Essays on the Economics and Policies of Deforestation in Brazil

by

Vilane Goncalves Sales

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

Birmingham Business School

College of Social Sciences

University of Birmingham

December 2018

Abstract

This thesis presents three essays on the economics and effectiveness of the environmental policies implemented in Brazil with the aim of taking into account possible barriers that contribute to the net forest loss. In the first essay we examine to what extent the Brazil's institutional environmental framework (IEF) has been successful in curbing deforestation by exploring the determinants of Brazilian deforestation for the period 2004-2015. Our results suggest that the creation of an IEF whilst a positive step, was undermined by weak enforcement such that the introduction of environmental laws as an instrument of the National Environmental Policy had little impact on deforestation. The second essay investigate trends in deforestation in the state of Maranhão. We study the state of Maranhão because of the uniqueness of the area. The artificial line of Legal Amazon crosses part of the state and proposes a natural experiment of deforestation in the under the Legal Amazon Maranhão and Cerrado Maranhão since the former has been subject to fundamentally different environmental policies compared to the latter. We suggest that cloud cover may have acted as an impediment to infringement detection via satellites,as is conducted by the environmental policy program. In our final essay, we investigate further the findings from the previous essay. We use event history analysis to confirm the previous findings by considering the heterogeneity within the region because not all sites have the same risk of deforestation. Our results show that forests inside the specific surveillance policy area had a lower probability of survival comparing to the area not covered by the environmental policy.

To my loving husband.

To my beautiful sister.

To my inspiring parents.

This is for us.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This thesis is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This thesis contains fewer than 80,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 100 figures.

by

Vilane Goncalves Sales

December 2018

Acknowledgements

In every way and everywhere I acknowledge this thesis with all gratitude. This thesis has been funded by the Coordination for the Improvement of Higher Education Personnel (CAPES) in Brazil and for that, I am deeply grateful.

I am especially indebted to professors Robert Elliott and Eric Strobl who have been so supportive, so patient and, so committed to my academic life. I am certain that I would never be able to achieve this much without them. For that, I am eternally grateful.

I wish to express my sincere thanks to the University of Birmingham for providing me the research facilities. It helped my thesis significantly. I would also like to recognise the support of the members of the departmental administration teams, Marleen Vanstockem, Olivia Crowther and, Anne Brookes. I thank the Economics department in the name of professor Anindya Banerjee.

I am grateful for the people I've met and the friends I've made while working on the thesis. My dear friends, Theophiline and Gladys, thank you for your amazing support in everything. We've become a family and your friendship is one of the rewards of this experience.

I am honour bound to my loving sister, Laisa, for her generous help, moral support, and valuable suggestions. I've learned so much from you my little sister! Thank you. I am thankful to my parents for their moral support and encouragement. Nobody has been more important to me in this pursuit of this thesis than the members of my family.

Most importantly, I wish to thank my loving and supportive husband, Arthur. I know the journey was not easy but you kept me going. Thank you.

Soli Deo gloria

Contents

List of Figures

List of Tables

Introduction

The natural process of evolution of mankind brought many changes. These changes are political, social, economic and especially environmental. Economic development has brought with it changes to the environment and these changes resulted in the improvements in the use of the environment and its degradation. This is the reason why we have seen an increase in studies that focus on the impact of global environmental change and the impact of natural hazards and problems relating to land-use and land-cover change.

In the context of climate change and sustainable development, deforestation remains the second leading cause of greenhouse gas emissions (Culas, 2014). The world's forests play an important role in the global carbon cycle. Forests are integral to any global carbon management and sequestration strategy and they play a major role in global climatic regulation as a sink and reservoir for carbon dioxide. The importance of forests in the process of climate change is reflected by the fact that despite the widespread deforestation in recent decades there is still more carbon in the world's forests than in the atmosphere. Therefore, a growing recognition that forests and climate change need to be treated as together (Buizer et al., 2014).

Deforestation was extensive in the temperate and sub-tropical areas during the 19th and 20th centuries but is no longer significant in the developed temperate countries. It can now be seen that temperate countries are recovering their forest area mainly in with upper middle and high income countries (see Figure 1).

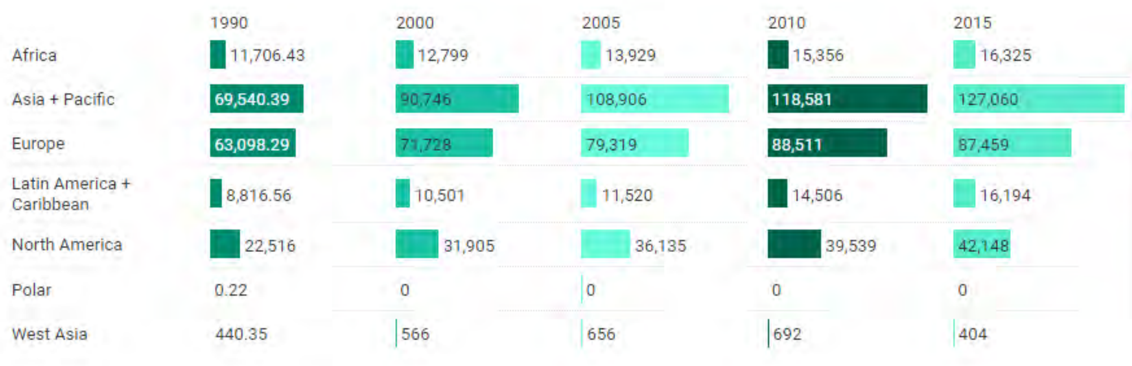

Figure 1 Planted Forest Area from 1990 to 2015 in thousand of hectares for afforestation or reforestation. Forest plantation is a forest established by planting and/or seeding in the process of afforestation or reforestation. It consists of introduced species or, in some cases, indigenous species. Forest plantation and natural forests are included in the term forest, a term that refers to land with a tree cover of more than 10 percent and area of more than 0.5 ha. Forests are determined both by the presence of trees and the absence of other predominant land uses. The trees should be able to reach a minimum height of 5 m. Young stands that have not yet reached, but are expected to reach, a crown density of 10m percent and tree height of 5 m are included under forest, as are temporarily unstocked areas. The term includes forests used for purposes of production, protection, multiple use or conservation (i.e. forest in national parks, nature reserves and other protected areas), as well as forest stands on agricultural lands (e.g. windbreaks and shelterbelts of trees with a width of more than 20 m) and rubberwood plantations and cork oak stands. The term specifically excludes stands of trees established primarily for agricultural production, for example fruit tree plantations. It also excludes trees planted in agroforestry systems. Source: (UNEP, 2018).

In contrast, tropical deforestation is a relatively modern event that gained momentum in the second half of the 20th century, more precisely, considerable deforestation happened during the period 1990-2015 and, was almost entirely confined to the tropical regions. Figure 2 shows the percentage of forested land in each region. Asia and Latin America had the highest net loss of forest during the period of 1990 to 2015.

Figure 3 shows that deforestation levels across Latin America concentrated in Meso America and South America. This is not surprising given that this area contains 86% of the total tropical forest area found in Latin America. In fact, Brazil accounts for 60%

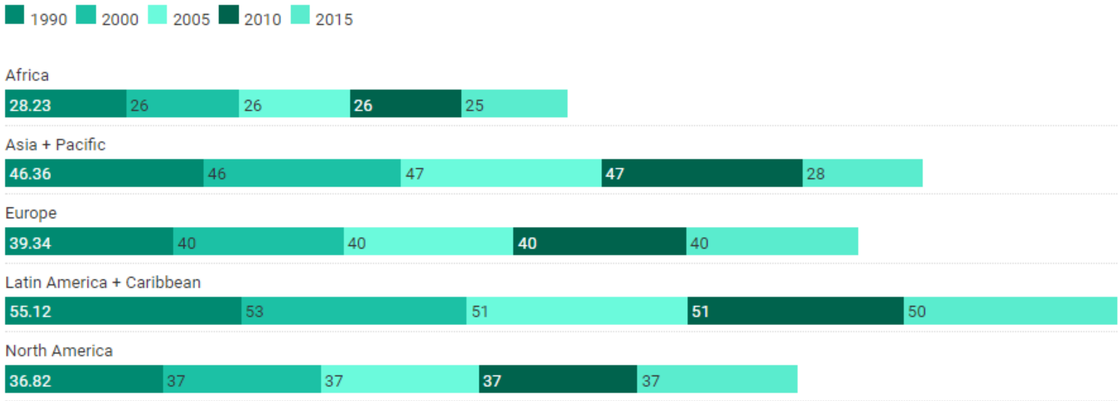

Figure 2 Proportion of land covered by forest from 1990 to 2015 in percentage. Forest: Land spanning more than 0.5 hectares with trees higher than 5 meters and a canopy cover of more than 10 percent, or trees able to reach these thresholds in situ. It does not include land that is predominantly under agricultural or urban land use. Explanatory notes 1. Forest is determined both by the presence of trees and the absence of other predominant land uses. The trees should be able to reach a minimum height of 5 meters in situ. Areas under reforestation that have not yet reached but are expected to reach a canopy cover of 10 percent and a tree height of 5 m are included, as are temporarily unstocked areas, resulting from human intervention or natural causes, which are expected to regenerate. 2. Includes areas with bamboo and palms provided that height and canopy cover criteria are met. 3. Includes forest roads, firebreaks and other small open areas; forest in national parks, nature reserves and other protected areas such as those of specific scientific, historical, cultural or spiritual interest. 4. Includes windbreaks, shelterbelts and corridors of trees with an area of more than 0.5 ha and width of more than 20 m. 5. Includes plantations primarily used for forestry or protection purposes, such as rubberwood plantations and cork oak stands. 6. Excludes tree stands in agricultural production systems, for example in fruit plantations and agroforestry systems. The term also excludes trees in urban parks and gardens. The term is mainly related to FRA 2005 National Reporting Table T1. Source: (UNEP, 2018).

of Latin America's tropical forests. Thus, the slowdown in deforestation in Brazil is largely responsible for the decline in overall tropical deforestation in Latin America. It is also possible to aggregate information from Figures 1, 2 and 3 in this process. Lower middle income and upper middle income tropical countries account for the decrease in the percentage of land covered by forests from 1990-2015. In other words, these economies were rapidly changing land use, by converting forests, woodlands and other natural habitat to agriculture and other land-based development activities (Barbier, 2004).

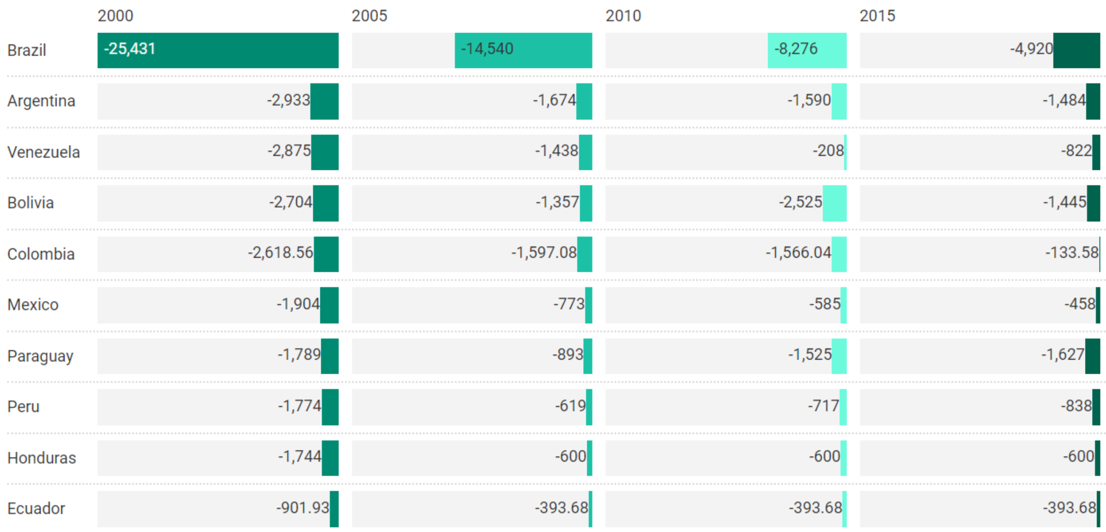

Figure 3 Forest Average Annual Change from 1990 to 2015 in Thousand Hectares per Year. Forest Average Annual Change – Total is the net change in forests and includes expansion of forest plantations and losses and gains in the area of natural forests. Total Forest includes natural forests and forest plantations. The term is used to refer to land with a tree cover of more than 10 percent and area of more than 0.5 ha. Forests are determined both by the presence of trees and the absence of other predominant land uses. The trees should be able to reach a minimum height of 5 m. Young stands that have not yet reached, but are expected to reach, a crown density of 10m percent and tree height of 5 m are included under forest, as are temporarily unstocked areas. The term includes forests used for purposes of production, protection, multiple use or conservation (i.e. forest in national parks, nature reserves and other protected areas), as well as forest stands on agricultural lands (e.g. windbreaks and shelterbelts of trees with a width of more than 20 m) and rubberwood plantations and cork oak stands. The term specifically excludes stands of trees established primarily for agricultural production, for example fruit tree plantations. It also excludes trees planted in agroforestry systems. Source: (UNEP, 2018).

As can be shown, deforestation, especially tropical deforestation, happens in countries where the status of development and welfare of the citizens are crucial factors in determining the extent of the forest loss. The accentuation of deforestation in low income countries is often explained by poverty, overpopulation and indebtedness. In this sense, the need for economic growth results in growing demand for agricultural and forest derived products. As a result, the governments of many developing countries believe that deforestation is one of the easiest and the most accessible ways of responding to ever increasing economic pressure (Culas, 2014).

In more recent years, studies are beginning to relate the major human causes of land-cover change in different geographical and historical contexts (Geist and Lambin, 2001). Direct and indirect causes of deforestation are concerned with the fact that some causes are direct in the sense that their occurrence or variation generates more or less deforestation through simple channels and other causes are indirect in operating through more complex channels (Combes Motel et al., 2009).

More specifically, direct drivers are identified as the hands-on land-management activities that lead to a land-use conversion or modification of ecosystems (Leemans, 2009). Geist and Lambin (2002) also define direct drivers as human activities or immediate actions at the local level, such as agricultural expansion, that originate from intended land use and directly impact forest cover. Indirect driving forces are the fundamental social processes that underpin the direct causes and either operate at the local level or have an indirect impact from the national or global level. In other words, indirect drivers influence the nature and strength of the direct drivers and include many diverse and often diffuse factors that often operate at different levels than the ensuing direct drivers. In addition, the drivers of deforestation must be defined in terms of their spatial and temporal dimensions, which are often clearly specified, but also defined in terms of institutions, social, cultural and, environment with their inherent levels (Leemans, 2009).

Based on several studies, it is widely accepted that the main direct causes for deforestation in Brazil, the most land converted country in Latin America, are agriculture expansion, cattle ranching, transport infrastructure and settlement expansion.[1] As indirect drivers of tropical deforestation, Brazil deals with the impact of economic growth and the changes in global demand for food, along with demographic factors, weak policies and institutional factors. To curb these actions, the implementation of social-political and environmental policies played an essential role.

Broadly, there are three major barriers to be able to execute effective policies to reduce deforestation. The first barrier is the profitability incentives, which is often in contradiction with forest conservation and sustainable forest management. Secondly, many direct and indirect drivers of deforestation lie outside of the forest sector and, thirdly, limited regulatory and institutional capacity and inadequate resources that constrain the ability of governments to implement forest and sectoral policies related to the forestry sector (Gerold et al., 2004; Nabuurs et al., 2007; Tacconi et al., 2003).

In this thesis we investigate the effectiveness of the environmental policies implemented in Brazil with the aim of taking into account possible barriers such as, limited regulatory and institutional capacity and inadequate resources that constrain the ability of governments to implement forest policies, the fact that many indirect drivers are outside of the forestry sector and, environmental factors that contribute to the net forest loss. Our analyses take an interdisciplinary approach to investigate these aforementioned issues. We combine the use of remote sensing technique and statistic methodologies that are rarely applied

[1] See references for Almeyda Zambrano et al. (2010); Arima et al. (2014); Barni et al. (2015); Barretto et al. (2013); Bhattarai and Hammig (2001); Cabral et al. (2012); Culas (2014); Davalos et al. (2014); Geist and Lambin (2001, 2002); Imori and Guilhoto (2015); Kuik (2013); Lambin and Geist (2006); Latawiec et al. (2014); Lopez-Carr and Burgdorfer (2013); Marcellus Caldas and Simmons (2007); Mendes (2009); Molina Vale (2014); Nepstad et al. (2014); Oestreicher et al. (2014); Olson et al. (2010); Pascale et al. (2010); Pfaff et al. (2007); Pfaff (1999, 1997); Richards (2015); Richards and VanWey (2015); Silva Costa et al. (2012); Soler et al. (2014); Stickler et al. (2013)

to deforestation in Brazil and, by making this statement, we acknowledge the potential contribution of this thesis.

In the first chapter we examine to what extent the Brazil's institutional environmental framework (IEF) has been successful in curbing deforestation by exploring the determinants of Brazilian deforestation for the period 2004-2015. Although previous studies (Borges de Lima and Buszynski, 2011; Dias et al., 2015; Nepstad et al., 2014; Oliveira, 2011; Pailler, 2018; Pascale et al., 2010) emphasise the role of institutions in different levels and policy intervention, little has been done concerning the impact of the environmental policy expansion on reducing forest loss and how their effectiveness may be affected by the presence of municipality level institutions for managing environmental policy while controlling for market expansion in products such as timber, cattle and soy.

The contribution of this chapter is to investigate the impact of both the expansion of the beef, soy and forest product markets and the effectiveness of Brazil's policies to protect the Legal Amazon conditional on the institutional environmental framework (IEF) in place at the time.[1] We provide a comprehensive analysis at the municipality level where we have detailed information on 562 municipalities where we are able to control for a wide range of possible determinants of deforestation. We employ a panel fixed effects regression model that takes into account environmental policies and prices, and the indirect institutional environmental framework. We expand the analysis by accounting for spatial effects and, quantify in which level of deforestation may have looked like had Brazil not implemented its current IEF.

Our results suggest that the creation of an institutional environmental framework (IEF) whilst a positive step, was undermined by weak enforcement such that the introduction of environmental laws as an instrument of the National Environmental Policy had little impact

[1]The Legal Amazon is an area that corresponds to 59% of the Brazilian territory and encompasses all eight states (Acre, Amapá, Amazonas, Mato Grosso, Pará, Rondônia, Roraima and Tocantins) and part of the State of Maranhão (west of the meridian Of 44ºW), totalling more than 5 million km^2.

on deforestation. However, municipalities who opened a dedicated environmental office did experience a reduction in the deforestation rate when combined with environmental policies. In our spatial analysis we find that the existence of an IEF within municipality has a negative spillover effect on neighbouring municipalities. Our counter-factual simulations show that the monetary benefits of avoiding deforestation and preserving the forest is worth the equivalent to 12 billion tonnes of stored CO_2.

In chapter 2, we investigate trends in deforestation in the biome most affected by human occupation over the last three decades in Brazil. Importantly, the Brazilian Cerrado has also been subject to spatially and temporally heterogeneous environmental policies discouraging such deforestation. We document what role such interventionist policies may have played in observed trends in deforestation so it can potentially provide a platform with which to assess future possible scenarios of deforestation in the Cerrado biome and the Amazon forest. We conduct this analysis by using remote sensing data and non-linear models to shape the trends of deforestation and its indirect drivers in the *Cerrado* region in the Brazilian state of Maranhão using the non-linear modelling approach of Generalized Additive Models (GAMs).

The state of Maranhão provides a particularly interesting context within which to study trends in deforestation and the possible role of environmental policy. More specifically, Maranhão is divided by an artificial line that separates it in two parts: the Legal Amazon Maranhão and the Cerrado Maranhão. This division, occurs at approximately 44° west of the meridian, and was established in 1953 due to the necessity to plan economic development in the region. This scenario provides a unique natural experiment of deforestation in the Legal Amazon Maranhão (LM) and Cerrado Maranhão (MA) since the former has been subject to fundamentally different environmental policies compared to the latter. More specifically, the tropical forest in the Legal Amazon Maranhão is under a surveillance environmental policy which detects deforestation or fire incidence in the region using satellite data and informs the occurrences to the environmental police so that they can fine or arrest the responsible persons (IBAMA, 2017). In contrast, this specific environmental policy is not applicable for the other biomes in the Maranhão state. We use this spatial division to determine how deforestation trends may have been different being the Legal Amazon Maranhão and Cerrado Maranhão.

Dealing with nonlinear processes trying to approximate to linear estimation methods may lead to inconsistent results guided by the linear estimator. There is an imposed assumption regarding the distribution and nature of the data. On the subject of the choice of distribution, it is doubtful that a satisfactory model could have been produced in this manner. This is part of the motivation for seeking to allow a more compact and flexible way of specifying smooth functional relationships within the models. Our use of non-linear modelling for the task at hand derives from the recognition in recent previous research that most ecological and climatic data represent complex relationships and thus that non-linear models, such as GAMs, may be particularly suited to capture confounding effects in trends; see (Antunez et al., 2017; Auderset Joye and Rey-Boissezon, 2015; Bell et al., 2015; Bio et al., 1998; de Souza et al., 2017; Halperin et al., 2016; Liu et al., 2018; Lusk et al., 2016; Moreno-Fernández et al., 2018; Pourtaghi et al., 2016). However, a review of the literature shows that such models have only been used sparsely to study deforestation and among trends of deforestation, this methodology has not been mentioned in recent analysis (Aleixandre-Benavent et al., 2018).

Here we apply a GAM with a negative binomial distribution and logarithmic link function. We capture deforestation by the construction of monthly time series from remote sensing sources (MODIS), given that high temporal resolution satellite products are particularly suitable to obtain detailed knowledge about the seasonal cycles of vegetation in biomes with strong seasonal contrast, such as the Cerrado biome and Ecotone forest (Bayma and Sano, 2015). We find that for the Legal Maranhão region most of the deforestation happened during the rainy season, while in the unprotected Cerrado Maranhão deforestation also occurred in the dry season. The fact that precipitation and solar incidence also played an important role in deforestation in the rainy season in the Legal Maranhão region, we suggest that cloud cover may have acted as an impediment to infringement detection via satellites, as is conducted by the environmental policy program. We further substantiated this claim by showing that for settlements located in both region but that are not the target of environmental policy deforestation mainly took place during both seasons as well.

In our final chapter, we investigate further the findings from Chapter 2. Given that, the environmental policy faded in areas of biome transition that coincided with the presence of cloudiness, we use event history analysis to confirm this findings by considering the heterogeneity within the region because not all sites within ecological tension zone have the same risk of deforestation. The uniqueness of the area proposes a natural experiment of deforestation in the Legal Amazon Maranhão and Cerrado Maranhão since the former has been subject to fundamentally different environmental policies compared to the latter. We estimate how the probability of transition between intact forest to disturbed forest, given risk factors and conditional on the time elapsed until the occurrence of the transition, is affected by cloud coverage. The results suggest that the presence of clouds has increased deforestation in the region covered by the satellite detection program, and thus is likely an active barrier to legal compliance.

Our results show that forests inside the specific surveillance policy area had a lower probability of survival comparing to the area not covered by the environmental policy. Forested pixels close to protected areas, which include conservation units and indigenous land, had a higher chance of being cleared comparing to forested pixels far from these special zones. Most importantly the presence of clouds was an active barrier to the legal compliance and, since the studied area has no systematic differences, we can agree on the lack of effectiveness due to cloud barrier.

From this economic and policy analyses of deforestation in Brazil it is possible to derive key policy implications about the potential ways to curb deforestation in the country. We've seen that institutional factors are an important instrument to protect or restrict deforestation in the Legal Amazon however the framework needs to be tightened up in terms of enforcement. We also follow the past trend of deforestation in two areas of great importance for the Brazilian biome, Amazon and Cerrado. From the trends we can deduce that the deforestation process was shifted to rainy periods and non surveilled areas. One policy implication is to expand the ecotonic/transition forests along the Amazon Forest. In addition, the monitoring system that can observe deforestation regardless of clouds and atmospheric barriers should be considered for the application of the environmental policy, for example the satellite Sentinel-2. Finally, since the studied areas are severely influenced by anthropic actions, one step further it would be extending the environmental policy to the Cerrado region.

The remainder of the thesis is organized as follows. The first essay, "Deforestation and the role of institutions and environmental policies in the Brazilian Legal Amazon" is presented in Chapter 1. The second essay, "Trends in Deforestation and Environmental Policy in Maranhão, Brazil" is presented in the following chapter. The third essay, "Satellite Monitoring of Deforestation and the Role of Clouds in Maranhão" is presented in Chapter 3. Finally, a brief conclusion is presented at the end of the thesis.

Chapter 1

Deforestation and the role of institutions and environmental policies in the Brazilian Legal Amazon

Abstract

In this paper we empirically investigate the impact of environmental legislation and environmental institutions on deforestation in the Brazilian legal Amazon at the municipality level for the period 2004 to 2015. Our results show that environmental policies, whilst giving the impression that they would curb deforestation, tend to be ineffective unless reinforced by a strong institutional environmental framework. A spatial econometric analysis to take into account spatial correlation collaborates our main findings. Counter-factual simulations indicate that the current institutional environmental framework has resulted in an decrease in deforestation of around 63%.

Keywords: Deforestation; Environmental Policy; Market Condition; Institutions

JEL classification: Q23; Q28.

1.1 Introduction

Forests are integral to any global carbon management and sequestration strategy and play a major role in global climate regulation. In the context of climate change, deforestation is recognised as the second major cause of greenhouse gas emissions. With Brazil being home to over 64 per cent of the Amazon rainforest it has experienced varying degrees of deforestation for more than five decades. It is therefore important to understand the drivers of deforestation and in particular the effectiveness of different government policies and how they are supported through new and existing institutional structures put in place to manage and prevent deforestation. Historically, the main causes of deforestation in Brazil have been identified as transport infrastructure (roads), settlement expansion, agricultural expansion (e.g. soybeans), and cattle ranching (Geist and Lambin (2002), Pfaff et al. (2007), Nepstad et al. (2014) and Richards (2015)). In an attempt to reduce deforestation in the legal Amazon, Brazil has enacted a series of environment related legislation and established a number of regulatory institutions.[1 2]

The purpose of this study is to examine the determinants of Brazilian deforestation between 2004 and 2015 with a view to investigating the extent to which Brazil's recently established institutional framework (IEF) has been successful in reducing deforestation. More specifically, our first contribution is, conditional on the IEF being in place at the time, to investigate the effectiveness of Brazil's policies to protect the Legal Amazon controlling also for the expansion of the cattle, soy and forest product markets. Our methodological approach is to estimate a panel fixed effects regression and spatial model (to control for spatial correlations) for 562 municipalities for the period 2004-2015. A second contribution, following Assunçao et al. (2015) is to quantify what the level of deforestation may have looked like had Brazil not implemented its current institutional environmental framework.

Both the rates and the reasons for deforestation have varied considerably over time. In the late 60's there was substantial investment in transport infrastructure with the aim of increasing the agricultural productivity of isolated regions (Baynard et al., 2012). In

[1]The Legal Amazon is an area that corresponds to 59% of the Brazilian territory and encompasses all eight states (Acre, Amapá, Amazonas, Mato Grosso, Pará, Rondônia, Roraima and Tocantins) and part of the State of Maranhão (west of the meridian Of 44ºW), totalling more than 5 million km^2. The Legal Amazon was established in 1953 and its territorial limits stem from the need to plan the economic development of the region and, therefore, are not limited to the ecosystem, which occupies 49% of the national territory and also extends into the territory of eight neighbouring countries (IPEA, 2008).

[2]A draft of this chapter had previously been submitted to the 10th International Research Meeting in Business and Management in France 2018.

addition, the government identified land settlement as a means to benefit both the landless poor and southern Brazilian agribusiness (Diprose and McGregor, 2009; Marcellus Caldas and Simmons, 2007). During the 1970s, Brazil's policy shifted from rapid frontier colonisation to agrarian reform as a strategy to reverse the tendency for land to become more concentrated into large landholdings resulting in the gradual internationalisation of the Amazon region. Part of the reform process involved the creation of the Brazilian Institute of Agrarian Reform (IBRA in Portuguese) and the National Institute of Agrarian Development (INDA in Portuguese), later replaced by the Institute for Rural Settlement and Agrarian Reform (INCRA in Portuguese). In the 1980s, the causes of deforestation changed again and were largely driven by the need for additional land to be cleared to satisfy the global demand for cattle and soy production (Cropper et al., 2001; Soler et al., 2014).

By the end of the 20th century, the Brazilian government, through INCRA, implemented a system of national land reform that promoted various colonisation schemes in areas close to newly built highways. During this process, for so-called agricultural pioneers, access to credit and subsidies for agriculture and pasture were increased especially along the deforestation line that ran next to major road arteries (for example, the BR 153 "Belem - Brasilia" and the BR 364 "Cuiaba - Porto Velho"). These policies had an indirect effect on deforestation that was further accentuated by the construction of federal roads, mining, and hydroelectric projects. Such actions eventually came to the attention of international organisations who, in turn, exerted pressure on the Brazilian government to reduce levels of deforestation. As a result, the government introduced a combination of environmental policies and environmentally focused institutions.

The focus on preventing deforestation resulted, in the early 2000s, in Brazil's National Environmental System (SISNAMA in Portuguese) changing the National Environmental Policy (Política Nacional do Meio Ambiente in Portuguese) on deforestation to make it a more decentralised operation in the hope that this would result in increases in both effectiveness and efficiency. To this end, the National Environmental Policy (NEP) aggregated federal, state, and local government bodies to form, what we call in this chapter, Brazil's institutional environmental framework (IEF). Two of the local institutional actions were the implementation of environmental laws and the establishment of environmental offices in municipality city halls. The explicit role of the personnel in these new offices was to formulate and execute policies related to the environment and to control, monitor, evaluate, and execute the management of the natural resources within a municipality in accordance with the local environmental legislation. Importantly, for our study, is that many municipalities chose not implement this action and would often designate this role to existing offices (often with different core agendas).

Although the NEP covers all of Brazil, given that much of the world's tropical forests lie within the region defined as the Legal Amazon, the focus of the policy tended to be on this region. In accordance with the NEP, in 2004 the government created the Action Plan for the Prevention and Control of Deforestation in the Legal Amazon (PPCDAm in Portuguese). The purpose of the PPCDAm is to plan development, control land use and ensure compliance with environmental laws and to promote sustainable practice. Crucially, in order to control land use and prevent further deforestation, the PPCDAm includes a satellite-based monitoring programme PRODES (Projeto de Estimativa do Desflorestameno da Amazônia in Portuguese) (INPE, 2017). Through PRODES, the government attempts to record incidents of deforestation that occur throughout the year within a given policy area. The premise is that the data gathered by the action plan are then used to enforce the PPPCDAm plan, which includes the issuing of fines for agents who clear or damage the forest, the organisation of embargoes for those areas in the process of being cleared and the confiscation of equipment, and restrictions on access to subsidised credit (Aubertin, 2015).[1] As part of the PPCDAm, a policy of Protected Areas (PA) was introduced with the

[1]In order to control for degradation of the forest by selective logging and forest fires, the government uses the DETER program. In addition, in 2007 two other systems were introduced: DEGRAD (Mapeamento

aim of limiting the expansion of cleared land where PAs represent ecological mosaics and corridors considered essential for conserving biodiversity and traditional communities.[1]

In terms of the existing literature, institutions tend to only be considered in relation to action plans at an aggregate level, and land tenure insecurity and property rights at the more disaggregated level (Oliveira, 2011; Pascale et al., 2010). For example, Borges de Lima and Buszynski (2011) examine the problem of deforestation in the Amazonian region and local environmental governance and public policies. The study identified some participatory-based, decentralized models of forest management and regulatory frameworks and how local environmental governance can be used to strike a balance between the use of natural resources, conservation and regional planning. Likewise, Nepstad et al. (2014) highlights the negative impact on deforestation from institutional factors such as the enforcement of laws, intervention in the soy and cattle supply chains, restrictions to credit, and expansion of protected areas. However, Dias et al. (2015) found that the deforestation rates were not significantly influenced by environmental governance. More recently, Pailler (2018) investigate how local political processes create incentives to manipulate forest resources and shows that the local electoral processes can lead to increased Brazilian deforestation. The main finding is that electoral deforestation cycles do not appear to be driven by changes in agricultural policy implementation but are linked to corruption and campaign financing, suggesting that weak institutions facilitate the electoral manipulation of forest resources.

da Degradação Florestal na Amazônia Brasileira in Portuguese), for mapping forest degradation in the Legal Amazon, and DETEX (Mapeamento da Cobertura Florestal na Amazônia Brasileira in Portuguese), for detecting logging operations in the Legal Amazon region (Pinheiro et al., 2016).

[1]The Law on the National Register of Conservation Units (Sistema Nacional de Unidades de Conservacao (SNUC)) dates from 2000 and defines 12 different types of conservation unit (CU). In general, CUs are territorial spaces, including their environmental resources, with relevant natural characteristics, which have the function of ensuring the preservation of the biological patrimony. In 2017 there were 334 CUs operating in the Legal Amazon, of which more than two thirds are classified as conservational sustainable use units which differ from standard CUs due to the presence of traditional populations. Such populations contribute to the development of sustainable economic activities. If we include the indigenous lands, nearly 50% of Amazonian land area is protected (Aubertin, 2015; MMA, 2017; MMA-CNUC, 2017). In this paper we classify indigenous land and conservational units as protected areas.

To briefly summarise our results, we find that the overall effectiveness of Brazil's institutional environmental framework, was undermined by weak enforcement to such an extent that the introduction of environmental laws as an instrument of the NEP had little impact on deforestation rates. However, we find that in those municipalities that had a dedicated environmental office did experience a reduction in the deforestation rate when it was combined with environmental policies. When we take account of spatial dependence (Hargrave and Kis-Katos, 2012) we find that the existence of an IEF within a municipality had a negative spillover effect on neighbouring municipalities. Our counter-factual simulations to quantify the amount of forest that would have been cleared, had there been no institutional environmental framework, show that deforestation would have been approximately 63% more than it was if there had been no institutional environmental framework. Following Assunçao et al. (2015), the monetary benefits of avoiding deforestation and protecting the forest show that the preserved forest area is worth the equivalent to 12 billion tonnes of stored CO_2 with an estimated value of US$ 62b.

The reminder of the essay is organized as follows. Section 2 describes the conceptual framework and outlines the main testable hypotheses. Section 3 describes the data and our empirical approach. Section 4 presents the results of our main regressions while Section 5 provides a series of counter-factual simulations. Section 6 concludes.

1.2 Conceptual Framework

1.2.1 Main Hypothesis

Our methodological approach is to include Brazil's institutional structure within the framework suggested by Kaimowitz and Angelsen (1998) and Hargrave and Kis-Katos (2012). As the Legal Amazon is often thought of as a frontier region it is often considered open access land in which deforestation depends positively on the expected profits from using the land for unsustainable activities, such as logging, cattle ranching or farming.[1] Moreover, in the Legal Amazon region, land use decisions are made taking into account the expected profit differential between sustainable (such as protected areas (PAs)) and unsustainable land use (Hargrave and Kis-Katos, 2012). We assume that the services from sustainable land use are public goods and thus are ignored by agents when they decide which land to clear (Kaimowitz and Angelsen, 1998). In this sense, agents can be considered to be profit maximizing and hence they choose the amount of land to clear subject to a number of constraints.

Expected profits from land use are determined by the prices of agricultural and forestry goods, access to markets, and other municipality specific conditions. The costs of deforestation are driven by the cost of clearing the land, agricultural and cattle ranching costs, the availability of credit and the costs associated with any fines imposed for being caught clearing the forest illegally (Hargrave and Kis-Katos, 2012). In turn, given the satellite monitoring system in place, means that agents may change their behaviour and clear land when it is cloudy (Calixto, 2016).

In the context of our paper, the term institutional framework is defined as a code of practice, behaviour, or relationship that has significant tangible or intangible social outcomes, such as socioeconomic and environmental sustainability (North, 1994). Often such relationships, behaviour, or practices are facilitated by physical organisations and/or actors, sanctions or rewards. In this sense, the concept of an institution incorporates policies and legislation. Strong institutions have the potential to reduce the rate of forest clearing through the enforcement of environmental laws and environmental management. For example, the risk of being penalised for illegally clearing a forest is higher when there

[1]According to Brazilian Forest Service data, by December 2015 there were 68.8 million hectares of forests awaiting allocation for the recognition of indigenous lands and traditional communities, as well as conservation units (BRASIL, Serviço Florestal Brasileiro, 2016; Brito, 2017).

is institutional compliance within a local area. However, the strength of any institution depends in part on political structures within a municipality including the political party, education of the mayor, the level of corruption, or the election process.[1]

If institutions are fragile, agents have a greater incentive to misappropriate revenues (Albuquerque and Ramos, 2006; Brollo et al., 2013; Bugarin and Meneguin, 2016; Garcia, 2003). In addition, how well the institutional apparatus works in a municipality is also likely to determine the preservation and maintenance of sustainable land (Rochedo et al., 2018). Our framework divides the expected profits into those drive by market conditions, policy pressure, predetermined conditions, and institutional factors.

1.2.2 Institutional Environmental Framework

The development of the Brazilian IEF began during the 1970s in response to pressure from a number of international organisations. As a result Brazil created the Special Secretariat for the Environment (Secretaria Especial de Meio Ambiente in Portuguese) which was given the remit to protect the environment and regulate the rational use of the natural resources within the territory under the supervision of the Interior Ministry. Ten years later

[1]When dealing with developing countries it is important to account for the fact that regulations may provide government officials with more levers for extracting rents. Hence, it is necessary to include controls for government efficiency. One approach is to capture the possibility of corruption within a system. For example, Brollo et al. (2013) identified that larger federal transfers from the FPM (Fundo de Participacao Municipal in Portuguese) increased observed corruption. Likewise, Bugarin and Meneguin (2016) found that the existence of commissioned workers in the Brazilian government lead to corruption in the Executive body. According to the Survey of Basic Municipal Information commissioners workers are employees, who are not effective in the City Hall, and whose only job is the commissioned position they carry out. Usually the position is given by the mayor or city councils in exchange for political favours and benefits (Bugarin and Meneguin, 2016). Corruption is defined as the attempt to obtain private gain at public expense such that corruption can be considered to arise whenever an agent, who is responsible for certain public responsibilities, and is influenced by the prospect of some reward, performs actions in favour of those who provide such reward, and therefore harms the public to whom such an individual should respond (Friedrich, 2002). The existence of corruption within the municipality governmental infrastructure is therefore related to the benefits and costs of an unlawful act since each agent is considered a utility maximiser. The costs of corruption are determined by the probability of detection and by the severity of the punishment, which may be either loss of popularity for politicians, wages, employment, and also the possibility of being prosecuted and imprisoned. The benefits are related to the financial gain (Albuquerque and Ramos, 2006; Bugarin and Meneguin, 2016; Garcia, 2003) or the possibility of being reelected (Brollo et al., 2013).

Brazil implemented the National Environmental System (SISNAMA in Portuguese) which provides the structure for national environmental management along with the introduction of the National Environmental Policy (NEP). At that time, municipalities did not have political and administrative autonomy and the introduction of a decentralised policy was considered to be an innovative initiative. [1]

The National Environmental System consists of six distinctive bodies and environmental institutions each with a specific role. At the highest level, the government council consists of all ministers and the Union's legal advisers who suggest to the chief in command on environmental subjects that enables the president to formulate national policy and set the rules for the correct use of the environment and its resources.

The level below is the environmental council that proposes policy on the environment and natural resources to the governing council, and deliberates on norms and standards compatible with an ecologically balanced environment. The environmental council is a representative body which consists of five groups, namely: federal, state and municipal bodies, the business sector and society. In this way, it is incumbent upon the environmental council to establish the federal standards and norms that must be observed by states and municipalities. However, states and municipalities also have the authority to introduce other standards, as long as they do not violate the standards established by the environmental council. As a central body, the Ministry of Environment has to plan, coordinate, supervise, and control the national environmental policy and the governmental directives set for the environment as federal level body (Mendes, 2017). The framework is shown in Figure 1.

At the executive level, the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA in Portuguese) also known as the environmental police (Hargrave and Kis-Katos, 2012) are responsible for carrying out actions of coordination, control,

[1]Pailler (2018) considers the National Environmental System to be a centralised central agency. For more details on the Brazilian decentralised environmental system see Scardua and Bursztyn (2003) and Sanches et al. (2017).

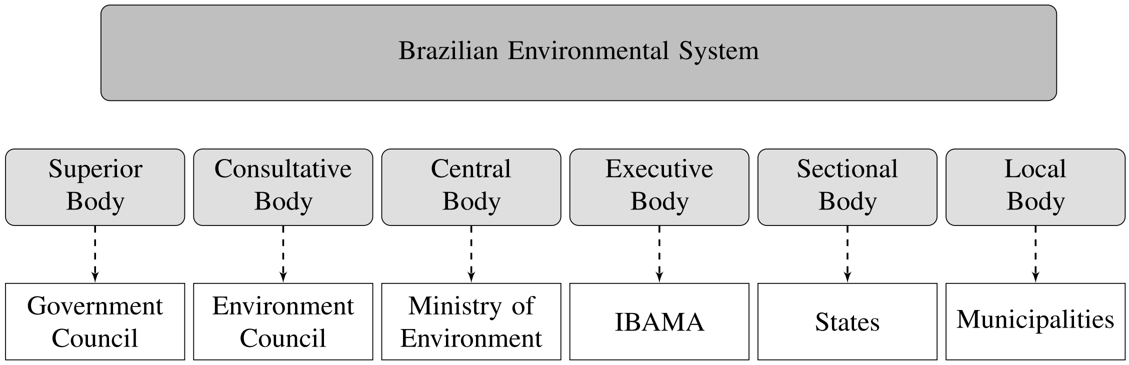

Figure 1.1 Organisational structure of the Brazilian Environmental System. Source: (MMA, 2018b)

supervision, monitoring, and orientation to ensure the execution of actions related to environmental emergencies using instruments such as fines and embargoes. Even though the duty to act in cases of environmental emergencies encompasses all of the entities of the federation (union, states and municipalities), given the common competence established in the Federal Constitution of 1988, many municipalities and states are not equipped with the technical capabilities and enforcement officers are often unable to attend environmental incidents within their jurisdiction. In such a case, the environmental police assist these bodies at the local level to prevent deforestation (Mendes, 2017; Moraes et al., 2016).[1]

In terms of enforcement more generally, of the entities within the federation, states and municipalities are the sectional body and local body, respectively, and are responsible for most of the environmental control activity. Therefore, each state and municipality is required to organise its environmental control agency and environmental laws according to its needs. Finally, municipalities are legally able to exercise environmental management within their territorial limits by creating specific environmental offices and laws. In addition, municipalities have environmental police powers, which legitimises them to apply appropriate sanctions and forbid or close establishments that do not comply with the legislation. (Mendes, 2017).

[1]The Federal Constitution from 1988 establishes that the environment is a fundamental right and it is classified as a public good. Therefore, the environment does not belong exclusively to a person or group, nor is it attributed to anyone who owns it.

The institutional framework described above underpins the NEP, the purpose of which is to regulate the various activities that involve the environment, including preservation, improvement, and recovery of environmental quality, and the promotion of sustainable development more generally. The NEP provides policies that are specific to different regions of Brazil through the creation of programs and action plans.

