m²) / Yearly	2013/ 2000 - today
Carnegie Landsat Analysis System (CLASlite)	Asner et al., 2009	Software	Calculates fractional cover of vegetation canopies, dead vegetation, and bare surface based on the Automated Monte Carlo Unmixing approach. The differences between time steps are used to classify pixels according to a set of rules that can be modified	Raster with bands representing the fractional cover and uncertainty, and a raster in which each of the cells indicates a high probability of deforestation	Satellite cell (either Landsat, Modis, Sentinel) / Custom	2009/ Custom
				Raster with bands representing the fractional cover and uncertainty, and a raster in which each of the cells indicates a high probability of degradation		
Monitoramento do Desmatamento da Floresta Amazônica Brasileira por Satélite (PRODES)*	INPE, 2019	End-product	Satellite images are prepared by applying contrast enhancements to highlight evidence of deforestation/degradation. Them, areas are mapped individually	Raster/shapefile in which each of the objects represents the occurrence of observed deforestation in "pristine" areas	6.25ha / Yearly	1988/ 1988 - today
Mapeamento da Degradação Florestal na Amazônia Brasileira				Raster/shapefile in which each of the objects represents the occurrence of observed degradation	6.25ha / Yearly	2008/ 2007 - 2016

(DEGRAD)						
Amazon Near Real-Time Deforestation Detection System (DETER-B)**	INPE, 2019; Diniz, 2015		Visual interpretation based on five main elements (color, hue, texture, shape and context) followed by the application of Spectral Mixing Linear Model (MLME) to pixel classification	Raster/shapefile in which each of the objects represents the occurrence of observed disturbance that is classified as an indication of either (i) deforestation with exposed soil, (ii) deforestation with remaining vegetation, (iii) mining, (iv) degradation, (v) fire scar, (vi) disordered selective logging, or (vii) geometric selective logging	1ha (publicly available at 6.25ha) / near real-time	2016/ 2016 - today

*Recently the monitoring program is also available for other Brazilian Biomes; ** Not include in the comparative analysis; MMA = Minimum mapped area; Pristine = areas which have not been deforested in recent decades

Asner, G.P., Knapp, D.E., Balaji, A., Paez-Acosta, G., 2009b. Automated mapping of tropical deforestation and forest degradation: CLASlite. J. Appl. Remote Sens. 3, 033543. https://doi.org/10.1117/1.3223675

Diniz, C.G., Souza, A.A.D.A., Santos, D.C., Dias, M.C., Luz, N.C. Da, Moraes, D.R.V. De, Maia, J.S.A., Gomes, A.R., Narvaes, I.D.S., Valeriano, D.M., Maurano, L.E.P., Adami, M., 2015. DETER-B: The New Amazon Near Real-Time Deforestation Detection System. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. https://doi.org/10.1109/JSTARS.2015.2437075

Hansen, M., Potapov, P., Moore, R., Hancher, M., Turubanova, S., Tyukavina, A., Thau, D., Stehman, S., Goetz, S., Loveland, T., Kommareddy, A., Egorov, A., Chini, L., Justice, C., Townshend, J., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. Science. 342, 850–853. https://doi.org/10.1126/science.1244693

INPE - Instituto Nacional de Pesquisas Espaciais, 2019. Monitoramento da Floresta Amazônica Brasileira por Satélite [WWW Document]. http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes

Supplementary material 2. Comparison among datasets.

SM 2 - Table 1. Area occupied by each vegetation type in Upper Xingu River Basin and the deforested or degraded area capture by each indicator.

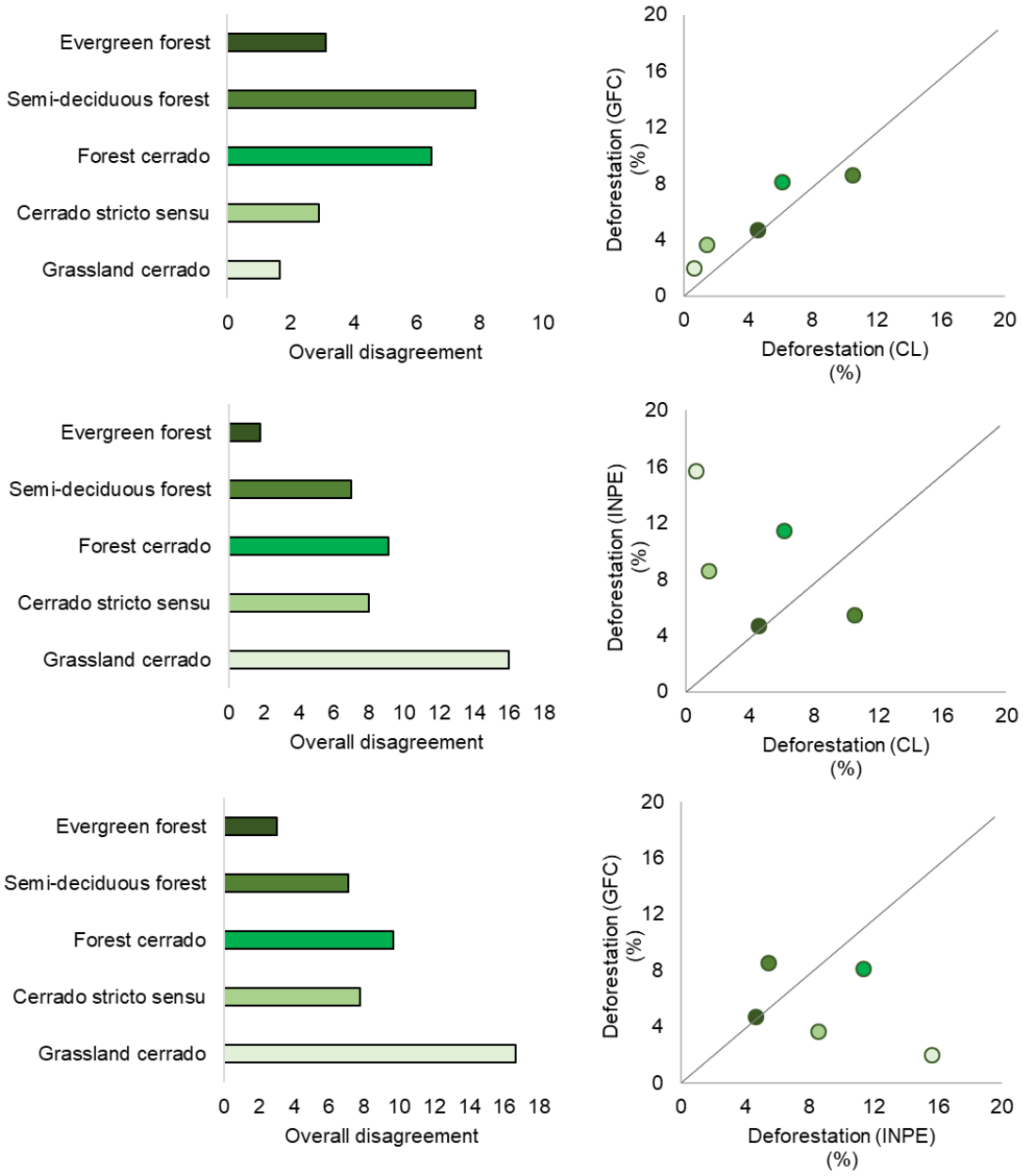
DEFORESTATION							
		Area (km^2)			Proportion (%)		
	Total	**CLASlite**	**GFW**	**INPE**	**CLASlite**	**GFW**	**INPE**
Grassland cerrado	7,748	55	32	221	0.71	0.42	0.99
Cerrado stricto sensu	27,097	293	218	454	1.08	0.80	0.99
Forest cerrado	14,460	345	215	377	2.39	1.49	2.61
Semi-deciduous forest	102,376	3,055	1,418	1,131	2.98	1.39	1.10
Evergreen forest	7,275	98	57	63	1.35	0.79	0.87
DEGRADATION							
		Area (km^2)			Proportion (%)		
	Total	**CLASlite**	**GFW**	**INPE**	**CLASlite**	**GFW**	**INPE**
Grassland cerrado	7,748	48	15	7	0.62	0.20	0.09
Cerrado stricto sensu	27,097	182	145	106	0.67	0.53	0.39
Forest cerrado	14,460	160	269	767	1.11	1.86	5.30
Semi-deciduous forest	102,376	1,242	2,394	9,062	1.21	2.34	8.85
Evergreen forest	7,275	92	47	457	1.26	0.65	6.28