In order to establish the attributions and actions of a municipal administrative unit of environment, the so-called environmental offices, a municipality may define criteria creating and using specific environmental legislation. According to the Brazilian Environmental System, every municipality has to replicate the National Environmental Policy at the local level to create means to consolidate its policy. In the legal structure, the law approved by the City Council may provide actions of local interest; the organisation of the municipal environmental system; environmental zoning; control of pollution (sound, water, visual and soil); preventive instrument such as environmental licensing and authorisation; infractions and penalties; protection of fauna and flora; and, urban preservation areas (UPAs) and protected areas (PAs). Sometimes environmental offices are created jointly to other agendas, for instance, tourism and agriculture. Given the nature of the environmental agenda, many of the offices are turned down during political transitions, which configure a deficiency of the institutional framework.

As previously discussed, one of the most important tropical forest preservation plans was the implementation of the PPCDAm which is divided in three strategic phases.[1] In the first phase (2004-2008), the main objectives were to combat deforestation using satellite monitoring and to create a series of protected areas. Satellite monitoring meant that it was possible to notify the Brazilian environmental police (IBAMA) of environmental emergencies related to forest clearing thus enabling the police to penalise the offenders and prevent further deforestation. The action plan also allowed for the creation of conservation

[1]During 2016, the Executive Board reviewed the 13 strategic objectives of Phase 3 (2012-2015) with the objective to listing strategic actions for the period 2016-2020 (BRASIL, 2016).

units and the ratification of indigenous lands as instruments of policy conservation. During the first phase, the plan also advised on territorial planning and prioritised the fight against the illegal takeover of public lands (Mello and Artaxo, 2017).

The second phase of the PPCDAm (2008-2011) included plans to value the forest in terms of biodiversity conservation, and how best to manage the exploitation of timber and non-timber products. In this phase incentives were included for the monitoring of agricultural practices in deforested areas, including technological innovation and sustainable production systems. According to Mello and Artaxo (2017), the second phase encouraged the implementation of the Rural Environmental Registry (Cadastro Ambiental Rural in Portuguese), an instrument through which environmental agencies geo-reference rural properties in order to qualify the remote monitoring and effectiveness of field inspection operations, as well as to guide the regularisation process for rural properties.[1]

The third phase (2012-2015) included rules to strengthen decentralisation for states and municipalities through partnerships between the Union, states and municipalities, to deepen integration and to further improve the prevention of environmental damage and the promotion of sustainable production systems. During the three phases, the PPCDAm plan analysed changes in municipality level deforestation with the aim of highlighting possible improvements at the local level to curb deforestation (Moraes et al., 2016).

[1]Many properties in the Legal Amazon did not complete the Rural Environmental Registry during the second phase of the PPCDAm delaying the process. It was eventually concluded at the end of 2015 during the third phase of the action plan.

1.3 Data and empirical approach

1.3.1 Data and controls

It is useful to begin by illustrating in Figure 2 the geography of the Legal Amazon which includes more that 5 million square kilometres across 9 states. Our data consists of yearly observations for 562 municipalities in the Brazilian Legal Amazon from 2004 to 2015.[1]

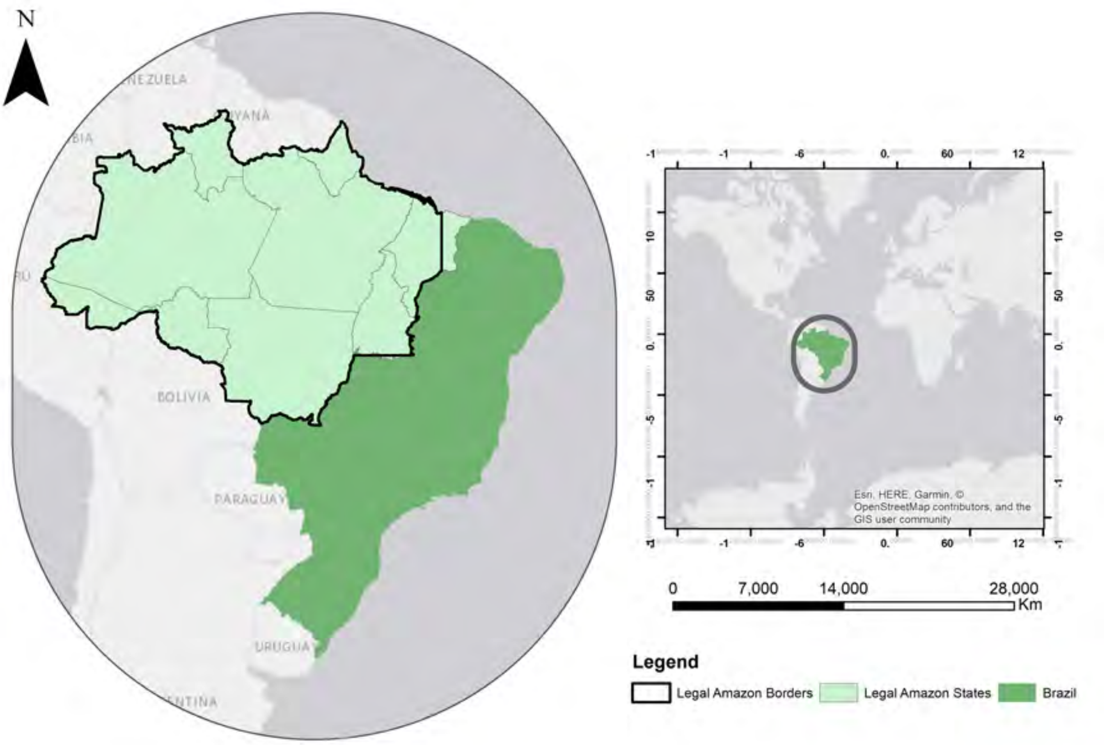

Figure 1.2 The Legal Amazon and Municipalities in the 9 states of Brazil: Rondônia, Acre, Amazonas, Roraima, Pará, Amapá, Tocantins, Maranhão e Mato Grosso. Sources: Own construction based on data from IBGE (2017); INPE (2017)

.

[1]The Legal Amazon officially includes 771 municipalities. However, in order to avoid bias introduced by municipalities with ecotone forest, we follow Hargrave and Kis-Katos (2012) and exclude those with less than 5% of rainforest cover. In robustness checks we re-estimate our results using 10 and 15% thresholds.

Data on deforestation at the municipality level comes from satellite based information from the Brazilian Institute for Space Research (INPE, 2017) under the project PRODES and measures yearly deforestation in square kilometres. The same data also provides information on cloud cover and areas that are unobserved by the satellite. Variables for cloud cover and unobserved areas are included to control for measurement error related to the ability of the satellites to detect changes in the land cover across all municipalities. More specifically, because the satellite used is incapable of detecting land cover changes when its view of land is obscured by clouds, detection will be delayed until skies are clear again and it reflects the frailty of the satellite. [1] [2]

Our dependent variable is the natural log of the extent in km^2 of yearly deforestation within a municipality measured between July and August of any given year. We adjust all numerical variables to match the July-August year.

To control for commodity markets, we collect soy and timber prices at the municipality level using data from the Brazilian Statistical Office (IBGE, 2017).[3] The beef price is reported at the state level by the Centre for Advanced Studies on Applied Economics (CEPEA, 2017). For municipalities with no commodity market data we include average local prices from neighbouring municipalities weighted by GDP at the municipality level. This differs from previous studies that impute zero prices for those municipalities with no commodity output values in order to get the direct effect of commodity prices on deforestation for commodities producing regions only (Gollnow and Lakes, 2014; Hargrave and Kis-Katos, 2012). We argue our approach takes into account transport costs from a municipality to neighbouring municipalities. All prices are deflated by IPCA which is the official Brazilian consumer price index. GDP is included at the municipality level to

[1]As a matter of fact, Assunçao et al. (2017) show that cloud coverage is an important predictor of the extent of deforestation fines issued within municipalities in the Brazilian amazon.

[2]According to Kintisch (2007) and Achard et al. (2010), the PRODES estimates are considered reliable by the national and international academic science.

[3]Timber prices come from the production of timber defined as firewood and logs of wood in cubic meters.

control for changes in overall economic activity (IBGE, 2017). We also contorl for federal paved roads that enables greater access to forests (Baynard et al., 2012; Cropper et al., 2001; Pailler, 2018; Pfaff et al., 2007).

Data on the IEF come from the Survey of Basic Municipal Information (IBGE, 2017). We therefore include dummy variables to capture whether a municipality has an environmental office and/or has introduced an environmental law. One important aspect that may influence whether a municipality has a well functioning IEF is the attitude and character of the local mayor. Hence, we also collect data from the Survey of Basic Municipal Information on whether the political party of the mayor is pro-farmer in which case may lead to greater deforestation, their level of education, gender, age, and the number of commissioned workers (as a proxy for corruption). We also gather data from BRASIL, Tribunal Superior Eleitoral (2018) to capture the mayoral re-election process. A recent study by Pailler (2018) shows that deforestation rates increased 8–10% in election years when an incumbent mayor ran for re-election.[1]

In terms of other economic and public policies, the number of settlements in a municipality and the settlement density are reported by the Brazilian Agency of Agrarian Reform (INCRA, 2017). We include settlement density to capture rural development in sparse regions.[2] In addition, we control for the presence of housing projects within a municipality using data from the Survey of Basic Municipal Information (IBGE, 2017).[3] We also collect data on the extent to which a municipality benefits from subsidised rural credits from the Brazilian central bank (BCB, 2017a) and development banks (BASA, 2017;

[1]More than 20 parties advocate a pro-farmer agenda. PMDB (Partido do Movimento Democrático Brasileiro), PP (Partido Progressista), PSDB (Partido da Social Democracia Brasileira) parties correspond to almost 44% of the pro-farmer seats - bancada ruralista in Portuguese - in the deputy chambers. These political parties represent the center-right political spectrum in Brazil with the largest number of affiliates (BRASIL, 2017).

[2]Settlement density refers to number of settled families divided by the area of the settlement.

[3]Our housing project dummy is intended to capture a set of interrelated and coordinated actions on housing with the aim of achieving specific objectives such as building houses within the budget limits over a given period of time. "Plano Nacional de Habitação (PlanHab)" and "Minha Casa Minha Vida (MCMV)" in Portuguese are examples of this type of housing project (Klintowitz, 2016; Krause et al., 2013).

BNB, 2017). Within the official credit system in each municipality a rural credit variable is included to capture the amount of credit provided to a municipality to encourage agriculture and pasture activities under the economic development policy called the National Program for Sustainable Family Agriculture (PRONAF in Portuguese).[1]

Finally, we want to control for other environmental related policies that may impact deforestation rates. First, we measure the size of conservation units (CUs) and the size of protected indigenous land at the municipal level using data from the Brazilian Environmental Ministry (MMA, 2018b). Currently, there are 886 federal, 729 state, and 147 municipal CUs that cover nearly 150 million ha. Sustainable use CUs (1214 sites, or 68.9% of all CUs), which have the goal of conserving ecosystems and habitats and cultural values and traditional natural resource management systems, are the most numerous and cover the largest area (about 100 million ha or 65.9% of all CU). Because of their area and diversity, CUs are essential to maintain biodiversity and ecological services, including carbon storage and sequestration. Considering indigenous land, in the early 1900s, the Brazilian Government began to offer protection to the indigenous population, treating Indians as wards of the state and guaranteeing protection of traditional lands. As most indigenous peoples resided in the nine states of the legal Amazon, by the beginning of 2000 indigenous lands accounted for 89 million hectares (17.5%) of the land area in the region, most of which maintained native forest cover despite its use for subsistence (BenYishay et al., 2017). Our main environmental policy variable is, then, the sum of these two areas called Protected Areas (PAs). Protected areas (PAs) have been the main strategy to safeguard biodiversity in Brazil and are key elements for biodiversity conservation and ecosystem services. Brazil. From 2000s, Brazil substantially expanded its environmental policy and expanded rapidly the implementation of the policy during the mid-1990s to the mid-2000s especially with the PPCDAm action plan.We choose to sum the values instead

[1]The National Program for Strengthening Family Agriculture (PRONAF) aims to stimulate income generation and improve the family farm production through the financing of rural agricultural and non-agricultural activities and services developed in a rural establishment or in regional community (BCB, 2017b).

of including them separately because of the high correlation between the two series. In robustness checks we include them separately. Second, we collect data on fines imposed at the municipality level for environmental violations using data provided by the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA, 2017). Fines reflect the amount of money that the agents within a municipality receive a penalty in any given year. In obtain the level of fines per municipality we sum the fines across agents for each municipality in any given year.

Table A.3.2 provides summary statistics for our explanatory variables. Table A.1.1 and A.1.2 of the appendix present the summary statistics for the first and final years of our sample and Table A.3.4 provides the source, level of aggregation and units of measurement for each of our variables. Our summary statistics show that nearly 45% of municipalities have an environmental law in place and nearly 75% have an environmental office. We also find that around 35% of mayors are over 50 years of age and that nearly 90% are male. Around 40% have an education level that is tertiary or above, nearly 2.2% of the employees in the city hall are commissioned and that 21% of mayors tend to be re-elected. On average, municipalities have 3 settlements within their limits and at least 2% of families live in a settlement. On average municipalities have approximately 20 thousands km^2 of their area protected and impose on average R$3.2 million in fines per municipality. For our sample period the average value of rural credits per municipality is R$20mi and 78% municipalities have at least one housing project in place.

Table 1.1 Summary Statistics - Averages of the sample

Variable	Mean	St. Dev.	Min.	Max.
Deforestation	20.200	57.183	0.1	1407.8
Fines	0.324	1.298	0	24.227
Protected Areas	2.188	5.354	0	44.877
Rural Credits	0.235	0.948	0	22.809
Environmental Law	0.450	0.497	0	1
Environmental Office	0.754	0.430	0	1
Housing Projects	0.780	0.413	0	1
Settlements	3.688	5.965	0	76
Settlements Density	0.018	0.034	0	1.016
GDP	0.057	0.331	0.001	9.850
Beef Price	0.852	0.472	0.518	2.419
Soy Price	0.630	0.126	0.201	1.761
Timber Price	0.104	0.080	0.002	0.916
Roads	0.019	0.029	0.001	0.089
Clouds	10.323	47.735	0	1315.286
No obs	0.151	1.507	0	49.786
Mayor Political Party (pro-farmer)	0.920	0.271	0	1
Mayor Gender (Male)	0.901	0.298	0	1
Mayor Age (% above 50)	0.356	0.479	0	1
Mayor Education	0.407	0.491	0	1
Corruption	0.022	0.397	0	25.167
Re-election	0.211	0.408	0	1

Note: Statistics refer to N=6178 observations for 562 municipalities for 11 years (2004 - 2015). In the table, the variable deforestation is not logged and represents the actual extent in km^2.

Before describing our empirical strategy we show the trend in deforestation and our key variables. Figure 1.3 shows that deforestation reached its highest level in 2004 (27,423 square kilometres) (MMA, 2018b). Since then there has been a fairly clear decreasing trend. Figure 1.3 also shows the steady increase in the area of land designated as a Protected

Area and the rapid increase in the amount of fines per area of land deforested. We also show the cumulative number of municipalities with environmental laws and environmental offices increased steadily with a distinct jump between 2007 and 2008. Similar trends are observed when we compare deforestation rates and changes in the market for commodities that are captured through the price of soy, timber and beef and the IEF captured by the existence of environmental laws and environmental offices in Figure 1.4.[1]

[1]See Hargrave and Kis-Katos (2012) and Gollnow and Lakes (2014) for a detailed discussion of the role of commodity markets on deforestation.

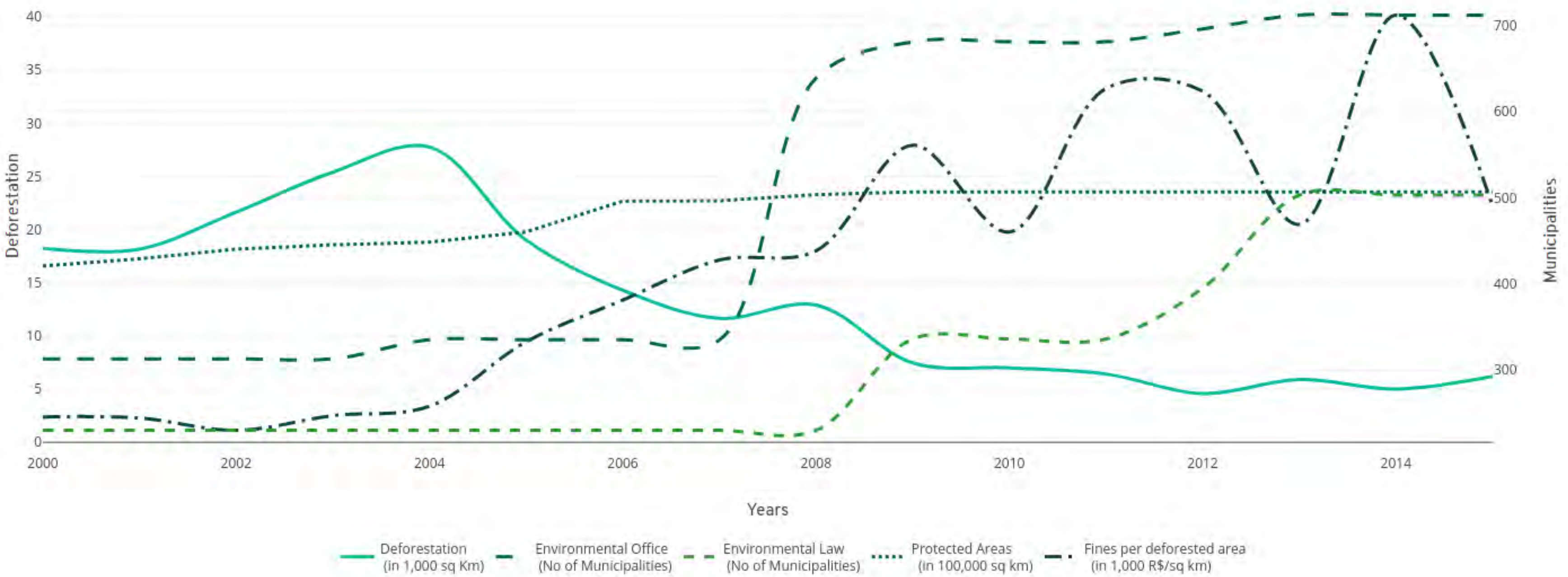

Figure 1.3 Deforestation in the Legal Amazon, Environmental Policy and the Institutional Environmental Framework (2004-2015). Left axis shows deforestation and right axis shows number of municipalities. Source: (IBAMA, 2017; INPE, 2017; MMA, 2018b)

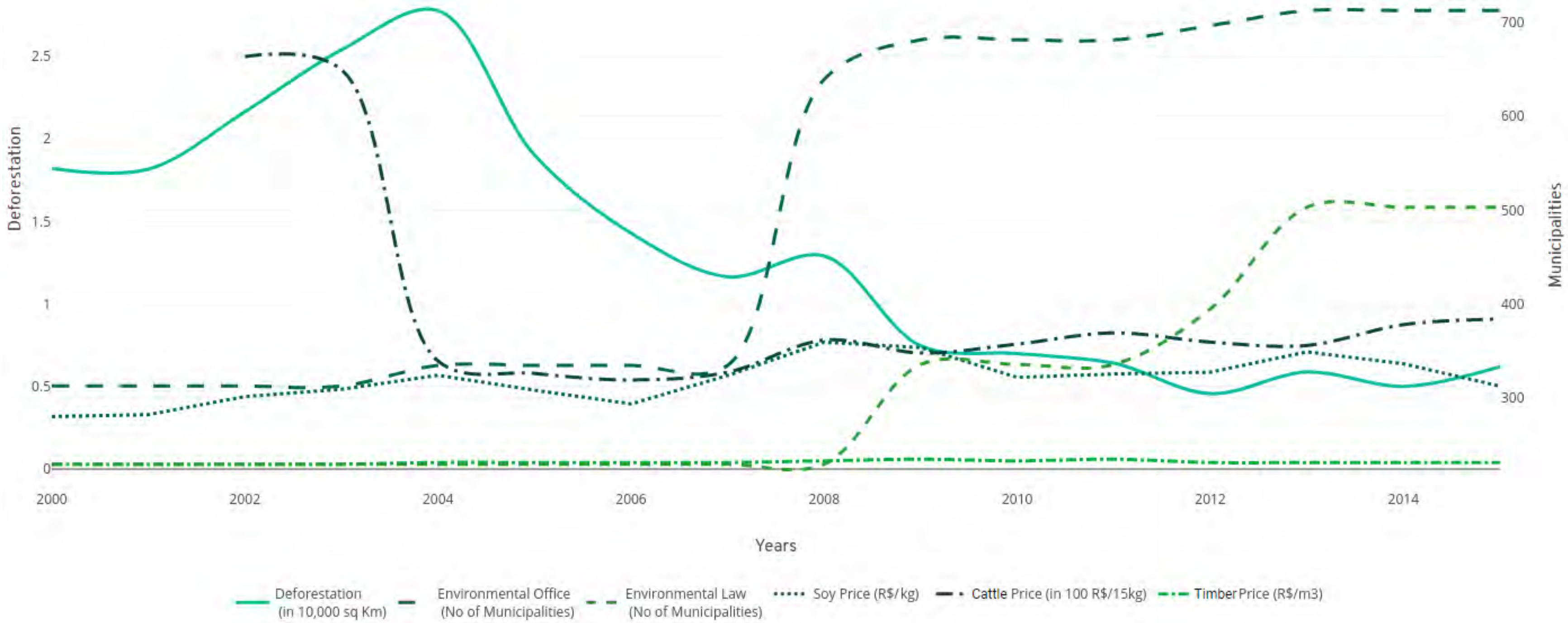

Figure 1.4 Deforestation in the Legal Amazon, Commodities price and the Institutional Environmental Framework (2004-2015). Left axis shows deforestation and right axis shows number of municipalities. Source: (CEPEA, 2017; IBGE, 2017; INPE, 2017)

1.3.2 Empirical strategy

In order to analyse the impact of different economic and environmental policies under different institutional settings our estimations are based on a panel of municipality level data between 2004 and 2015. Our baseline regression is given by:

$$\ln D_{i,t} = \mu_i + \mathrm{X}_{i,t-1}\beta_1 + \mathrm{X}_{i,t}\beta_2 + \lambda_t + \varepsilon_{i,t} \tag{1.1}$$

where the dependent variable, $lnD_{i,t}$, is the log of yearly deforestation in municipality i in year t. We include time dummies to capture the effects of overall changes in the explanatory variables in deforestation common to all municipalities, λ_t. The vector of controls, $X_{i,t-1}$, includes economic activity (GDP), agricultural prices at municipality level (Soy Price and Timber Price) and state level (Beef Price), access to markets (Roads), the number of settlements within each municipality (Settlements) and number of families per size of settlement (Settlement Density), the total amount of rural credits granted to recipients in a given municipality (Rural Credits) and an indicator variable for the existence of housing projects (Housing Projects) within a municipality. Variables are also included to control for political factors (Mayor political party, Mayor education, Mayor gender, Mayor Age) and the amount of commissioned workers standardised by population at municipality level (Corruption), and whether the mayor is re-elected or not (Re-election). Our policy variables include fines per municipality in any given year (Fines), the sum of the area of conservational units plus indigenous land per municipality level (Protected Areas). Finally, we include dummy variables for when a municipality has an environmental law and an environmental office (with a value of 1 from the year it was in place and zero otherwise.). Finally, we include a variable to capture the degree of cloud cover (Clouds) and area of a municipality that is unobserved by satellite monitoring (No obs), $\mathrm{X}_{i,t}\beta_2$.

As we are dealing with grouped data, the errors are inter-related and hence we include Driscoll and Kraay (1998) standard errors that are robust to autocorrelation, heteroskedasticity and cross-sectional dependence. The year and municipality level dummies are included to capture time and municipality invariant effects (e.g. climate, infrastructure or population), respectively.

In terms of our right hand side variables in $X_{i,t-1}$, one can conjecture a number of ex-ante expectations of their role in deforestation. The commodity prices of beef, soy and timber are included to capture the influence of market conditions which reflect local variations in prices after average municipality differences and general trends have been removed. For example, an increase in the prices of beef or soy may lead to greater pressure for further deforestation to clear land for cultivation or grazing. In the case of the price of timber, the effect is ambiguous as on the one hand, as prices rise it may lead to a greater incentive to harvest. On the other hand, higher prices may lead to a greater awareness of the value of forests and lead to greater efforts to protect what is now a more valuable resource.

Turning to the environmental policies, the expectation is that any policy increases the risk of being caught which should reduce expected profits. Hence, we expect a negative sign on the coefficient of our environmental fines variable. The expected sign on our protected areas variable is ambiguous. On the one hand, if an area is protected in terms of additional enforcement patrols for example, this should reduce the area in a municipality that can be easily deforested. On the other hand, the decision to designate an area as protected may have been in response to previous deforestation Hargrave and Kis-Katos (2012) and may simply reflect that deforestation problems are more severe in that particular area. . Hence, the impact of the introduction of a PA might simply be to shift deforestation activities to areas just outside of the PA but still within the boundary of a given municipality. More than this, considering that federal and state government regulates most PAs, the role of municipal offices might be scant or insignificant to the environmental policy in place. (Girardi, 2017).[1]

In terms of other public policies, we expect housing projects to have a negative impact on deforestation as the provision of affordable housing should reduce the incentive for agents to move to the frontier and clear forest as a means of finding somewhere to live. The sign and significance of rural credit variable is uncertain. On the one hand, the availability of official subsidised credit may fuel deforestation by helping to fund clearing through investment in machinery, tools and fertilizers. On the other hand, if credit enables the development of a functioning market in forestry products it may improve forest management practices and reduce deforestation pressure. For our settlement variables, we expect that the number and density of settlements to result in an increase in the demand for deforestation although we expect this relationship to be non-linear (hence we include the squared terms of each variable).

Finally, our controls for cloud cover and unobserved area variables are expected to have positive and negative coefficients respectively. For unobserved areas, we expect that the greater the area not observed within a municipality, the more difficult it will be to detect the deforestation process. As for cloud cover, we argue that agents are aware of the policy monitoring program that employs satellite surveillance and hence may use cloudy days as a cover for their deforesting actions when they know that the satellite cannot detect deforestation and hence will reduce their chances of being caught and fined. Hence, the more cloudy days recorded in a given year is expected to be a positive determinant of deforestation within that year.

[1] According to ICMBIO (2017) and FUNAI (2017) the signalling process of a PA is through signboards (but not fences). These signboards are placed in strategic locations and act as a guide to citizens. Other protection procedures (for herds / planting areas or other protection needs) are implemented when requested by the community, but this happens in isolated and specific cases and it is not part of the delimitation of protected areas specifying their physical limits.

One of the main contributions of our study is to investigate the interaction between environmental institutions and environmental policies taking into account the expansion in commodity markets captured through the inclusion of commodity prices.[1]

Hence, we estimate (3.2) given by:

$$\begin{aligned} \ln D_{i,t} = \mu_i + \mathrm{X}_{i,t}\beta_1 + \mathrm{X}_{i,t-1}\beta_2 + (\text{IEF * Env. Policies})_{i,t-1}\beta_3 + \\ (\text{IEF * Prices})_{i,t-1}\beta_4 + \lambda_t + \varepsilon_{i,t} \end{aligned} \tag{1.2}$$

This equation allows us to estimate the extent to which the effectiveness of environmental policies depends on the IEF that operates within a municipality. We expect to see a reduction in those municipalities that no only have policies in place but they are reinforced by having proper enforcement of the law through a well functioning institutional framework. Likewise, in those municipalities where the institutional framework is weak then environmental policies may not be introduced and hence may increase deforestation rates if loggers are attracted to or pushed away from well regulated municipalities. One possible driver of the implementation and enforcement of environmental policies are the characteristics and behaviour of the mayor of a given municipality. Although we have no priors on age and gender we might expect education to have a negative influence of deforestation while we might expect a mayor that supports pro-farmer policies to have a positive influence on deforestation. A related issue is the level of corruption within a municipality which is expected to have a positive impact on deforestation. Finally, mayors running for re-election might be expected to exploit forests for economic reasons in the belief that this will create jobs and hence votes.

One concern with regard to the estimating equation 3.2 and equation 3.1 is the potential endogeneity of many of the explanatory variables, and hence their interpretation in terms

[1]The analysis conducted here does not limit the possibility of further studies using different approaches such as difference-in-difference techniques.

of causality. Since it would be difficult to find plausible instruments for many, if not for all, of our independent variables, we instead take a three pronged approach to alleviate any such concerns. First, we control for municipality fixed effects, allowing us to purge all time invariant unobservables from the specifications. Secondly, using an extensive set of other controls we hope that we have been able to capture all likely determinants of deforestation. Finally, we lag all control variables by one period, so that under assumption that, after controlling for fixed effects all confounding shocks are only contemporaneous in nature, we are left with only the exogenous variation in the elements of $X_{i,t-1}$.

1.4 Results

1.4.1 The effect of environmental policies and market condition on deforestation

Results from our estimation of equation 3.2 are presented in Table 1.2. Our baseline model includes 562 municipalities. In Columns (1) through (7) we experiment with including separately environmental policies, such as fines and protected areas, the prices of soy, timber and beef, and our IEF variables given by our environmental law and environmental office dummies. Taking our institutional variables first (dummies for an environmental office and an environmental law) we find that neither is significant when included separately in Columns (3) or (4) or together in Column (8). This suggests that having an office and an environmental law in place is necessary but not sufficient to reduce deforestation.

Turning to our other controls, in Column (8) we include all the variables together. One immediate observation from Column(8) is that our fines variable has no significant effect on the rate of deforestation despite the rapid growth in fines levied over this period. One explanation relates to enforcement. The environmental police often have limited financial resources and manpower to issue fines and more importantly often have insufficient legal means for collecting the fines once they have been issued (Araújo et al., 2017). For the PA variable we find that an increase in one unit of protected areas (10,000 km^2) is, everything else equal, associated with a 2.7% increase in deforestation the following year. To put this number into context if we use the mean value of deforestation for our sample (20.2 km^2) we find that this translates into a mean average increase in deforestation of 0.55 km^2. This intuitively unexpected results may be because PAs are established in response to previous deforestation pressure which is then displaced to neighbouring areas just outside the PAs and is therefore capturing the presence of active groups of loggers in a municipality. A second explanation is that it may be related to limited public monitoring and enforcement in sustainable use PAs where certain economic activities are allowed (Rico et al., 2017).[1]

For our market demand variables, we find that the estimated coefficient in Column (8) for the price of timber is negative and significant. Everything else equal, an increase in the timber price of 10% leads to an average decrease in deforestation of 2.7%. In terms of our sample, this percentage would mean average decreases in deforestation of approximately 0.6km^2. The coefficients on the price of beef and soy are not significant either together or individually.

In terms of our other economic policy variables, we find that the greater the number of settlements within a municipality the greater the degree of deforestation (a positive liner term) but at a decreasing rate (a negative squared term). Increasing the number of settlements will lead to, ceteris paribus, an increase in deforestation of nearly 6%. The turning point is within the sample range. We estimate the turning point to be 16 settlements within a municipality and from our sample we know that only five percent of

[1]Rico et al. (2017) examines the dynamics of forest loss and governance in PA's in the Peruvian Amazon.

our municipalities sample have passed the turning point which suggests that for the vast majority of municipalities additional settlements lead to increased deforestation.[1]

For our settlement density variable we find that an increase in the density of settlements within a municipality leads to higher rates of deforestation but that the effect is non-linear so that deforestation increases but at a decreasing rate. This suggests that more densely populated settlements deforest more presumably as there is greater pressure on land required for expansion. The turning point is 0.62 and within our sample only one settlement (Altamira in Para) has surpassed that value and is one of the more economically developed settlements. Hence, increasing the number of families per settlement within a municipality in 10% is associated with a decreasing rate of 0.6%. Likewise, as expected, housing projects have a negative effect on deforestation since the incentives for agents living in poor quality housing to clear land to live is reduced. The estimated coefficient implies that the existence of housing projects within a municipality decreases logged deforestation by 15%. To put this number into context, we use the mean value of deforestation for our sample (20.2 km^2) which shows that the average decrease in deforestation is around 3 km^2.

Turning to the political factors, only mayor age and the level of corruption have a significant effect on deforestation. We find that an increase in age above the average (50 years) results in an average decrease in deforestation of 5.6%. At the mean value of deforestation for our sample, this represents an average decrease in deforested areas of 1.1km^2. One explanation is that younger politicians who are keen to remain in power are more likely to allow activities that are more harmful to the environment. Corruption, captured by the number of commissioned workers per capita, is found to lead to higher deforestation rates within a municipality. An increase in the number of commissioned workers per capita of one unit is associated with an increase of 2.4% in deforestation. For our sample, this would represents an increase of 0.48 km^2 per each

[1] Conceição do Araguaia - Pará, Itupiranga - Pará, Marabá - Pará, Novo Repartimento - Pará and Zé Doca - Maranhão represent the municipalities with the highest number of settlements.

additional commissioned worker averaged by the population. Our measure of corruption is positive and significant in columns (1) through (8), which reinforces the need for some political control. Gender, being a pro-farmer politician, and having an education level of tertiary of above have no effect on deforestation rates.

Table 1.2 Effects of Environmental Policies and Commodities Prices on Deforestation

Sample	Dependent: Δ *ln* Deforestation							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sett$_{-1}$	0.080***	0.070***	0.078***	0.079***	0.074***	0.079***	0.078***	0.064***
	(5.36)	(4.59)	(5.61)	(5.50)	(4.87)	(5.50)	(5.08)	(3.98)
Sett$_{-1}$ sq	-0.002**	-0.002**	-0.002***	-0.002***	-0.002**	-0.002**	-0.002**	-0.002**
	(3.04)	(2.90)	(3.16)	(3.18)	(2.97)	(2.94)	(3.08)	(2.48)
Sett Dens$_{-1}$	4.056**	4.510***	4.159**	4.120**	3.866**	4.397***	4.116**	4.685***
	(2.99)	(3.36)	(2.89)	(2.97)	(2.97)	(3.13)	(2.92)	(3.73)
Sett Dens$_{-1}$ sq	-3.639***	-3.212**	-3.620***	-3.581***	-3.283***	-3.841***	-3.581***	-3.353***
	(3.51)	(2.88)	(3.41)	(3.50)	(3.60)	(3.70)	(3.47)	(3.34)
Rural Credits$_{-1}$	0.025	0.025*	0.024	0.024	0.024	0.024	0.028	0.028
	(1.75)	(1.81)	(1.77)	(1.77)	(1.71)	(1.67)	(1.61)	(1.51)
Housing Projects$_{-1}$	-0.147**	-0.150**	-0.147**	-0.147**	-0.150**	-0.147**	-0.149**	-0.153**
	(2.92)	(2.83)	(2.91)	(2.92)	(2.91)	(2.93)	(2.80)	(2.71)
GDP$_{-1}$	-0.237	-0.249	-0.234	-0.248	-0.232	-0.261	-0.232	-0.274
	(1.50)	(1.56)	(1.46)	(1.53)	(1.50)	(1.60)	(1.52)	(1.66)
GDP$_{-1}$ sq	0.027	0.028	0.027	0.027	0.025	0.031	0.026	0.031
	(1.48)	(1.51)	(1.46)	(1.49)	(1.48)	(1.68)	(1.51)	(1.74)
Mayor Party$_{-1}$	0.041	0.039	0.039	0.042	0.037	0.039	0.042	0.036
	(0.90)	(0.86)	(0.86)	(0.93)	(0.85)	(0.87)	(0.89)	(0.81)
Mayor Education$_{-1}$	0.010	0.016	0.011	0.011	0.012	0.019	0.011	0.024
	(0.31)	(0.57)	(0.35)	(0.38)	(0.41)	(0.68)	(0.38)	(0.96)
Mayor Age$_{-1}$	-0.065**	-0.062**	-0.063**	-0.064**	-0.060**	-0.062**	-0.060**	-0.056**
	(2.46)	(2.38)	(2.40)	(2.46)	(2.41)	(2.29)	(2.61)	(2.39)
Mayor Gender$_{-1}$	0.025	0.025	0.023	0.021	0.020	0.030	0.022	0.030
	(0.69)	(0.66)	(0.61)	(0.55)	(0.53)	(0.79)	(0.58)	(0.78)
Corruption$_{-1}$	0.026**	0.026**	0.026**	0.025**	0.026**	0.026**	0.025**	0.024**
	(2.94)	(2.98)	(2.78)	(2.78)	(2.85)	(2.88)	(2.70)	(2.90)
Re-election$_{-1}$	0.004	0.002	0.003	0.003	-0.004	0.003	-0.000	-0.008
	(0.17)	(0.11)	(0.14)	(0.14)	(0.21)	(0.11)	(0.01)	(0.35)
Roads$_{-1}$	-0.614	1.218	-0.954	-0.697	1.134	-0.646	-5.817	-2.007
	(0.08)	(0.15)	(0.12)	(0.09)	(0.17)	(0.08)	(0.34)	(0.15)
Clouds	0.001	0.000	0.001	0.001	0.001	0.001	0.001	0.000
	(0.61)	(0.47)	(0.60)	(0.59)	(0.63)	(0.59)	(0.61)	(0.47)
NoObs	-0.004	-0.002	-0.004	-0.003	-0.003	-0.003	-0.003	0.000
	(0.47)	(0.25)	(0.50)	(0.42)	(0.37)	(0.41)	(0.42)	(0.09)
Fines$_{-1}$	-0.015							-0.018
	(0.73)							(0.83)
PAs$_{-1}$		0.027***						0.027***
		(4.10)						(3.89)
Env. Law$_{-1}$			0.022					0.022
			(0.95)					(1.00)
Env. Office$_{-1}$				-0.054				-0.058
				(1.55)				(1.51)
Soy Price$_{-1}$					0.630			0.549
					(1.27)			(1.18)
Timber Price$_{-1}$						-1.331*		-1.285*
						(2.02)		(2.11)
Beef Price$_{-1}$							1.664	1.180
							(0.58)	(0.49)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	6,178	6,178	6,178	6,178	6,178	6,178	6,178	6,178

*, **, *** denote significance at 10%, 5% and 1% levels, respectively. Robust standard error values are indicated in parentheses under the coefficients. PAs stand for Protected Areas.

1.4.2 Environmental policies and market condition conditioning on institutional environmental framework

In Table 1.3 we consider the impact of our policy variables conditional on the institutional environmental framework including the same set of controls from Table 1.2 with the year fixed effects (Columns 1-8). In Column 9 we show that agents from municipalities that receive fines when there is an environmental law in place do not reduce deforestation. One explanation is that in the Legal Amazon many local environmental laws are not enforced even though the municipalities have the power to do so. Thus, it seems more likely that the offender will be punished at the state or federal level rather than at the municipality level, which contributes to the limited effectiveness of fines as a punitive instrument.[1]

In contrast, we find that fines do deter deforestation if the municipality has an environmental office. More specifically, the estimated average drop in deforestation associated with the fines policy when there is an environmental office all else equal is around 1.8 km^2. This finding illustrates that a coordinated process of implementation of the national environmental system (SISNAMA) that decentralises forces to assist the protection of the environment can help to reduce deforestation. Hence, environmental offices appear to act as a primary partner to the SISNAMA together with IBAMA, which operates at the national level as an administrative arm and the environmental police working through the Brazilian Ministry of Environment. This coalition appears to make the institutional framework work more effectively at the local level.

However, we also find that municipalities with an environmental office and protected areas within their territory tend to increase deforestation by approximately 3%. Using the mean value of deforestation for our sample (20.2 km^2) we translates into an average

[1]We regress the baseline model as pooled regression and, incrementally, add time fixed effects, municipality fixed effects and two-way fixed effects so we can observe the contribution to the results. The table is shown in the Appendix A.2. From the results, the model improves significantly when adding time and municipality fixed effects.

increase in deforestation of 0.6 km^2. This unexpected result may be explained by the fact that environmental offices are often set up to work together with other offices, such as agriculture, tourism, education, and infrastructure offices. In this sense environmental management may be associated with other themes and the creation of joint offices can lead to conflict with the agendas of others. For example, given that the expansion of agriculture is an important determinant of deforestation, having a joint office for the environment and agriculture could lead to a conflict of interest. Regarding protected areas, conditioned on the existence of environmental laws at the municipality level, we find no significant effect. This suggests that protected areas are not effectively monitored at the local level since a substantial number of these units are regulated at the state and federal level.

Turning to our commodity market variables, we find that the IEF plays an important role when we condition on commodity prices. An increase in soy prices, conditioned on the existence of environmental law, of 10% corresponds to an average decreases in deforestation of almost 5%. Much of this effect is likely to be the result of the moratorium implemented in 2006.[1] In this sense, the IEF was used to help compliance with the soy moratorium. However, when looking to the overall effect we find that the soy price does have a positive impact on deforestation. An increase in the soy price by 10% at the mean values for the existence of environmental law is, all else equal, associated with an average increase in deforestation of approximately 15.8%. We find no significant effect for other commodities when conditioning on the existence of an environmental law. [2]

In contrast, timber (negative) and beef (positive) prices are significant when conditioned on the existence of an environmental office. This is consistent with our assumptions that

[1]The moratorium establishes that companies purchasing grain and its derivatives after 2008 cannot deal with those farmers who grew the grain in deforested areas, within indigenous lands or that are on a slave labour list. The moratorium prohibited soy producers from negotiating the sale of production originating from deforested areas in the legal Amazon for a period of two years.

[2]Recently, Caviglia-Harris (2018) addressed the existence of a nonlinear relationship with the intensification of cattle production in part of the Legal Amazon and found that as farms become more intensive the demand for newly cleared land increases, but this then decreases with further intensification.

environmental management offices are associated with other strategic themes and that an institutional framework on its own fails to protect the environment. From our sample, forestry prices have a negative impact on deforestation and beef market has a positive impact on deforestation. Taking the mean value of deforestation (20.2 km^2) we find out that an increase in the timber price of 10% results in an average decreases in deforestation of 0.3 km^2 and for the beef market the result is an average increases in deforestation of 2.1 km^2.