SM 2 - Table 2. Agreement among dataset in raw data format for both deforestation and degradation in each phytophysiognomy or vegetation type in Upper Xingu River Basin.

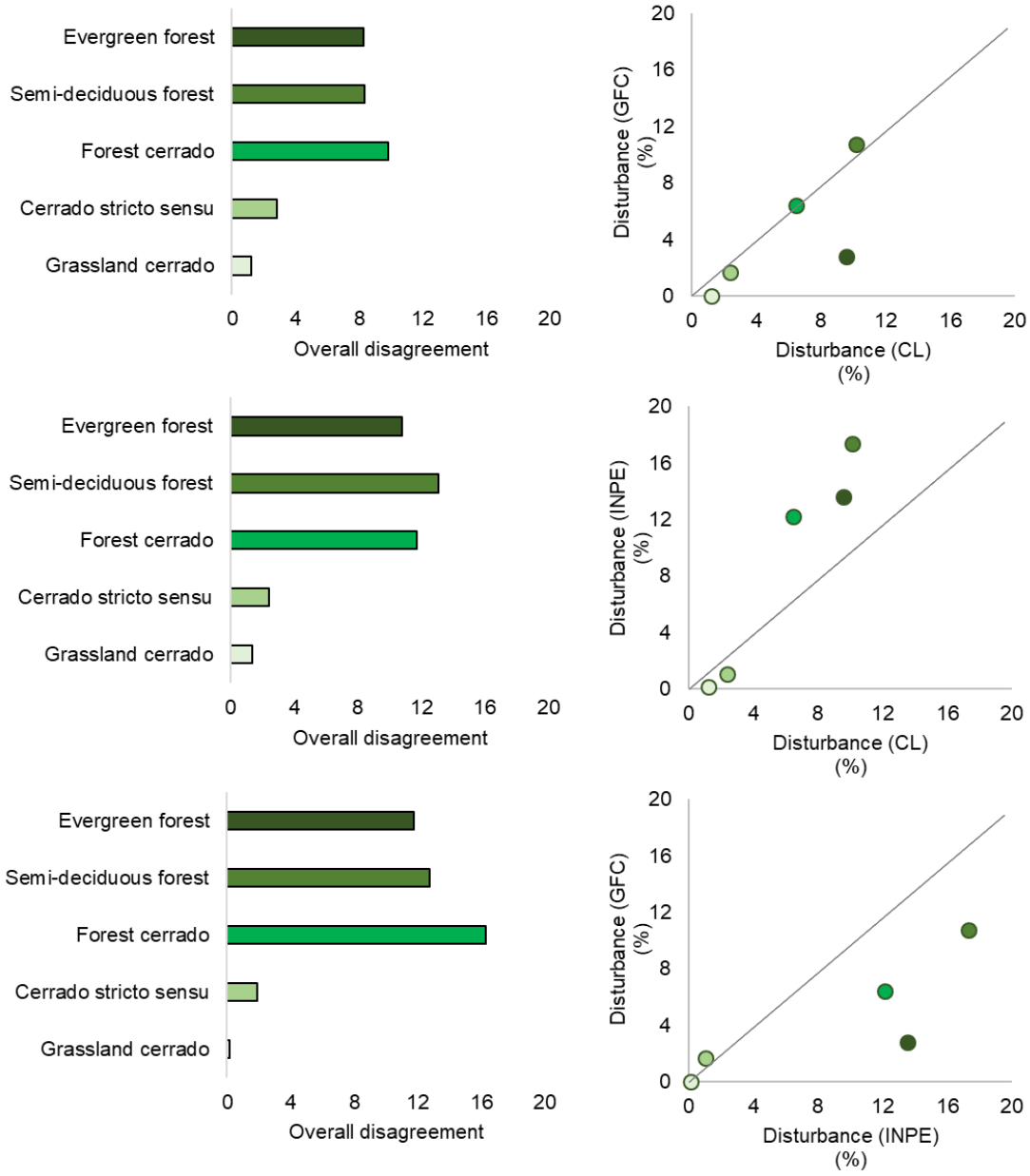
	DEFORESTATION					DEGRADATION				
Grassland cerrado										
	Quantity	Exchange	Shift	Overall	Jaccard	Quantity	Exchange	Shift	Overall	Jaccard
CL - GFW	0.29	0.64	0.00	0.94	0.99	0.43	0.38	0.00	0.81	0.99
CL - INPE	2.10	1.37	0.00	3.47	0.97	0.55	0.17	0.00	0.71	0.99
GFW - INPE	2.39	0.65	0.00	3.04	0.97	0.11	0.16	0.00	0.28	1.00
Cerrado stricto sensu										
	Quantity	Exchange	Shift	Overall	Jaccard	Quantity	Exchange	Shift	Overall	Jaccard
CL - GFW	0.31	1.27	0.00	1.58	0.98	0.15	1.02	0.00	1.18	0.99
CL - INPE	0.57	1.87	0.00	2.44	0.98	0.30	0.73	0.00	1.04	0.99
GFW - INPE	0.88	0.80	0.00	1.68	0.98	0.15	0.70	0.00	0.85	0.99
Forest cerrado										
	Quantity	Exchange	Shift	Overall	Jaccard	Quantity	Exchange	Shift	Overall	Jaccard
CL - GFW	1.04	1.84	0.00	2.88	0.97	0.75	2.14	0.00	2.89	0.97
CL - INPE	0.14	3.72	0.00	3.86	0.96	4.43	1.82	0.00	6.25	0.94
GFW - INPE	1.17	1.76	0.00	2.93	0.97	3.68	3.25	0.00	6.93	0.93
Semi-deciduous forest										
	Quantity	Exchange	Shift	Overall	Jaccard	Quantity	Exchange	Shift	Overall	Jaccard
CL - GFW	1.58	1.42	0.00	3.00	0.97	1.11	1.87	0.00	2.98	0.97
CL - INPE	1.87	0.95	0.00	2.81	0.97	7.59	1.22	0.00	8.81	0.91
GFW - INPE	0.28	1.40	0.00	1.68	0.98	6.48	1.89	0.00	8.37	0.92
Evergreen forest										
	Quantity	Exchange	Shift	Overall	Jaccard	Quantity	Exchange	Shift	Overall	Jaccard
CL - GFW	0.54	0.56	0.00	1.10	0.99	0.72	1.07	0.00	1.79	0.98
CL - INPE	0.46	0.65	0.00	1.11	0.99	4.51	2.12	0.00	6.63	0.93
GFW - INPE	0.08	0.60	0.00	0.68	0.99	5.23	0.79	0.00	6.02	0.94