Table 1.3 Effects of Environmental Policies on Deforestation: Baseline Model

Sample	Dependent: Δ *ln* Deforestation								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Fines$_{-1}$	0.079	-0.019	-0.018	-0.017	-0.018	0.080*	0.083*	0.069	0.069
	(1.73)	(0.89)	(0.84)	(0.80)	(0.84)	(1.82)	(1.90)	(1.64)	(1.62)
Protected Areas$_{-1}$	0.027***	-0.004	0.026***	0.026***	0.027***	-0.004	-0.005	-0.003	-0.003
	(3.99)	(0.38)	(3.73)	(3.79)	(3.91)	(0.35)	(0.43)	(0.30)	(0.30)
Env. Law$_{-1}$	0.024	0.008	0.478***	0.079*	0.072*	0.010	0.470***	0.532***	0.538***
	(1.04)	(0.36)	(3.38)	(2.12)	(2.19)	(0.39)	(3.27)	(3.18)	(3.22)
Env. Office$_{-1}$	-0.041	-0.094*	0.122	0.131**	-0.101	-0.077	0.125	0.293	0.287
	(1.03)	(2.06)	(0.63)	(2.38)	(1.27)	(1.61)	(0.63)	(1.45)	(1.45)
Soy Price$_{-1}$	0.551	0.558	1.068**	0.553	0.549	0.561	1.109***	1.114***	1.188**
	(1.17)	(1.21)	(3.06)	(1.20)	(1.18)	(1.21)	(3.20)	(3.12)	(3.10)
Timber Price$_{-1}$	-1.263*	-1.282*	-1.294*	0.855	-1.280*	-1.261*	-1.270*	0.686	0.639
	(2.08)	(2.19)	(2.12)	(1.13)	(2.11)	(2.16)	(2.17)	(0.90)	(0.81)
Beef Price$_{-1}$	1.089	1.109	1.232	0.791	1.200	1.014	1.060	0.723	0.685
	(0.44)	(0.47)	(0.53)	(0.33)	(0.50)	(0.42)	(0.46)	(0.31)	(0.29)
Env. Law$_{-1}$* Fines$_{-1}$	-0.009					-0.012	-0.014	-0.012	-0.012
	(0.47)					(0.71)	(0.91)	(0.74)	(0.73)
Env. Office$_{-1}$* Fines$_{-1}$	-0.101**					-0.101***	-0.103***	-0.089**	-0.088**
	(3.03)					(3.12)	(3.18)	(2.79)	(2.77)
Env. Law$_{-1}$* PAs$_{-1}$		0.007				0.008	0.008	0.007	0.007
		(1.29)				(1.44)	(1.50)	(1.40)	(1.38)
Env. Office$_{-1}$* PAs$_{-1}$		0.029**				0.029**	0.029**	0.028**	0.028**
		(2.95)				(2.94)	(2.93)	(2.82)	(2.83)
Env. Law$_{-1}$* Soy Price$_{-1}$			-0.716***				-0.721***	-0.736***	-0.729***
			(3.55)				(3.40)	(3.47)	(3.16)
Env. Office$_{-1}$* Soy Price$_{-1}$			-0.308				-0.346	-0.339	-0.447*
			(1.09)				(1.21)	(1.25)	(1.90)
Env. Law$_{-1}$* Timber Price$_{-1}$				-0.483*				-0.442	-0.450
				(2.19)				(1.65)	(1.69)
Env. Office$_{-1}$* Timber Price$_{-1}$				-2.120***				-1.941***	-1.891***
				(6.92)				(5.67)	(5.35)
Env. Law$_{-1}$* Beef Price$_{-1}$					-0.060**				-0.011
					(2.54)				(0.50)
Env. Office$_{-1}$* Beef Price$_{-1}$					0.049				0.070**
					(1.26)				(2.26)
Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	6,178	6,178	6,178	6,178	6,178	6,178	6,178	6,178	6,178

*, **, *** denote significance at 10%, 5% and 1% levels, respectively. Robust standard error values are indicated in parentheses under the coefficients. PAs stand for Protected Areas.

1.4.3 Spatial Analysis

One concern with our previous analysis is the possibility of spatial dependence i.e. deforestation in one municipality affects neighbouring municipalities and when there is a strong institutional environment in one municipality if may lead to a shift in deforestation to nearby municipalities with lower enforcement capabilities. In order to test this assumption, we conduct a spatial analysis following Anselin (1988) and Elhorst (2012). The existence of spatial autocorrelation can occur if a given event in a given place has a significant impact on neighbouring regions such that there may be spillover effects from certain activities or factors that may affect both the characteristics of the data and the nature of the events. Spatial econometric models with panel data capture the relationship between variables over time and arising for the existence for spatial autocorrelation. We assume that the area studied remains constant over our sample period (Almeida, 2012). Our spatial analysis is estimated on a balanced panel which reduces our sample from 562 to 456 municipalities which still represents 65% of total Legal Amazon municipalities. We perform LM tests to establish the dependence factor (see Table A.1.7).The test results point to the spatial lag model specification being the preferred specification. In view of testing procedure we ran two models, a spatial auto-regressive model (SAR) and a spatial error model (SEM) (not shown here). The SAR is preferred in all specifications.

To incorporate the autocorrelation of spatial lag type of dependent variable, our fixed effects model needs to be modified, generating the Spatial Autoregressive model with fixed effects:

$$\begin{gathered}\ln D_{i,t} = \mu_i + \rho W_1 y_t + \mathrm{X}_{i,t}\beta_1 + \mathrm{X}_{i,t-1}\beta_2 + (\text{IEF * Env. Policies})_{i,t-1}\beta_3 + \\ (\text{IEF * Prices})_{i,t-1}\beta_4 + \lambda_t + \varepsilon_{i,t}\end{gathered} \tag{1.3}$$

where $W_1 y_t$ is the spatial lag of the dependent variable. The spatial weight matrix W is defined according to various criteria and is kept unchanged for all years of the panel. The calculation of the spatial weight matrices used the coordinates from the municipality's centroid. We calculate three different weight matrices. Our baseline weight matrix (W1) finds the 5 nearest neighbours and constructs a spatial weight matrix based on this number of neighbours, normalised to have row-sums of unity. Alternatively, we create a distance based spatial weight matrix (W2) with bands using latitude and longitude coordinates a the Great Circle distance formula. The bands used were 1km to 100km. Finally, we construct a distance based spatial weight matrix (W3) using latitude and longitude coordinates and the Great Circle distance formula (Lacombe's Method) (Elhorst, 2012). Using these techniques we gradually increase the number of neighbours.

Since it is not possible to compare the coefficient estimates in the non-spatial model, our baseline, with their counterparts in the spatial models SAR, we use the direct and indirect effects estimates as a result of impacts passing through neighbouring municipalities and back to the municipalities themselves. These feedback effects are partly due to the coefficient on the spatially lagged dependent variable. Our approach is to follow LeSage and Pace (2009) and present the direct, indirect and total effect of the coefficients separately.

Table 1.4 reports our results for the three different matrices. All results are consistent with the literature (Lambin and Geist, 2006; Lambin and Meyfroidt, 2011). We first consider the direct effect of our estimates. We find that one of the institutional apparatus,environmental law, has no direct effect when conditioning on protected areas and, timber and beef prices. Our spatial correlation is positive and significant and can be interpreted as the impact of changes in municipality on deforestation. The impact is magnified by 1.2 and 1.6 through the spatial lag in the model. For our baseline weight matrix (W), we can see that the direct effect of fines, from the SAR model, on deforestation was underestimated by a factor of 0.75. Since the direct effect of fines is 0.517 and its coefficient estimate 0.596, its feedback effect is around 0.07 or 15% of the direct effect. For environmental law and having an environmental office, using the second weighting method, the effects amount to 0.24 and 0.15 or 79% and 14% of the direct effect respectively. Looking to the indirect effect on the third model (W3), we can clearly see that the indirect effect of a change in the amount of fines issued appears to be 62% of the direct effect. If the amount of fines issued increase in one municipality, the change in neighbouring municipalities to the change in the municipality itself is in the proportion of approximately 1 to 1.6 in case of increasing fines.

Overall, we observe that the IEF, when established in a municipality, tends to lead to reduced deforestation in neighbouring municipalities. As an example, in all of our specifications, conditioning the existence of the IEF and the exercise of fines, we have a negative effect on neighbouring municipalities. We see the same effect for the market prices conditioned on the IEF which is consistent with our main findings 1.3.

Table 1.4 Spatial Analysis - Baseline Model

Dependent: Δ *ln* Deforestation			
Sample	W1	W2	W3
Fines_{-1}	0.596 (10.549)	0.595 (10.532)	0.599 (10.742)
$\text{Protected Areas}_{-1}$	-0.025 (-1.891)	-0.025 (-1.890)	-0.030 (-2.281)
Env. Law_{-1}	0.542 (2.369)	0.543 (2.375)	0.578 (2.560)
Env. Office_{-1}	1.193 (4.684)	1.191 (4.677)	1.221 (4.854)
Soy Price_{-1}	1.175 (3.080)	1.174 (3.078)	1.261 (3.348)
Timber Price_{-1}	4.429 (6.034)	4.428 (6.032)	4.200 (5.794)
Beef Price_{-1}	1.963 (1.627)	1.962 (1.626)	1.752 (1.470)
Env. Law_{-1}* Fines_{-1}	-0.174 (-5.164)	-0.174 (-5.152)	-0.174 (-5.216)
Env. Office_{-1}* Fines_{-1}	-0.203 (-3.767)	-0.202 (-3.755)	-0.213 (-4.015)
Env. Law_{-1}* PAs_{-1}	-0.001 (-0.113)	-0.001 (-0.118)	0.001 (0.163)
Env.Office_{-1}*PAs_{-1}	0.014 (1.073)	0.014 (1.072)	0.016 (1.264)
Env. Law_{-1}* Soy Price_{-1}	-0.823 (-2.296)	-0.825 (-2.301)	-0.846 (-2.390)
Env. Office_{-1}* Soy Price_{-1}	-1.036 (-2.504)	-1.035 (-2.500)	-1.094 (-2.676)
Env. Law_{-1}* Timber Price_{-1}	0.707 (1.317)	0.708 (1.317)	0.522 (0.984)
Env. Office_{-1}* Timber Price_{-1}	-2.442 (-3.171)	-2.440 (-3.168)	-2.350 (-3.089)
Env. Law_{-1}* Beef Price_{-1}	-0.050 (-0.506)	-0.050 (-0.507)	-0.059 (-0.607)
Env. Office_{-1}* Beef Price_{-1}	-0.034 (-0.355)	-0.034 (-0.352)	-0.026 (-0.267)
σ^2	2.081	2.081	2.030
ρ	0.168 (8.904)	0.167 (8.847)	0.389 (14.711)
(Pseudo) R^2	0.44	0.45	0.46
LogLik	-9541.60	-9541.62	-9484.62
AIC	19155	19155	19042
BIC	19393	19393	19280

t-values are indicated in parentheses next to the coefficients. PAs stands for Protected Areas.

Table 1.5 Spatial Analysis - Baseline Model - Direct Effects

Dependent: Δ *ln* Deforestation			
Sample	W1	W2	W3
Direct Effects			
Fines$_{-1}$	0.599 (10.501)	0.598 (10.485)	0.603 (10.840)
Protected Areas$_{-1}$	-0.024 (-1.800)	-0.024 (-1.800)	-0.030 (-2.292)
Env. Law$_{-1}$	0.550 (2.386)	0.552 (2.393)	0.597 (2.633)
Env. Office$_{-1}$	1.201 (4.612)	1.200 (4.606)	1.238 (4.881)
Soy Price$_{-1}$	1.197 (3.150)	1.196 (3.148)	1.287 (3.381)
Timber Price$_{-1}$	4.454 (6.036)	4.453 (6.035)	4.225 (6.012)
Beef Price$_{-1}$	1.964 (1.594)	1.964 (1.593)	1.724 (1.398)
Env. Law$_{-1}$* Fines$_{-1}$	-0.173 (-5.157)	-0.173 (-5.161)	-0.174 (-5.179)
Env. Office$_{-1}$* Fines$_{-1}$	-0.205 (-3.784)	-0.205 (-3.773)	-0.214 (-3.924)
Env. Law$_{-1}$* PAs$_{-1}$	-0.001 (-0.115)	-0.001 (-0.121)	0.001 (0.121)
Env.Office$_{-1}$*PAs$_{-1}$	0.013 (1.011)	0.014 (1.010)	0.016 (1.253)
Env. Law$_{-1}$* Soy Price$_{-1}$	-0.843 (-2.360)	-0.846 (-2.366)	-0.870 (-2.454)
Env. Office$_{-1}$* Soy Price$_{-1}$	-1.052 (-2.491)	-1.051 (-2.488)	-1.114 (-2.668)
Env. Law$_{-1}$* Timber Price$_{-1}$	0.696 (1.304)	0.696 (1.305)	0.515 (0.968)
Env. Office$_{-1}$* Timber Price$_{-1}$	-2.422 (-3.091)	-2.420 (-3.088)	-2.350 (-3.183)
Env. Law$_{-1}$* Beef Price$_{-1}$	-0.045 (-0.458)	-0.046 (-0.459)	-0.059 (-0.619)
Env. Office$_{-1}$* Beef Price$_{-1}$	-0.032 (-0.324)	-0.032 (-0.322)	-0.026 (-0.277)

t-values are indicated in parentheses next to the coefficients. PAs stands for Protected Areas.

Table 1.6 Spatial Analysis - Baseline Model - Indirect Effects

Dependent: Δ *ln* Deforestation			
Sample	W1	W2	W3
	Indirect Effects		
Fines_{-1}	0.119 (6.203)	0.119 (6.180)	0.375 (7.063)
$\text{Protected Areas}_{-1}$	-0.004 (-1.736)	-0.005 (-1.735)	-0.018 (-2.210)
Env. Law_{-1}	0.109 (2.282)	0.109 (2.287)	0.372 (2.520)
Env. Office_{-1}	0.240 (3.866)	0.238 (3.857)	0.771 (4.234)
Soy Price_{-1}	0.239 (2.862)	0.237 (2.859)	0.803 (3.121)
Timber Price_{-1}	0.889 (4.675)	0.884 (4.665)	2.631 (4.955)
Beef Price_{-1}	0.393 (1.560)	0.391 (1.559)	1.070 (1.380)
Env. Law_{-1}* Fines_{-1}	-0.034 (-4.281)	-0.034 (-4.268)	-0.108 (-4.584)
Env. Office_{-1}* Fines_{-1}	-0.041 (-3.378)	-0.041 (-3.367)	-0.133 (-3.582)
Env. Law_{-1}* PAs_{-1}	-0.000 (-0.114)	0.000 (-0.120)	0.001 (0.119)
Env.Office_{-1}*PAs_{-1}	0.002 (0.996)	0.003 (0.994)	0.010 (1.239)
Env. Law_{-1}* Soy Price_{-1}	-0.168 (-2.281)	-0.167 (-2.285)	-0.541 (-2.364)
Env. Office_{-1}* Soy Price_{-1}	-0.210 (-2.312)	-0.209 (-2.308)	-0.694 (-2.536)
Env. Law_{-1}* Timber Price_{-1}	0.139 (1.273)	0.138 (1.274)	0.320 (0.962)
Env. Office_{-1}* Timber Price_{-1}	-0.483 (-2.849)	-0.480 (-2.845)	-1.464 (-2.958)
Env. Law_{-1}* Beef Price_{-1}	-0.009 (-0.458)	-0.009 (-0.459)	-0.037 (-0.612)
Env. Office_{-1}* Beef Price_{-1}	-0.006 (-0.321)	-0.06 (-0.318)	-0.016 (-0.274)

t-values are indicated in parentheses next to the coefficients. PAs stands for Protected Areas.

Table 1.7 Spatial Analysis - Baseline Model - Total Effects

Dependent: Δ *ln* Deforestation			
Sample	W1	W2	W3
	Total Effects		
Fines_{-1}	0.718 (10.253)	0.717 (10.239)	0.978 (9.886)
$\text{Protected Areas}_{-1}$	-0.029 (-1.798)	-0.029 (-1.798)	-0.048 (-2.277)
Env. Law_{-1}	0.660 (2.385)	0.662 (2.392)	0.969 (2.615)
Env. Office_{-1}	1.441 (4.575)	1.438 (4.568)	2.009 (4.744)
Soy Price_{-1}	1.436 (3.135)	1.433 (3.133)	2.091 (3.326)
Timber Price_{-1}	5.344 (5.965)	5.337 (5.964)	6.856 (5.781)
Beef Price_{-1}	2.358 (1.594)	2.355 (1.594)	2.795 (1.396)
Env. Law_{-1}* Fines_{-1}	-0.207 (-5.146)	-0.207 (-5.135)	-0.283 (-5.097)
Env. Office_{-1}* Fines_{-1}	-0.246 (-3.771)	-0.246 (-3.760)	-0.347 (-3.863)
Env. Law_{-1}* PAs_{-1}	-0.001 (-0.115)	-0.001 (-0.121)	0.002 (0.121)
Env.Office_{-1}*PAs_{-1}	0.016 (1.011)	0.016 (1.010)	0.027 (1.252)
Env. Law_{-1}* Soy Price_{-1}	-1.011 (-2.363)	-1.013 (-2.369)	-1.411 (-2.441)
Env. Office_{-1}* Soy Price_{-1}	-1.262 (-2.480)	-1.260 (-2.477)	-1.808 (-2.644)
Env. Law_{-1}* Timber Price_{-1}	0.835 (1.303)	0.835 (1.304)	0.836 (0.968)
Env. Office_{-1}* Timber Price_{-1}	-2.906 (-3.083)	-2.900 (-3.080)	-3.814 (-3.136)
Env. Law_{-1}* Beef Price_{-1}	-0.054 (-0.458)	-0.055 (-0.460)	-0.097 (-0.618)
Env. Office_{-1}* Beef Price_{-1}	-0.038 (-0.324)	-0.038 (-0.322)	-0.042 (-0.276)

t-values are indicated in parentheses next to the coefficients. PAs stands for Protected Areas.

1.4.4 Robustness checks

In order to deal with concerns surrounding the delimitation of the Legal Amazon that is touched on in a number of studies (Hargrave and Kis-Katos (2012) and Nepstad et al. (2014)), we re-estimate Column (9) of Table 1.3 taking into account remaining forest cover.[1] In studies including municipalities with low levels of forest, it is possible that deforestation rates are low because there is less forest to be cleared. We check whether the dynamics of deforestation presented in the results from Column (9) of table 1.3 change when we use a set of different percentages of forest cover. Table 1.8 shows the results. Column (1) shows that municipalities with at least 5% of remaining forest in any given year and columns (2) and (3) represent municipalities with at least 10% and 15% of remaining forest cover in any given year. The results confirm that our main findings hold. Having a protected area within municipality when there is an environmental office still increases deforestation. In Column (3) we find that an increase in protected areas by one unit (10,000 km^2) at the mean value for the existence of environmental office is linked to an approximate average increase in deforestation of 2%. This result suggest that the existence of an IEF changes the results when we reduce the sample to those municipalities with a considerable amount of forest remaining. More precisely, fines conditional on there being an environmental office no longer have a negative effect on deforestation. This may indicate that municipalities with high levels of remaining forest tend to concentrate more in isolated areas with limited access to environmental police. Commodity prices conditioned on IEF are still significant. An increase in the timber price of 10%, at the mean value for the existence of environmental office, is associated with an average decrease in deforestation of 0.2 km^2. The beef price conditioned on environmental offices has no effect when we have municipalities with at least 15% of forest cover in any given year although we find a negative impact on deforestation when conditioning beef prices on environmental law. The decrease in deforestation within a municipality associated with a 10% increase in the beef price is on average 1.8 km^2. For soy, the overall impact of a 10% increase in the soy price leads to, ceteris paribus, an average increase in deforestation of 3.2 km^2 when 15% of forest remains.

[1]We re-estimate Table 1.3 adding a lagged dependent variable since we might expect that the current level of deforestation be heavily determined by its past level. In that case, not including the lagged variable of deforestation would lead to omitted variable bias and the results could be unreliable. However, the results presented in the Appendix A.1 exclude the possibility of a dynamic panel or dependence of the model. In this case, the baseline model stated is preferred to others.

Table 1.8 Environmental policies and municipal institutional framework - Robustness Check for Forest Cover

Sample	Dependent: Δ *ln* Deforestation		
	(1)	(2)	(3)
Fines$_{-1}$	0.069	0.073	0.095
	(1.62)	(1.59)	(1.29)
Protected Areas$_{-1}$	-0.003	-0.002	-0.001
	(0.30)	(0.16)	(0.10)
Env. Law$_{-1}$	0.538***	0.609***	0.662***
	(3.22)	(3.30)	(3.54)
Env. Office$_{-1}$	0.287	0.297	0.256
	(1.45)	(1.72)	(1.21)
Soy Price$_{-1}$	1.188**	1.342***	1.278**
	(3.10)	(3.22)	(2.97)
Timber Price$_{-1}$	0.639	0.880	0.485
	(0.81)	(0.99)	(0.46)
Beef Price$_{-1}$	0.685	1.408	1.464
	(0.29)	(0.46)	(0.42)
Env. Law$_{-1}$* Fines$_{-1}$	-0.012	-0.021	-0.021
	(0.73)	(1.27)	(1.49)
Env. Office$_{-1}$* Fines$_{-1}$	-0.088**	-0.086*	-0.102
	(2.77)	(2.13)	(1.55)
Env. Law$_{-1}$* PAs$_{-1}$	0.007	0.006	0.002
	(1.38)	(1.35)	(0.37)
Env.Office$_{-1}$*PAs$_{-1}$	0.028**	0.025**	0.023***
	(2.83)	(3.05)	(3.11)
Env. Law$_{-1}$* Soy Price$_{-1}$	-0.729***	-0.831***	-0.790***
	(3.16)	(3.33)	(3.20)
Env. Office$_{-1}$* Soy Price$_{-1}$	-0.447*	-0.389*	-0.315
	(1.90)	(1.81)	(1.39)
Env. Law$_{-1}$* Timber Price$_{-1}$	-0.450	-0.178	-0.227
	(1.69)	(0.61)	(1.10)
Env. Office$_{-1}$* Timber Price$_{-1}$	-1.891***	-2.286***	-2.288***
	(5.35)	(5.41)	(5.31)
Env. Law$_{-1}$* Beef Price$_{-1}$	-0.011	-0.041	-0.082***
	(0.50)	(1.76)	(3.93)
Env. Office$_{-1}$* Beef Price$_{-1}$	0.070**	0.063*	0.070
	(2.26)	(2.10)	(1.78)
Forest Cover	5%	10%	15%
Set of Controls	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes
Number of Observations	6,178	5,241	4,476

*, **, *** denote significance at 10%, 5% and 1% levels, respectively. Robust standard error values are indicated in parentheses under the coefficients. PAs stand for Protected Areas.

In many studies protected areas are captured via two distinct variables, namely conservation units and indigenous land. Since we are summing the areas of indigenous land and conservation units within a municipality, we need to check whether we are underestimating the true value of this policy. Based on Table 1.3, we re-estimate our results excluding protected areas from the analysis and including indigenous land and conservation units as separate variables. Column (9) from Table 1.9 shows that indigenous land is insignificant when conditioned on the institutional framework and conservation units continue to influence deforestation when a municipality has an environmental institutional environmental framework system. An increase in protected areas of one unit (10,000 km^2) at mean values for the existence of an environmental office leads to an increase in deforestation of 3% or 0.6km^2 holding everything else constant. One explanation is that indigenous lands tend to be regulated and managed by the federal government and there is no interference or influence from the local governmental body. A second explanation is that the enforcement of indigenous land within a municipality has no effect on deforestation, as shown by BenYishay et al. (2017), so the IEF triggers no effect. In this setting, fines, conditioned on there being an environmental office, increasing by one unit ($10,000,000 BRL) leads to an almost 1.5% decrease in deforestation. Putting this percentage in numbers, we consider our deforestation sample (20,2 km^2) and find out that it represents to approximately 1.7 km^2.

For commodities prices we observe the same patterns that we find in our baseline model. Increasing the price of soy and beef by 10%, holding everything else constant and at mean values of the environmental institutional framework, we see an increase in deforestation by 0.06% and 62% respectively. Likewise, an increase in 10% of the price for forestry products leads to a decrease in deforestation by 1.5% or 0.3 km^2 when considering our sample.

Table 1.9 Environmental policies and municipal institutional framework - Robustness Check for Protected Areas

Sample	Dependent: Δ *ln* Deforestation								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Fines$_{-1}$	0.080	-0.019	-0.019	-0.018	-0.019	0.078	0.081*	0.067	0.067
	(1.75)	(0.89)	(0.86)	(0.83)	(0.86)	(1.77)	(1.83)	(1.59)	(1.57)
Conservation Units$_{-1}$	0.026***	-0.006	0.025***	0.025***	0.026***	-0.006	-0.007	-0.005	-0.005
	(3.84)	(0.51)	(3.58)	(3.67)	(3.77)	(0.47)	(0.54)	(0.42)	(0.42)
Indigenous Land$_{-1}$	0.181	0.268*	0.161	0.182	0.174	0.262*	0.248*	0.251*	0.254*
	(1.75)	(2.07)	(1.55)	(1.77)	(1.57)	(1.98)	(1.85)	(2.03)	(2.02)
Env. Law$_{-1}$	0.024	0.012	0.476***	0.078*	0.072**	0.014	0.475***	0.536***	0.542***
	(1.01)	(0.52)	(3.35)	(2.08)	(2.26)	(0.54)	(3.20)	(3.12)	(3.17)
Env. Office$_{-1}$	-0.039	-0.078	0.121	0.134**	-0.101	-0.062	0.129	0.297	0.291
	(0.78)	(1.48)	(0.58)	(2.13)	(1.11)	(1.14)	(0.59)	(1.37)	(1.37)
Soy Price$_{-1}$	0.552	0.544	1.067**	0.555	0.551	0.548	1.086***	1.093***	1.166***
	(1.18)	(1.19)	(3.07)	(1.21)	(1.18)	(1.19)	(3.22)	(3.14)	(3.13)
Timber Price$_{-1}$	-1.258*	-1.281*	-1.289*	0.872	-1.275*	-1.259*	-1.268*	0.682	0.637
	(2.08)	(2.19)	(2.13)	(1.15)	(2.11)	(2.16)	(2.17)	(0.89)	(0.80)
Beef Price$_{-1}$	1.122	1.085	1.254	0.817	1.228	1.005	1.049	0.717	0.680
	(0.46)	(0.46)	(0.54)	(0.34)	(0.52)	(0.42)	(0.45)	(0.31)	(0.29)
Env. Law$_{-1}$* Fines$_{-1}$	-0.010					-0.013	-0.014	-0.012	-0.012
	(0.53)					(0.71)	(0.90)	(0.74)	(0.73)
Env. Office$_{-1}$* Fines$_{-1}$	-0.102**					-0.099***	-0.101***	-0.087**	-0.086**
	(3.08)					(3.16)	(3.20)	(2.82)	(2.80)
Env. Law$_{-1}$* Conservation Units$_{-1}$		0.007				0.008	0.008	0.008	0.007
		(1.48)				(1.63)	(1.75)	(1.63)	(1.60)
Env. Office$_{-1}$* Conservation Units$_{-1}$		0.031***				0.031***	0.031***	0.029**	0.030**
		(3.29)				(3.22)	(3.19)	(3.07)	(3.08)
Env. Law$_{-1}$* Indigenous Land$_{-1}$		-0.014				-0.013	-0.014	-0.015	-0.015
		(0.56)				(0.48)	(0.54)	(0.59)	(0.60)
Env.Office$_{-1}$* Indigenous Land$_{-1}$		-0.125				-0.109	-0.101	-0.091	-0.089
		(1.57)				(1.49)	(1.33)	(1.19)	(1.17)
Env. Law$_{-1}$* Soy Price$_{-1}$			-0.714***				-0.722***	-0.737***	-0.728***
			(3.51)				(3.33)	(3.41)	(3.12)
Env. Office$_{-1}$* Soy Price$_{-1}$			-0.304				-0.330	-0.324	-0.432*
			(1.07)				(1.12)	(1.17)	(1.80)
Env. Law$_{-1}$* Timber Price$_{-1}$				-0.477*				-0.433	-0.443
				(2.18)				(1.61)	(1.65)
Env. Office$_{-1}$* Timber Price$_{-1}$				-2.136***				-1.938***	-1.889***
				(6.90)				(5.72)	(5.38)
Env. Law$_{-1}$* Beef Price$_{-1}$					-0.061**				-0.013
					(2.71)				(0.61)
Env. Office$_{-1}$* Beef Price$_{-1}$					0.051				0.071**
					(1.28)				(2.28)
Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	6,178	6,178	6,178	6,178	6,178	6,178	6,178	6,178	6,178

*, **, *** denote significance at 10%, 5% and 1% levels, respectively. Robust standard error values are indicated in parentheses under the coefficients.

Finally, we explore the assumption that environmental offices operating with other offices and strategic themes negatively influence the environmental agenda. Hence, we substitute environmental offices combined with other key themes to look only at environmental offices that deal with the environmental. Column (9) of Table 1.10, shows that at the 10% significance level, policy variables conditioned on the IEF do not affect the path of deforestation. This suggests that environmental offices that have a narrow environmental agendas do not facilitate the execution of environmental policies.[1]

The results also show that combining environmental offices with other strategic agendas increases deforestation. For commodities prices, we observe that, ceteris paribus, forestry products respond negatively to environmental offices with an environmental agenda. A 10% increase in the price of forestry products results in a decrease in deforestation by 0.8 km^2 at mean values for the existence of institutional environmental framework. For soy, we see that with only environmental offices the impact on deforestation is insignificant. We also observe that beef market has no longer impact when we consider only environmental offices with a narrow environmental agenda. These results reinforce our finding that environmental offices working with different kind of strategic agendas might fail to prevent deforestation.

[1]In Assunçao et al. (2017) they correlate the annual number of environmental fines applied at the municipality level to a more stringent monitoring and law enforcement regarding institutional apparatus such as environmental offices. In this sense, we re-estimate Table 1.3 excluding the variable fines in order to see if the model changes significantly given the assumption stated by the authors. As can be seen from Table A.1.6, the results indicate that fines has a positive impact on curb deforestation and the model without the variable seems to have no improvement of the solely existence of environmental offices.

Table 1.10 Environmental policies and municipal institutional framework - Robustness Check for Environmental Office

Sample	Dependent: Δ *ln* Deforestation								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Fines$_{-1}$	-0.235	-0.192	-0.191	-0.162	-0.180	-0.203	-0.222	-0.276	-0.272
	(0.80)	(0.91)	(0.91)	(0.76)	(0.84)	(0.72)	(0.80)	(1.03)	(1.02)
Protected Areas$_{-1}$	0.027***	0.017*	0.026***	0.027***	0.027***	0.017*	0.016*	0.016*	0.016*
	(3.79)	(1.84)	(3.77)	(3.68)	(3.81)	(1.85)	(1.73)	(1.74)	(1.75)
Env. Law$_{-1}$	0.021	-0.000	0.449***	0.077*	0.057*	0.002	0.438***	0.489***	0.492***
	(0.88)	(0.02)	(3.46)	(2.06)	(1.81)	(0.08)	(3.61)	(3.50)	(3.55)
Env. Office$_{-1}$	-0.003	-0.012	0.397	0.104***	0.082**	-0.016	0.376	0.453	0.473
	(0.15)	(0.62)	(1.35)	(5.31)	(2.53)	(0.87)	(1.28)	(1.57)	(1.75)
Soy Price$_{-1}$	0.552	0.562	0.987**	0.553	0.550	0.564	1.007**	1.006**	0.998**
	(1.19)	(1.21)	(2.54)	(1.20)	(1.18)	(1.21)	(2.57)	(2.53)	(2.49)
Timber Price$_{-1}$	-1.295*	-1.308*	-1.290*	-0.790	-1.287*	-1.307*	-1.300*	-0.819	-0.813
	(2.09)	(2.14)	(2.05)	(1.12)	(2.09)	(2.14)	(2.10)	(1.19)	(1.19)
Beef Price$_{-1}$	1.107	1.050	1.252	1.087	1.186	1.051	1.202	1.188	1.204
	(0.47)	(0.45)	(0.54)	(0.46)	(0.51)	(0.45)	(0.53)	(0.52)	(0.53)
Env. Law$_{-1}$* Fines$_{-1}$	-0.048					-0.102	-0.100	-0.060	-0.062
	(0.27)					(0.70)	(0.71)	(0.43)	(0.44)
Env. Office$_{-1}$* Fines$_{-1}$	0.170					0.148	0.148	0.218	0.212
	(1.36)					(1.15)	(1.09)	(1.65)	(1.68)
Env. Law$_{-1}$* PAs$_{-1}$		0.010				0.010*	0.010*	0.009*	0.009
		(1.70)				(1.86)	(1.95)	(1.80)	(1.79)
Env.Office$_{-1}$*PAs$_{-1}$		0.007				0.007	0.008	0.007	0.007
		(1.33)				(1.22)	(1.39)	(1.31)	(1.26)
Env. Law$_{-1}$* Soy Price$_{-1}$			-0.670***				-0.677***	-0.678***	-0.689***
			(3.63)				(3.79)	(3.89)	(3.63)
Env. Office$_{-1}$* Soy Price$_{-1}$			-0.604				-0.604	-0.571	-0.516
			(1.30)				(1.29)	(1.23)	(1.02)
Env. Law$_{-1}$* Timber Price$_{-1}$				-0.510**				-0.444	-0.440
				(2.24)				(1.76)	(1.77)
Env. Office$_{-1}$* Timber Price$_{-1}$				-0.954***				-0.952***	-0.975***
				(4.06)				(4.22)	(4.47)
Env. Law$_{-1}$* Beef Price$_{-1}$					-0.047*				0.003
					(2.05)				(0.17)
Env. Office$_{-1}$* Beef Price$_{-1}$					-0.098**				-0.065
					(2.98)				(1.00)
Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	6,178	6,178	6,178	6,178	6,178	6,178	6,178	6,178	6,178

*, **, *** denote significance at 10%, 5% and 1% levels, respectively. Robust standard error values are indicated in parentheses under the coefficients. PAs stand for Protected Areas.

1.5 Counter-factual Simulations

Using the estimates from column 9 of table 1.3 we conduct counter-factual simulations in the spirit of Assunçao et al. (2015) to quantify the contribution of the existence of the IEF in the Legal Amazon in terms of avoided deforestation associating with CO_2 emissions over our sample period. More specifically, the predicted impact on the natural logarithm of deforestation for each sample municipality is defined as:

$$\begin{aligned} \ln\widehat{D}_{i,t} = \widehat{\mu}_i + \mathrm{X}_{i,t}\widehat{\beta}_1 + \mathrm{X}_{i,t-1}\widehat{\beta}_2 + (\text{IEF * Env. Policies})_{i,t-1}\widehat{\beta}_3 \\ +(\text{IEF * Prices})_{i,t-1}\widehat{\beta}_4 + \widehat{\lambda}_t \end{aligned} \tag{1.4}$$

in which $\ln\widehat{D}_{i,t}$ is the predicted log of deforestation, calculated using the estimated coefficients illustrated by the hatted parameters. Given the estimated parameters, we are able to recalculate each $\ln\widehat{D}_{i,t}$ under the alternative condition Environmental Law = 0 and Environmental Office = 0. This calculation gives the log of annual deforestation in a scenario of no environmental institutional framework. We follow the Assunçao et al. (2015) methodology by accumulating $\ln\widehat{D}_{i,t}$ across all 562 sample municipalities and all sample years to calculate total predicted deforestation under the hypothetical scenario. Table 1.11 shows the total observed deforestation trend from 2004 to 2015. It also includes the counterfactual analysis for the hypothetical scenario. We can see that total observed deforestation from 2004 to 2015 was approximately 124,841 km^2. Under the hypothetical scenario of no institutional instrument the level of deforestation would have been 341 thousand km^2 or 2.75 times higher. The results thus indicate that the presence of institutional environmental framework avoided over 63% of forest clearing that would have occurred in the absence of such a frawework.

Table 1.11 Counterfactual Simulations - Observed and predicted deforestation

Year	Deforestation in sample municipalities	
	Observed	Predicted
2004	27,423	85,708
2005	23,490	75,555
2006	10,489	23,156
2007	11,332	23,755
2008	13,192	28,229
2009	6,547	16,335
2010	6,324	25,859
2011	5,627	15,571
2012	4,495	9,583
2013	5,435	12,262
2014	5,157	13,789
2015	6,182	11,365
Total deforestation, 2004-2015	124,841	341,167
Avoided deforestation, 2004-2015	-	216,325
Avoided deforestation, 2004-2015 (as % of predicted deforestation)	-	63%

Sample includes 562 municipalities located in the Legal Amazon region. Regressions include year effects. Analysis is based on a municipality level panel data set covering from 2004 to 2015. The counterfactual simulation is based on model 3.1 and uses the specification presented in column (9) of table 1.3.

To provide a basic cost-benefit analysis, we estimate that the preserved forest area is equivalent to 12 billion tonnes of stored CO_2 with a value of US$ 62b.[1] Considering the Environmental Police's (IBAMA) budget and PRODES project Institute (INPE) through our sample period was around 8 billion USD (at 2010 prices), we account that any price in the carbon market set above 0.69 USD/tCO_2 would compensate the cost of environmental monitoring and law enforcement (similar findings were found in Assunçao et al. (2017)). Taking into account that the price used in current empirical studies is 5 USD/tCO_2, the monetary gains would have the potential to surpass the costs by 6.2 times (51 billion USD net monetary gains). Although this is a large amount, we emphasise that the hypothetical scenario is set against there being no institutional framework in the Legal Amazon, which would be characterised as an open access land with no legal and institutional articulation. We understand that much of this avoided deforestation also considers the indirect effects

[1] Following Assunçao et al. (2015), the estimations are based on conversion factors of 10,000 tC/km^2 which figures 36,700 tCO_2/km^2 and 5 USD/tCO_2/km^2 established by the Ministry of Environment (MMA, 2018b). Fearnside (2016) considers this conversion factor a gentleman's agreement' rather than a realistic measure. He argues that the value was deliberately chosen to be conservative. We took this approach in order to make a comparative analysis according to the Ministry of Environment (MMA).

occurring through changes in environmental policies and market condition. Nevertheless, this result suggests the importance of having a well functioning institutional environmental framework.

1.6 Conclusion

In this study we quantify the determinants of deforestation in the Brazilian Legal Amazon introducing an institutional approach into the discussion. More specifically, we analyse whether environmental policies and market prices conditioned on IEF were responsible for the decrease in deforestation at municipal level in the Legal Amazon since 2004. In terms of data, we use municipality level with yearly frequency for environmental policies and market condition. A spatial analysis was conducted in order to account for spatial dependence. The results show that the existence of an IEF conditioned on policies and prices reduced deforestation and has negative spillovers to neighbouring municipalities. Counter-factual simulations indicate that the existence of institutional environmental framework avoided 63% of total forest clearing that would have occurred had the institutional framework not been instrumented and this preserved forest corresponds to 12 billion tonnes of stored CO_2 with the value of US$ 62b.

Our main findings show that environmental law as an instrument of institutional environmental framework can have an impact on the path of deforestation, but only through market conditions. In other words, we find evidence that environmental legislation affected soy production through the indirect effect of the moratorium mechanism. The existence of an environmental office has an effect on reduce forest clearing is used in conjunction with the application of fines. Environmental offices are also responsible inducing the office to adapt agendas that may be in conflict with environmental actions since in the Legal Amazon environmental offices can have a number of strategic themes. We observe this pattern with regard to the prices of timber and beef which gives an incentive for forest management in the private sector and thus reducing the effect on deforestation and prioritising the beef market role in deforestation. However, the Brazilian institutional and regulatory capacity of government urgently needs to strengthen its environmental actions and to improve policies to allow a more homogeneous approach among local authorities considering that the Legal Amazon as a heterogeneous region.

Our results suggest important policy implications regarding the institutional mechanisms that could be used to curb tropical deforestation in the Legal Amazon. In an unfavourable institutional environment, the protection of the environment is largely ignored. It is imperative that the Brazilian institutional and regulatory capacity of government is strengthened and improvements in policies to allow a more homogeneous approach among local authorities are made. We concluded that the Brazilian institutional environmental framework requires tightening and narrowing of environmental policies at local level.

Appendix A.1

Table A.1.1 Summary Statistics - Year 2004

Variable	Mean	St. Dev.	Min.	Max.
Deforestation	53.342	107.654	0.1	1082.5
Fines	0.082	0.307	0	4.290
Protected Areas	1.501	3.428	0	36.930
Rural Credits	0.264	0.830	0	9.649
Environmental Law	0.339	0.473	0	1
Environmental Office	0.479	0.500	0	1
Housing Projects	0.728	0.444	0	1
Settlements	2.570	4.652	0	65
Settlements Density	0.016	0.048	0	1.016
GDP	0.014	0.082	0.001	1.659
Beef Price	2.409	0.035	2.286	2.419
Soy Price	0.703	0.138	0.272	1.034
Wood Price	0.074	0.049	0.002	0.401
Roads	0.018	0.029	0.001	0.081
Clouds	0.284	1.459	0	18.557
No obs	0.260	2.333	0	49.530
Mayor Political Party (pro-farmer)	0.809	0.393	0	1
Mayor Gender (Male)	0.923	0.265	0	1
Mayor Age (% above 50)	0.357	0.479	0	1
Mayor Education	0.347	0.476	0	1
Corruption	0.007	0.007	0	0.056
Re-election	0	0	0	0

Note: Statistics refer to N=6178 observations for 562 municipalities.

Table A.1.2 Summary Statistics - Year 2015

Variable	Mean	St. Dev.	Min.	Max.
Deforestation	11.449	31.028	0.1	308.6
Fines	0.342	1.642	0	24.227
Protected Areas	23.134	57.679	0	448.777
Rural Credits	0.919	0.364	0	3.281
Environmental Law	0.696	0.460	0	1
Environmental Office	0.942	0.232	0	1
Housing Projects	0.822	0.382	0	1
Settlements	4.239	6.460	0	76
Settlements Density	0.018	0.027	0	0.277
GDP	0.110	0.530	0.002	9.850
Beef Price	0.877	0.006	0.866	0.889
Soy Price	0.722	0.060	0.540	0.899
Wood Price	0.105	0.087	0.005	0.649
Roads	0.206	0.028	0.001	0.083
Clouds	9.019	35.686	0	488.1
No obs	0	0	0	0
Mayor Political Party (pro-farmer)	0.944	0.229	0	1
Mayor Gender (Male)	0.866	0.340	0	1
Mayor Age (% above 50)	0.368	0.482	0	1
Mayor Education	0.505	0.500	0	1
Corruption	0.042	0.539	0	12.534
Re-election	0.194	0.396	0	1

Note: Statistics refer to N=6178 observations for 562 municipalities.

Table A.1.3 Data Description - Scope, Sources and Unit of Measurement

Variable	Level	Source	Unit of Measurement
Deforestation	Municipality	INPE	ln(square km)
Fines	Municipality	IBAMA	in $10,000,000 BRL
Conservation Units	Municipality	MMA/IBAMA	in 10,000 sq km
Indigenous Land	Municipality	FUNAI	in 10,000 sq km
Rural Credits Grants	Municipality	BASA, BNB, BACEN	in $10,000,000,000 BRL
Environment Law	Municipality	IBGE	Indicator of environmental law
Environmental Agency	Municipality	IBGE	Indicator of environmental office
Housing Projects	Municipality	IBGE	Indicator of housing project
Settlements	Municipality	INCRA	Number of settlements
Settlements Density	Municipality	INCRA	Ratio of families per settlement
GDP	Municipality	IBGE	in $100,000,000 BRL
Beef Price	State	CEPEA	per 15kg, $100 BRL
Soy Price	Municipality	IBGE	per 60kg, BRL
Wood Price	Municipality	IBGE	per cubic m, BRL
Roads	State	DNIT	Number of ‘new‘ paved kms averaged by state size
Mayor Political Party	Municipality	IBGE	Dummy if "pro-farmer" party
Mayor Education	Municipality	IBGE	Dummy if mayor has tertiary education
Mayor Age	Municipality	IBGE	Dummy if age is higher than the average of 50
Clouds	Municipality	INPE	in 52 weeks
No Obs	Municipality	INPE	Area not observed per year
Corruption	Municipality	IBGE	Number of Commissioned Workers per City Hall averaged by population
Re-election	Municipality	TSE	Dummy if mayor is running for re-election in election years

Note: All prices and economic values are expressed in constant 2010 Brazil Reais (BRL).
Exchange Rate 1.00 BRL to 0.602 US Dollar.

Table A.1.4 Effects of Environmental Policies on Deforestation: Baseline Model - Robustness Check for Two-way Fixed Effects

Sample	Dependent: Δ *ln* Deforestation			
	(1)	(2)	(3)	(4)
Fines_{-1}	0.600***	0.587***	0.055	0.069
	(-10.01)	(-10.39)	(-1.39)	(-1.87)
$\text{Protected Areas}_{-1}$	0.01	0.01	-0.022*	-0.003
	(-1.03)	(-0.54)	(-1.9)	(-0.3)
Env. Law_{-1}	0.484**	0.414*	0.589***	0.538***
	(-2.09)	(-1.89)	(-3.9)	(-3.84)
Env. Office_{-1}	1.092***	1.264***	-0.274*	0.287
	(-4.32)	(-5.19)	(-1.7)	(-1.88)
Soy Price_{-1}	1.548***	1.448***	0.758***	1.188***
	(-4.49)	(-3.99)	(-3.45)	(-5.06)
Timber Price_{-1}	3.824***	4.985***	-1.690***	0.639
	(-5.37)	(-7.40)	(-3.07)	(-1.23)
Beef Price_{-1}	0.396***	6.031***	0.064	0.685
	(-5.54)	(-5.46)	(-1.42)	(-0.93)
Env. Law_{-1}* Fines_{-1}	-0.153***	-0.158***	0.003	-0.012
	(-4.25)	(-4.69)	(-0.11)	(-0.56)
Env. Office_{-1}* Fines_{-1}	-0.158***	-0.134**	-0.101***	-0.088**
	(-2.77)	(-2.48)	(-2.72)	(-2.57)
Env. Law_{-1}* PAs_{-1}	-0.01	-0.01	0.005	0.007
	(-1.52)	(-1.34)	(-0.87)	(-1.22)
Env. Office_{-1}* PAs_{-1}	0.02	0.02	0.031***	0.028**
	(-1.05)	(-1.55)	(-3.42)	(-3.33)
Env. Law_{-1}* Soy Price_{-1}	-1.117***	-0.801**	-0.933***	-0.729***
	(-3.09)	(-2.33)	(-4.04)	(-3.4)
Env. Office_{-1}* Soy Price_{-1}	-2.350***	-1.480***	-0.688***	-0.447*
	(-5.72)	(-3.70)	(-2.67)	(-1.82)
Env. Law_{-1}* Timber Price_{-1}	1.153**	1.280**	-0.869**	-0.45
	(-2.14)	(-2.52)	(-2.23)	(-1.25)
Env. Office_{-1}* Timber Price_{-1}	-1.526**	-1.845**	-0.817	-1.891***
	(-2.00)	(-2.56)	(-1.55)	(-3.85)
Env. Law_{-1}* Beef Price_{-1}	0.05	-0.04	0.039	-0.011
	(-0.52)	(-0.39)	(-0.65)	(-0.19)
Env. Office_{-1}* Beef Price_{-1}	0.231**	0.04	0.243***	0.070**
	(-2.32)	(-0.46)	(-4)	(-1.23)
Set of Controls	Yes	Yes	Yes	Yes
Year Fixed Effects	No	Yes	No	Yes
Municipality Fixed Effects	No	No	Yes	Yes
Number of Observations	6,178	6,178	6,178	6,178

*, **, *** denote significance at 10%, 5% and 1% levels, respectively. Robust standard error values are indicated in parentheses under the coefficients using the conventionally derived variance estimator for generalized least-squares regression.
PAs stand for Protected Areas.

Table A.1.5 Effects of Environmental Policies on Deforestation: Baseline Model - Robustness Check for Lagged Deforestation

Sample	Dependent: Δ *ln* Deforestation	
	(1)	(2)
Fines_{-1}	0.069*	0.068*
	(-1.870)	(-1.840)
$\text{Protected Areas}_{-1}$	-0.003	-0.004
	(-0.300)	(-0.330)
Env. Law_{-1}	0.538***	0.532***
	(-3.840)	(-3.800)
Env. Office_{-1}	0.287*	0.278*
	(-1.880)	(-1.810)
Soy Price_{-1}	1.188***	1.168***
	(-5.060)	(-4.970)
Timber Price_{-1}	0.639	0.648
	(-1.230)	(-1.250)
Beef Price_{-1}	0.685	0.642
	(-0.930)	(-0.870)
Env. Law_{-1}* Fines_{-1}	-0.012	-0.012
	(-0.560)	(-0.560)
Env. Office_{-1}* Fines_{-1}	-0.088**	-0.087**
	(-2.570)	(-2.550)
Env. Law_{-1}* PAs_{-1}	0.007	0.007
	(-1.220)	(-1.200)
Env. Office_{-1}* PAs_{-1}	0.028***	0.028***
	(-3.330)	(-3.310)
Env. Law_{-1}* Soy Price_{-1}	-0.729***	-0.720***
	(-3.400)	(-3.350)
Env. Office_{-1}* Soy Price_{-1}	-0.447*	-0.433*
	(-1.820)	(-1.770)
Env. Law_{-1}* Timber Price_{-1}	-0.450	-0.453
	(-1.250)	(-1.260)
Env. Office_{-1}* Timber Price_{-1}	-1.891***	-1.886***
	(-3.850)	(-3.840)
Env. Law_{-1}* Beef Price_{-1}	-0.011	-0.010
	(-0.190)	(-0.180)
Env. Office_{-1}* Beef Price_{-1}	0.070	0.071
	(-1.230)	(-1.250)
$ln\ \text{Def}_{-1}$		0.011
		(-0.970)
Set of Controls	Yes	Yes
Year Fixed Effects	Yes	Yes
Municipality Fixed Effects	Yes	Yes
Number of Observations	6,178	6,177

*, **, *** denote significance at 10%, 5% and 1% levels, respectively. Robust standard error values are indicated in parentheses under the coefficients using the conventionally derived variance estimator for generalized least-squares regression.
PAs stand for Protected Areas.

Table A.1.6 Environmental policies and municipal institutional framework - Robustness Check for Fines

Sample	Dependent: Δ *ln* Deforestation	
	(1)	(2)
$Fines_{-1}$	0.069	
	(-1.87)	
Protected $Areas_{-1}$	-0.003	-0.004
	(-0.3)	(-0.37)
Env. Law_{-1}	0.538***	0.550***
	(-3.84)	(-3.94)
Env. $Office_{-1}$	0.287	0.256
	(-1.88)	(-1.68)
Soy $Price_{-1}$	1.188***	1.159***
	(-5.06)	(-4.94)
Timber $Price_{-1}$	0.639	0.743
	(-1.23)	(-1.43)
Beef $Price_{-1}$	0.685	0.77
	(-0.93)	(-1.04)
Env. Law_{-1}* $Fines_{-1}$	-0.012	
	(-0.56)	
Env. $Office_{-1}$* $Fines_{-1}$	-0.088**	
	(-2.57)	
Env. Law_{-1}* PAs_{-1}	0.007	0.006
	(-1.22)	(-1)
Env. $Office_{-1}$* PAs_{-1}	0.028***	0.028***
	(-3.33)	(-3.38)
Env. Law_{-1}* Soy $Price_{-1}$	-0.729***	-0.744***
	(-3.4)	(-3.47)
Env. $Office_{-1}$* Soy $Price_{-1}$	-0.447*	-0.404*
	(-1.82)	(-1.65)
Env. Law_{-1}* Timber $Price_{-1}$	-0.45	-0.478
	(-1.25)	(-1.34)
Env. $Office_{-1}$* Timber $Price_{-1}$	-1.891***	-2.018***
	(-3.85)	(-4.13)
Env. Law_{-1}* Beef $Price_{-1}$	-0.011	-0.011
	(-0.19)	(-0.2)
Env. $Office_{-1}$* Beef $Price_{-1}$	0.07	0.073
	(-1.23)	(-1.29)
Set of Controls	Yes	Yes
Year Fixed Effects	Yes	Yes
Municipality Fixed Effects	Yes	Yes
Number of Observations	6,178	6,178

*, **, *** denote significance at 10%, 5% and 1% levels, respectively. Robust standard error values are indicated in parentheses under the coefficients using the conventionally derived variance estimator for generalized least-squares regression.
PAs stand for Protected Areas.

Table A.1.7 LM Tests

	W1	
Elhorst		
LM test no spatial lag, probability	0.0000	1.000
robust LM test no spatial lag, probability	314.0441	0.000
LM test no spatial error, probability	288.3646	0.000
robust LM test no spatial error, probability	602.4087	0.000
Lacombe		
LM lag test for omitted spatial lag in panel data	111.2545	0.000
LM error test for spatial errors in panel data	103.5038	0.000
Robust LM lag test for omitted spatial lag in panel data	10.1198	0.001
Robust LM error test for spatial errors in panel data	2.3692	0.123
	W2	
Elhorst		
LM test no spatial lag, probability	0.0000	1.000
robust LM test no spatial lag, probability	314.3242	0.000
LM test no spatial error, probability	288.6314	0.000
robust LM test no spatial error, probability	602.9555	0.000
Lacombe		
LM lag test for omitted spatial lag in panel data	111.2895	0.000
LM error test for spatial errors in panel data	104.1005	0.000
Robust LM lag test for omitted spatial lag in panel data	9.7553	0.001
Robust LM error test for spatial errors in panel data	2.5663	0.109
	W3	
Elhorst		
LM test no spatial lag, probability	0.0000	1.000
robust LM test no spatial lag, probability	921.5621	0.000
LM test no spatial error, probability	1370.3817	0.000
robust LM test no spatial error, probability	2291.9439	0.000
Lacombe		
LM lag test for omitted spatial lag in panel data	326.6028	0.000
LM error test for spatial errors in panel data	334.1047	0.000
Robust LM lag test for omitted spatial lag in panel data	21.3014	0.000
Robust LM error test for spatial errors in panel data	28.8033	0.000

Note that the results satisfy the condition tha LM spatial Lag + Robust LM spatial error = LM spatial error + robust LM spatial lag (Anselin, 1988).

Chapter 2

Trends in Deforestation and Environmental Policy in Maranhão, Brazil

Abstract

This study investigates deforestation trends in the Brazilian *Cerrado* region in Maranhão, Brazil, which provides a unique natural experiment in that there were spatially hetereogenous environmental policies to combat deforestation in place. The analysis applies the non-linear estimation approach of Generalized Additive Models (GAMs) on satellite data derived measures of deforestation, where validation is conducted using features of neural networks. The GAMs confirmed that deforestation is related to climatic factors in that increased during high levels of precipitation and low levels of solar incidence. More importantly, the results revealed that there are substantially differences in trends between seasons across regions according to their policy distinction. This was further substantiated by showing that deforestation happened during both seasons for settlements which were not target of the environmental policy, but only during the rainy season for the protected areas, likely due to the lower rate of remotely sensed detection during cloud cover.

Keywords: Deforestation trends, Generalized Additive Models, Remote Sensing Analysis.

2.1 Introduction

Tropical deforestation is an event that gained momentum in the second half of the 20^{th} century. As a matter of fact, the considerable deforestation observed globally during 1990-

2010 was almost entirely confined to the tropical regions (Culas, 2014). In this regard, the Brazilian Cerrado is perhaps the biome which has arguably been most affected by human occupation over the last three decades, mainly due to the increasing pressures for opening up of new areas for the production of meat, grains and ethanol, mostly at the sacrifice of forested areas (Bayma and Sano, 2015; MMA, 2018a). Importantly, the Brazilian Cerrado has also been subject to spatially and temporally heterogeneous environmental policies discouraging such deforestation. Understanding what role such interventionist policies may have played in observed trends in deforestation could potentially provide a platform with which to assess future possible scenarios of deforestation in the Cerrado biome and the Amazon forest in general (Nepstad et al., 2013). In this paper we use remote sensing data and non-linear models to model the trends of deforestation and its underlying drivers in the *Cerrado* region in the Brazilian state of Maranhão using the non-linear modelling approach of Generalized Additive Models (GAMs).

Arguably, the state of Maranhão provides a particularly interesting context within which to study trends in deforestation and the possible role of environmental policy. More specifically, Maranhão is divided by an artificial line that separates it in two parts: the Legal Amazon Maranhão and the Cerrado Maranhão. This division, occurring approximately 44° west of the meridian, was established in 1953 due to the necessity to plan economic development in the region. This scenario provides a unique natural experiment of deforestation in the Legal Amazon Maranhão (LM) and Cerrado Maranhão (MA) since the former has been subject to fundamentally different environmental policies compared to the latter.[1] More specifically, the tropical forest in the Legal Amazon Maranhão is under a surveillance environmental policy, called DETER, which detects deforestation or fire incidence in the region using satellite data and informs the occurrences to the environmental police (IBAMA in Portuguese) so that they can fine or arrest the responsible

[1]The Legal Amazon is an area that corresponds to 59% of the Brazilian territory and encompasses all eight states (Acre, Amapá, Amazonas, Mato Grosso, Pará, Rondônia, Roraima and Tocantins) and part of the State of Maranhão (west of the meridian Of 44ºW), totalling more than 5 million km^2 (IPEA, 2008).

persons (IBAMA, 2017). In contrast, DETER is not applicable for the other biomes in the Maranhão state.[1] We use this spatial division to determine how deforestation trends may have been different being the Legal Amazon Maranhão and Cerrado Maranhão.

Our use of non-linear modelling for the task at hand derives from the recognition by recent previous research that most ecological and climatic data represent complex relationships and thus that non-linear models, such as GAMs, may be particularly suited to capture confounding effects in trends; see (Antunez et al., 2017; Auderset Joye and Rey-Boissezon, 2015; Bell et al., 2015; Bio et al., 1998; de Souza et al., 2017; Halperin et al., 2016; Liu et al., 2018; Lusk et al., 2016; Moreno-Fernández et al., 2018; Pourtaghi et al., 2016). However, a review of the literature shows that such models have only been used sparsely to study deforestation. For example, Chaves et al. (2008) modelled how deforestation affected incidence rates of a disease in Cuba. Also, Green et al. (2013) used a binomial GAM model to account for forest and habitat losses in protected areas on the Eastern Arc Mountains of Tanzania. More recently, Bebber and Butt (2017) studied the impact of protected areas on global carbon emissions in America, Africa and Asia. For Brazil, Mendes and Junior (2012) observed the relationship between deforestation, corruption, and economic growth in the region of Legal Amazon, in Brazil. Here we apply a GAM with a negative binomial distribution and logarithmic link function. To capture deforestation we construct monthly time series from remote sensing sources (MODIS), given that high temporal resolution satellite products are particularly suitable to obtain detailed knowledge about the seasonal cycles of vegetation in biomes with strong seasonal

[1]The DETER is a rapid survey of alerts of changes in forest cover in the Legal Amazon made by National Institute os Spatial Research (INPE in Portugues) since May 2004, with data from Terra's MODIS sensor, with a spatial resolution of 250 m. DETER was developed as an alert system to support the environmental police to curb illegal deforestation. With this system, it is possible to detect only changes in the forest cover with an area larger than 25 ha. Due to cloud cover not all changes are identified by DETER. The lower resolution of the sensors used by DETER is compensated by the daily observation capacity, which makes the system an ideal tool to quickly inform the inspection bodies about new changes DETER operates daily and delivers deforestation alert maps to environmental policy in a five days after the date of the MODIS image (BRASIL, 2018).

contrast, such as the Cerrado biome and Ecotone forest (Bayma and Sano, 2015). Our climatic covariates are derived from data from meteorological stations.

The results from our GAMs revealed that for the Legal Maranhao region most of the deforestation happened during the rainy season, while in the unprotected Cerrado Maranhão deforestation occurred in the dry season as well. The fact that precipitation and solar incidence also played an important role in deforestation in the rainy season in the Legal Maranhao region suggests that cloud cover may have acted as an impediment to infringement detection via satellites, as is done in the DETER program. We further substantiated this claim by showing that for settlements that were not the target of environmental policy deforestation mainly took place during both seasons as well.

2.2 Study Context

2.2.1 Study Location

Maranhão, or *'flowing river'* in the indigenous language (Girardi, 2015), is one of the ten largest States of Brazil with more than 330 thousand square kilometers and located in the northeast part of Brazil. However, while it is considered one of the richest regions in biodiversity in the country (Batistella et al., 2014), it has historically ranked among the states with the worst social and economic indicators (Celentano et al., 2017). The State's political boundaries encompass natural hydrological barriers, where within these borders 64.1% of the area of the territory is in the Cerrado/Savana biome, 34.8% in the Amazon biome, and only 1.1% in the *Caatinga* biome (Stella, 2011). The Cerrado represents the largest ecosystem in Maranhão, and is located from the Northeast to the southern region of the State, covering about 60% of its surface, occurring in approximately 55 of the total

217 municipalities. Of these, 23 are almost exclusively covered by this type of vegetation (Batistella et al., 2013b).

In terms of vegetation, in the centre of Maranhão there is a contact area between the Amazonian and Cerrado biomes of around 21.228km^2, where it is possible to observe a mosaic of savanna vegetation physiognomies concomitant with the ombrophilous forest formations (open and dense forest). This area is known as an Ecological Tension Zone (ETZ) of Maranhão and is home to a transitional vegetation which usually occurs with intermediate characteristics of the two formations, with species common and distinct to both, providing great biodiversity (Rossatto et al., 2013). In Cerrado Maranhão it is also possible to distinguish between ecotonic areas and the presence of secondary vegetation in the central region of the state. Ecotone is defined as the transition area between two or more distinct habitats or ecosystems, which may have characteristics of both or their own. Secondary Vegetation includes the various stages of natural succession in areas where there was human intervention for land use, whether for mining, agricultural or livestock purposes, or discharging the primary vegetation (Santos-Filho et al., 2013).

There is a strong debate over the definition of the ecotone forest and secondary vegetation in the state of Maranhão. Considering the Cerrado/savanna territory, almost 10% represents the transition and secondary forest. Reis and Conceição (2010) state that the Cocais Forest (*"Mata dos Cocais"* in Portugues) is considered a characteristic landscape of the State, although it develops in the transition between several biomes. The authors pointed out that the Cocais Forest associates with the open fields, towards the North, with the Cerrado vegetation to the South and East, and gradually joins the forest towards the west.

More recently, Garcia et al. (2017) studied part of the Maranhão region and defined forest as a combination of riparian forest, transitional forest, and Cerrado woodland - the latter defined as having higher tree density and Cerrado, representing physiognomies with a wood layer but lower tree density. Their results showed intense conversion and fragmentation of native vegetation and, consequently, the impoverishment of the quality of native cover, reassuring the possibility of conducting studies considering ecotone forest,such as *Mata dos Cocais*, as a transition forest instead of secondary vegetation.

As noted in the introduction, in addition to the natural barriers, Maranhão is also divided into two parts: the Legal Amazon Maranhão and the Cerrado Maranhão, with 209 municipalities located in the latter and 138 municipalities in the former. This division occurs approximately 44° west of the meridian and was established in 1953 in order to plan the economic development of the region comprised of the tropical forest areas of Maranhão state. We depict this delineation in Figure 2.1

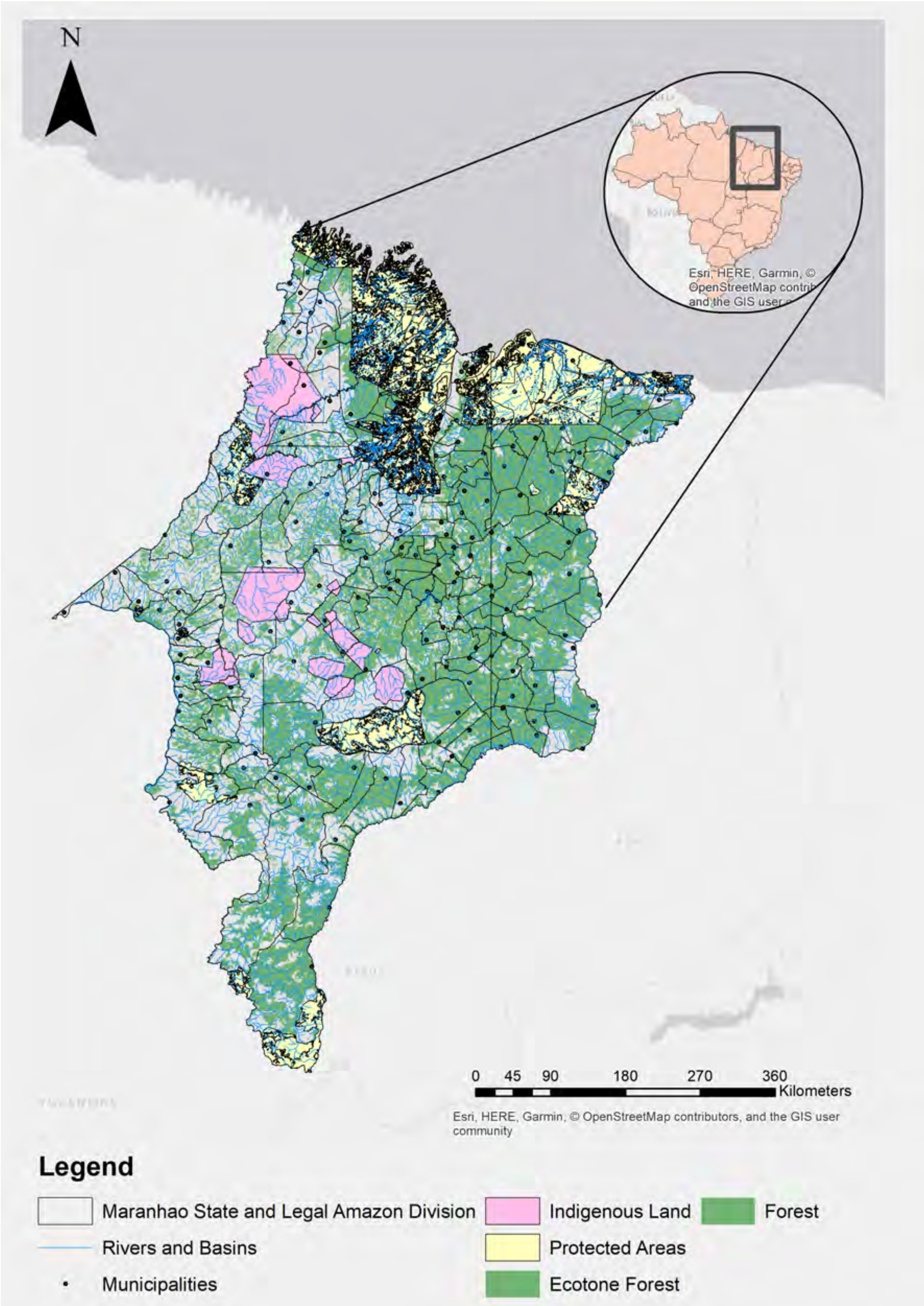

Figure 2.1 Maranhão State and the Legal Amazon delimitation. The map includes municipalities centre, rivers and basins, protected areas and indigenous land. Source: (EMBRAPA, 2018; MMA, 2018b; Núcleo Geoambiental - NUGEO, 2018).

2.2.2 Deforestation in Maranhão

Large-scale deforestation in the Maranhão Amazon forest began in the 1960s, when the military government promoted the occupation of this territory through the construction of highways and provided incentives for large farming projects on public lands and logging centres. In the 1980s, with the implantation of the iron mining project in the in the neighbouring state of Pará (Carajas Project), a railroad linking the mine to the port in Maranhão was built. Moreover, many pig iron facilities were installed in the Maranhão Amazon region, demanding large quantities of charcoal, which increased the pressure on forest resources (Celentano et al., 2017).

While many existing studies have focused on the Amazon tropical forests (Almeyda Zambrano et al., 2010; Arima et al., 2014; Bhattarai and Hammig, 2001; Celentano et al., 2017; Geist and Lambin, 2001, 2002; Kuik, 2013; Lambin and Geist, 2006; Nepstad et al., 2014; Olson et al., 2010; Pfaff et al., 2007; Pfaff, 1999, 1997; Richards, 2015; Richards and VanWey, 2015; Soler et al., 2014; Stickler et al., 2013) due to the fact the ample environmental information was available through specific environmental policies, such as DETER, studying the Cerrado biome and, consequently, transition forests remains precarious. The first obstacle in monitoring the Cerrado biome is due to the high heterogeneity of the forests (open and dense forest, for example) which are substantially influenced by the climatic seasonality (Bayma and Sano, 2015). The second challenge is related to the fact that there is no environmental policy in place to prevent rampant deforestation. Nevertheless, in the context of the Amazon region, it is arguably crucial to understand the dynamic of Cerrado and its potential to influence adjacent forests of Amazonia since it provides a valuable endpoint from which climate and anthropogenic related aspects in the Amazon forest my be better understood (see Figure 2.2).

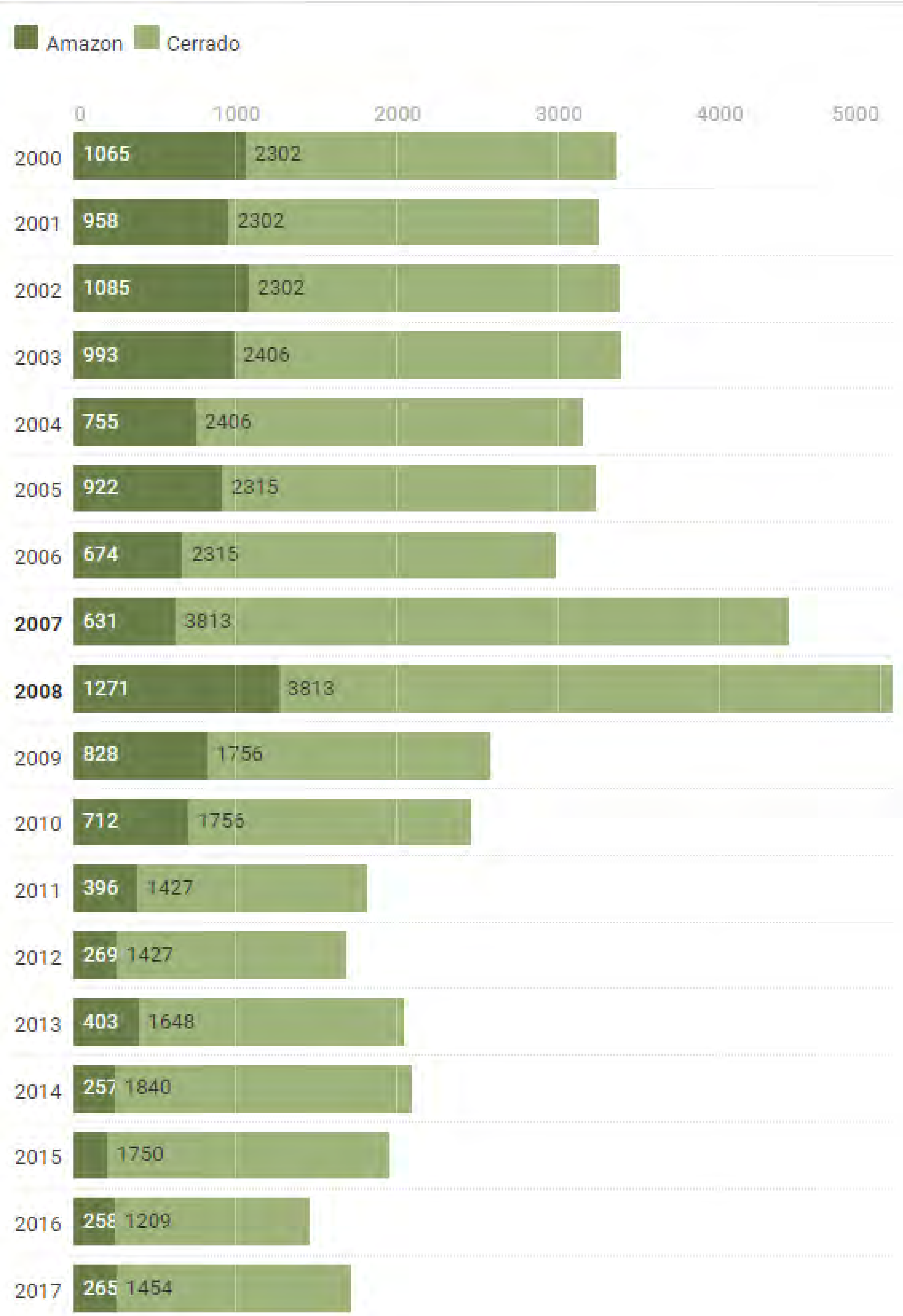

Figure 2.2 Estimated Deforestation in Maranhão 2000-2017 (sq Km). Source: (MMA, 2018b).

In the past many aspects have arguably contributed significantly to deforestation in the Cerrado Maranhão, such as the agricultural occupation of the Cerrado biome and the unconstrained use of mechanisation, propitiated by the predominantly flat relief of the region and the existence of good depth soil conditions and good water supply, making it possible to practice rainfed agriculture (Bayma and Sano, 2015). In contrast, more recent forest losses have been due to settlement projects, illegal logging, pasture, subsistence agriculture and commodities (Celentano et al., 2017; Costa, 2018), and these drivers are deeply connected to the process of the state's development. More precisely, during the 40's, almost 85% of the population was living in rural areas with low rate of population density (3,81). At that time, Maranhão had more than 200.000 km^2 uninhabitated, which included transition forest, Cerrado, and pre-Amazon forest. This "territorial gap" (/textitfundos territoriais in Portuguese) favoured the creation of several settlement projects along with the creation of federal roads and the Carajas railway project (Ferreira, 2008). Moreover, the increasing global demand for commodities affected the economic region significantly. For instance, fiscal incentives increased the production of soy, where the planted area increased from 42,6 km^2 in 1983/84 to 3,940 km^2 in 2004/05. The subsequent ecological tension zone coincides with the Brazilian agricultural frontiers, known as the deforestation arc, and is an area of intense exploitation. [1]

As can be seen from the graph 2.2, deforestation in the Amazon biome of Maranhão has decreased over the years. Great part of this reduction is likely due to protectionist policy enforcement in the region. In this regard, the national environmental policy established in 2004 involved the creation of the Action Plan for the Prevention and Control of Deforestation in the Legal Amazon (PPCDAm in Portuguese). In order to control land use and prevent further deforestation, the PPCDAm also included the satellite-based monitoring programme called DETER, which alerts in real time the environmental police

[1]The expansion of soybean cultivation in Brazil has shifted the agricultural frontier to an area known as MAPITOBA. This area includes the states of MAranhão,PIauí, TOcantins, and BAhia,, and has maintained its expansion across the Cerrado This led to deforestation and degradation, conservation conflicts and conflicts over land, increased burning, and displacement of traditional populations (Mustin et al., 2017)

of illegal logging and deforestation. Importantly, until July 2018 there was no systematic satellite monitoring program for the other parts of Maranhão, such as the transitional forest of Cerrado and *Caatinga*. But, in August of the same year, the National Institute of Spatial Research (INPE in Portuguese) together with several other institutions published an annual dataset covering 18 years of Cerrado biome deforestation and with this data it was possible to show the trends of deforestation in the Cerrado biome of Maranhão. As shown in Figure 2.2, deforestation in great Cerrado, which includes the transitional forest, has been two times higher than in the Amazon region of Maranhão.

2.3 Material and Methods

2.3.1 Remote Sensing

The use of satellite time series along with statistical analysis can be helpful in understanding the characteristics of vegetation dynamics. More precisely, since vegetation has a unique spectral feature (e.g., reflectance) it is possible to identify its unique characteristics from an optical remote sensor on a satellite. In such vegetation mapping, incorporating the spectral radiances in the red and near-infra-red regions into the spectral vegetation indices (VI) gives the possibility to estimate forage quantity and quality of grass prairie, for example (Xie et al., 2008). Earlier studies coarse spatial resolution data from the Advanced Very High Resolution Radiometer (AVHRR) was used to mainly monitor land cover changes at regional and global scales, however, since 2000, the availability of Moderate Resolution Imaging Spectroradiometer (MODIS) data with superior features relative to AVHRR has provided an improved basis for regional and global mapping (Huang and Friedl, 2014).

MODIS

The MODIS sensor is flown on two spacecrafts. The Terra satellite is on an AM overpass, whereas the Aqua platform provides complementary observations in the afternoon. The instrument on-board NASA's Terra satellite is a scanning radiometer system with 36 spectral bands, extending from the visible to the thermal infrared wave-lengths, and has a viewing swath width of 2330km by 10km. The Terra orbital configuration and MODIS viewing geometry produce full global coverage every one to two days, except for the equatorial zone where the repeat frequency is approximately 1.2 days (Setiawan et al., 2014; Zhan et al., 2002a). The high temporal resolution of MODIS is a determining factor in phenological studies and spectral discrimination, and can be used to obtain detailed knowledge about the seasonal cycles of vegetation in biomes with strong seasonal contrast, such as the Cerrado biome and Ecotone forest.

Of the many data products derived from MODIS observations, we use two extensively used here: MCD12Q1 and MOD13Q1. The MODIS Land Cover Type Product (MCD12Q1) provides 13 science data sets (SDSs) that map global land cover at 500m spatial resolution at annual time steps for six different land cover legends from 2001-2016. In contrast, the MCD12Q1 product is created using supervised classification of MODIS reflectance data and includes 5 legacy classification schemes such as the University of Maryland classification (UMD), which recognises 17 classes, covering natural vegetation (11 classes), mosaic lands (2 classes), and non-vegetated lands (4 classes). A complete list of the classes and their definitions is given in Table 2.2 (Setiawan et al., 2014; Sulla-Menashe and Friedl, 2018).

The MODIS Vegetation Indices (VI) (MOD13Q1) product consists of time series comparisons of global vegetation conditions that can be used to monitor the Earth's terrestrial change detection. The two vegetation indices thatr we derive from these are the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI). The NDVI is a normalized transformation of the NIR (Near Infrared) to the red reflectance ratio standardized to range from -1 to 1. Ratana et al. (2005) notes that this index is sufficiently stable to permit meaningful comparisons of seasonal, inter-annual, and long-term variations of vegetation structure, phenology, and biophysical parameters:

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \tag{2.1}$$

where ρ_{red} and ρ_{NIR} are the surface bidirectional reflectance factors for MODIS bands 1 (620-670nm) and 2 (841-876nm).

To optimise the vegetation signal and minimise atmospheric effect and soil background noise, the EVI index has been reported to be more responsive to canopy structural variations including canopy type. The EVI formula is written as:

$$EVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + C_{1\rho_{red}} - C_{2\rho_{blue}} + L}(G) \tag{2.2}$$

where ρ_{red} and ρ_{NIR} and ρ_{blue} are the reflectance in MODIS bands 1,2 and 3 (459-479nm) and, C_1 and C_2 are the atmospheric resistance coefficients. L and G are the canopy background adjustment and the gain factor, respectively. The coefficients adopted for the MODIS EVI algorithm are, L=1, C_1 =6, C_2 =7.5 and G=2.5. The Enhanced Vegetation Index differs from NDVI by attempting to correct for atmospheric and background effects. In addition, EVI is superior in discriminating subtle differences in areas of high vegetation density than NDVI because the latter tends to saturate (Didan et al., 2015; Ratana et al., 2005).

2.3.2 Study area characterization

We compare deforestation trends within the vincinity of both sides of artificial border of the Legal Amazon. To this end we experiment with three bandwidths of 25km, 50km and 100km in both east and west direction from the line giving a total of 6 sampled areas. The buffer zone is characterised by intense presence of Ecotone Forest and covers the east region (MA, hereafter) and west centre region (LM, hereafter) of the Maranhão state. As can be seen in Figure 2.3 the study area comprises more than one third of the State which represents our 100km buffer to east and west of the territory.

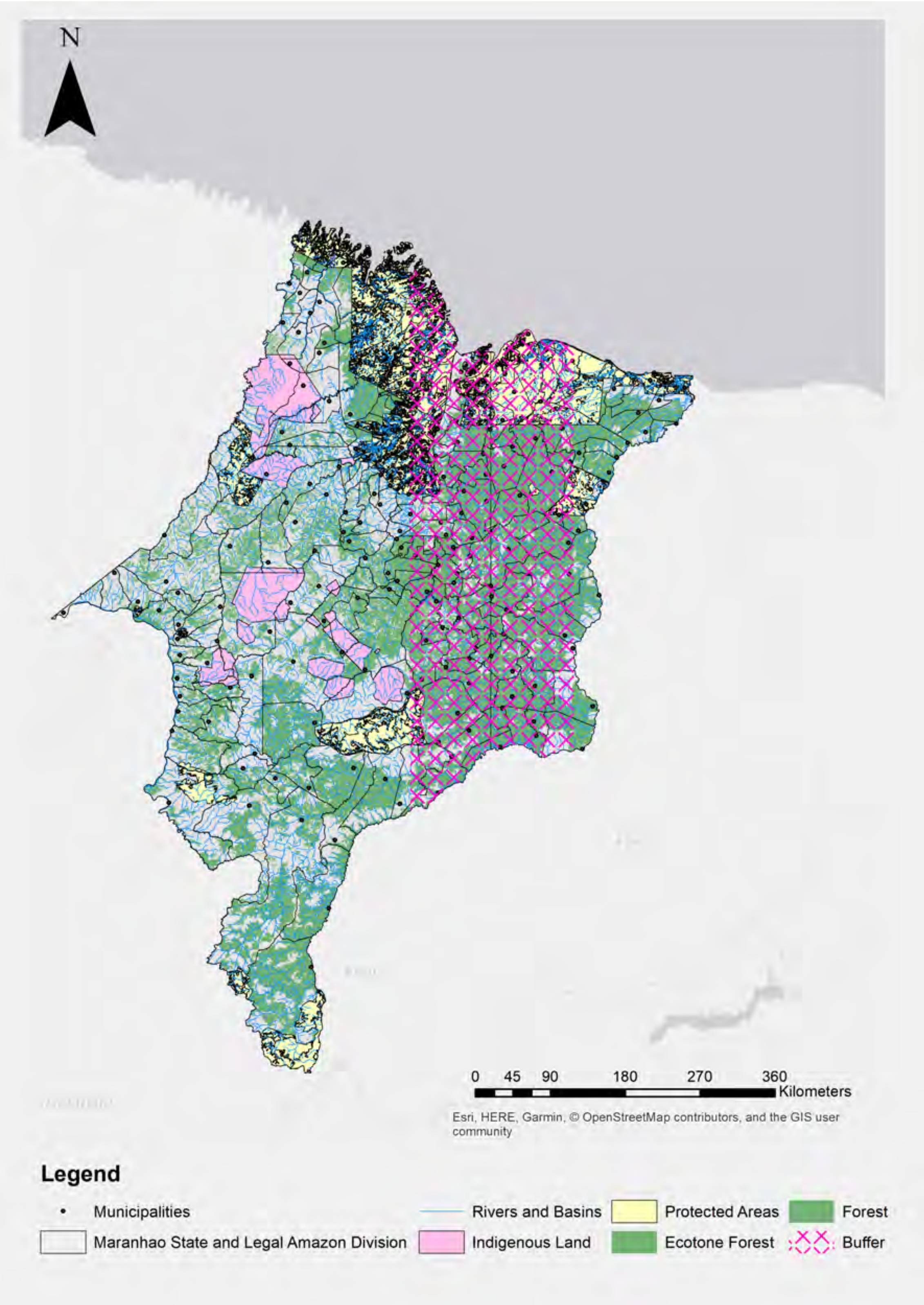

Figure 2.3 Maranhão state and 100km buffer departing from the Legal Amazon line to the east and west portion of the territory. Source: (EMBRAPA, 2018; MMA, 2018b; Núcleo Geoambiental - NUGEO, 2018).

The Study area is characterised by the occurrence of a rainfall regime with two well defined seasons. The rainy season, which is concentrated from December to June, reaching the highest peaks of rain in the month of March. The sample region presents itself as a large sloping platform in the south-north direction, with a low dip to the Atlantic Ocean. The relief is classified into two large units: plains, which are subdivided into smaller units, and plateaus. Plains are considered to be surfaces with dimensions of less than 200 meters. The plateaus are areas with heights above 200 meters, restricted to the south-central areas of the studied region (Batistella et al., 2013a; Feitosa, 2006). Information on the geology of the ecotonic region were extracted from the official map published in 2011 by the Brazilian Institute of Geography and Statistics (IBGE) on the scale 1: 1,400,000.

Solar radiation on the terrestrial surface has direct implications for local meteorology, especially in the studies on climate variability, interfering in satellite image analysis (Cohen et al., 2002; da Silva, 2004; Pereira et al., 2017). The study region here is privileged in its energetic potential, since it is located completely in the region bounded by the Tropics of Cancer and Capricorn, and close to the Equator, a condition that favors high rates of solar irradiation. The State of Maranhão presents an average annual global irradiation value of approximately 5.0 kWh / m2 (Pereira et al., 2017). In the ecological tension zone, i.e. studied area, the municipalities of Caxias (5.4 kWh / m2) (MA), Chapadinha (5.3 kWh / m2) (MA), Bacabal (4.9 kWh / m2)(LM) and São Luís (4.9 kWh / m2) (LM) are distinct for having the highest solar irradiation rates.

2.3.3 Data preparation

Handling and preparing spatial data requires specific softwares. I used ArcMap 10.4.1, ArcPy 10.4.1, and the extensions Geostatistical Analyst, Spatial Analyst and Spatial Statistics from ArcToolbox (ESRI, 2016a,b), and MATLAB R2017a and its Statistics and Machine Learning and Image Processing Toolbox (MATLAB, 2017). For statistical analysis and modelling I worked with R (R Core Team, 2018) and several packages specially 'MASS'(Venables and Ripley, 2002), 'mgcv' (Wood, 2017, 2003, 2004, 2011) and 'gratia' (Simpson, 2018).

Vegetation Indices

Two remotely sensed datasets were used – Vegetation Indices 16-Day L3 Global 250m MODIS13Q1 and Land Cover Type Yearly L3 Global 500m MODIS12Q1. These products were retrieved from the online Application for Extracting and Exploring Analysis Ready Samples (AppEEARS) tool courtesy of the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, (Didan, 2015; Didan et al., 2015; Sulla-Menashe and Friedl, 2015, 2018).

In the AppEEARS tool it is possible to define the region of interest by uploading a polygon file in shapefile format and the output file format is in Georeferenced Tagged Image File Format (GeoTIFF). When selecting GeoTIFF, one GeoTIFF will be created for each feature in the input polygon file for each layer by observation. After defining the area of interest, the tool uploads the input polygon and reproject the input file to the source projection for each data product using the Geospatial Data Abstraction Library (GDAL) ('gdalwarp' function and the PROJ.4 definition) for each data collection (USGS, 2018). In this manner the MODIS images from MODQ131 and MCD12Q1 products were acquired in GeoTIFF format and the projection chosen was the geographic datum WGS84 – EPSG:4326. Two shapefile with the same coordinate system were used to extract the location LM and MA. LM (Legal Maranhão) refers to the area under the surveillance program to the west of the Legal Amazon line and, the MA (Cerrado Maranhão) refers to the area comprising the east portion of the buffer zone (see Figure 2.3).

Next, a bounding box for each feature in the MODIS file was determined using the minimum and maximum latitude and longitude values, with a one pixel buffer applied to each corner. For each feature, the tool determines which spatial tiles intersect with the bounding box, and the tiles are then extracted from OPeNDAP (Cornillon et al., 2003) and mosaiced into a single image. The process is repeated and the MODIS images are ultimately configured into a time series image stack for each feature in the file. Reprojection is performed using nearest neighbour resampling technique and the latitude and longitude of the sample region are maintained in the conversion. Nearest neighbour resampling was selected to ensure that categorical data sets including quality data layers are able to be transformed (USGS, 2018).

A total of 776 images of the Vegetation Indices product MOD13Q1 were downloaded, from February 2000 to December 2016, and 16 images of the product Land Cover MCD12Q1 for the years 2001 to 2016. Also from the product bands MOD13Q1 the composite band day of the year was obtained, which provides the date (day of the year: 1 to 366) of acquisition of each pixel that composes the image; the band pixel reliability summary quality assurance - QA, which provides a summary of the quality of the pixels; and VI Quality detailed - QA band, which provides detailed pixel quality information. The product bands MCD12Q1 was downloaded the Land Cover type quality check – QC along with five different types of land cover data set.