SM 2 - Table 3. Agreement among dataset in hotspot ($p \geq 0.95$) format for both deforestation and degradation in each phytophysiognomy or vegetation type in Upper Xingu River Basin.

	DEFORESTATION					DEGRADATION				
Grassland cerrado										
	Quantity	Exchange	Shift	Overall	Jaccard	Quantity	Exchange	Shift	Overall	Jaccard
CL - GFW	1.27	0.40	0.00	1.66	0.98	1.21	0.00	0.00	1.21	0.99
CL - INPE	14.85	1.16	0.00	16.01	0.84	1.06	0.31	0.00	1.37	0.99
GFW - INPE	13.59	3.11	0.00	16.70	0.83	0.15	0.00	0.00	0.15	1.00
Cerrado stricto sensu										
	Quantity	Exchange	Shift	Overall	Jaccard	Quantity	Exchange	Shift	Overall	Jaccard
CL - GFW	2.18	0.71	0.00	2.89	0.97	0.73	2.07	0.00	2.80	0.97
CL - INPE	7.23	0.79	0.00	8.02	0.92	1.36	1.07	0.00	2.43	0.98
GFW - INPE	5.05	2.74	0.00	7.79	0.92	0.64	1.29	0.00	1.93	0.84
Forest cerrado										
	Quantity	Exchange	Shift	Overall	Jaccard	Quantity	Exchange	Shift	Overall	Jaccard
CL - GFW	1.93	4.56	0.00	6.49	0.94	0.81	9.02	0.00	9.83	0.90
CL - INPE	5.19	3.96	0.00	9.15	0.91	5.65	6.05	0.00	11.70	0.88
GFW - INPE	3.26	6.45	0.00	9.71	0.90	6.46	9.83	0.00	16.29	0.98
Semi-deciduous forest										
	Quantity	Exchange	Shift	Overall	Jaccard	Quantity	Exchange	Shift	Overall	Jaccard
CL - GFW	1.95	5.92	0.00	7.87	0.92	0.45	7.87	0.00	8.32	0.92
CL - INPE	5.04	1.96	0.00	7.00	0.93	7.13	5.94	0.00	13.07	0.87
GFW - INPE	3.10	3.99	0.00	7.09	0.93	6.68	6.11	0.00	12.79	0.87
Evergreen forest										
	Quantity	Exchange	Shift	Overall	Jaccard	Quantity	Exchange	Shift	Overall	Jaccard
CL - GFW	0.06	3.05	0.00	3.11	0.97	6.88	1.40	0.00	8.28	0.92
CL - INPE	0.08	1.70	0.00	1.78	0.98	3.16	7.61	0.00	10.77	0.89
GFW - INPE	0.02	3.00	0.00	3.02	0.97	10.04	1.74	0.00	11.78	0.88



SM 2 - Figure 1. Overall disagreement (left) among indicators and their area as a proportion (right) of the vegetation type in which they are observed. The graphs show a comparison among the Global Forest Change (GFC), CLASlite (CL) and PRODES (INPE) datasets for deforestation detected between 2010 and 2015, and are calculated based in the hotspot format.

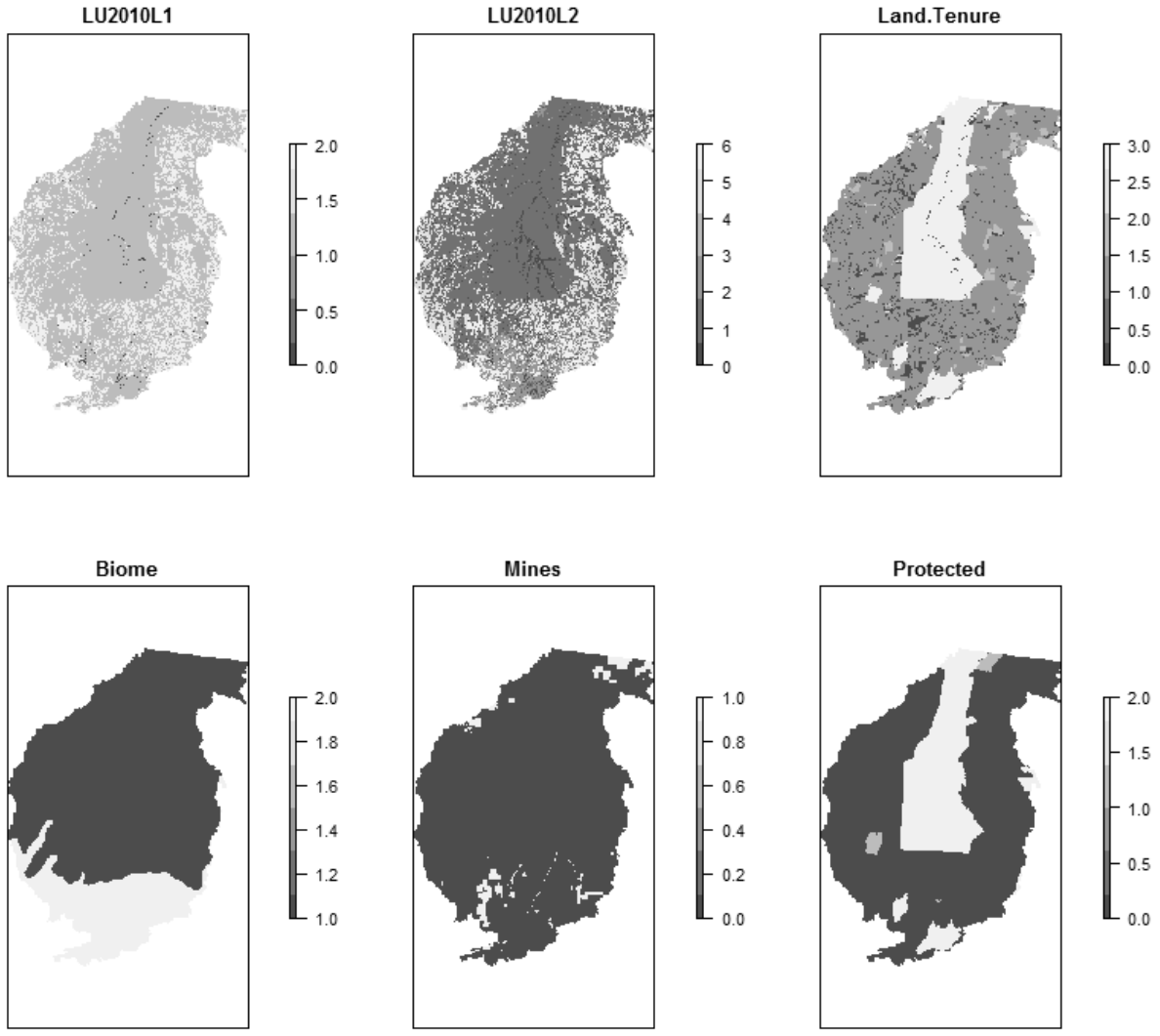


SM 2 - Figure 2. Overall disagreement (left) among indicators and their area as a proportion (right) of the vegetation type in which they are observed. The graphs show a comparison among the Global Forest Change (GFC), CLASlite (CL) and DEGRAD (INPE) datasets for degradation detected between 2010 and 2015, and are calculated based on a hotspot data format.

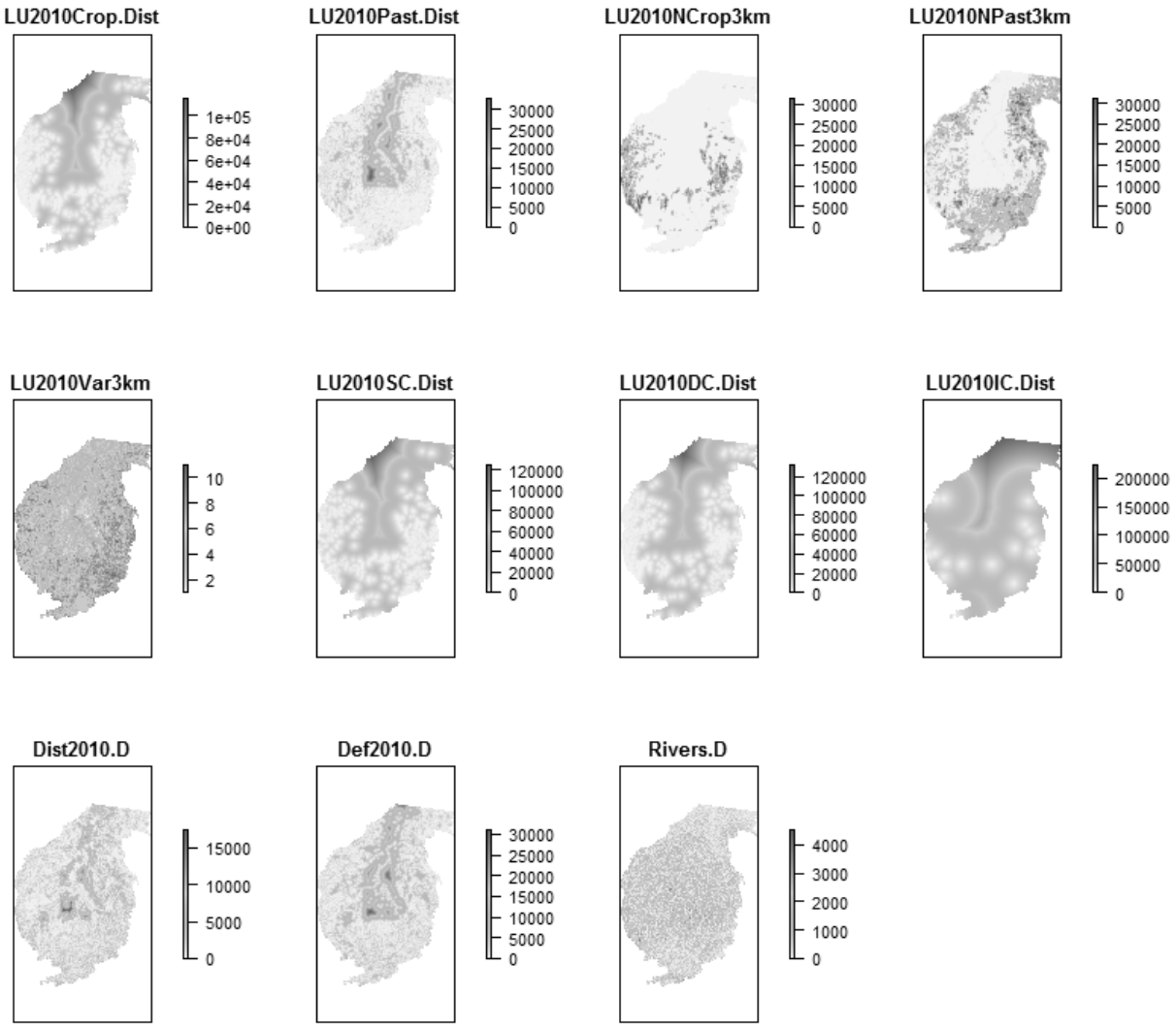
Appendix C

Supplementary material to support Chapter 4: Model selection for deforestation, degradation, and vegetation loss suggests different causations and future scenarios