Climatic variables

The climatic data were obtained from the Meteorological Database for Teaching and Research of the National Meteorological Institute (BDMEP – INMET in Portuguese), which stores historical series of several conventional meteorological stations of the INMET station network. The access is through registration but is freely available for academic purposes (BDMEP, 2018). Each conventional weather station (see Table 2.1 for the 9 stations in the sample area is composed of several isolated sensors that continuously record the meteorological parameters (e.g., temperature, precipitation, humidity, and solar radiation), which are annotated by an observer that sends the measurements to a collection center. In the historical series the maximum temperature is taken at 00 Universal Time Coordinated (UTC) of the day and the minimum temperature is collected at 12 UTC of the day. Precipitation is calculated by accumulating the last 24 hours collected at 12 UTC and solar radiation equals the number of hours the sun shines directly onto the surface as long as it is not blocked by clouds or any other obstacles. The relative humidity is obtained by the readings of the wet bulb temperature and dry bulb temperature at 12.00, 18.00, and 24.00 UTC (Vianello, 2011). We use the monthly average maximum temperature (MAMxT), monthly average compensated temperature (MACT), monthly average minimum temperature (MAMT), monthly average precipitation (MAP), monthly average relative humidity (MARH) and number of hours of sunlight in a month as total solar radiation (TS) from February 2000 to December 2016.

Table 2.1 INMET Metereological Stations

Station Number (ID)	Latitude	Longitude	Altitude	Name	Area
82571	-5.5	-45.23	153	BARRA DO CORDA	LM
82970	-9.5	-46.2	285	ALTO PARNAIBA	LM
82460	-4.21	-44.76	25	BACABAL	LM
82765	-7.33	-47.46	193	CAROLINA	LM
82376	-3.26	-45.65	45	ZE DOCA	LM
82476	-4.86	-43.35	104	CAXIAS	MA
82382	-3.73	-43.35	104	CHAPADINHA	MA
82676	-6.03	-44.25	180	COLINAS	MA
82280	-2.53	-44.21	51	SAO LUIS	MA

Note: Source: BDMEP (2018); INMET (2018).

Cross Validation

Cross validation is a necessary approach when dealing with remote sensing data. MODIS Vegetative Cover Conversion (MOD44A) was acquired from the Global Land Cover Facility - University of Maryland (GLCF, 2018) to conduct a validation process regarding the response variables (NDVI and EVI). The VCC product is used as an indicator of change and not as a means to measure change. It is available for vegetation burning and anthropogenic deforestation types of land cover conversion. In this sense, using this product as an indicator of accuracy is useful and reliable as states Zhan et al. (2002b).

As part of the validation process, we used a finer resolution data set to check the accuracy of the algorithm applied to the MOD13Q product. Hansen et al. (2013) provided results from time-series analysis of Landsat images in characterising the global forest extent and change from 2000 through 2017. The scenes utilised for the analysis contained forest losses during the period 2000–2016, defined as a stand-replacement disturbance, or a change from a forest to non-forest state. Encoded as either 0 (no loss) or else a value in the range 1–16, representing loss detected primarily for the years 2001–2016 (GFC, 2017).[1] Moreover, in 2018 Brazil's Spatial Research Institute (INPE), in accordance with the Brazil's Envinromental Ministry, published a data set covering the forest loss in the Cerrado Biome. This data consists of bi-annually images from 2000 to 2012 and yearly images from 2012 to 2016. Several sensors were used to create the composite data set, such as TM/Landsat5, ETM+/Landsat7, OLI/Landsat8, and LISS-III/IRS2. At a finer resolution, this product is a justified comparison between national and international land change products proving to be an acceptable validation procedure (Brito et al., 2018).

[1] Data Source: Hansen/UMD/Google/USGS/NASA. Data available on-line from: http://earthenginepartners.appspot.com/science-2013-global-forest.

2.3.4 Data Exploration and Interpretation

After the selection of the variables there is a refinement process which configures the most important part of the research in order to achieve usable answers for statistical modelling. In this session, we will describes the steps taken to conduct the statistical analysis and the statistical method applied to the study.

Response Variable

In order to perform the analysis, NDVI and EVI images were imported onto MATLAB and scaled to the valid range of -0.2 to 1. Two images per month for each year were uploaded - excluding October and November during leap years because they had only one image within the month. The VI Quality detailed - QA band for each scene was converted to unsigned16bit according to the VI User Guide (Didan et al., 2015) and it was used to create a Goodness mask to exclude pixels with clouds and not produced due to other reason than clouds.

Before filtering the NDVI and EVI scenes with the VI mask, it was necessary to condition NDVI and EVI values to a specific threshold so to avoid values not related to forest. The criterion was taken following Geerken (2009) and Bayma and Sano (2015) parameters to characterise forest in a transitional biome. Next, the VI indices are filtered and only values retained that have good quality according to the Goodness mask under the elimination of cloud coverage.

With the Land Cover product MCD12Q1 it was required to resize and interpolate each image to the corresponding sizes of NDVI and EVI images. The interpolation method utilised followed a deterministic method called Nearest Neighbourhood (NN) or Thiessen method. The nearest method was considered because there is no extrapolation of the data, which would not have been suitable for categorical data and because it showed to be the fastest computation method with modest memory requirements (MATLAB, 2017; Sluiter, 2009). After the interpolation, a land cover mask was produced to select only pixels in the images presenting forest classification. Following Sulla-Menashe and Friedl (2018), the University of Maryland classification which corresponded to Land Cover Type 2 in the MCD12Q1 product was selected. The mask included different types of forests with at least 40% of tree cover and canopy higher than 2m (see table 2.2 detailing each class definition). Forests presenting less than 40% of tree cover were excluded because this does not characterise a transitioning forest being predominantly assigned to Cerrado biome only (Bayma and Sano, 2015).

Table 2.2 University of Maryland (UMD) legend and class definitions

Name	Class	Description
Water	0	At least 60% of area is covered by permanent water bodies
Evergreen Needleleaf forest	1	Needleleaf Forests 1 Dominated by evergreen conifer trees (canopy >2m). Tree cover >60%.
Evergreen Broadleaf forest	2	Dominated by evergreen broadleaf and palmate trees (canopy >2m). Tree cover >60%.
Deciduous Needleleaf forest	3	Dominated by deciduous needleleaf (larch) trees (canopy >2m). Tree cover >60%.
Deciduous Broadleaf forest	4	Dominated by deciduous broadleaf trees (canopy >2m). Tree cover >60%.
Mixed forest	5	Dominated by neither deciduous nor evergreen (40-60% of each) tree type (canopy >2m). Tree cover >60%.
Closed shrublands	6	Dominated by woody perennials (1-2m height) >60% cover.
Open shrublands	7	Dominated by woody perennials (1-2m height) 10-60% cover.
Woody savannas	8	Tree cover 30-60% (canopy >2m).
Savannas	9	Tree cover 10-30% (canopy >2m).
Grasslands	10	Dominated by herbaceous annuals (<2m).
Permanent Wetlands	11	Permanently inundated lands with 30-60% water cover and >10% vegetated cover.
Croplands	12	At least 60% of area is cultivated cropland.
Urban and built-up	13	At least 30% impervious surface area including building materials, asphalt, and vehicles.
Cropland/Natural Vegetation Mosaics	14	Mosaics of small-scale cultivation 40-60% with natural tree, shrub, or herbaceous vegetation
Non-Vegetated Land	15	At least 60% of area is non-vegetated barren (sand, rock, soil) or permanent snow and ice with less than 10% vegetation.
Unclassified	255	Has not received a map label because of missing inputs

Note: Source: Sulla-Menashe and Friedl (2018).

Finally, to obtain a certain variation within each month, values of NDVI and EVI of the first period, with 15 first days of the month, were compared to the second period, with 30 days of the month. The assumption considered is explained in Table 2.3. In this sense, the final scene/image would present pixels assuming the highest quality and no cloud coverage. At the end of the process, pixels were selected within the bandwidths of 100km, 50km, and 25km - measured departing from the artificial Legal Amazon line to the west and east portion of the State. When the pixel had a variation greater than 0.1 within a month, it was considered a disturbance.

Table 2.3 Algorithm Assumption for NDVI and EVI values

$NDVI_1 > NDVI_2$	-> $NDVI_1 - NDVI_2$	Numbers (1) and (2) refer to the order of the period of the month
$NDVI_1 <= NDVI_2$	-> $NDVI_1 = NDVI_2$	Numbers (1) and (2) refer to the order of the period of the month. The second equation assumes that values did not change within the month and then the value assigned is from the last observation
$EVI_1 > EVI_2$	-> $EVI_1 - EVI_2$	Numbers (1) and (2) refer to the order of the period of the month
$EVI_1 <= EVI_2$	-> $EVI_1 = EVI_2$	Numbers (1) and (2) refer to the order of the period of the month. The second equation assumes that values did not change within the month and then the value assigned is from the last observation

The approach above was undertaken for all the images corresponding to NDVI and EVI values for each month of each year, giving 406 final image results. For the leap year, the process stopped at the land cover mask filtering process. To compose a panel for each monthly period over 17 years, we took the sum of pixels signalled as disturbed for each of 406 images.

Covariates Variables

The climatic variables needed a more complex processing system since the data from weather stations were sparse. First, all the data was in a tabular format, i.e. all variables in one table, and geographic locations in the form of latitude and longitude coordinates and z-coordinates, such as elevation values, was created for each weather station using ArcMap. After localising the x,y,z coordinates, a shapefile for each station was created. Then the shapefiles were selected and extracted by year, using ArcPy environment.

Next, an interpolation method was used to deal with areas with no data available. The method chosen was ordinary (point) Kriging which has been argued as the best interpolation technique available for sparse data (Sluiter, 2009). Ordinary kriging is part of the probabilistic methods in which the concept of randomness is incorporated into the analysis. This method is the basic form of Kriging, where the prediction relies on a linear combination of the measured values and the spatial correlation between the data, determining the weights. As the mean is unknown we assume that intrinsic stationary exists in the data. This assumption may fail for our data set since this type of data are usually not stationary. To overcome this issue we used different sizes and shapes of neighbourhood to adjust the kringing ordinary model (Sluiter, 2009).

After the interpolation, the images were converted to raster, resampled to the size of the response variable, and exported to MATLAB environment. Generally, for temperature, precipitation, and other climate data, the best way to interpret and study these phenomenons is using anomaly measurements which corresponds to the difference between measurement and mean. In this sense, the average value of the variable of each image for each month and each year was computed, giving a total of 408 images analysed for both regions (MA and LM), and for each climatic variable, a total of 2,448 images. Following this procedure, the number of pixels with values higher and lower than the average value of the variable was extracted to a table. At the end, the table contained above and below values compared to the mean for each variable translating into 12 variables. A summary of the response and explanatory variables of this study is presented at Table 2.4.

Table 2.4 Summary Statistics - Response and Covariate Variables

Variable	Mean	St. Dev.	Min.	Max.
NDVI 25Km	626	731	0	4803
NDVI 50Km	1279	1415	0	9185
NDVI 100Km	2087	2175	0	12660
EVI 25Km	8277	11353	0	73734
EVI 50Km	2554	3861	0	25250
EVI 100Km	8277	11353	0	73734
Below Min Temp 25km	83711	57037	0	218320
Below Min Temp 50km	159092	112926	0	436290
Below Min Temp 100km	250820	171007	0	633370
Above Max Temp 25km	105094	69290	0	218320
Above Max Temp 50km	201308	134299	0	436300
Above Max Temp 100km	287792	190863	0	633370
Below Sunlight 25km	79678	47638	0	169150
Below Sunlight 50km	167845	98578	0	434760
Below Sunlight 100km	261061	149782	0	633370
Below Humidity 25km	98957	50230	0	213400
Below Humidity 50km	192560	97230	0	324620
Below Humidity 100km	289185	140379	0	633370
Above Precipitation 25km	90697	37384	0	191170
Above Precipitation 50km	186745	77623	0	411370
Above Precipitation 100km	292985	118665	0	633370

Note: Statistics refer to N=204 observations for 17 years (2000 - 2016). The below and above nomenclature refers to the mean of each variable.

Modelling deforestation trends

Many recent studies of land cover changes focus explicitly on taking account of the trends and changes in the rates of environmental transformation in terms of their driving forces. More precisely, these studies try to identify the major causes of land-cover change within different geographical and historical contexts (Geist and Lambin, 2001). To this end proximate and underlying causes of deforestation models are concerned by the fact that some causes are direct in the sense that their occurrence or variation generates more or less deforestation through simple channels, while other causes are indirect in that they impact on the sources of deforestation through more complex channels (Combes Motel et al., 2009). In this regard, the physical environment strongly influences where agents deforest, where many studies provide evidence that forests in drier, flatter, higher-fertility areas, with adequate drainage and thus more suitable for agriculture are more likely to be cleared (Kaimowitz and Angelsen, 1998). In contrast, poor soil quality is also reported to lead to relatively high deforestation, since scant soil endowment fuels accelerated clearing for other activities, such as pasture (Geist and Lambin, 2001; Silva Costa et al., 2012).

Environmental factors and biophysical drivers are also increasingly being recognised as not only playing a role but being fundamental in deforestation (Geist and Lambin, 2001). To cite, Barni et al. (2015) showed that, independent of the rate and magnitude of deforested areas, the areas affected by forest fires were dependent on the forest type and climate factors. Zones with ecotone influence tended to be deforested more than zones without ecotone influence, i.e., the more dense a forest is the less deforested will be. In addition, the largest occurrence of forest fires was observed in the zones with ecotone influence in years with *El Nino* events, such as Maranhão state. The analysis also indicated that the areas most affected by forest fires during the studied period were associated with strong climatic events and the occurrence of these fires was amplified in the zones with ecotone influence (Barni et al., 2015). These facts strongly suggest that it is pertinent to control for climatic aspects in ecotone zones when studying trends in deforestation.[1]

Statistical Modelling

[1] In this study, ecological tension zone, ecotone zones and transitional forest have the same meaning.

The most common approach used to model deforestation would be a linear model with a single or multiple terms. However, things are rarely this simple, and the model might interact in a complex way. The four approaches presented earlier consider the linear model as the first premise of the analysis but they are challenging when dealing with non-linear relationship found between variables (Griffin, 2012). In environmental analysis the data are seldom modelled adequately by linear regression models.

Dealing with nonlinear processes trying to approximate to linear estimation methods may lead to inconsistent results guided by the linear estimator. There is an imposed assumption regarding the distribution and nature of the data. On the subject of the choice of distribution, it is doubtful that a satisfactory model could have been produced in this manner. This is part of the motivation for seeking to allow a more compact and flexible way of specifying smooth functional relationships within the models.

For most ecological and climatic data sets at least some of the assumptions underlying a linear regression model are unlikely to be valid (Zuur, 2011).[1]. To address this issue we here employ a generalized additive model (GAM). A literature review shows that GAMs are not extensively applied to deforestation models. Chaves et al. (2008) showed that social factors appear to play a critical role that may ultimately determine disease risk when evaluated with environmental and climatic factors. They modelled incidence rates of a disease in Cuba as the response variable and deforestation as one of the exploratory variables. Mendes and Junior (2012) observed the relationship of deforestation, corruption and economic growth in the region of Legal Amazon, in Brazil and found no statistical evidence for the existence of a Kuznets curve. Green et al. (2013) used a binomial GAM model to account for forest and habitat losses in protected areas on the Eastern Arc Mountains of Tanzania. More recently, Bebber and Butt (2017) studied the impact of

[1] An empirical illustration of the innovation of GAMs by comparing estimates using the same data of the method chosen and the previously best-practice old methods, such as additive moving averages and loess smooth functions is shown in the Appendix A.2. It is possible to observe that both methods give unwieldy results and do not allow for data flexibility.

protected areas on global carbon emissions in America, Africa and Asia. They used splines regressions in GAMs and suggested that tropical PAs overall reduced deforestation carbon emissions by 4.88 Pg, or around 29%, between 2000 and 2012.

A GAM is a generalized linear model with a linear predictor involving a sum of smooth functions of covariates. Mathematically, GAM is an additive modelling technique where the impact of the predicted variables is captured through smooth functions (Larsen, 2015). In general, the model has a structure defined as:

$$g(\mu_i) = A_i\theta + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + \dots \tag{2.3}$$

where $\mu i \equiv \mathbb{E}(Y_i)$ and $Y_i \sim EF(\mu_i, \phi)$. Y_i is a response variable, $EF(\mu_i, \phi)$ denotes an exponential family distribution with mean μi and scale parameter, ϕ, A_i is a row of the model matrix for any strictly parametric model components, θ is the corresponding parameter vector, the f_j are smooth functions of the covariates, x_k, and i refers to the unit of analysis (Wood, 2017). This model allows for flexible specification of the dependence of the response on the covariates because the smooth functions are nonparametric. The smooth function f_j is represented by basis expansions for each smooth, each with an associated penalty function controlling smoothness. According to Wood (2017, 2004), the estimation can be carried out by penalised regression methods, and the appropriate degree of smoothness for f_j can be estimated from data using cross validation or marginal likelihood maximisation.

The smooth function f is composed of the sum of basis functions b and their corresponding regression coefficients β, written in the form of:

$$f(x) = \sum_{j=1}^{k} b_j(x)\beta_j, \tag{2.4}$$

where k is the basis dimension (Wood (2017), p.162).

It is possible to regularise the smoothness of the predictor functions to prevent overfitting using the generalized cross validation score. Technically, GAM is an additive modelling technique where the impact of the predictive variables is captured through smooth functions, but provides a regularised, automatic and interpretable solution. Considering an additive model, the interpretation of the marginal effects of a single variable does not depend on the values of the other variables in the model. Also, predictor functions are automatically derived during model estimation (Larsen, 2015).

Autocorrelation

In principle, when the nature of the data is time series, the timing of one period may depend on the timing of the previous year. This means that one should check for the possibility of autocorrelation and if necessary take into account of such the auto-correlation in the data. In GAMs it is possible to include an ARMA error structure. More precisely, the ARMA model has two parameters defining its order with the number of auto-regressive parameters (p) and the number of moving average parameters (q). Model 2.3 now can be expressed as

$$g(\mu_i) = A_i\theta + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + \cdots + \varepsilon_i \tag{2.5}$$

where $\varepsilon_i = \phi\varepsilon_{i-1} + \phi\varepsilon_{i-p} + \eta_i$ is modelled as a function of the residuals of the p previous time points and white noise, and $\varepsilon_i = \theta\eta_{i-1} + \theta\eta_{i-q} + \eta_i$ is modelled as a function of the disturbance term and a past value of this disturbance term (Zuur et al., 2014).

Quantifying deforestation

The generalized additive model (GAM) with an exponential family distribution has been the most widely applied method to measure and quantify the non-linear association between phenology and covariates, such as meteorological conditions, mainly because it allows for non-parametric adjustments of non-linear confounding effects of seasonality and trends; (Antunez et al., 2017; Auderset Joye and Rey-Boissezon, 2015; Bell et al., 2015; Bio et al., 1998; de Souza et al., 2017; Halperin et al., 2016; Liu et al., 2018; Lusk et al., 2016; Moreno-Fernández et al., 2018; Pourtaghi et al., 2016). In an attempt to quantify forest disturbance as proxy for deforestation, we apply a GAM with a negative binomial distribution and a logarithmic link function. The negative binomial distribution suitable for this study since the variance of deforestation is much larger than the mean, which is frequent feature of ecological data (Zuur, 2011). More precisely, the means of NDVI and EVI are to 0.04% and 0.01% of their variance, respectively. Hence, the response variable is negative binomial distributed. The full description is as follows:

$$\begin{aligned} &Y_{is} \sim \mathrm{NB}(\mu_i,\mathrm{k}) \\ &E(Y_i) = \mu_i, and \quad var(Y_i) = \mu_i + \frac{\mu_i^2}{k} \\ &\log(\mu_i) = \alpha + f_j(X_{i1}) + \cdots + f_j(X_{iq}) \qquad \text{or} \qquad \mu i = e^{\alpha + f_j(X_{i1}) + \cdots + f_j(X_{iq})} \end{aligned} \tag{2.6}$$

Y_i is the response variable at observation i. The notation $f_j(X_{i1})$ stands for 'smoothing function of the covariate variable X', and *NB* is a negative binomial distribution with mean μ_i and dispersion parameter k. In general, negative binomial distributions are used to model overdispersed count data or Poisson data.

The geometric distribution is a negative binomial with overdispersion parameter, k, set to 1. In this sense, the variance increases as a quadratic function of the mean (Zuur et al., 2014). Correcting the data for overdispersion with the geometric distribution, the model is stated as

$$Def_{(NDVI_i,EVI_i)} = \alpha + f_{Year}(Year) + f_{Precip}(aPrecipitation) + f_{MaxTemp}(aMaxTemperature)$$

$$+ f_{MinTemp}(bMinTemperature) + f_{Sunlight}(bSunlight) + f_{Humidity}(bHumidity) \quad (2.7)$$

where $Def_{(NDVI_i,EVI_i)}$ is the response variable that can assume NDVI and EVI values for three different bandwidths (25km, 50km, 100km) in each month i, and $i = 1, \ldots, 204$. The remaining variables are the intercept α andthe additive smoothing functions of the explanatory variables Year, and the covariates Precipitation, Humidity, Max and Min temperature and sunlight. a and b refers to the sum of pixels above and below the mean, respectively.

The model selection followed the forwarding approach of Zuur et al., 2014, p.391. The model started with a GAM that used one variable, then fitted 13 different models and a different set of smoothers (penalised splines "ps", cubic splines "cr", and cyclic splines "cc") and compared their Akaike information criterion (AIC). The model with the lowest AIC was elected as the main model and, then it was fitted to 12 different models, each with the addition of the variable with the lowest AIC. The forward selection stopped at the moment the main model had the best AIC value comparing to the remaining models. After the model selection, an autocorrelation test was conducted but none of the models appeared to be autocorrelated. The model, including the splines takes the form of:

$$Def_{(NDVI_i,EVI_i)} = \alpha + f_{Year}(Year, bs = cc) + f_{Precip}(aPrecipitation, bs = cr) +$$

$$f_{MaxTemp}(aMaxTemperature, bs = cr) + f_{MinTemp}(bMinTemperature, bs = cc) + \quad (2.8)$$

$$f_{Sunlight}(bSunlight, bs = cc) + f_{Humidity}(bHumidity, bs = ps)$$

Due to the fact that his method is relatively recent, it is important to acknowledge that the algorithms available for choosing the optimal smoothing parameter are not yet well developed and can generate misleading results if care is not taken. Furthermore, use of such criteria can often lead to over-fitting and deliver implausible associations. The choice of smoothing parameters for smoothing splines in GAM should therefore always be accompanied by a graphical verification of functional associations with the outcome to verify clinical plausibility (Moore et al., 2011).

Model Validation

Validating the results from the algorithm applied to the MOD13Q1 and MCD12Q1 images required features of the machine learning domain. In summary, machine learning algorithms can figure out how to perform important tasks by generalizing from examples. This is often feasible and cost-effective where manual programming is not (Domingos, 2012).

There are different types of machine learning algorithms, where the most mature and widely used one is classification. According to Domingos (2012), a classifier is a system that inputs a vector of discrete and/or continuous feature values and outputs a single discrete value, the class. For this study, the filter classifies pixels into deforested or not deforested and its input may be a Boolean vector $x = (x_1, \ldots, x_j, \ldots, x_d)$ where $x_j = 1$ if the jth pixel is deforested and $x_j = 0$ otherwise. A learner inputs a training set (x_i, y_i), where $x_i = (x_{i,1}, \ldots, x_{i,d})$ is an observed input and y_i is the corresponding output, and output is a classifier. The test of the learner is whether this classifier produces the correct output y_t for future examples x_t. A feasible classification validation is the confusion matrix.

The confusion matrix is a two by two table that contains four outcomes produced by a binary classifier. The classification scheme divides the data randomly into a training set, a test set and a validation set. The training method is the Scaled Conjugate Gradient (SCG), which is a supervised learning algorithm for feed-forward neural networks (Mor, 1993). In order to optimise the performance, an iterative random sampling approach is applied. The Cross-Entropy approach is based on sampling and updating an underlying distribution function over the set of feasible solutions (Hu and Hu, 2009). Finally, calculations for the confusion matrix are done based on minimum excluded (MEX) calculations (MATLAB, 2017).

The four outcomes produced by the confusion matrix are true positive, true negative, false positive, and false negative. The true positives (TP) refer to the number of positives divided by all the positive outcomes and the same applies to true negatives (TN). False positives (FP) indicate the number of pixels assigned as positive but are, in fact, negative, divided by all the positive outcome. In turn, false negatives (FN) follow the same interpretation as false positives.

To validate the response variable results, it was used the MODIS Vegetative Cover Conversion (VCC) for the available period (2000-2005). The product is further divided in Deforestation product (MOD44A_C) and Burn product (MOD44A_B) and both were used to compute the validation test.

The method for the deforestation product is derived from the original space partitioning method (Zhan et al., 2002a) and relies on a decision tree classification algorithm (Gordon et al., 1984) to determine antecedent vegetation condition and compares this to current vegetation condition. Change due to burning product is derived using the difference Normalized Burn Ratio (dNBR) methodology from two scenes a year apart, as proposed by van Wagtendonk et al. (2004). Tests were computed per season (raining season and dry season) and per vegetation index (NDVI and EVI). The confusion matrix showed 100% true positives and true negatives, which gives high stability to the algorithm created and applied to the NDVI and EVI indices.

Checking the results with other datasets was an alternative approach taken in this study. A confusion matrix was applied to the Hansen et al. (2013) and Brazilian (INPE) data sets (Brito et al., 2018). The results showed no difference from the results presented in the VCC validation method. The confusion matrices are provided in the Appendix 2.6.

The covariates variables were validated using cross-validation processes during the interpolation procedure. Cross-validation uses all the data to estimate the trend and autocorrelation models. It removes each data location one at a time and predicts the associated data value. This is also known as leaving-one-out, and can be computed for all or a subset of the data locations (ESRI, 2016a).

In the kriging method, the cross-validation produced other results that helped evaluate the best interpolation method. More specifically, the Average Standard Errors (ASE) and Root Mean Square Standardized Error (RMSE) were computed. If ASE from the model are close to the RMSE then the model is correctly assessing the variability in prediction. If ASE are greater than RMSE then the model is overestimating the variability of the prediction and, finally, if the ASE are less than RMSE, the model is underestimating the variability in predictions. For the covariates analyses, ASE were on average 95% of the value of the RMSE, proving to be a reasonable interpolation method with valid results.

In terms of statistics, model validation with additive modelling was visual rather than numeric after the model selection phase. The steps taken included plotting the residuals against fitted values to identify violation of homogeneity, and plotting the residuals against each variable in the model and check for patterns. Also, a histogram of the residuals was examined to verify normality.

2.4 Results

2.4.1 Deforestation trend in a ecotone zone of Maranhão

Our baseline model includes 204 monthly observations of NDVI and EVI values changing over the years with five influencing covariates (Precipitation, Max Temperature, Min Temperature, Sunlight and Humidity). The baseline model was applied to three different distance spans (25km, 50km, 100km) considering the Legal Amazon line.[1]

In general, the deviance explains the models close to the artificial line better. At large distances (100 km), most of the covariates do not have a significant effect, and thus will be omitted from this part of analysis. Table 2.5 gives the summary of the results including deviance, AIC, p-value, degrees of freedom, and the estimated value of the function. The best way to understand and interpret GAMs is through visual representation. Considering that the results produced several models, we define the name of these models according to the location status, whether in Cerrado Maranhão (MA) or Legal Maranhão (LM), the bandwidth or distance from the Legal Amazon line, in which are 25km,50km and, 100km, and, finally, regarding each response variable that in our case is related to NDVI (n) and EVI (e) values. We add an indicative variable to indicate raining (r) and dry (d) seasonality. Plotting the smoothing functions, it is possible to check for the path of deforestation through the years and the climatic state during that period. The blue line refers to positive changes in deforestation or increments, and the red line indicates negative changes in deforestation or decreases (Simpson, 2018).

With respect to the model *ma25n*, the explained deviance is 15%, the variance was set to 1 in the geometric negative binomial distribution. The explanatory variable year is

[1]It is important to acknowledge that the numerical results of the model should not be interpreted in the same manner as the linear regression results. According to Wood (2011) in Zuur et al. (2014), p-values close to 0.05 can be around half of their correct value when the null hypothesis is true. This means that smoothers with p-values smaller than 0.001 can be trusted but p-values of 0.02 to 0.05 need to be viewed with caution.

Table 2.5 Models Output of GAMs

Model	AIC	Deviance Explained	AIC	Deviance Explained	AIC	Deviance Explained
Baseline Model			Raining Season		Dry Season	
ma25n	4592.974	15%	3008.928	53.8%	3029.610	52.3%
lm25n	4753.342	30.9%	3082.583	44%	3330.420	54.6%
ma25e	5382.682	20.1%	3486.908	52.1%	3533.256	50.9%
lm25e	5450.606	36.4%	3369.937	63.1%	3812.126	65.6%
ma50n	5035.296	17.3%	3402.803	60%	3390.411	42.2%
lm50n	5010.875	24.5%	3241.013	55.8%	3498.748	35.1%
ma50e	6123.032	20.5%	4154.889	62.6%	4051.907	47%
lm50e	5921.597	29.3%	3886.190	63.2%	4112.314	48.5%
ma100n	5312.836	16.6%	3723.872	62.2%	3562.066	45.2%
lm100n	5318.571	30.5%	3552.311	52.7%	3701.532	47.2%
ma100e	6097.770	23.6%	4176.352	59.4%	4051.806	45.3%
lm100e	5901.040	33.4%	3906.004	59.5%	4097.055	46.7%

Note: Models Output of GAMs with Akaike's Information Criteria (AIC) and Deviance goodness-of-fit statistic for each statistical model significant at 1% level, and the smoothers significant at 1% level are *Min Temperature*, *Max Temperature*, *Humidity* and *Sunlight*. The degrees of freedom for the smoothers are 2.4, 6.8, 4 and 7.9. In summary, for MA at 25km, deforestation had a positive effect after 2010 and, through the years, deforestation decreased during periods of low humidity (high thermal oscillation). It is also deduced from the model that deforestation decreased during periods of more hours of sunlight. There were more deforested pixels in periods where temperature declined.

For the model *lm25n*, the explained deviance is 30.9%, the explanatory variable year is significant at 5% level, the smoothers significant at 1% level are *Max Temperature*, *Humidity* and *Sunlight*. The degrees of freedom for these smoothers are 8.6, 6.1 and 1.4. The results are similar to the *ma25n* model by showing that deforestation also decreased during periods of less humidity and extreme higher levels of precipitation. Examining deforestation as a function of the year showed that there was a positive effect, i.e., increments on forest loss, during the beginning of the 2000's.

Models with EVI values were considered better in terms of cross validation. Model *ma25e* shows that deforestation increased over time with a positive peak after 2010. Deforestation also increased when the covariates sunlight and minimum temperature decreased. For maximum temperature, the negative effect is greater than the positive effect but, in essence, deforestation decreased with higher temperatures. On the Legal Maranhão side, the results of the model *lm25e* show that all variables are significant at the 1% level. The model explains 36% of the changes in EVI values, i.e., deforestation. From 2007 to 2012, deforestation increased in the region. The positive effect happened during periods of higher temperature and low humidity.

2.4.2 The effect of seasonality in the deforestation trend in a ecotone zone of Maranhão

Given the results in Section 2.4.1, it is clear that seasonality is a key factor for the deforestation trend in the transition forest of Maranhão. Low values for solar incidence, low levels of humidity, high levels of precipitation, and reduced values of temperature indicate possible differences in the trend for the winter and summer season. As explained in Section 2.3.2, the ecotone forest presents two well defined seasons: the rain period (summer) and the dry period (winter). In an attempt to refine the analysis, we divided the sample according these two seasons. The wet season starts in December and continues until the end of June. From July the dry season starts remaining until the end of November. Thus there are 102 observations for each season. To allow further comparison, the model took the same approach given by equation 2.8.

Raining Season

Subsetting the data and applying GAM, model *ma25nr* explains 53.8% of the deforestation path in the Maranhão eastern side. From the plot, deforestation increased in year cycles, i.e., for 2000-2002, 2006 and 2008. Also, increased forest loss happened with less available sunlight and increased precipitation levels. The deforestation trend shows a decrease when temperatures reached extreme high and low values. The same pattern is observed for the humidity covariate. At 25km in the LM area, deforestation over time had only two positive effects, from 2000-2001 and from 2011-2012. Model *lm25nr* with a 44% of deviance explained shows that clear-cutting expanded in periods of high levels of precipitation, and less exposure to sunlight.

The model with EVI values as the response variable shows a different trend comparing to NDVI values. Model *ma25er* indicates that deforestation increased in the MA region over the years. Accordingly, deforestation took place when precipitation levels were higher than the average and during lower hours of sunshine. For the *lm25er* model, 63.1% of the model explains the deforestation process. Cutting down the trees was more prominent during 2000 to 2005 and 2010 to 2015. The removal of trees increased with high levels of precipitation, and low levels of temperature and sunlight.

Using the areas within 50km of the artificial line, model *ma50nr* followed the same arrangement shown in model ma25nr, except for revealing that forest losses icnrease through time. There is also a clear cycle apparent from the Year plot. A similar cycle is also shown in model *lm50nr* for the explanatory variable year. Decreasing deforestation is associated with high levels of precipitation in the MA region.

Improving the model by deviance explained, model *ma50er* also exhibits a cycle pattern for deforestation in the MA region, with no singularities compared to the NDVI model (*ma50nr*). Model *lm50er* shows that deforestation increased before and after 2005 and when maximum temperatures were even higher than the maximum average. The decreasing process happened when precipitation levels were much higher than the average as well.

For the models with the largest buffer area in the MA ecotone region, model *ma100nr* still showed some deforestation cycle, with a peak right after 2005. At 100km, deforestation was positive when temperature dropped more than the minimum average and when sun exposure presented less number of direct sun hours. At the LM 100km-border, the model *lm100nr* provides evidence of an increasing path of deforestation through time, with the highest peak during 2009. Forest losses took place when precipitation levels were higher than the average up to a certain limit.

Relative to the EVI values analysis, the model *ma100er* presented similar results from the NDVI model. It is noticeable, however, that the EVI model shows an increasing path of deforestation during 2006 - 2010, unlike under the NDVI model (*ma100nr*). The EVI model for the LM side also presented similarities to the NDVI model. Looking over the years, deforestation had a positive effect until 2005, then again in 2008 to 2010, and once more recent in the years 2014-2016. However, these cycles were not different from what was already seen in the NDVI model (*lm100nr*).

Dry Season

The analysis of the dry season with the GAM model proposed in 2.8 shows that the deviance explained by the *ma25nd* model was 52.3%. All the terms were highly significant at the 0.1% level. The deforestation trend had three positive peaks during 2005-2006, 2011-2013, and after 2015. Forest loss increments appeared in periods of high precipitation level accompanied by high temperatures. The model looking at a 25km bandwidth on the LM side has 54.6% of deviance explained. However, only two variables are significant at 1% level, namely *Humidity* and *Sunlight*. In the dry season, the Legal Maranhão decreased deforestation when humidity levels were low and increased deforestation when sunlight was below the average. Apparently, deforestation did not changed over time during the dry season.

Turning to the second response variable, the model *ma25ed* has all the terms highly significant at the 0.1% level and the deforestation cycle through time is evident. The same pattern, presented in the model with the NDVI response variable, is seen in this model. Positive values of forest losses happened in periods of high precipitation accompanied by high temperatures with less hours of sun. Differently from model *lm25nd*, model *lm25ed* improved significantly with 65.6% deviance explained. One can observe four positive peaks of deforestation during the studied period 2004, 2006, 2011 and 2015, even though the overall path shows a significant decrease in 2005, which essentially compensates for the positive peaks.

For the 50km bandwidth on the MA side, the explanatory variable and the covariates are highly significant at 0.1%. In general, an increment on deforestation from the model *ma50nd* occurred during 2005 and 2006 and from 2013 until 2016. In general, the process occurred during high levels of precipitation and low temperatures. In the LM region (model *lm50nd*), during the dry season negative changes in the deforestation took place when humidity levels were low and temperatures were high. There appears to be no deforestation trend over the years.

Model *ma50ed* shows an interesting deforestation trend for the Maranhão region. Until 2005, deforestation was decreasing over time but, after 2005, the deforestation process increased especially after 2013. This process took place during periods of high precipitation rates and lower temperatures. Looking across the artificial line, for EVI values, the *lm50ed* model showed only one period of changes during 2005 (decreasing) to 2007 (increasing). After that, there appears to be no trend of deforestation through the years. The process captured during that period followed low temperatures.

Reflecting the same pattern from previous models with smaller buffers, the model *ma100nd* presents the explanatory variable and the covariates as highly significant at 0.1%. Deviance explained 45.2% of the model and a modest trend is seen with an increasing peak after 2013. With 47.2% of deviance explained, model *lm100nd* does not show a deforestation trend over the years, demonstrating accordance with the previous models (*lm25nd* and *lm50nd*). Less humidity in the period of forest change is observed for the *lm100nd* model.

The deforestation trend for model *ma100ed* is very similar to model *ma100nd*, showing no significant difference in the peaks. The variables are highly significant at the 0.1% level and the trend is observed with high levels of precipitation and lower temperatures. Decreased deforestation was detected during periods of low humidity levels. In terms of the 100km buffer, the LM region did not experience a trend in deforestation over time.

Deviance explained 46.7% of the model and forest change happened during periods of low humidity and low solar radiation.

2.4.3 Settlements

Thus far we have found clear differences the trends, and the factors driving them, in deforestation across the two regions examined here, regardless of their spatial definition. To further investigate this we also look specifically examine deforestation trends within settlements. Settlements areas are allocated and supervised by Brazil's Special Secretary of Agrarian Development, and there is virtually no law enforcement these plots, resulting in low levels of environmental compliance (Schneider and Peres, 2015).[1] The Brazilian Environmental Police (Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis—IBAMA) repeatedly fined the federal Agrarian Agency (Instituto Nacional de Colonização e Reforma Agrária—INCRA) for environmental violations in the settlements. Usually, the fines are sent to INCRA and not to the settler. When the agency receives the fines applied by the environmental policy, the justice system frequently takes the side of the agency or even annuls the fines because, from a legal point of view, the agency does not commit an environmental crime, but rather acknowledges the presence of preserved forest within settlements. In this sense the INCRA complies with the legislation because it leaves part of the forest of the whole settlement intact, but cannot oblige the granted slots to deforest.

[1] Due to the intense debate on the issue and the commitment of other Latin American countries with the implementation of agrarian reform in the decade of 60, the government included it as one of its priorities. An amendment that allowed the Union to promote the expropriation of social interest, upon payment of prior and fair compensation in special government bonds, was drafted and approved. Shortly thereafter, Law 4504 was enacted, which provides for the Land Statute (Estatuto da Terra, 1964). The Brazilian Federal Land Reform program was launched in 1964 to bring *people without land to land without people*. This applied not only to the poor and landless *peasants*, but also to the expanding Southern Brazilian agribusiness. At the same time, the Brazilian Institute of Agrarian Reform (IBRA in portuguese) and the National Institute of Agrarian Development (INDA in portuguese) were created, and replaced by the Institute for Rural Settlement and Agrarian Reform (INCRA). Brazil was provided with legal and institutional framework that would start a national land reform program (Ezzine-de Blas et al., 2011).

Deforestation in settlements is committed by the landowner or landholder. In most cases, there are loggers who lease lots and even press the small farmer to clear. They may also threaten and kill the leaderships that hinder the timber business. Moreover, there is pressure from the local commerce and from sawmills, who buy this wood. Therefore, it is not the agrarian reform policy that is the cause of deforestation. Under pressure by public opinion, INCRA established in 2012 the 'Green Settlement Program' to deal with the environmental debt of settlements. This policy, however, is still not implemented and there is no feasibility that could endorse the effectiveness of this policy Pacheco (2009); Schneider and Peres (2015). Under these circumstances, and in an attempt to corroborate with the finding results, this study took from the observed region, settlements areas from both sides MA and LM (see figure A.2.49 since they are not directly affected by any surveillance monitoring policy in order to check whether the trends of deforestation differed from the trends presented within settlements.

The analysis of deforestation in settlements follows the same strategy as presented in Section 2.4.1 and 2.4.2. The sample includes 204 monthly observations of Vegetation Indices values changing over the years including covariates, such as, *Precipitation*, *Max Temperature*, *Min Temperature*, *Sunlight* and *Humidity*. The baseline model presented in equation 2.8 is applied to the entire studied area since the environmental surveillance policy is not applied to the settlers. Plotting the smoothing functions, it is possible to check for the path of deforestation through the years and the climatic state during the studied period.

The MA model shows that lower levels of deforestation occurred when the settlements had lower presence of sunlight during the day. Also, deforestation decreased significantly before 2007 and no significant changes were observed after that period. Comparing the indices, the EVI model showed more variability than the NDVI model. Deforestation increased before 2004 and after 2010, and happened with significant changes in temperature, where forest losses took place with higher levels of temperature and preceded by lower levels of temperature. In general, these results resemble the findings for the Maranhão side. Deviance explained between 33% to 35% of the model.

Looking at the LM model, significant deforestation appeared when the settlements had lower presence of sunlight during the day, in contrast with MA model. It can be seen from the model that deforestation happened at a constant level having no significant changes throughout the studied years. No differences were found when comparing the two vegetation indices. The results here also do not provide evidence of significant differences when compared to the results found for the 100km threshold (see 2.6). Deviance in this model explained between 23% to 26% of the model.

When sub-sampling the dataset into rainy season, one observes that the MA model presents a cyclical deforestation trend, in congruence with the main findings. For both indices the deforestation cycle within settlements had less variability and the number of pixels deforested did not increase as it did for the findings in 2.4.2. The deviance explained about 72% of the model. The settlements in the Legal Amazon side experience similar results to the rainy season sample from the main findings. Deforestation followed a scenario of higher precipitation levels, less hours with sun visibility, and low temperatures. Deviance explained 49% to 58% of the model.

The dry season for the settlements within the Maranhão side reveals a similar path compared to the results found at 2.4.2. Inside settlements deforestation increased during high levels of precipitation and low levels of temperature. This model also states that deforestation did not change significantly before 2005, in contrast to the main findings for this period. Deviance explained 51% to 53% of the model. The deforestation path within settlements in the Legal Amazon side shows a completely different path when compared to the main finding results for this region. There is clearly a cycle path of deforestation within settlements. This result is apparent considering both vegetation indices, but is more prominent for NDVI. Deviance ranged from 36% to 41%. The results obtained for this region supported the three outcomes highlighted in Section 2.5. More precisely, the findings here suggest that much of the forest loss activity happened similarly between the two regions. The second outcome, that cloud cover might benefit deforestation because it inhibits the forest cover detection from satellites, is not observed for the settlements area. An explanation may be that in the main sample the artificial line was based on a political and policy barrier, but for settlements this barrier did not have a policy significance since the settlements are not constrained by it.

2.5 Discussion

Previous studies considering GAM models for modelling deforestation are scarce. Looking at the pure ecological analyses, there are studies related to changes in phenology using GAM, for example Antunez et al. (2017). The study here demonstrated that GAM-based models using satellite derived data could be useful in checking and understanding deforestation trends in an ecotonic region.

In the results, the GAMs confirmed that deforestation is related to year and covariates, but also revealed that there are substantially differences in trends between seasons. For the Legal Maranhão region, most of the deforestation happened during the rainy season. The results indicate that this event holds across the 25km and 50km buffers. In essence, the best description of deforestation for this region includes high levels of precipitation, reaching a threshold that corresponded to 6.3% (25km), 4.4% (50km) and, 3.9% (100km) of the observations of the whole sample, low incidence of direct sunlight in hours, reaching a threshold that corresponded to 9.3% (25km), 7.8% (50km) and, 6.8% (100km) of the observations of the whole sample. During the dry season, several models were not able capture a trend for deforestation, indicating that no significant changes were picked up by the models in this season. One exception was seen with model lm50ed. The model could capture the following year after the establishment of the environmental policy surveillance monitoring in that region.

The Maranhão side GAM models behaved very differently during the dry season compared to the Legal Maranhão area. It was confirmed that there was a trend for deforestation during this season. In fact, the results indicate that the Maranhão region has a well-defined deforestation trend for both seasons. Notably, the deforestation process increased during the dry season from 2005 onwards. The characterisation and environment that shaped deforestation for that region and was constant were low temperatures and low availability of sunlight with no thresholds.

These results seem to suggest that deforestation in the Amazon region (LM) was diminished during the dry season while in the Cerrado region (MA) the clearing process increased significantly. It is not possible to conclude that there was a shift in this period for the regions. It is also not possible to conclude the same process during the rainy season because both regions were characterized by positive increments for deforestation. There appears to be no spillover effect from the environmental enforcement executed in the LM region to the MA region during the raining season. One plausible piece of evidence is that deforestation remained during both seasons with distinct cycles. An interaction with precipitation and sunlight shows the different paths between seasons for both regions (see A.2.50, A.2.51 and A.2.52).

With these results, three possible outcomes emerges: i) since the region is a transitional zone, the two areas don't differ in biota aspects. In a sense, anthropic actions were responsible for apparent changes in the deforestation trends and, these changes ii) happened during high levels of precipitation and low levels of solar incidence which in turn shows that, in general, cloud cover might be a benefit for clear-cutting practices, keeping in mind that iii) the artificial line divides the two regions but many of the political boundaries of municipalities remain in both sides of the region (MA and LM), which could interfere in the deforestation path between the seasons.

The first outcome is not supported entirely by the results of the models. In fact, deforestation is a human activity and the oscillation process is caused by individuals' choices. The findings, show, however that much of this activity happened differently between the two regions. As it was mentioned earlier in section 2.3.2, the LM region is under an environmental monitoring policy (DETER) that uses satellites images to detect deforestation in the tropical and transitional forest and punish those found at fault. In this sense, the results inform us that deforestation took place during the raining season, which suggest an explanation for the second outcome.

The existence of clouds for satellite images is an impediment to detect vegetation changes. The presence of high levels of precipitation and low levels of direct sunlight might indicate also the existence of clouds as natural barriers. The second outcome states that cloud cover might benefit deforestation because it inhibits forest cover change detection as seen from satellites. Possibly human activity was displaced from dry seasons with clear sky to rainy season with cloudy days. In other words, human behavior changed due to the environmental monitoring policy.

Finally, as an artificial line, no concrete boundaries exist in the studied area and many of the political boundaries of municipalities remain on both sides of the regions (MA and LM). In other words, municipalities and provinces are split by this artificial line, so that much of the deforestation process during the dry season was displaced from the Legal Maranhão (LM) to the Maranhão side (MA), since there was no political or economic deterrent to these anthropic actions. This can be inferred from the findings of the models that showed deforestation trends in the MA region increasing, especially after 2010.

2.6 Conclusions

In this study, a new approach was taken to study deforestation trends in areas of ecological tension. Generalized Additive Modelling was implemented in the Maranhão state of in Brazil to detect the path of deforestation in the transitional forest of Amazon and Cerrado biomes. The technique applied because it allowed non linear relationship with several response variables. Images from satellite images combined to climatology weather station dataset formed the database used in this study. It was created an algorithm to capture Vegetation Indices changes over time in order to create a proxy for deforestation as a response variable. Climatologic variables were converted and resampled to adjust to the response variable and were used as covariates of the model. An artificial line, called the Legal Amazon line was used to divide the analysis in terms of two regions, Legal Maranhão (LM) and Maranhão (MA), according to their differing treatment in deforestation protection.

Model validation and cross validation was were taken with the use of neural networks and artificial intelligence (AI). It was found that models with EVI values as the response variable were a better fit for the deforestation trends, confirming the assumptions made by Bayma and Sano (2015); Didan et al. (2015); Ratana et al. (2005) in that ecotone forests respond better to EVI than NDVI values. Graphical results of the GAMs not only revealed the trend or turning points of regression, but also showed the possible limits within which the optimum forest loss of each Vegetation Index value could occur.

In the results, the GAMs confirmed that deforestation is related to year and covariates, but also revealed that there are substantially difference of trends between seasons and regions. For the Legal Maranhão region, most of the deforestation happened during the rainy season. In terms of the Maranhão side the GAM models behaved very differently during the dry season, compared to the Legal Maranhão area, in that there was a trend for deforestation during that season. In fact, the results indicate that the Maranhão region has a well-defined deforestation trend for both seasons. In particular, the deforestation process increased during the dry season from 2005 on wards. These findings were further validated by showing that deforestation happened during both seasons for settlements which were not target by the environmental policy.

In general, the models employed here arguably served the purpose of the study well, but this does not rule out exploring other tools in the future, such as applying an analysis including cloud cover as an indicator of human changing behavior. In addition to the aforementioned analysis, a natural experiment with regions completely isolated from the artificial line could be helpful to examine how and why much of the deforestation process during the dry season was displaced from the Legal Maranhão (LM) to the Maranhão side (MA), since there was no political or economic deterrent to these anthropic actions.

There are of course a number of limitations to the analysis undertaken here. Following Murase et al. (2009) approach, possible errors in modeling could be taken into account. First of all, in terms of predicting the trend of deforestation based on a list of variables, the model implicitly assumes that the predicted range or potential space is fully occupied by forest, which in reality might not be true. Additionally, the spatial distribution of the vegetation indices may exhibit dynamic behavior over time, so that a potential area may or may not be sparsely vegetated for a certain period (e.g., during sampling) due to progressive succession of forest. Or a temporary absence could be due to natural causes, such as, attack of pests or diseases or inter-species competition. Secondly, the regional environmental conditions follow changing trends of different duration, so it is possible that in certain cases an observed value may be declining due to regional changes rather local changes, but the prediction model does not detect this dynamic behavior. Additionally, the study was based on coarse image resolution which could neglect local changes in the sample area. The results could also feasibly suffer from overfiting since more data is needed to optimize the smoothing algorithms. Finally, our results may not be generalizable to other areas, such as dense tropical forest and open fields.

Appendix A.2

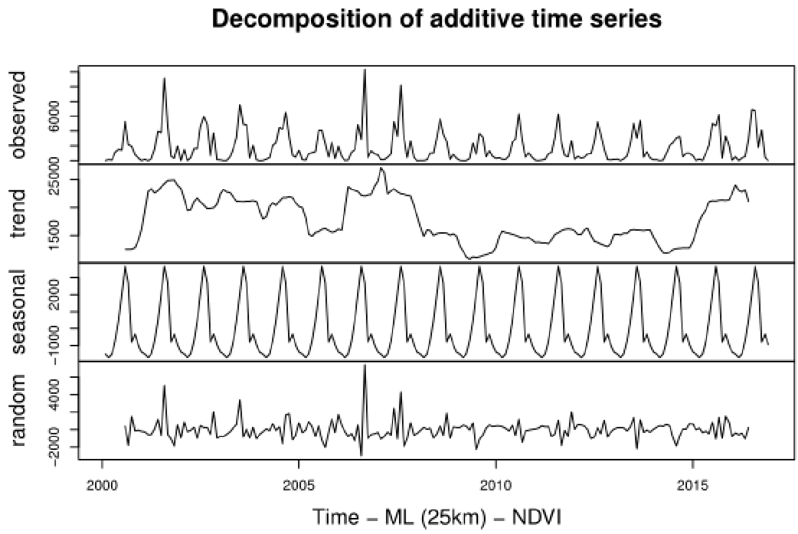

Figure A.2.1 Moving Averages Additive Model. The method is the moving averages additive model. In this method, there are the trend and the seasonal parameters defined separately. The ML analysis is conducted taking into consideration the NDVI values for band 25km. It is possible to check that there is no smoothness in the trend.

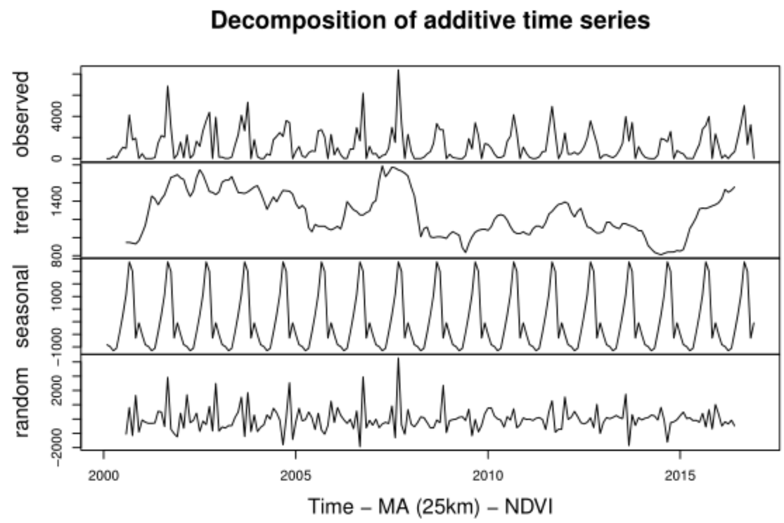

Figure A.2.2 Moving Averages Additive Model. The method is the moving averages additive model. In this method, there are the trend and the seasonal parameters defined separately. The MA analysis is conducted taking into consideration the NDVI values for band 25km. It is possible to check that there is no smoothness in the trend.

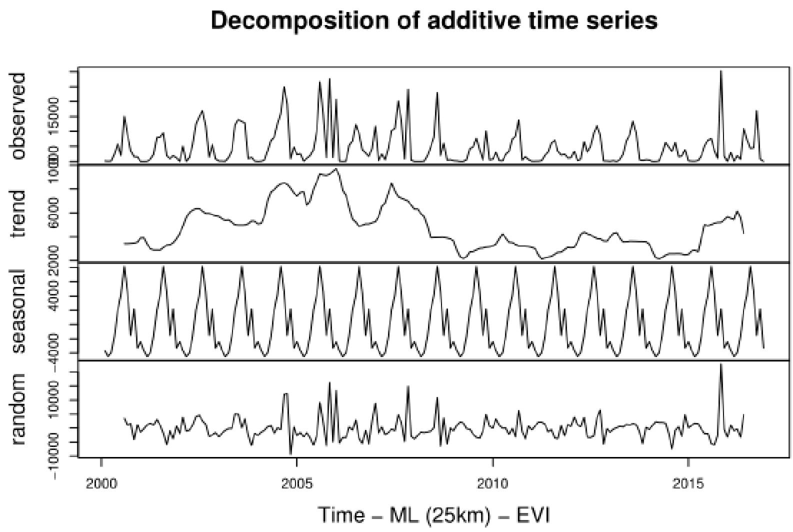

Figure A.2.3 Moving Averages Additive Model. The method is the moving averages additive model. In this method, there are the trend and the seasonal parameters defined separately. The ML analysis is conducted taking into consideration the EVI values for band 25km. It is possible to check that there is no smoothness in the trend.

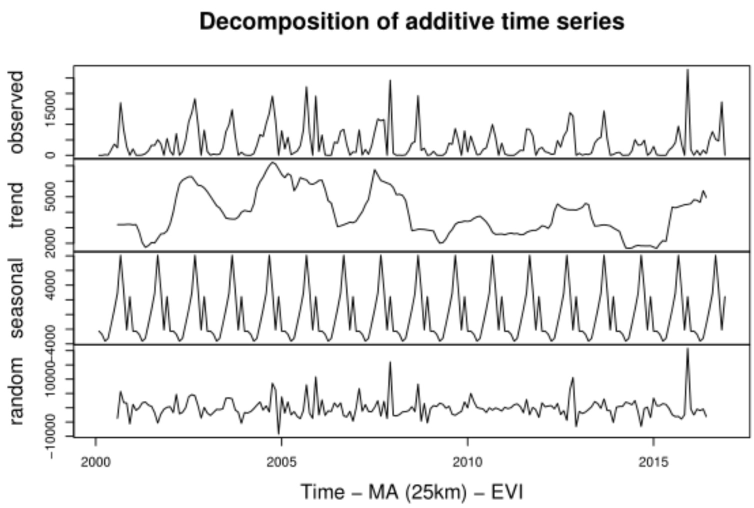

Figure A.2.4 Moving Averages Additive Model. The method is the moving averages additive model. In this method, there are the trend and the seasonal parameters defined separately. The MA analysis is conducted taking into consideration the EVI values for band 25km. It is possible to check that there is no smoothness in the trend.

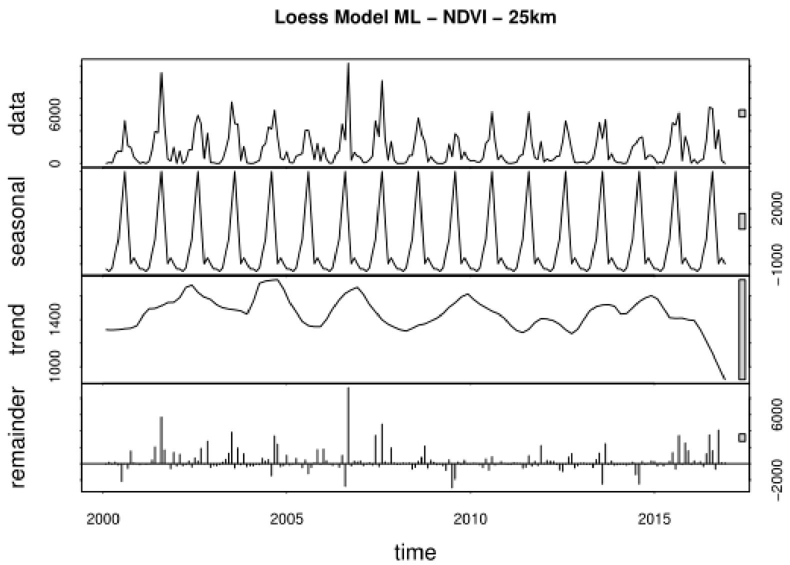

Figure A.2.5 Loess Model. The loess model is a filtering procedure for decomposing a time series into trend, seasonal, and remainder components. The ML analysis is conducted taking into consideration the NDVI values for band 25km. It is possible to check that there is no smoothness in the trend.

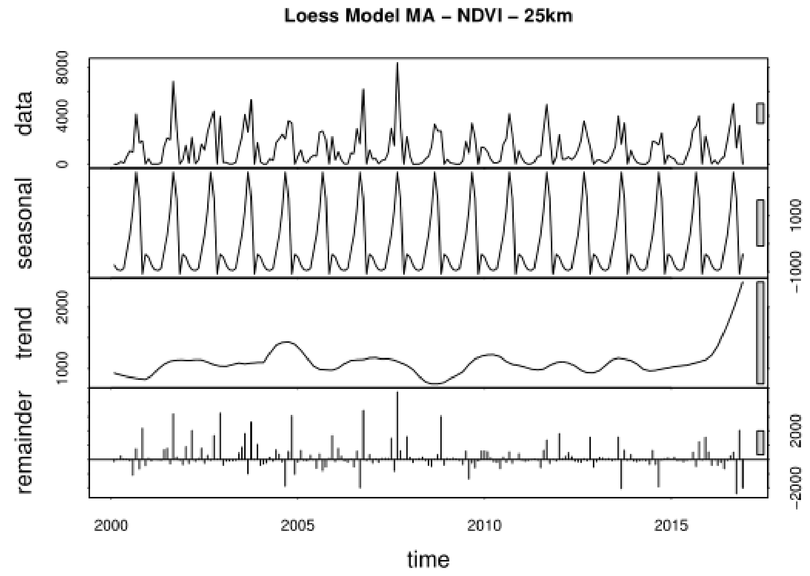

Figure A.2.6 Loess Model. The loess model is a filtering procedure for decomposing a time series into trend, seasonal, and remainder components. The MA analysis is conducted taking into consideration the NDVI values for band 25km. It is possible to check that there is no smoothness in the trend.

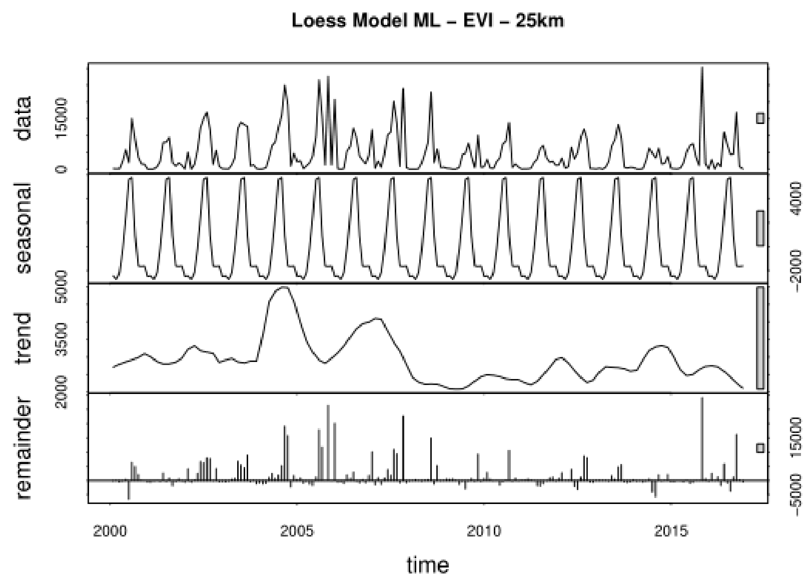

Figure A.2.7 Loess Model. The loess model is a filtering procedure for decomposing a time series into trend, seasonal, and remainder components. The ML analysis is conducted taking into consideration the EVI values for band 25km. It is possible to check that there is no smoothness in the trend.

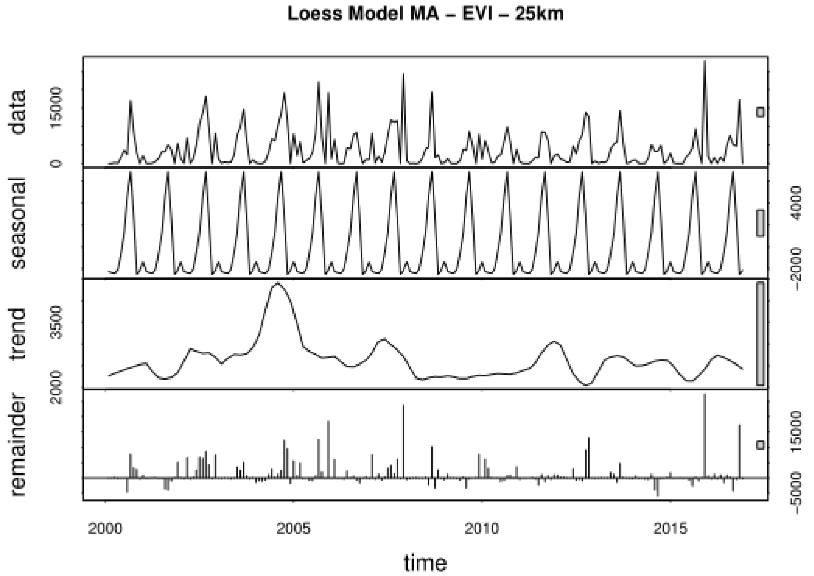

Figure A.2.8 Loess Model. The loess model is a filtering procedure for decomposing a time series into trend, seasonal, and remainder components. The MA analysis is conducted taking into consideration the EVI values for band 25km. It is possible to check that there is no smoothness in the trend.

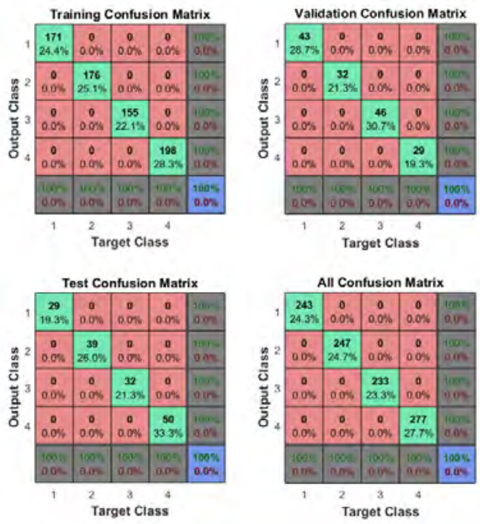

Figure A.2.9 Confusion Matrix Results for 2002 in Cerrado Maranhão. Accuracy measure using Hansen et al. (2013) dataset. Final confusion matrix is at the bottom right. The rows correspond to the predicted class and the columns correspond to the true class. The diagonal cells correspond to observations that are correctly classified. Both the number of observations and the percentage of the total number of observations are shown in each cell. The column on the far right of the plot shows the percentages of all the examples predicted to belong to each class that are correctly and incorrectly classified. The row at the bottom of the plot shows the percentages of all the examples belonging to each class that are correctly and incorrectly classified. The cell in the bottom right of the plot shows the overall accuracy (MATLAB, 2017).

Figure A.2.10 Confusion Matrix Results for 2002 in Legal Maranhão. Accuracy measure using Hansen et al. (2013) dataset. Final confusion matrix is at the bottom right. The rows correspond to the predicted class and the columns correspond to the true class. The diagonal cells correspond to observations that are correctly classified. Both the number of observations and the percentage of the total number of observations are shown in each cell. The column on the far right of the plot shows the percentages of all the examples predicted to belong to each class that are correctly and incorrectly classified. The row at the bottom of the plot shows the percentages of all the examples belonging to each class that are correctly and incorrectly classified. The cell in the bottom right of the plot shows the overall accuracy (MATLAB, 2017).

Figure A.2.11 Confusion Matrix Results for 2012 in Cerrado Maranhão. Accuracy measure using Hansen et al. (2013) dataset. Final confusion matrix is at the bottom right. The rows correspond to the predicted class and the columns correspond to the true class. The diagonal cells correspond to observations that are correctly classified. Both the number of observations and the percentage of the total number of observations are shown in each cell. The column on the far right of the plot shows the percentages of all the examples predicted to belong to each class that are correctly and incorrectly classified. The row at the bottom of the plot shows the percentages of all the examples belonging to each class that are correctly and incorrectly classified. The cell in the bottom right of the plot shows the overall accuracy (MATLAB, 2017).

Figure A.2.12 Confusion Matrix Results for 2012 in Legal Maranhão. Accuracy measure using Hansen et al. (2013) dataset. Final confusion matrix is at the bottom right. The rows correspond to the predicted class and the columns correspond to the true class. The diagonal cells correspond to observations that are correctly classified. Both the number of observations and the percentage of the total number of observations are shown in each cell. The column on the far right of the plot shows the percentages of all the examples predicted to belong to each class that are correctly and incorrectly classified. The row at the bottom of the plot shows the percentages of all the examples belonging to each class that are correctly and incorrectly classified. The cell in the bottom right of the plot shows the overall accuracy (MATLAB, 2017).

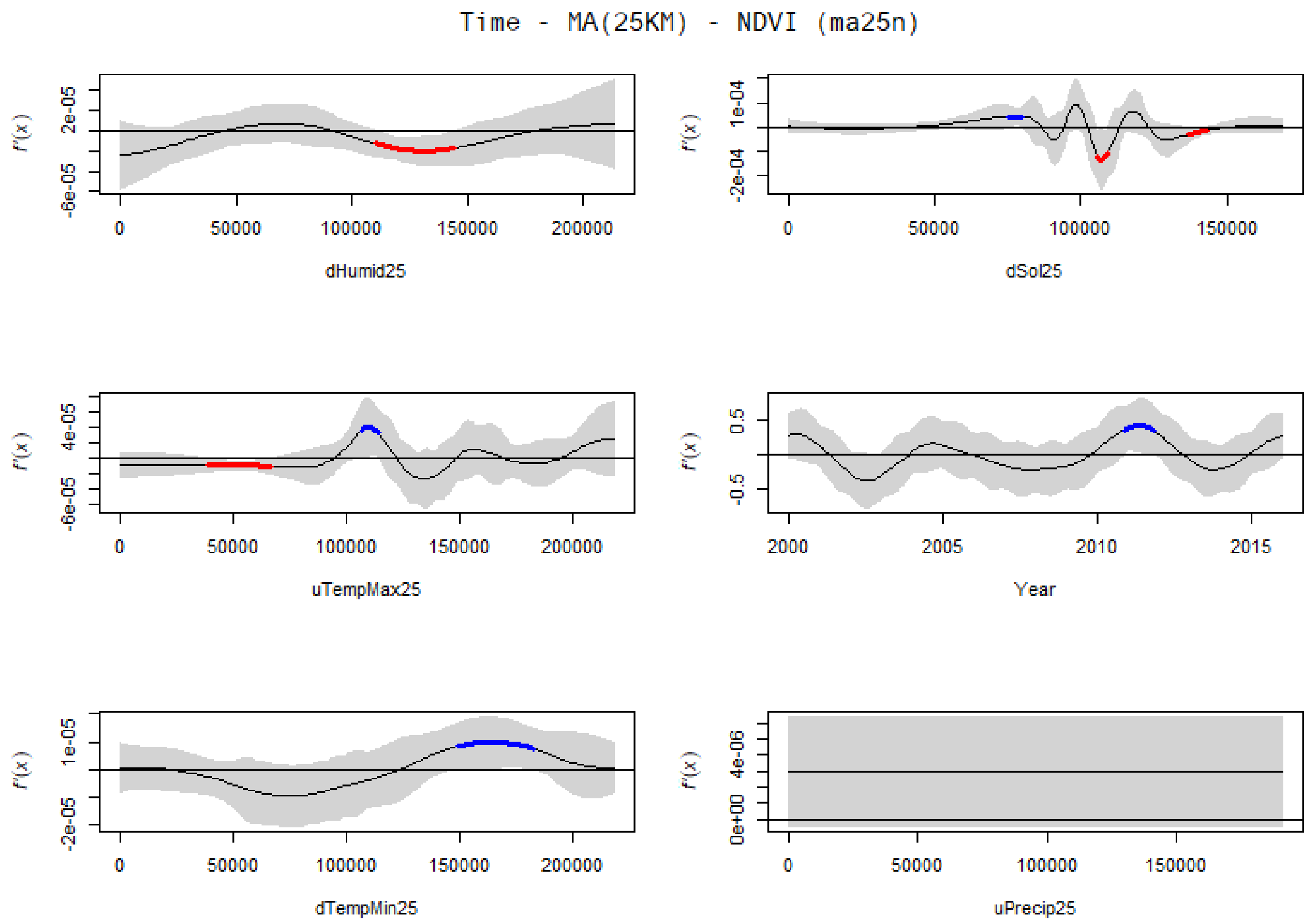

Figure A.2.13 Model ma25n. Model Maranhão Cerrado for 25km bandwidth using NDVI values. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

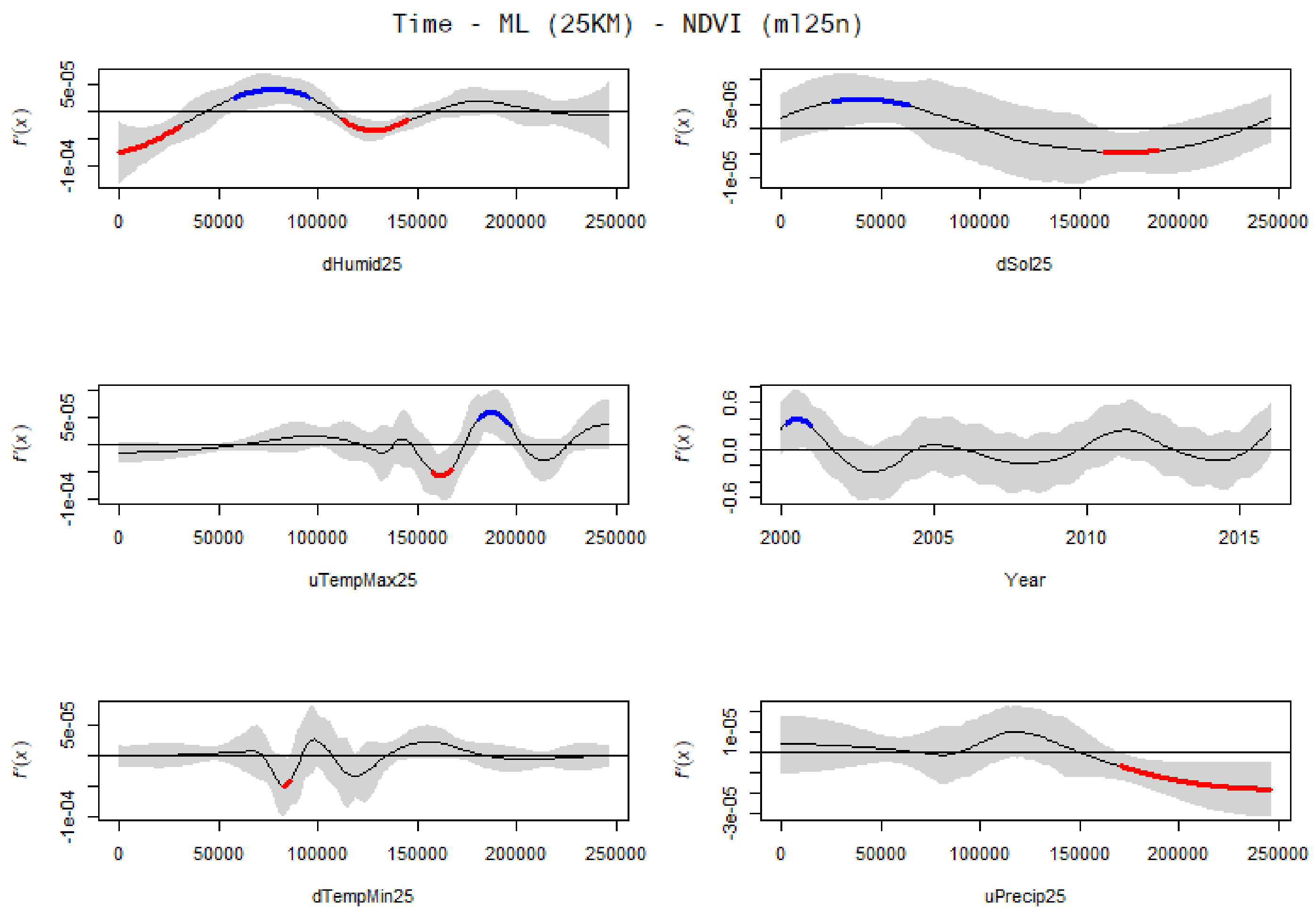

Figure A.2.14 Model lm25n. Model Legal Maranhão for 25km bandwidth using NDVI values. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

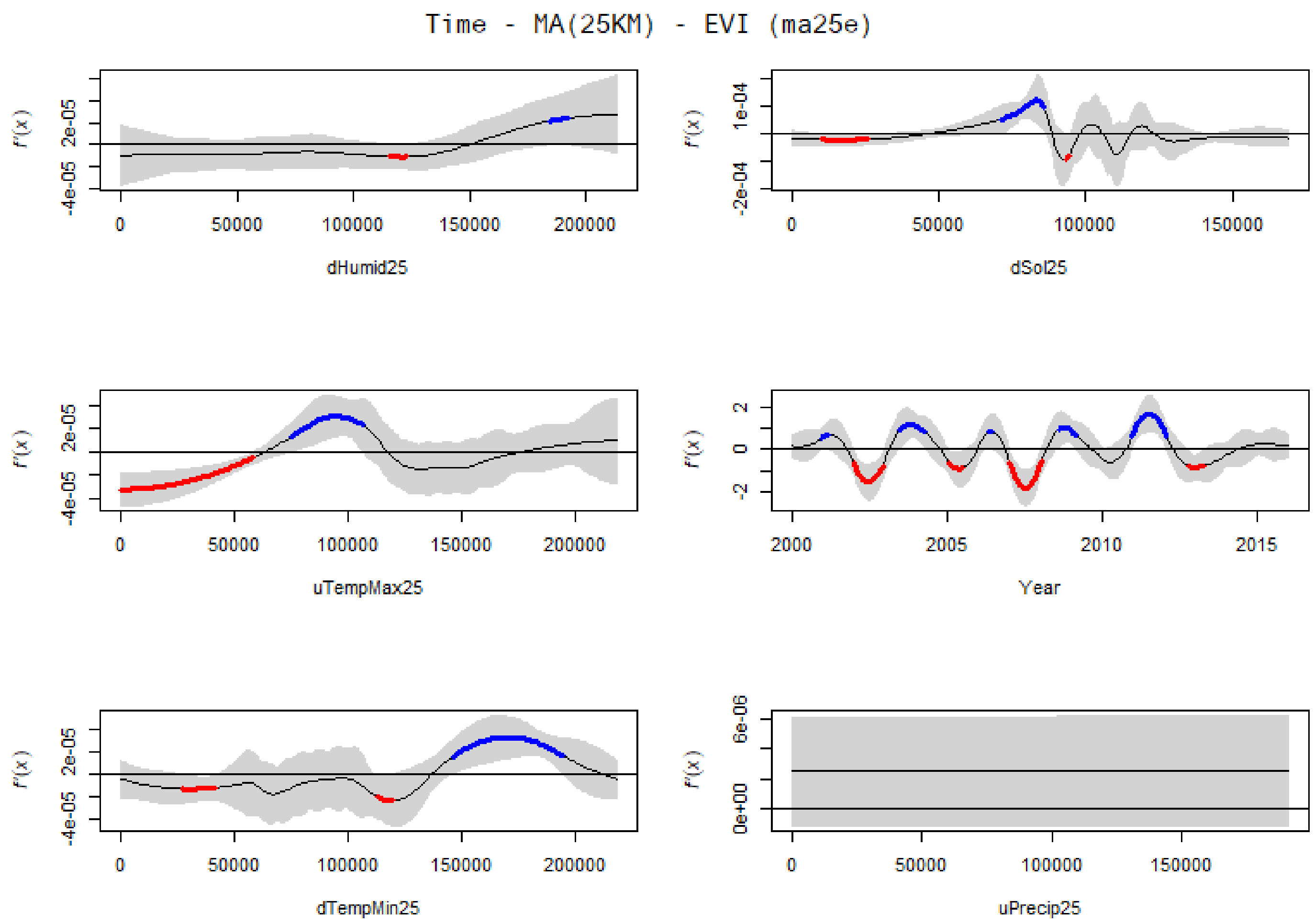

Figure A.2.15 Model ma25e. Model Maranhão Cerrado for 25km bandwidth using EVI values. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

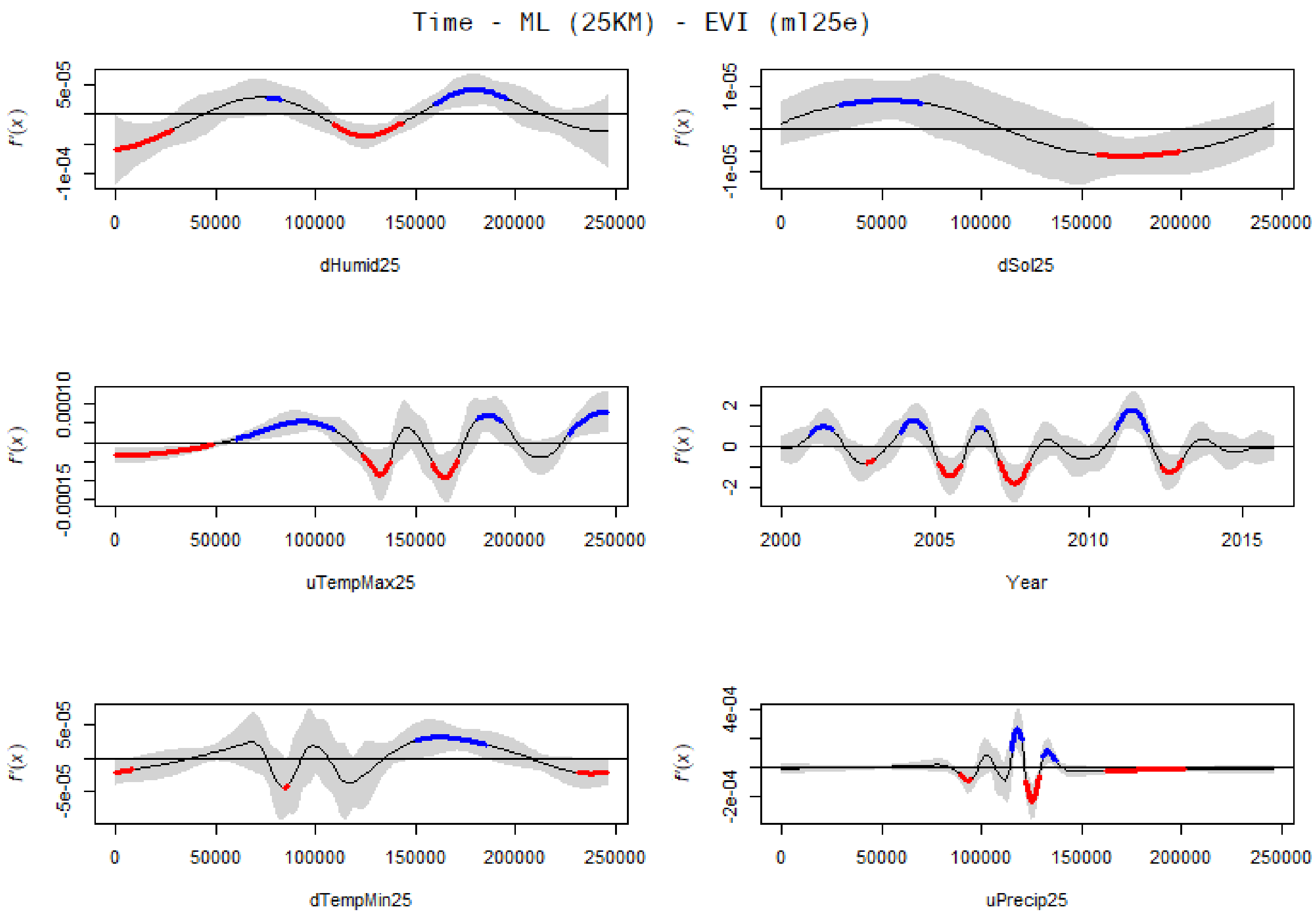

Figure A.2.16 Model lm25e. Model Legal Maranhão for 25km bandwidth using EVI values. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

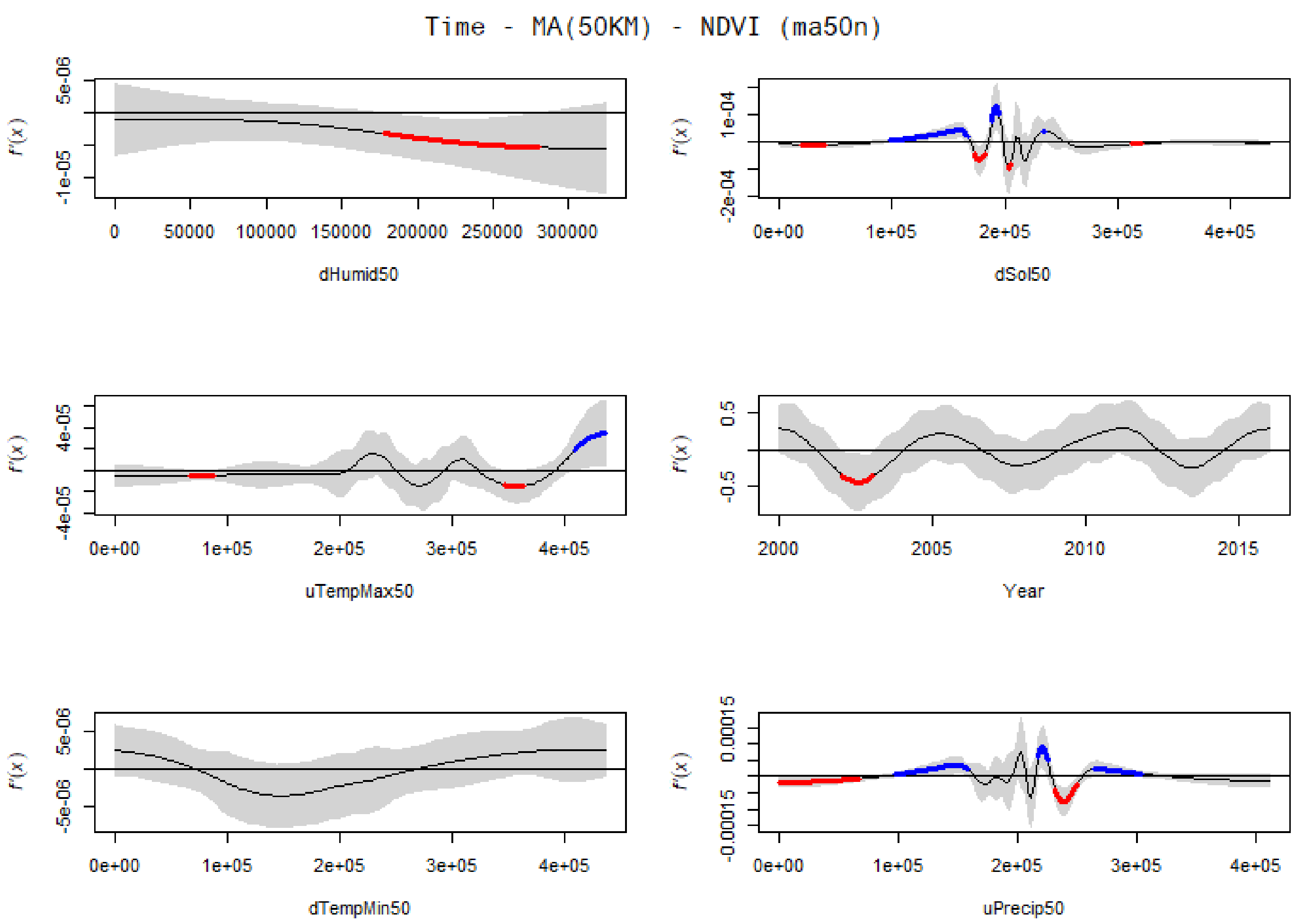

Figure A.2.17 Model ma50n. Model Maranhão Cerrado for 50km bandwidth using NDVI values. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