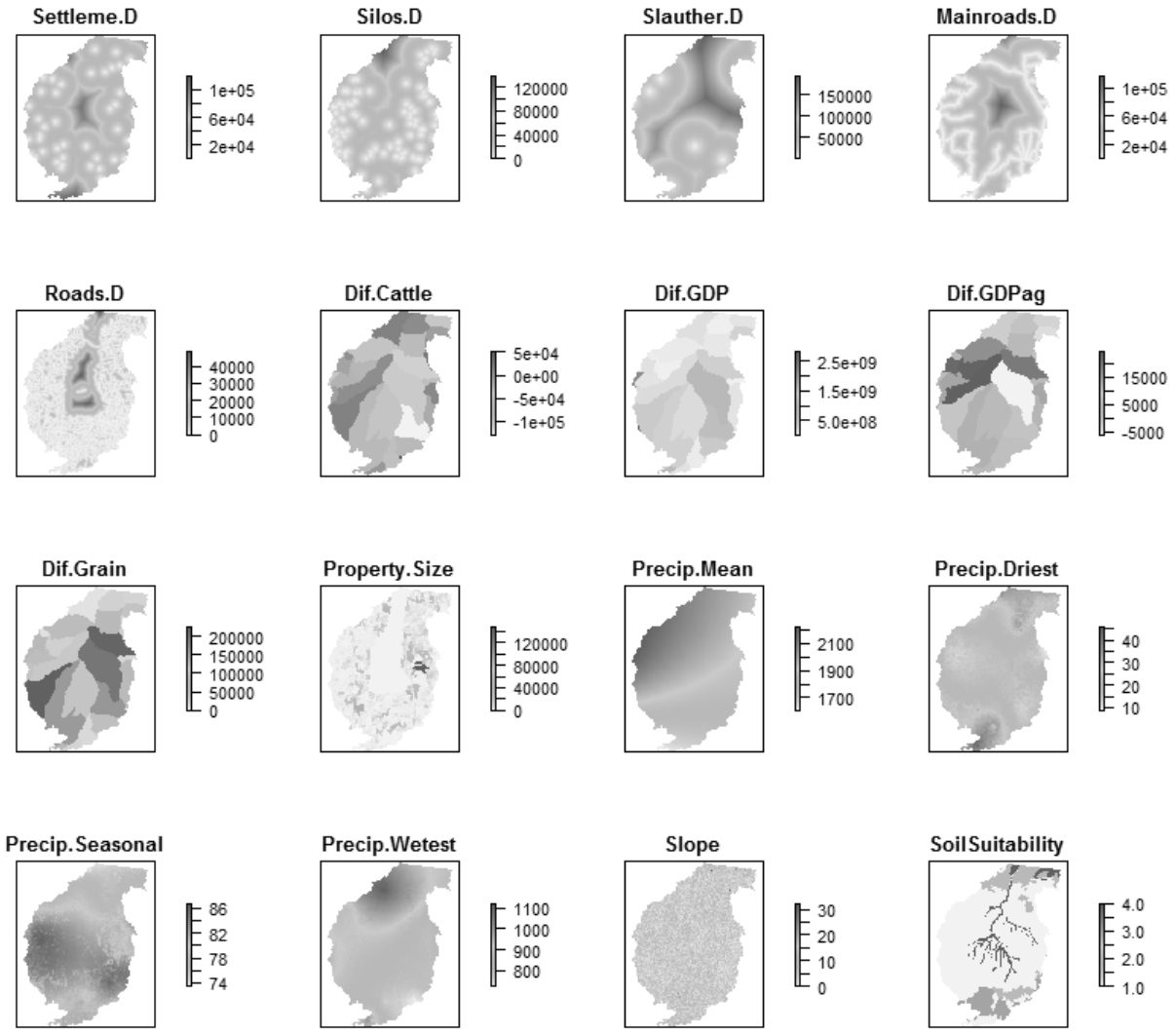
Supplementary material 1. Plot of all datasets analyzed in this chapter. Variables are separated in sets just to facilitate the plotting process and text flow.



SM 1 - Figure 1. Categorical variables.