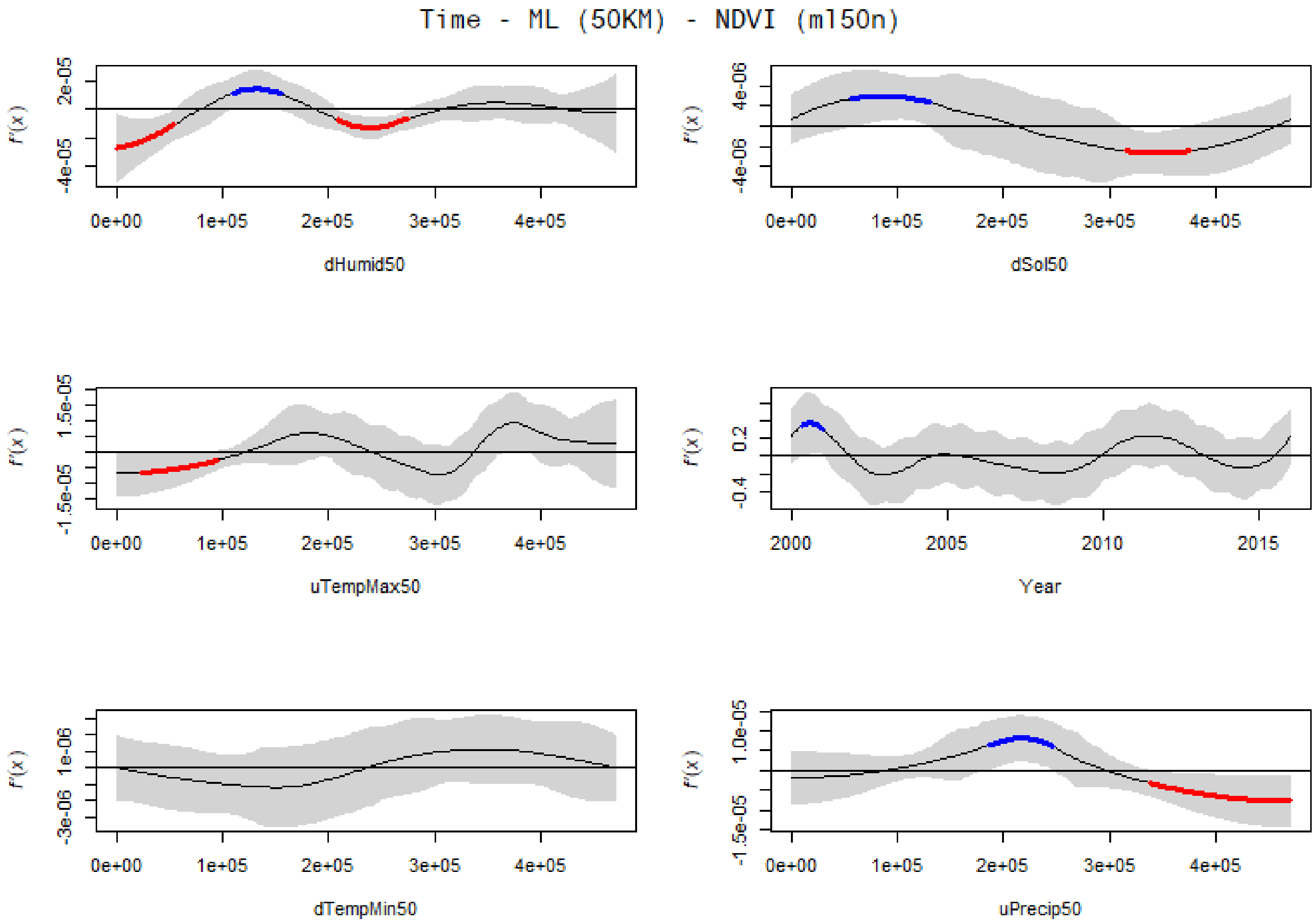

Figure A.2.18 Model lm50n. Model Legal Maranhão for 50km bandwidth using NDVI values. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

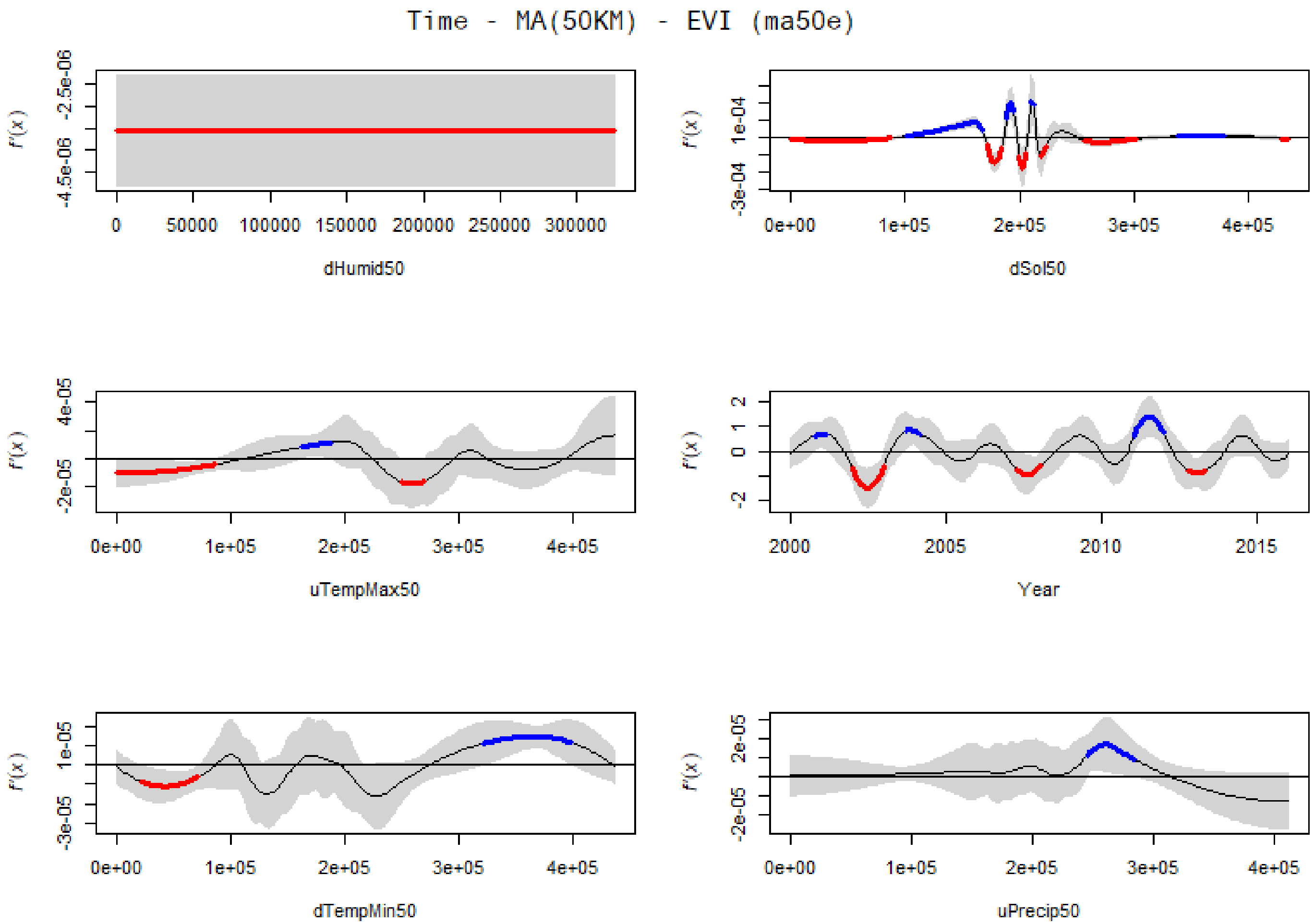

Figure A.2.19 Model ma50e. Model Maranhão Cerrado for 50km bandwidth using EVI values. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

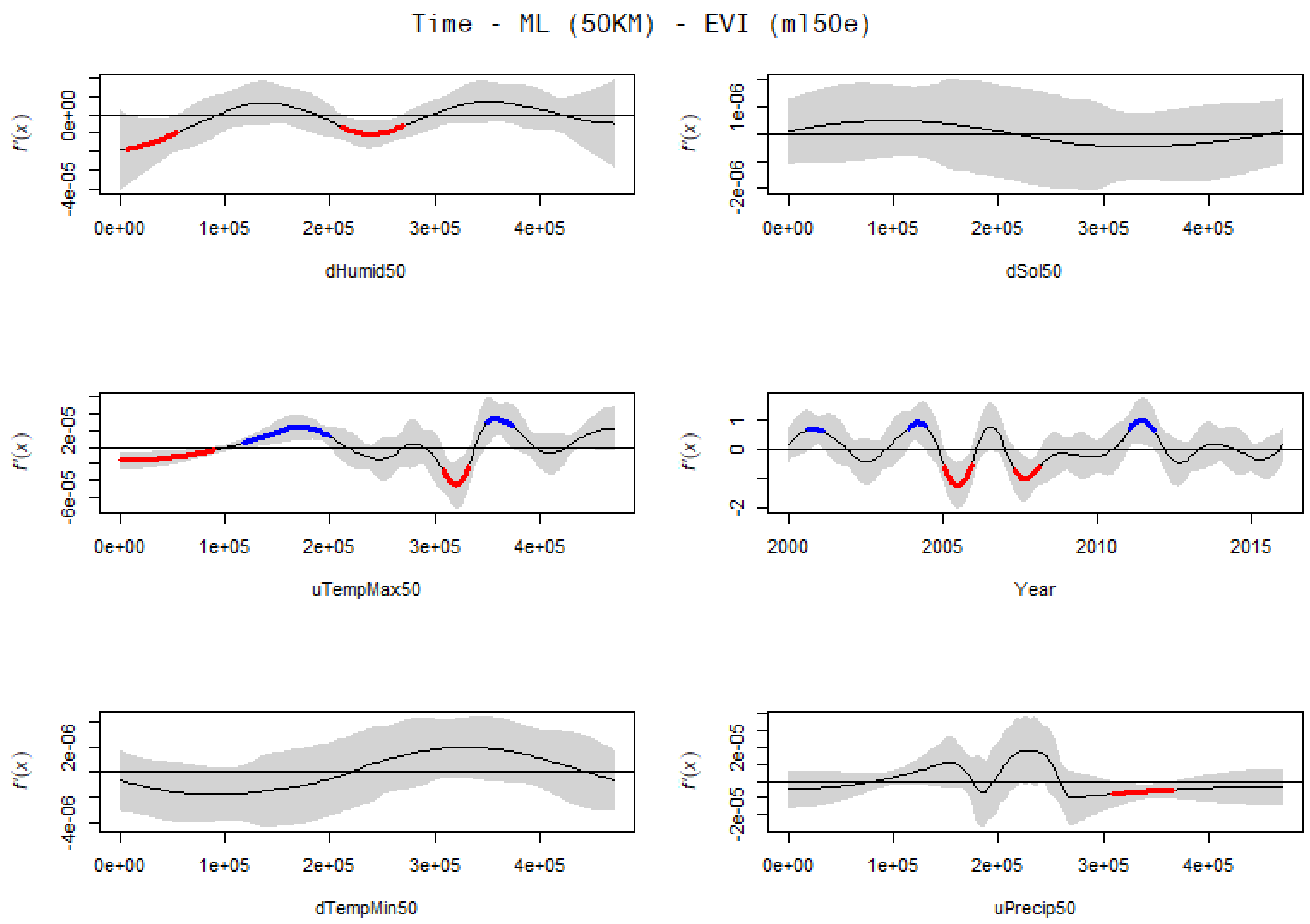

Figure A.2.20 Model lm50e. Model Legal Maranhão for 50km bandwidth using EVI values. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

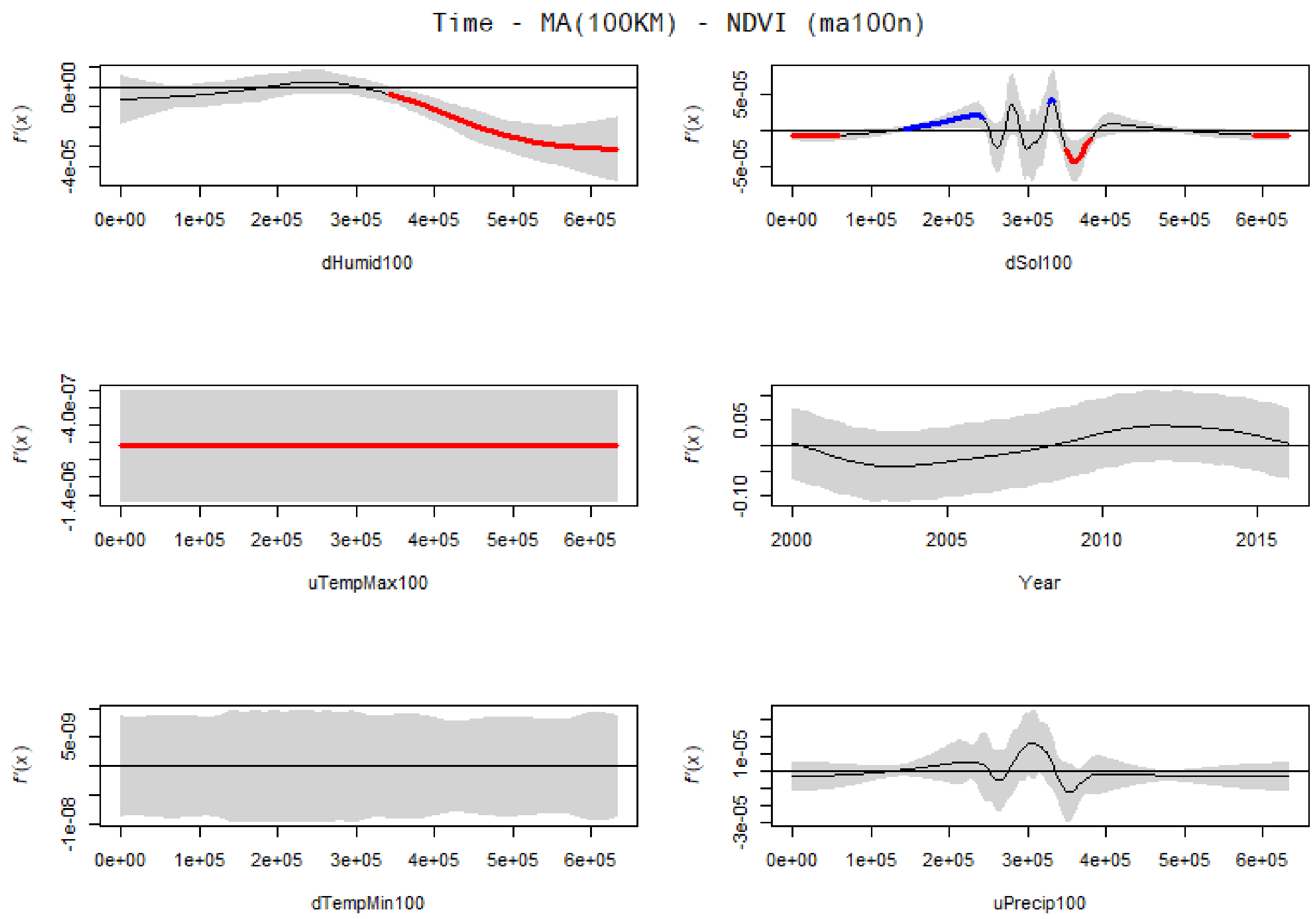

Figure A.2.21 Model ma100n. Model Maranhão Cerrado for 100km bandwidth using NDVI values. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

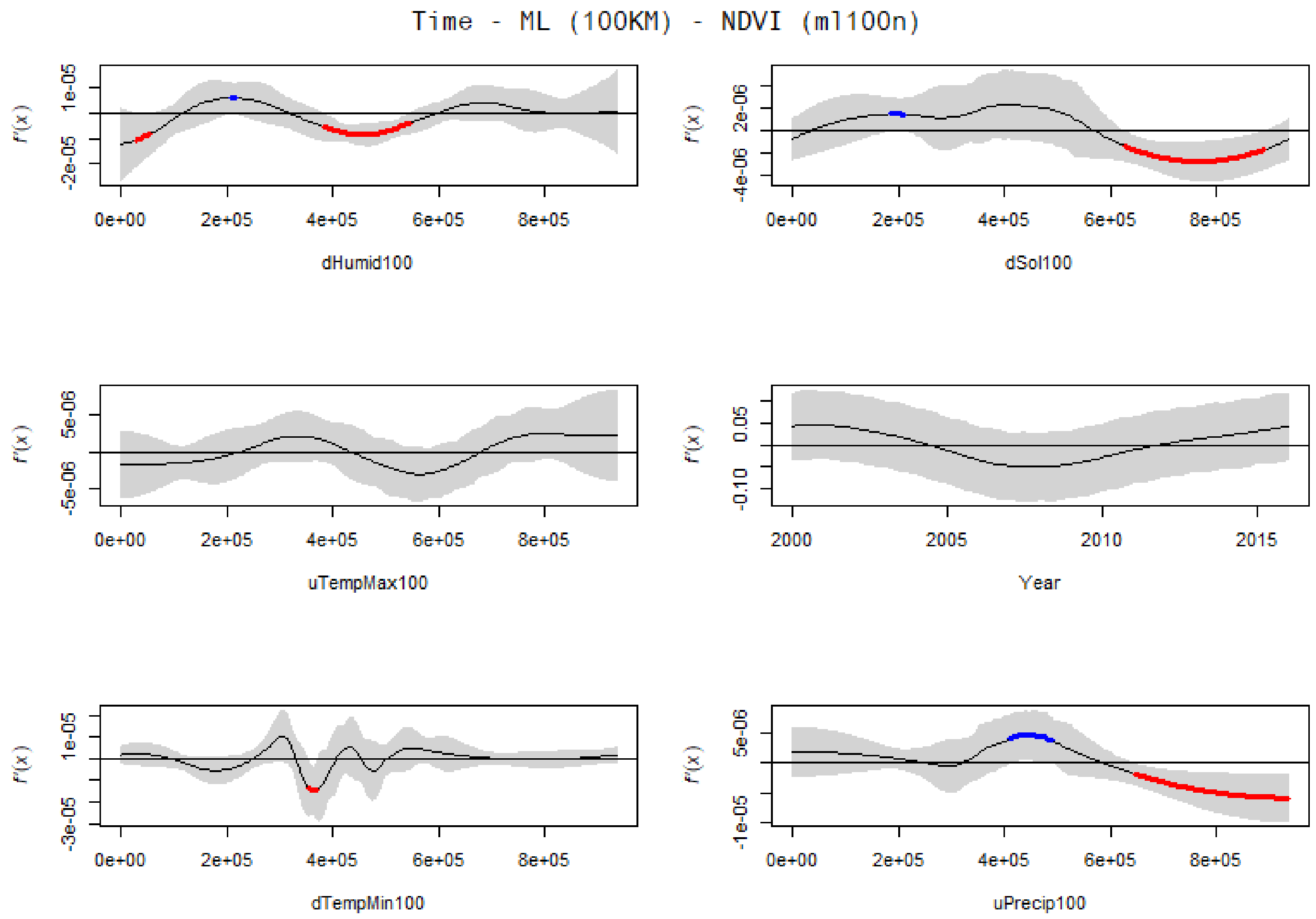

Figure A.2.22 Model lm100n. Model Legal Maranhão for 100km bandwidth using NDVI values.First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

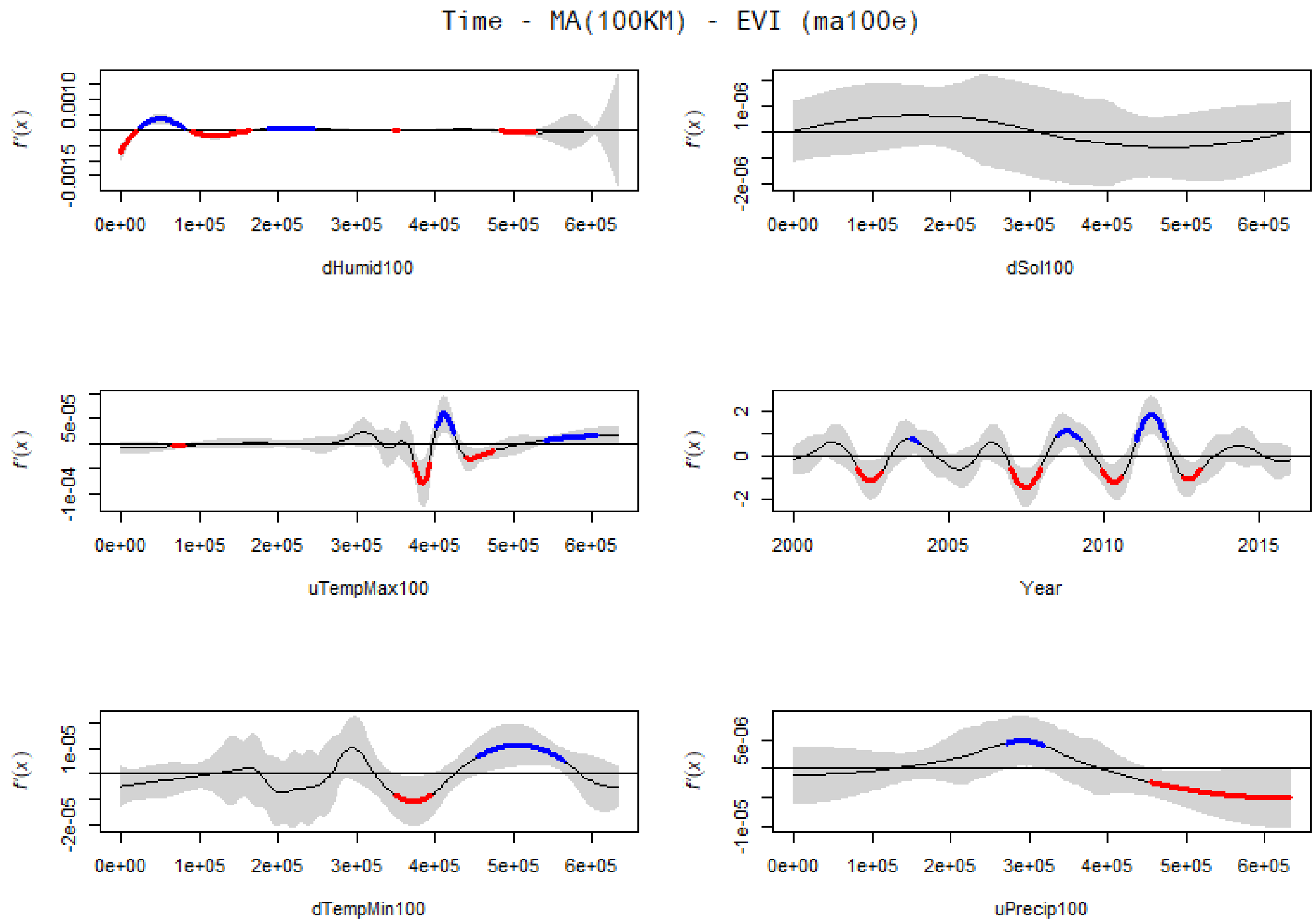

Figure A.2.23 Model ma100e. Model Maranhão Cerrado for 100km bandwidth using EVI values. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

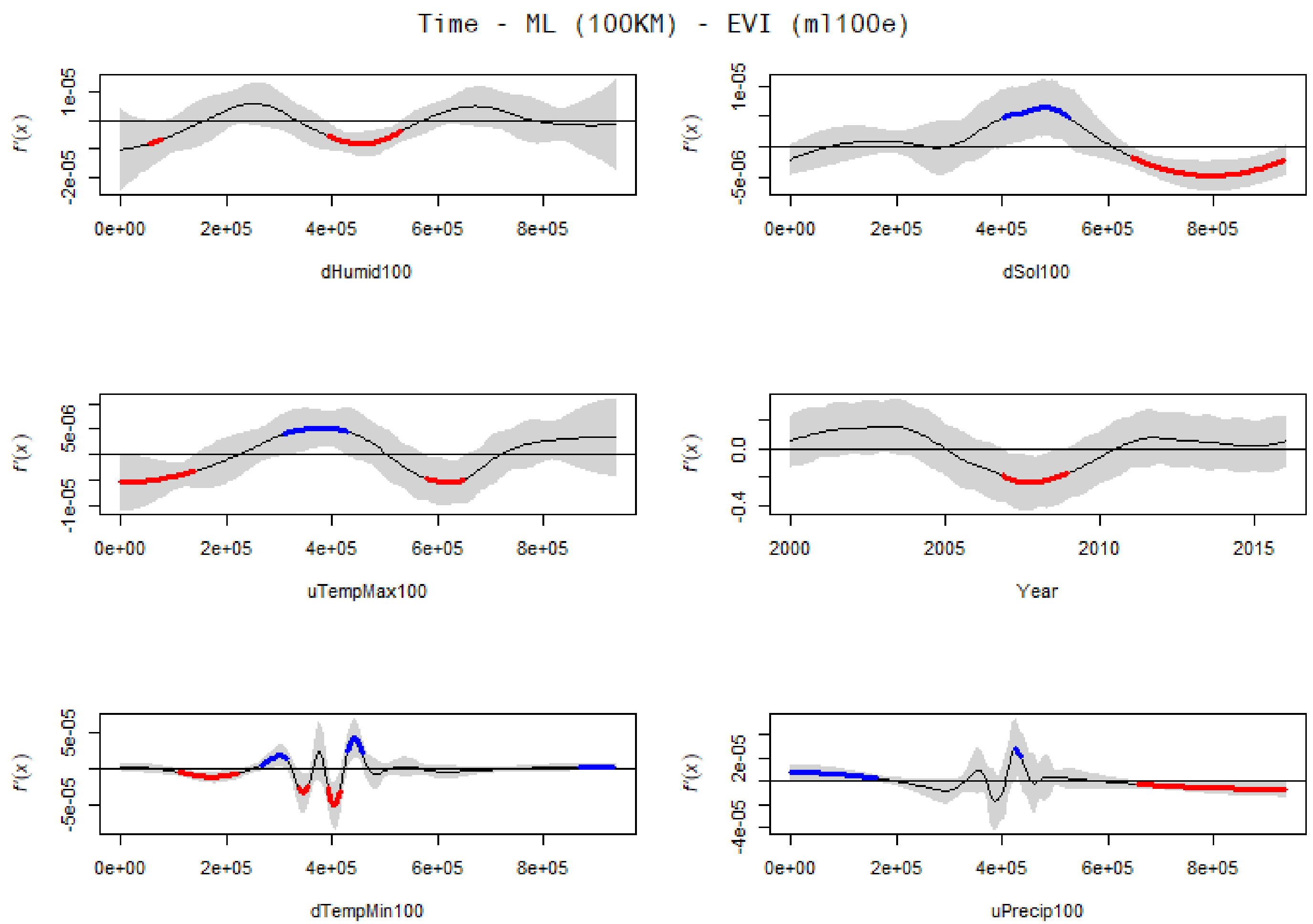

Figure A.2.24 Model lm100e. Model Legal Maranhão for 100km bandwidth using EVI values. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

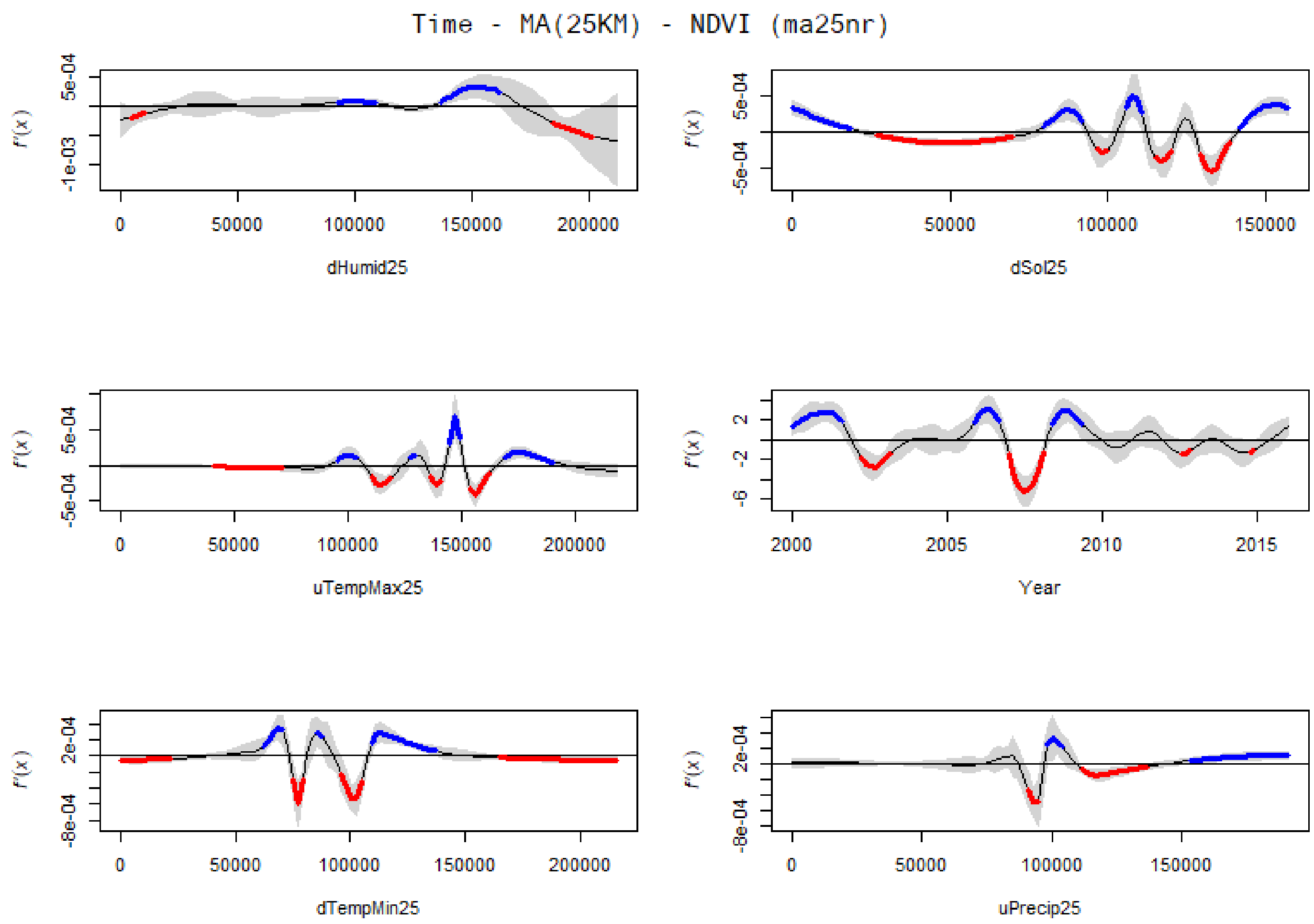

Figure A.2.25 Model ma25nr. Model Maranhão Cerrado for 25km bandwidth using NDVI values for raining season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

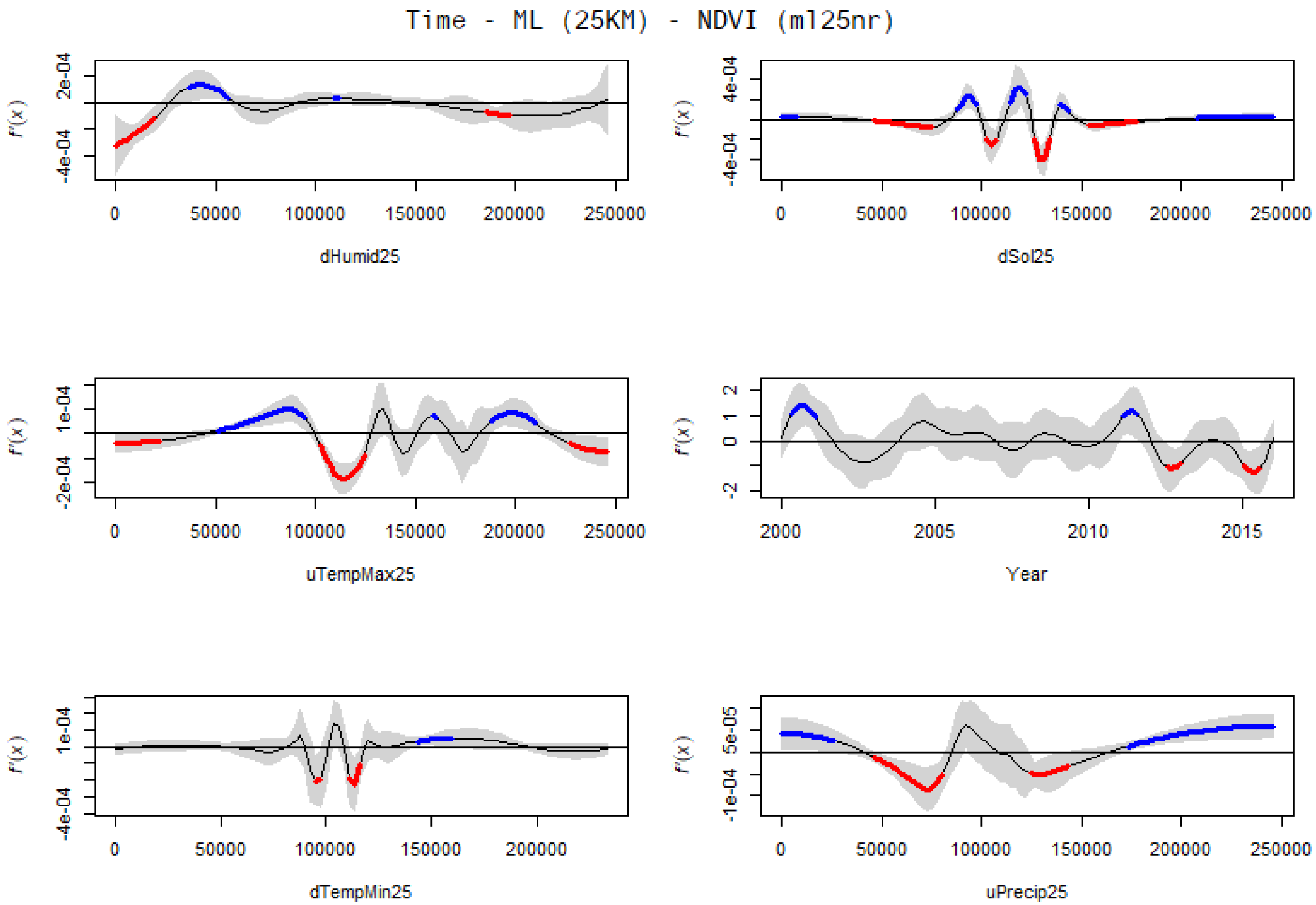

Figure A.2.26 Model lm25nr. Model Legal Maranhão for 25km bandwidth using NDVI values for raining season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

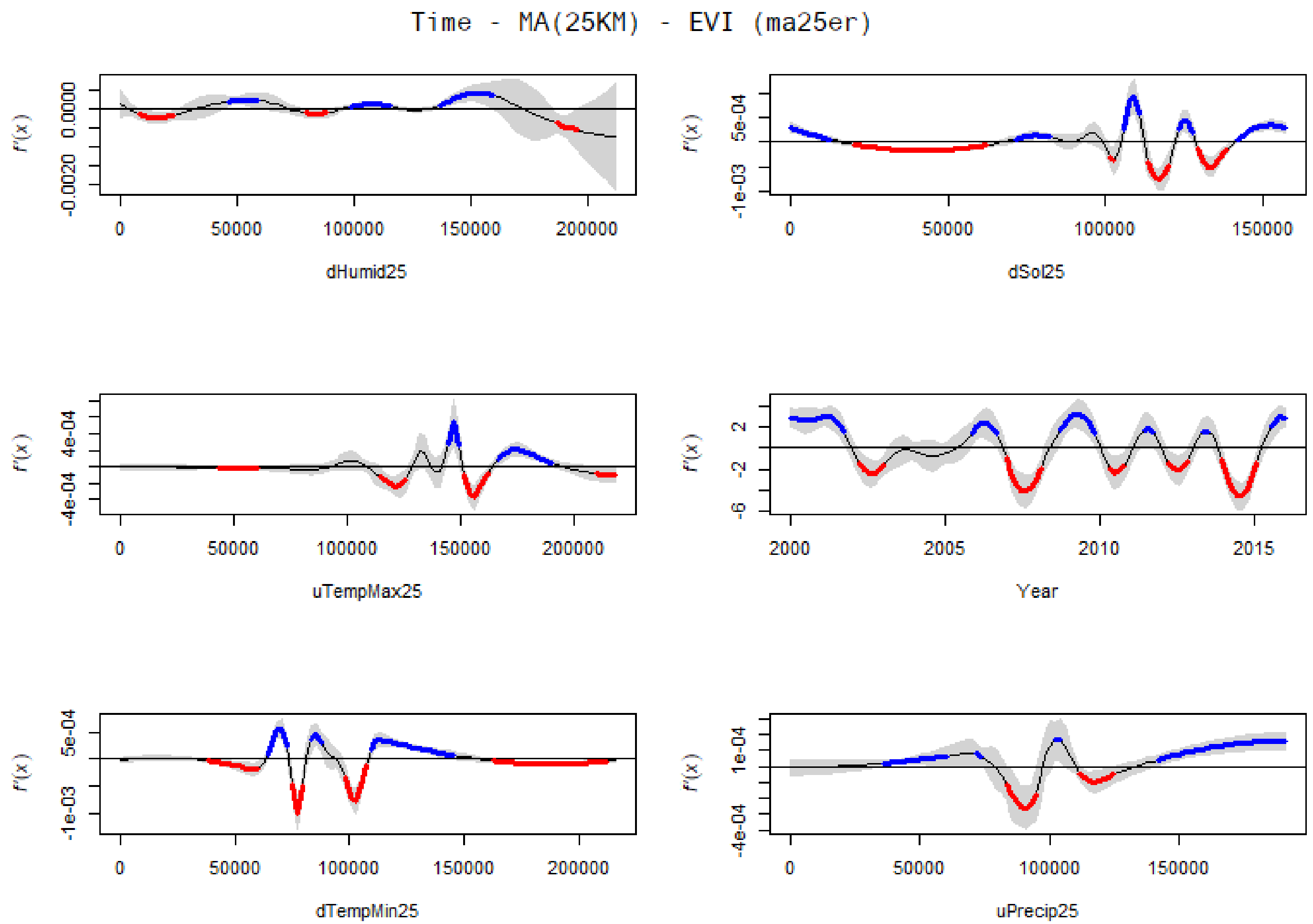

Figure A.2.27 Model ma25er. Model Maranhão Cerrado for 25km bandwidth using EVI values for raining season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

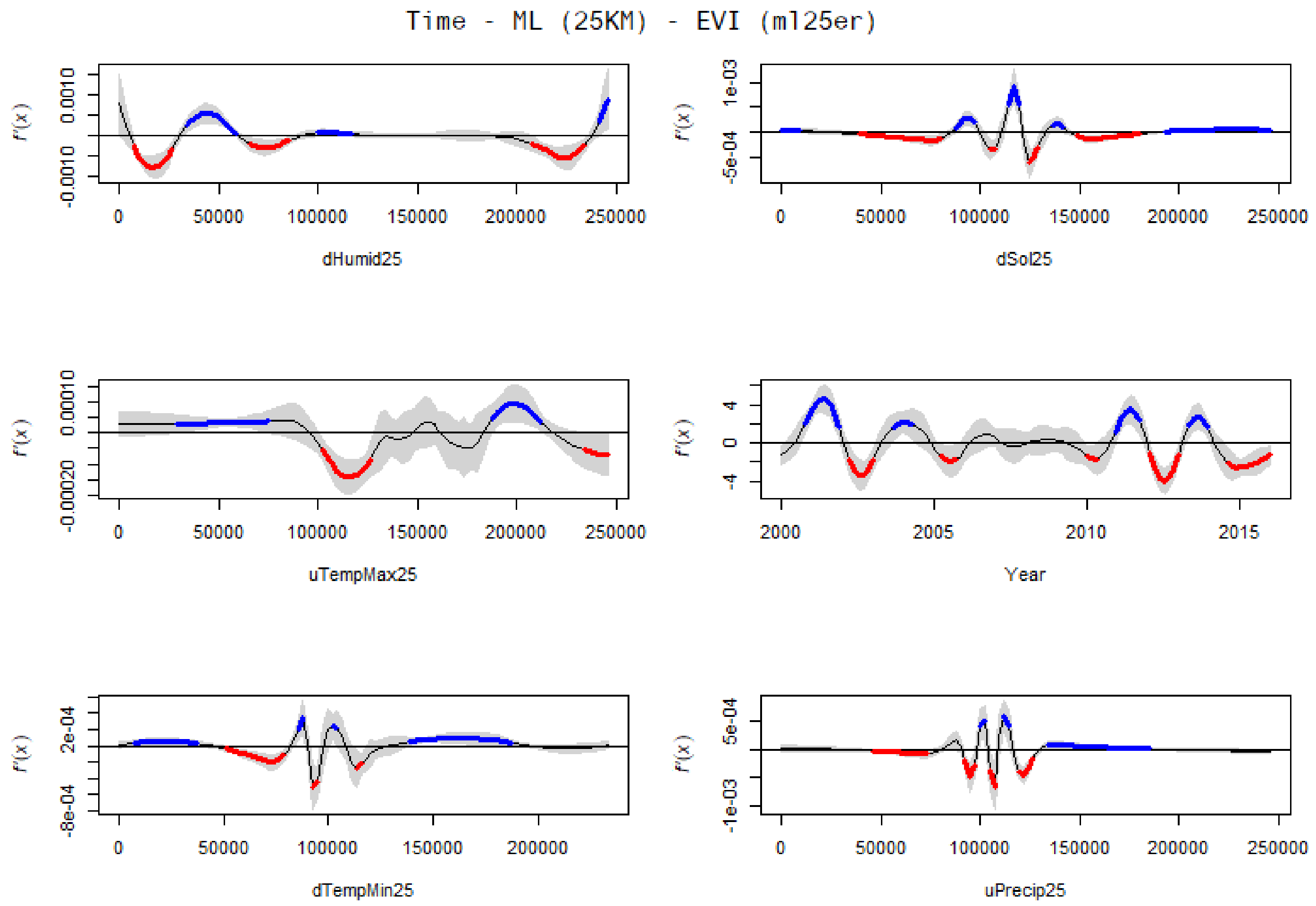

Figure A.2.28 Model lm25er. Model Legal Maranhão for 25km bandwidth using EVI values for raining season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

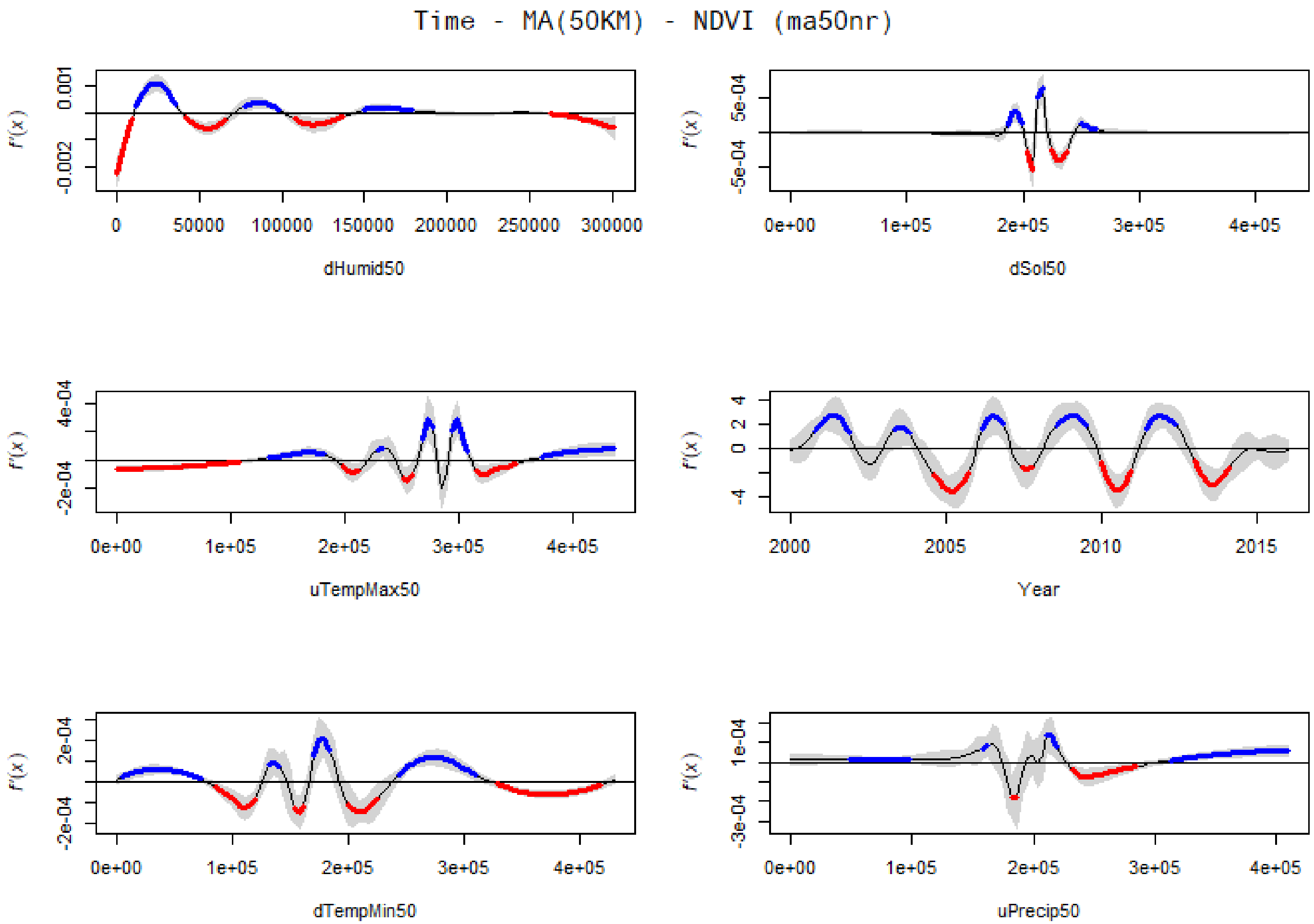

Figure A.2.29 Model ma50nr. Model Maranhão Cerrado for 50km bandwidth using NDVI values for raining season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