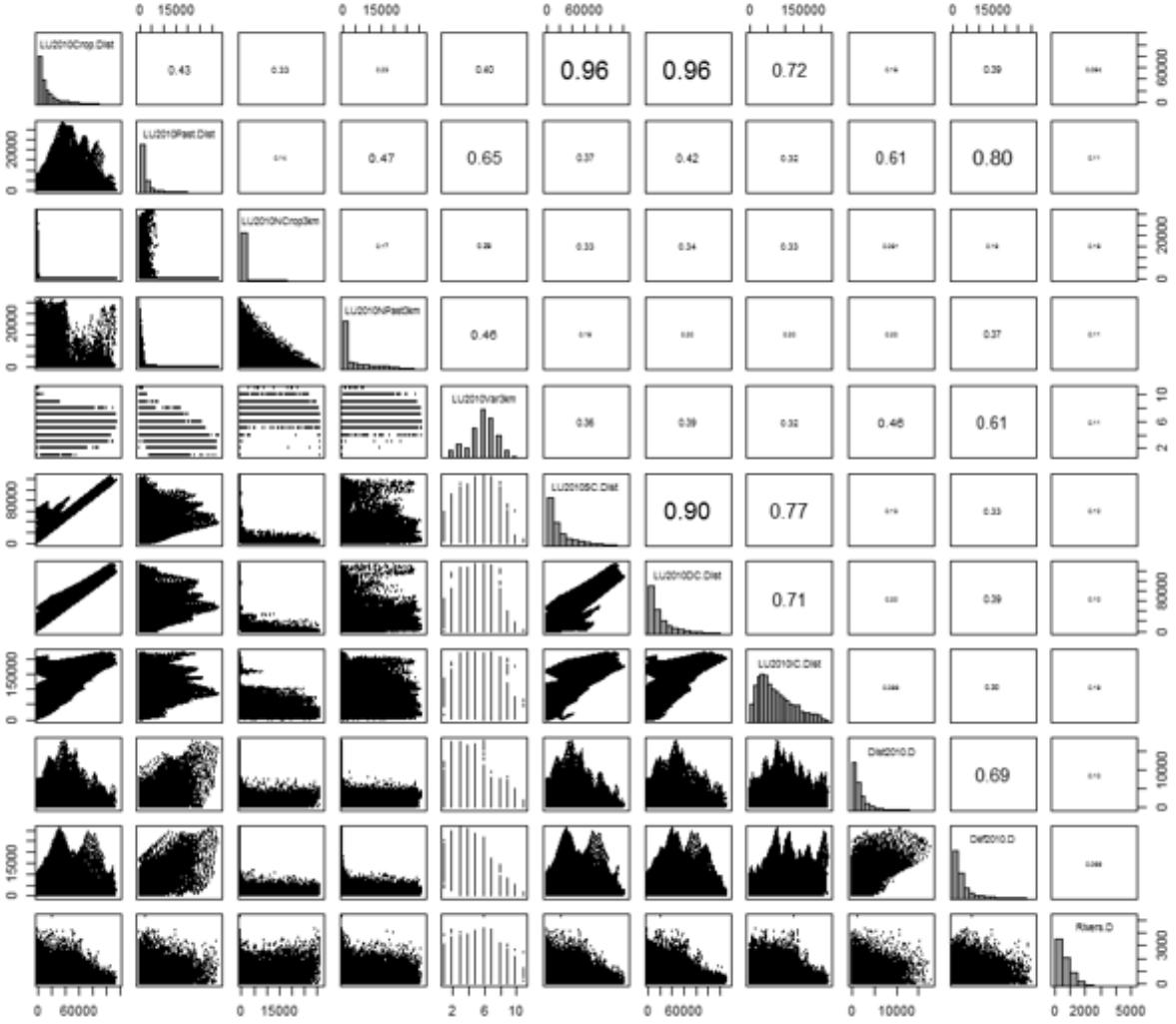


SM 1 - Figure 2. Continuous variables (set a).

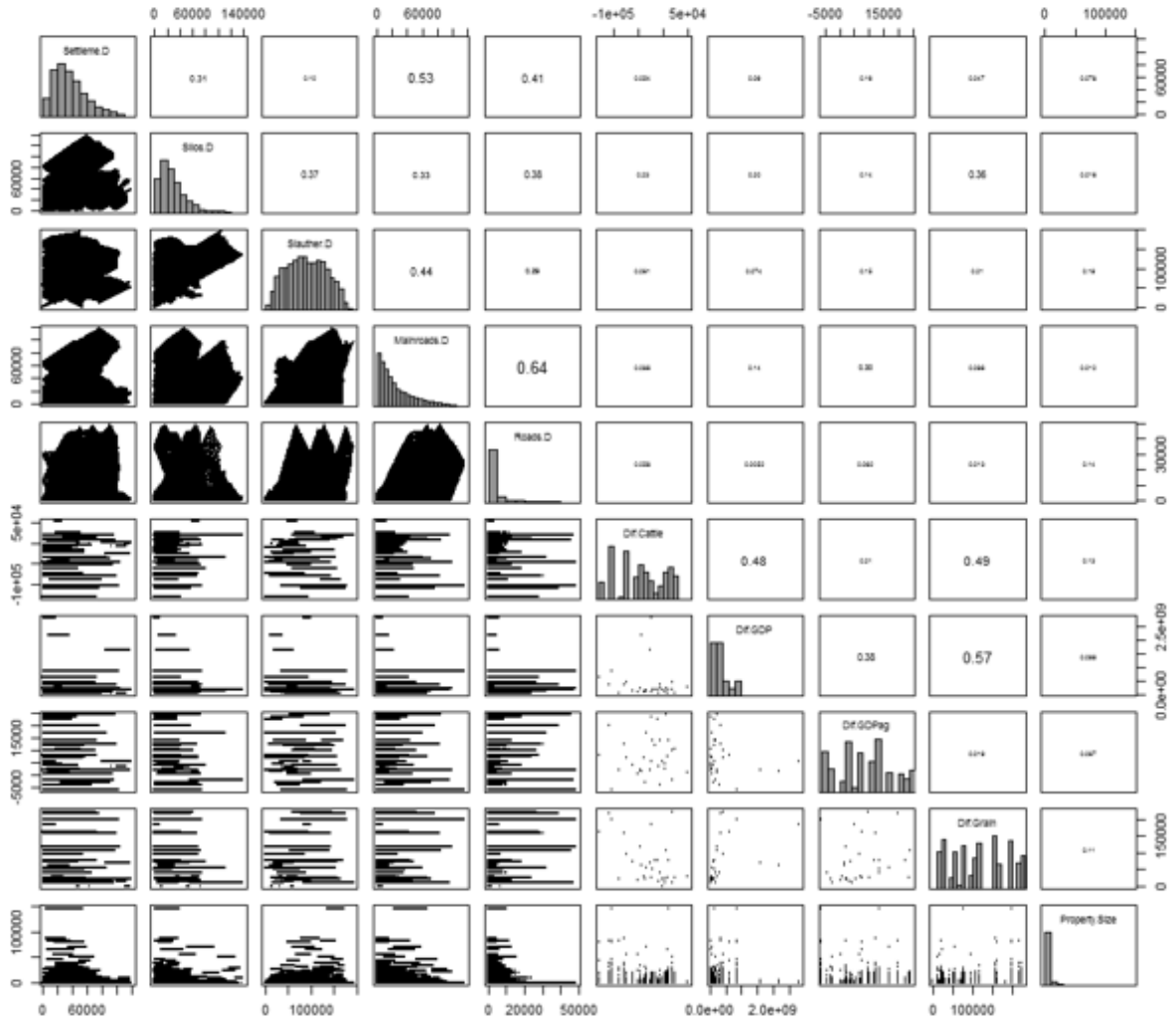


SM 1 - Figure 3. Continuous variables (set b).

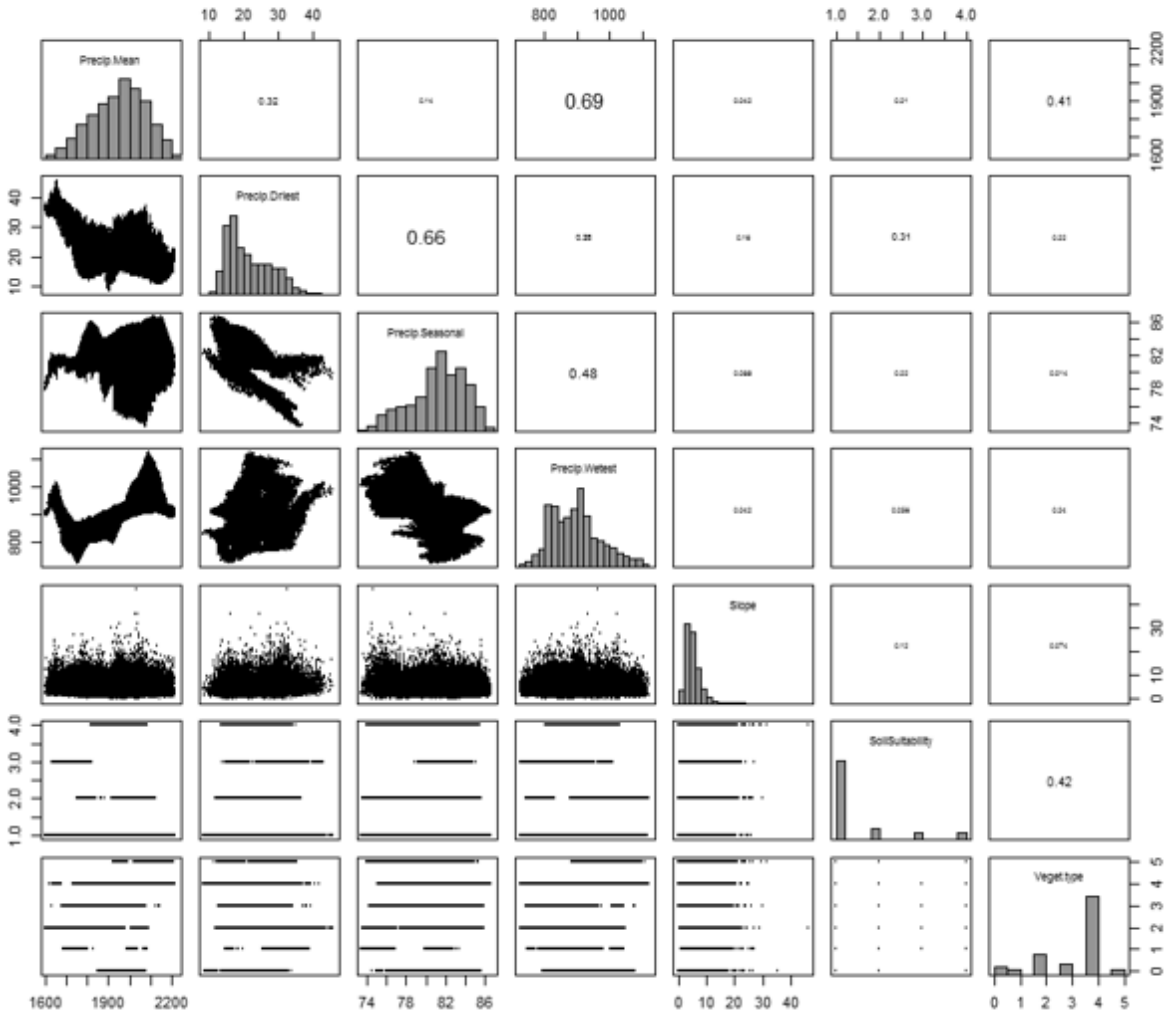
Supplementary material 2. Correlation and association among variables.



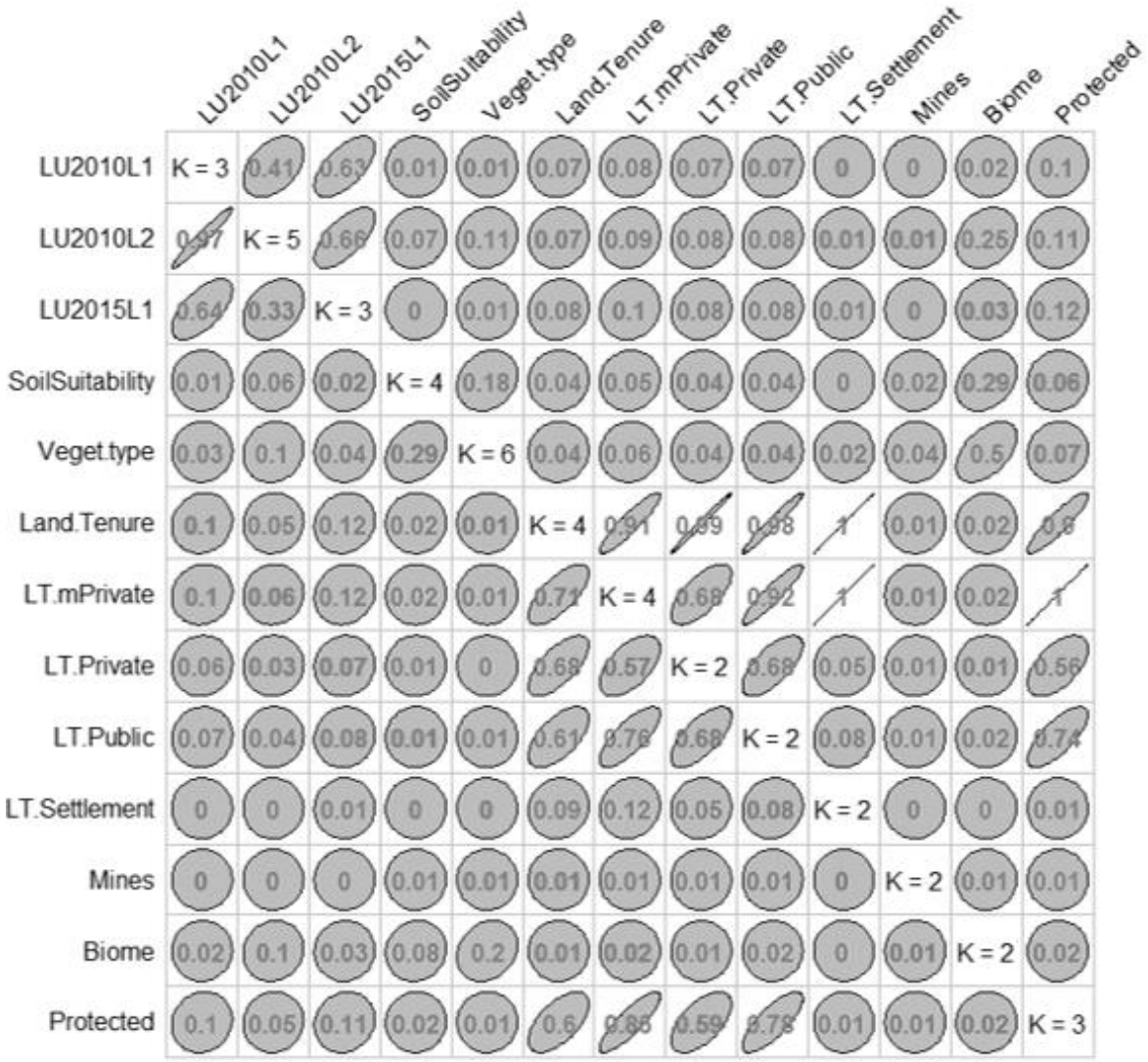
SM 2 - Figure 1. Correlation among variables. Indexes closer to 1 or -1 indicate large correlation.



SM 2 - Figure 2. Correlation among variables. Indexes closer to 1 or -1 indicate large correlation.

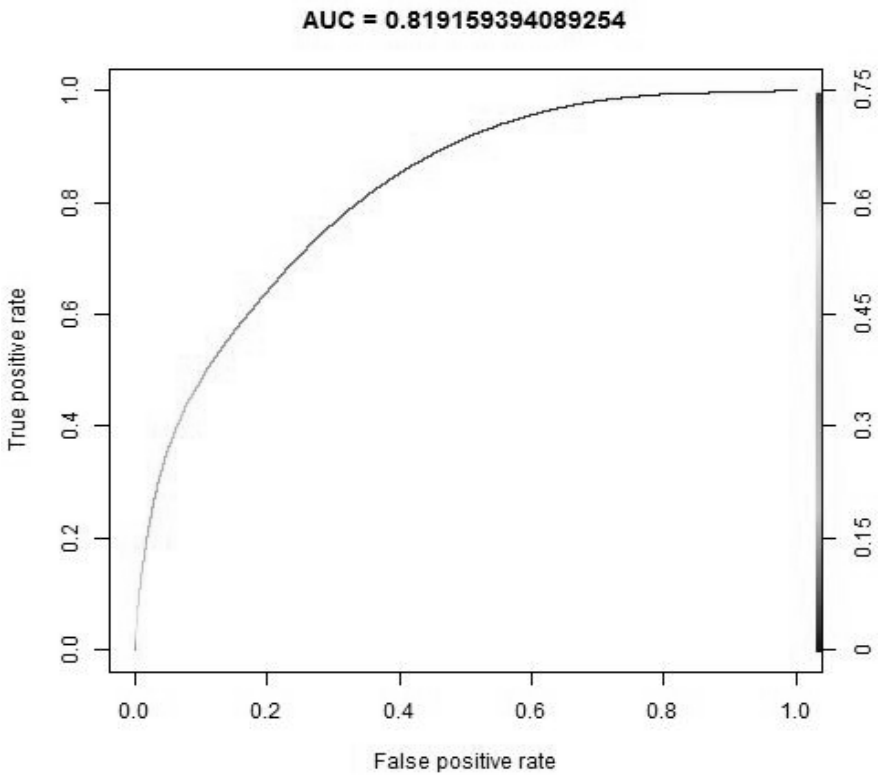


SM 2 - Figure 3. Correlation among variables. Indexes closer to 1 or -1 indicate large correlation.

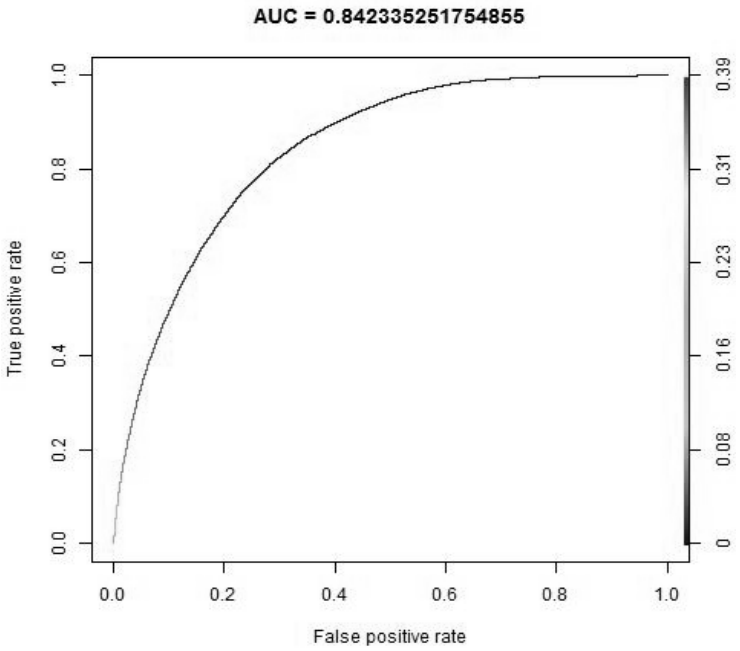


SM 2 - Figure 4. Correlation among variables. Indexes closer to 1 or -1 indicate large correlation.

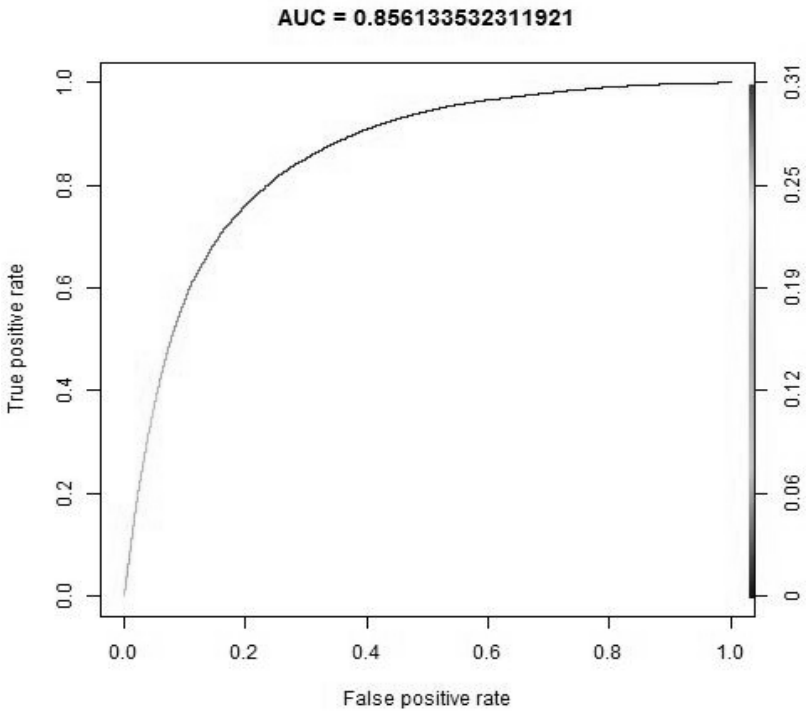
Supplementary material 3. Measure of fattiness for selected models.



SM 3 - Figure 1. Receiver Operating Characteristics curve (ROC) for native vegetation loss - a probability curve between false and true positive rates of classification. The calculated Area Under the Curve (AUC) indicates the separability power of the model. AUC equal to 0.5 (half of its potential) means the model has no separability power, while an AUC equal to 1 means the model has full separability power.



SM 3 - Figure 2. Receiver Operating Characteristics curve (ROC) for deforestation - a probability curve between false and true positive rates of classification. The calculated Area Under the Curve (AUC) indicates the separability power of the model. AUC equal to 0.5 (half of its potential) means the model has no separability power, while an AUC equal to 1 means the model has full separability power.



SM 3 - Figure 3. Receiver Operating Characteristics curve (ROC) for degradation - a probability curve between false and true positive rates of classification. The calculated Area Under the Curve (AUC) indicates the separability power of the model. AUC equal to 0.5 (half of its potential) means the model has no separability power, while an AUC equal to 1 means the model has full separability power.