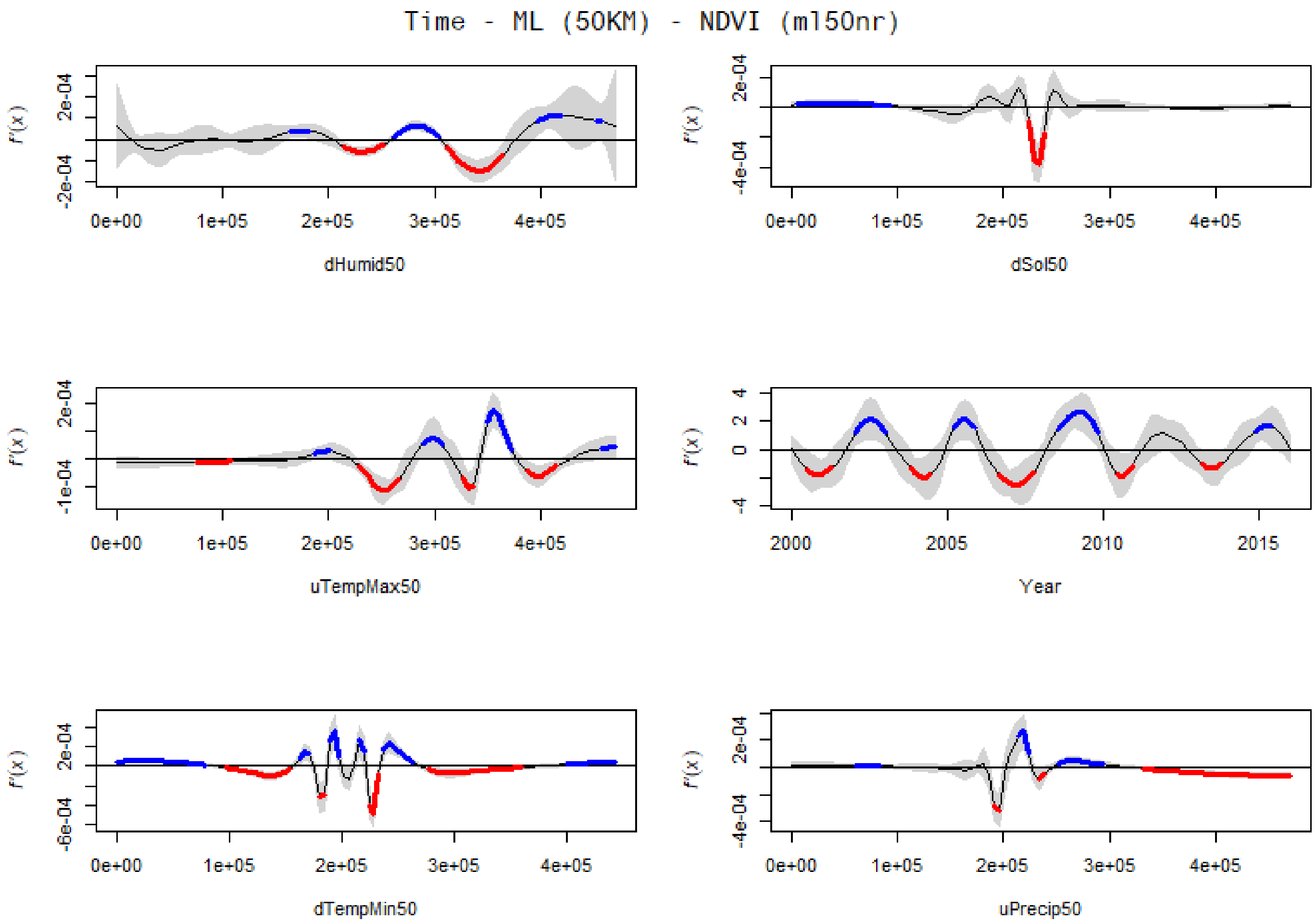

Figure A.2.30 Model lm50nr. Model Legal Maranhão for 50km bandwidth using NDVI values for raining season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

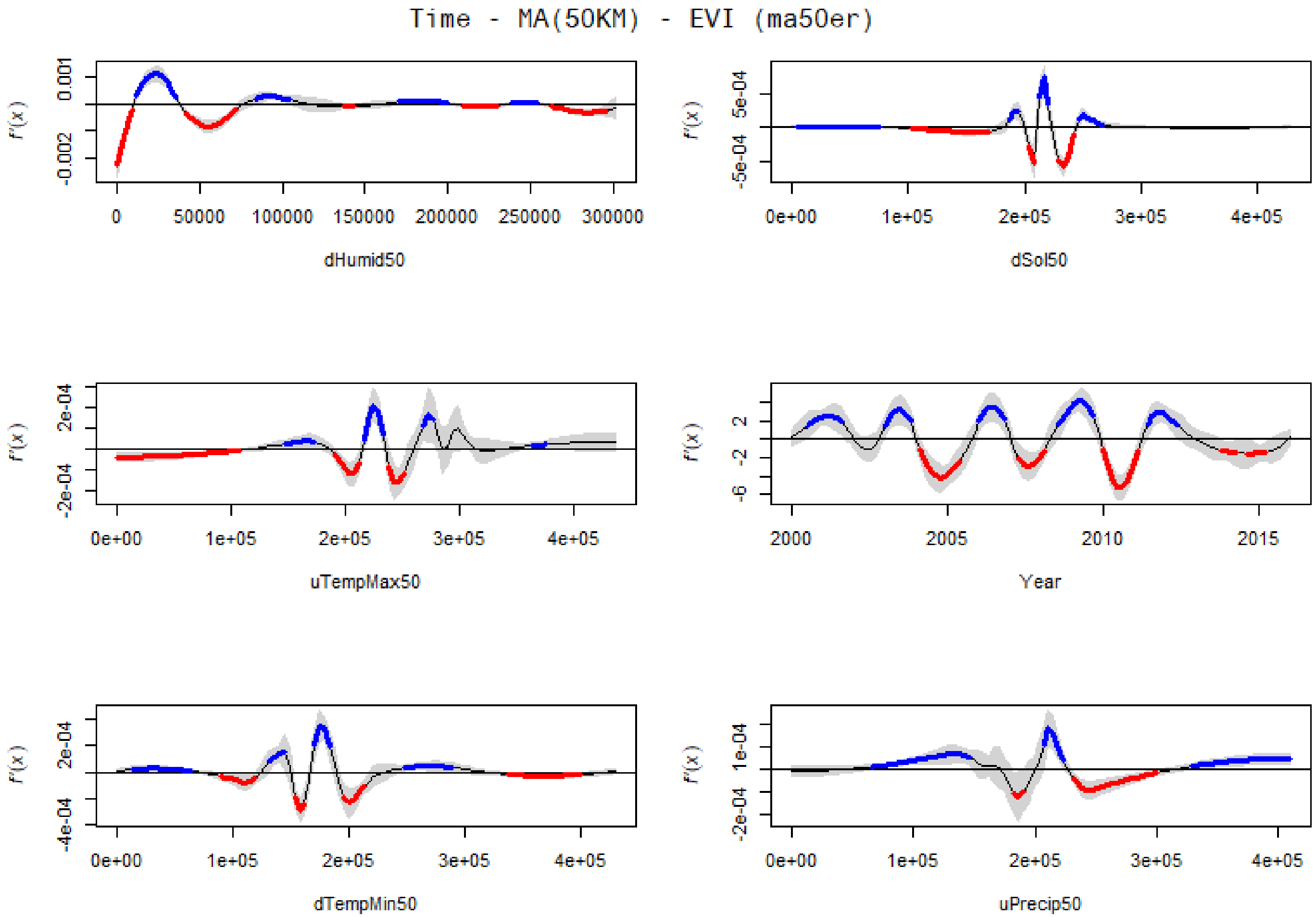

Figure A.2.31 Model ma50er. Model Maranhão Cerrado for 50km bandwidth using EVI values for raining season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

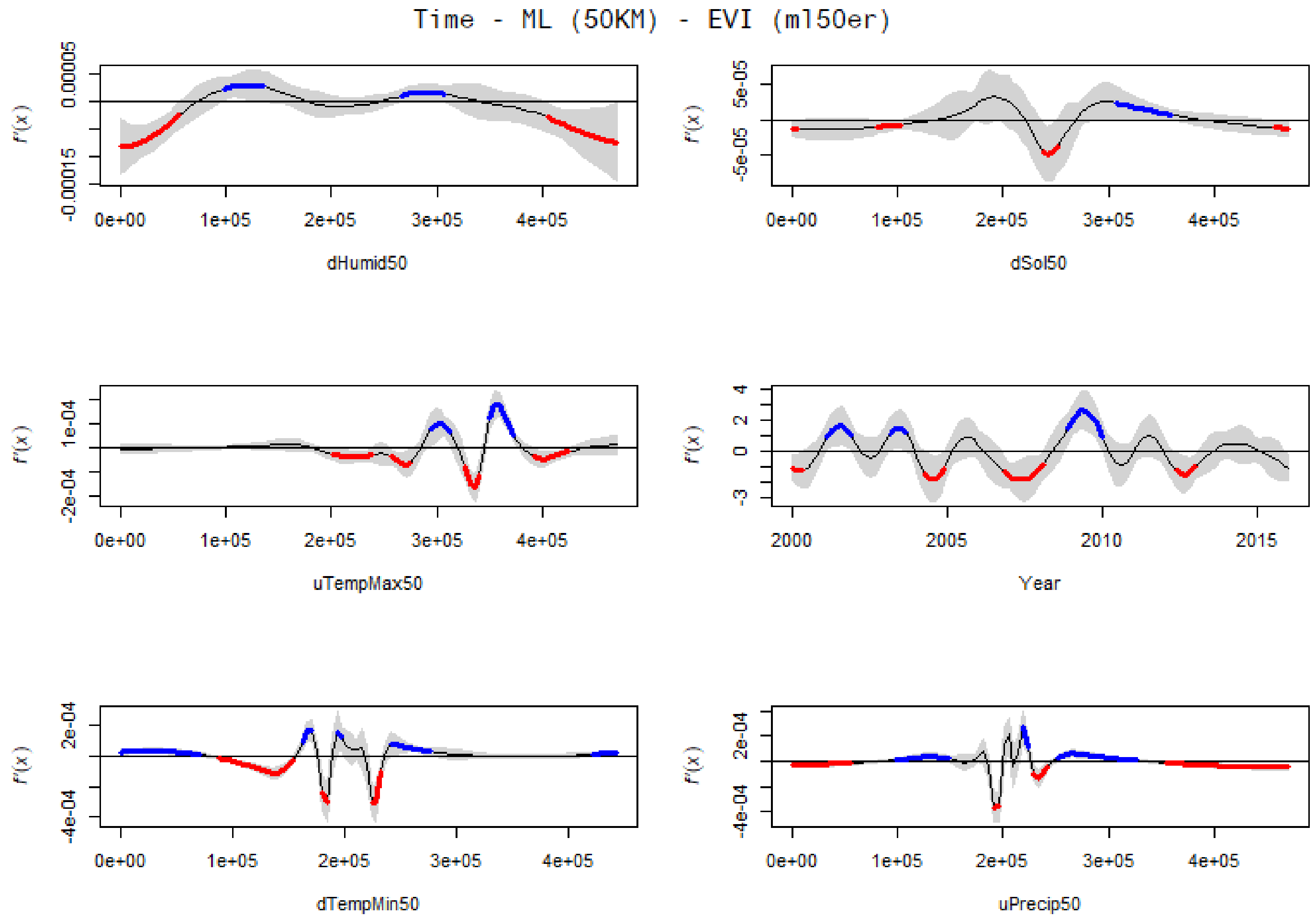

Figure A.2.32 Model lm50er. Model Legal Maranhão for 50km bandwidth using EVI values for raining season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

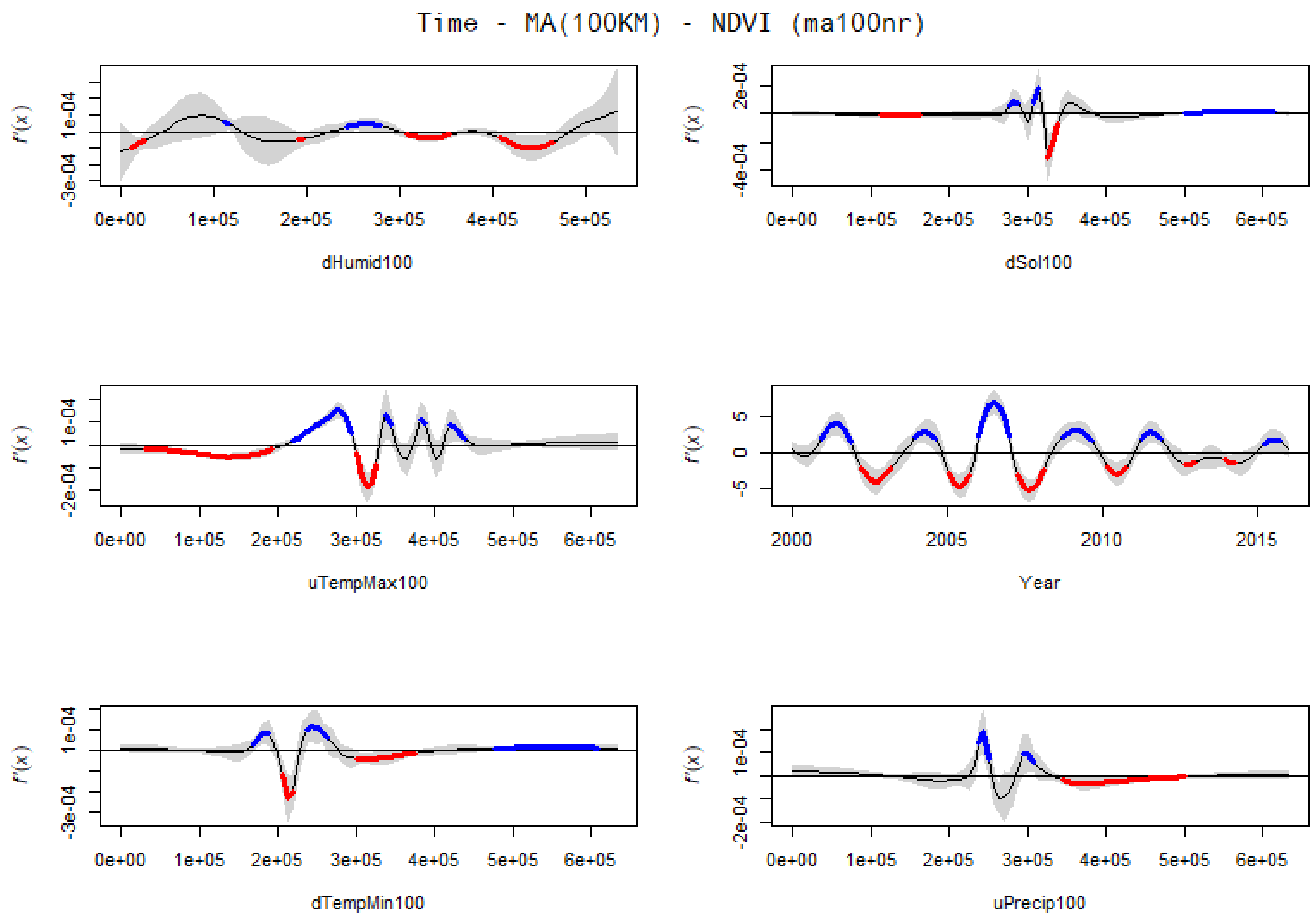

Figure A.2.33 Model ma100nr. Model Maranhão Cerrado for 100km bandwidth using NDVI values for raining season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

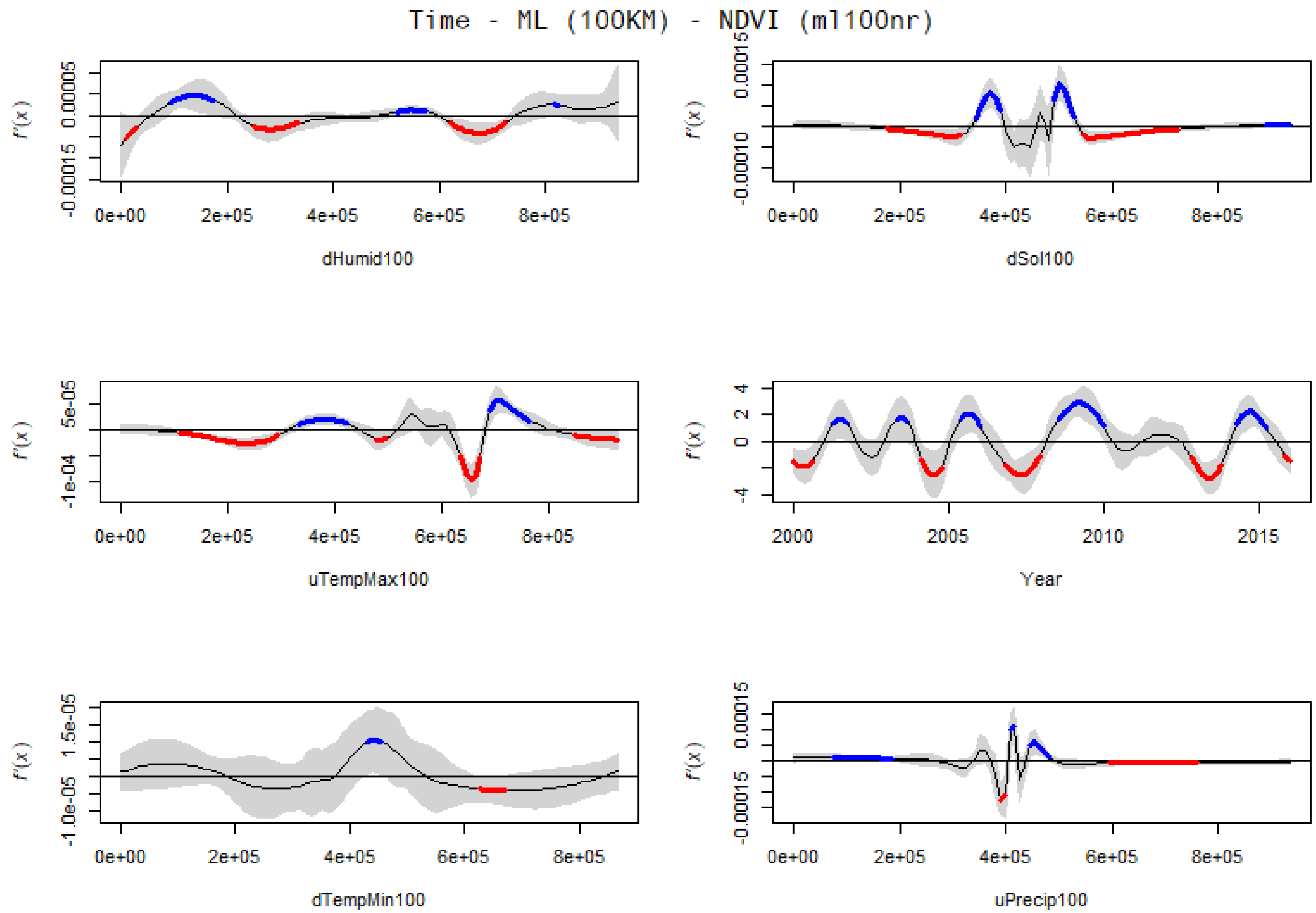

Figure A.2.34 Model lm100nr. Model Legal Maranhão for 100km bandwidth using NDVI values for raining season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

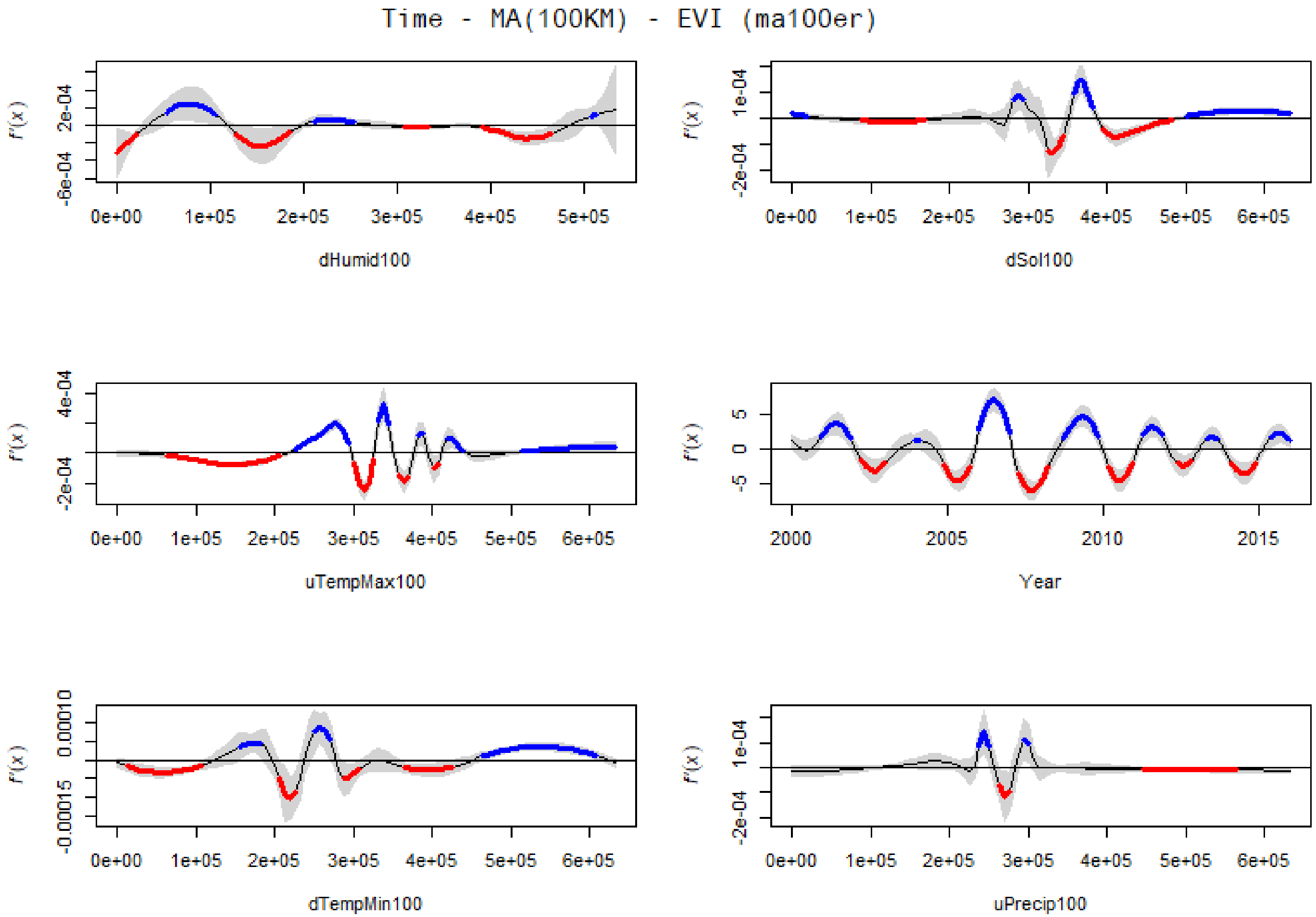

Figure A.2.35 Model ma100er. Model Maranhão Cerrado for 100km bandwidth using EVI values for raining season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

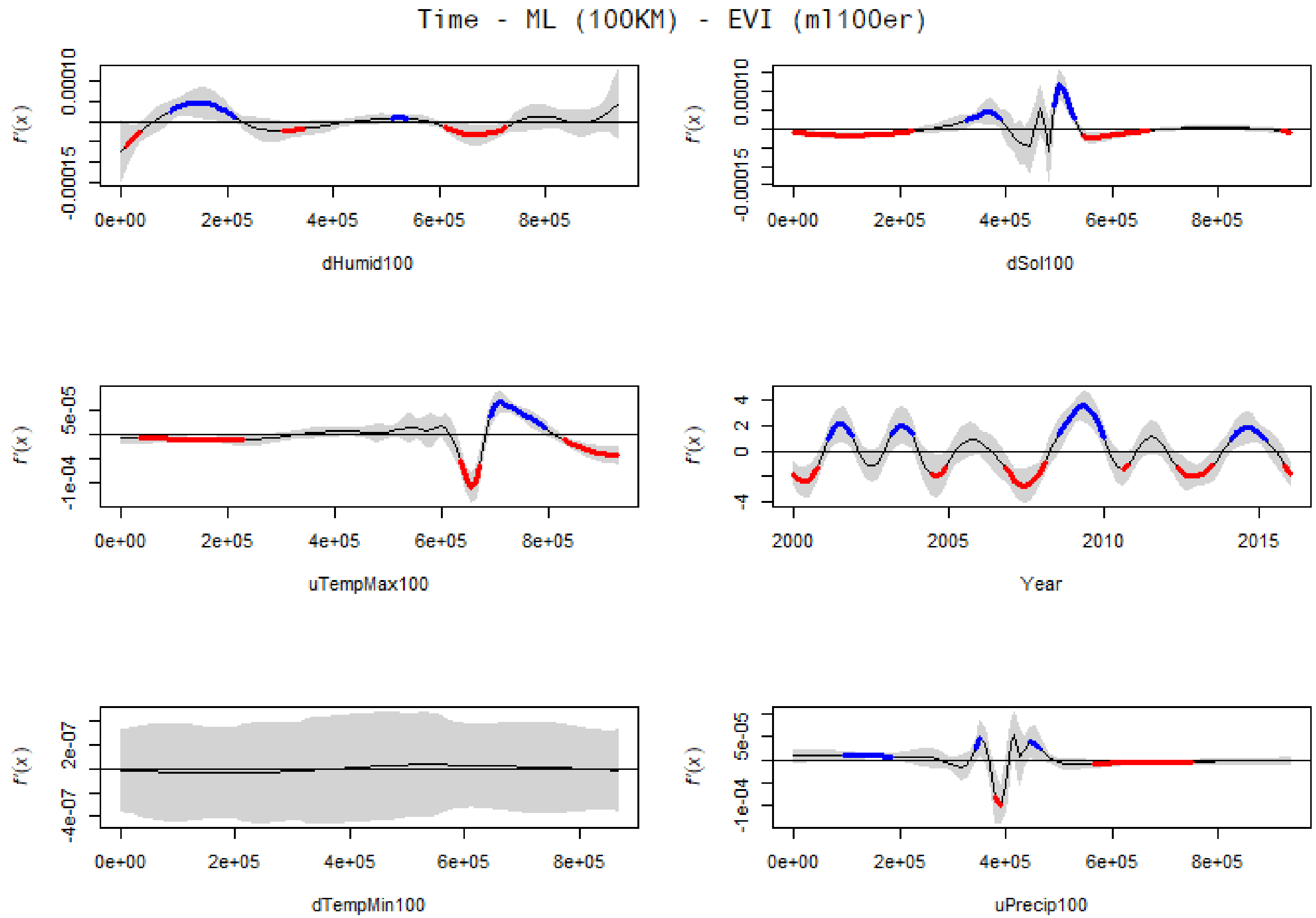

Figure A.2.36 Model lm100er. Model Legal Maranhão for 100km bandwidth using EVI values for raining season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

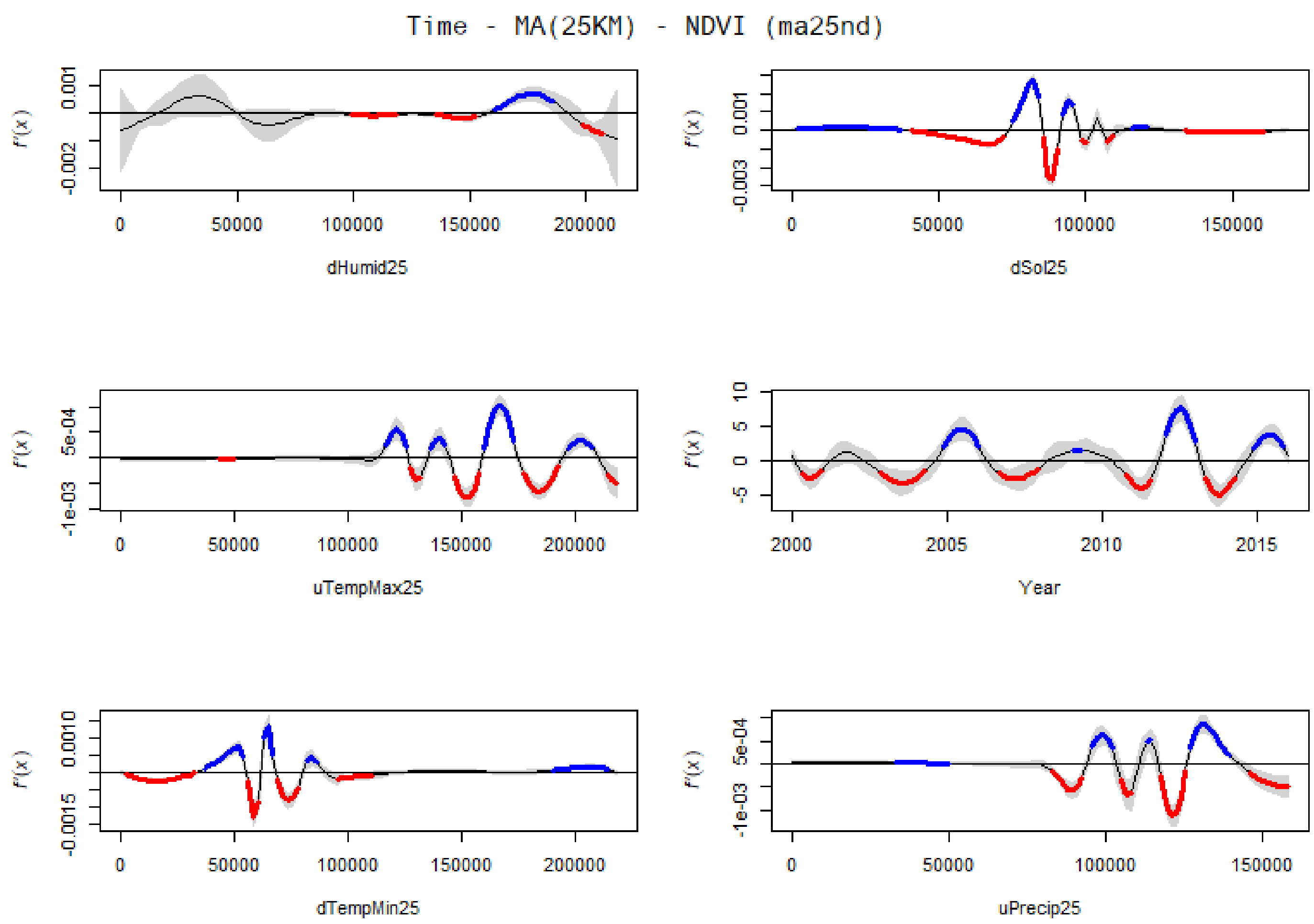

Figure A.2.37 Model ma25nd. Model Maranhão Cerrado for 25km bandwidth using NDVI values for dry season.First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

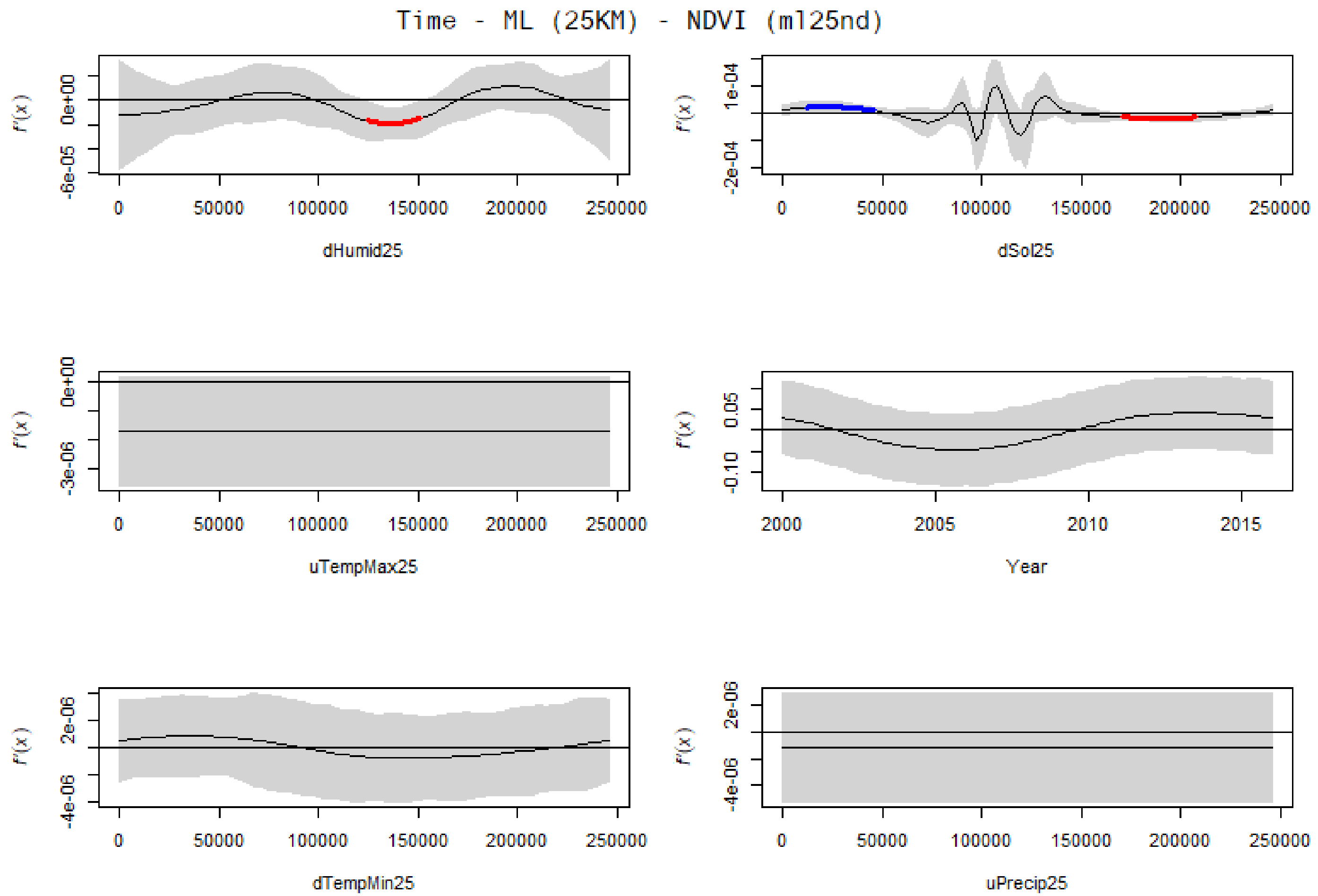

Figure A.2.38 Model lm25nd. Model Legal Maranhão for 25km bandwidth using NDVI values for dry season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

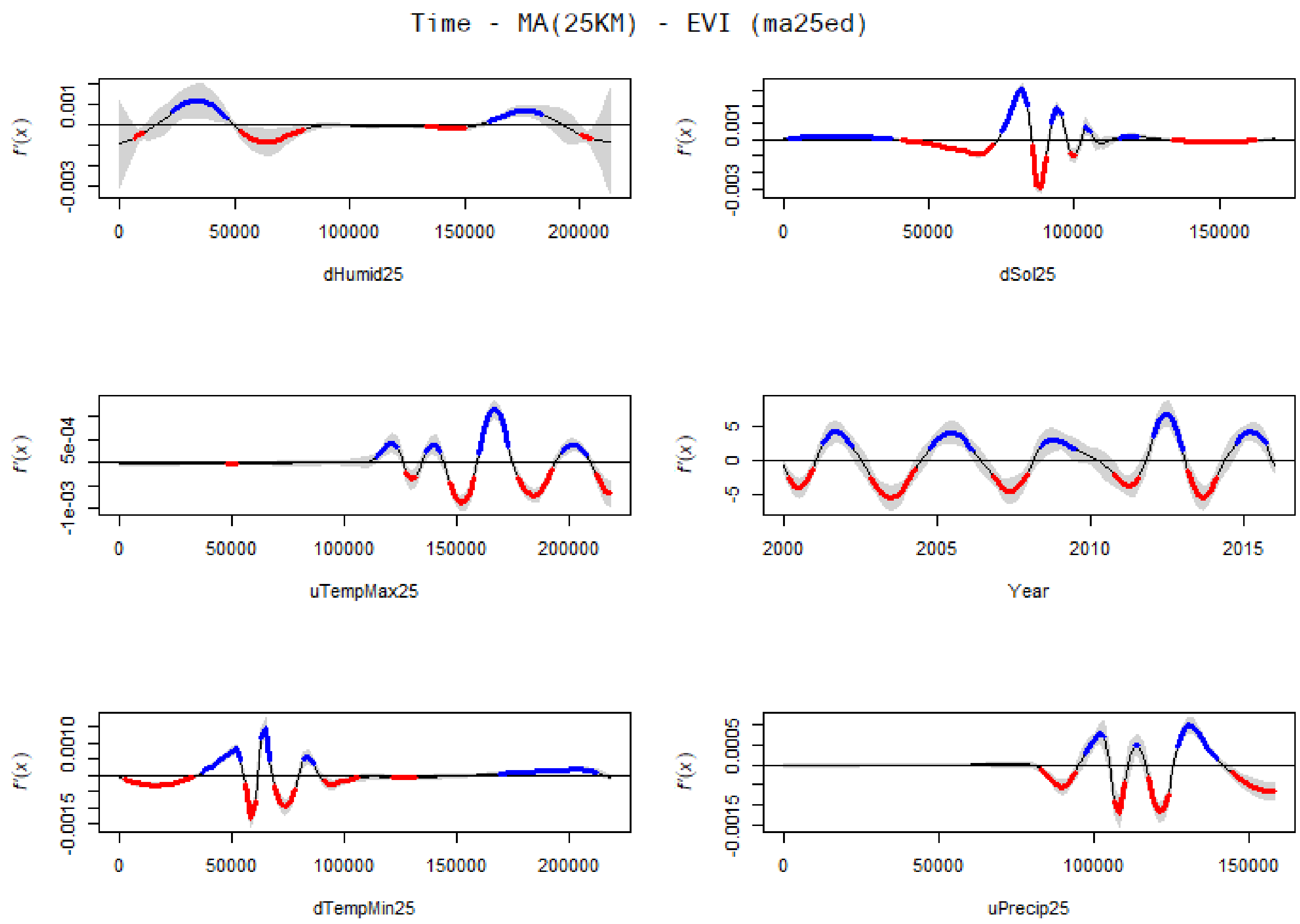

Figure A.2.39 Model ma25ed. Model Maranhão Cerrado for 25km bandwidth using EVI values for dry season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

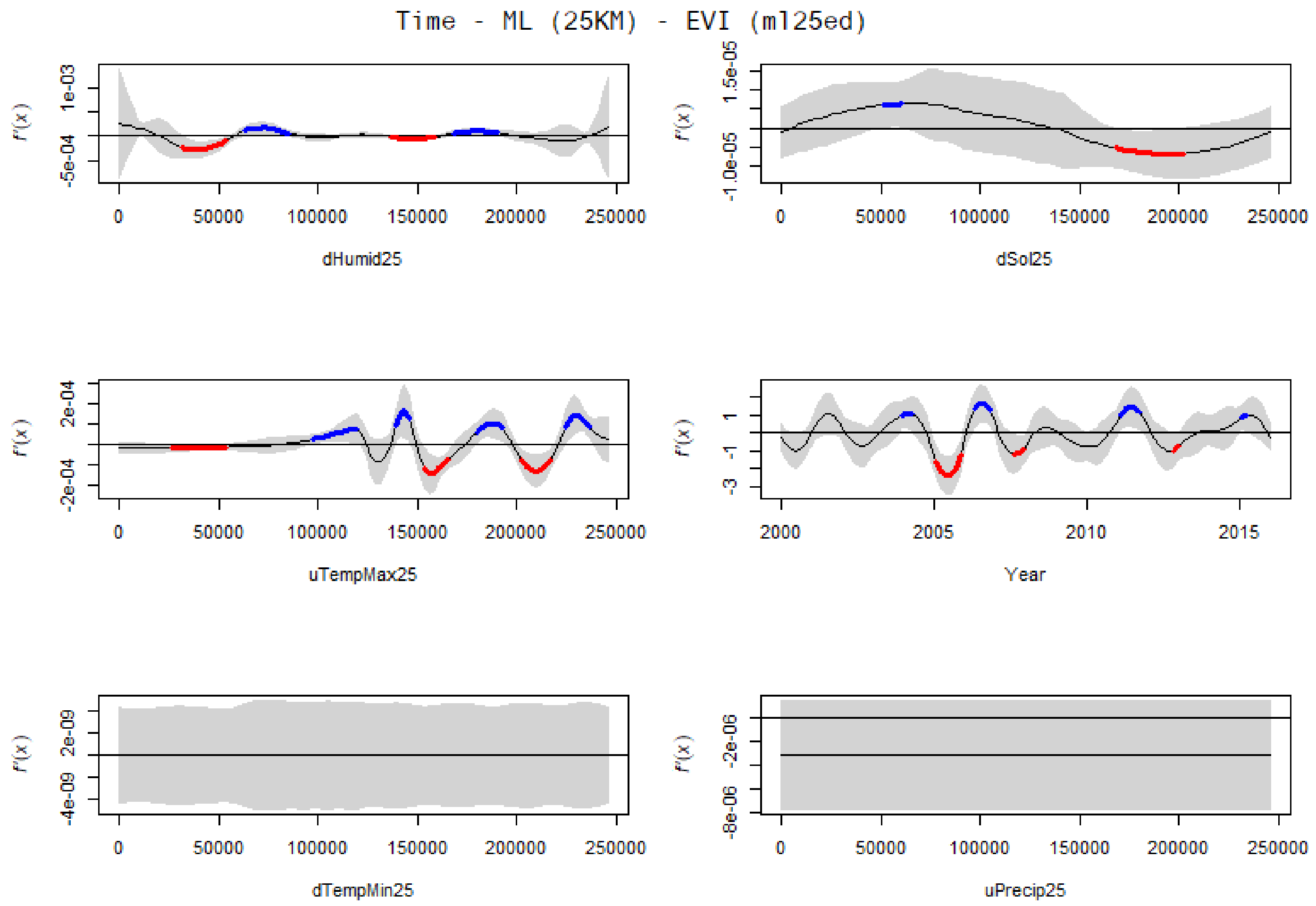

Figure A.2.40 Model lm25ed. Model Legal Maranhão for 25km bandwidth using EVI values for dry season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

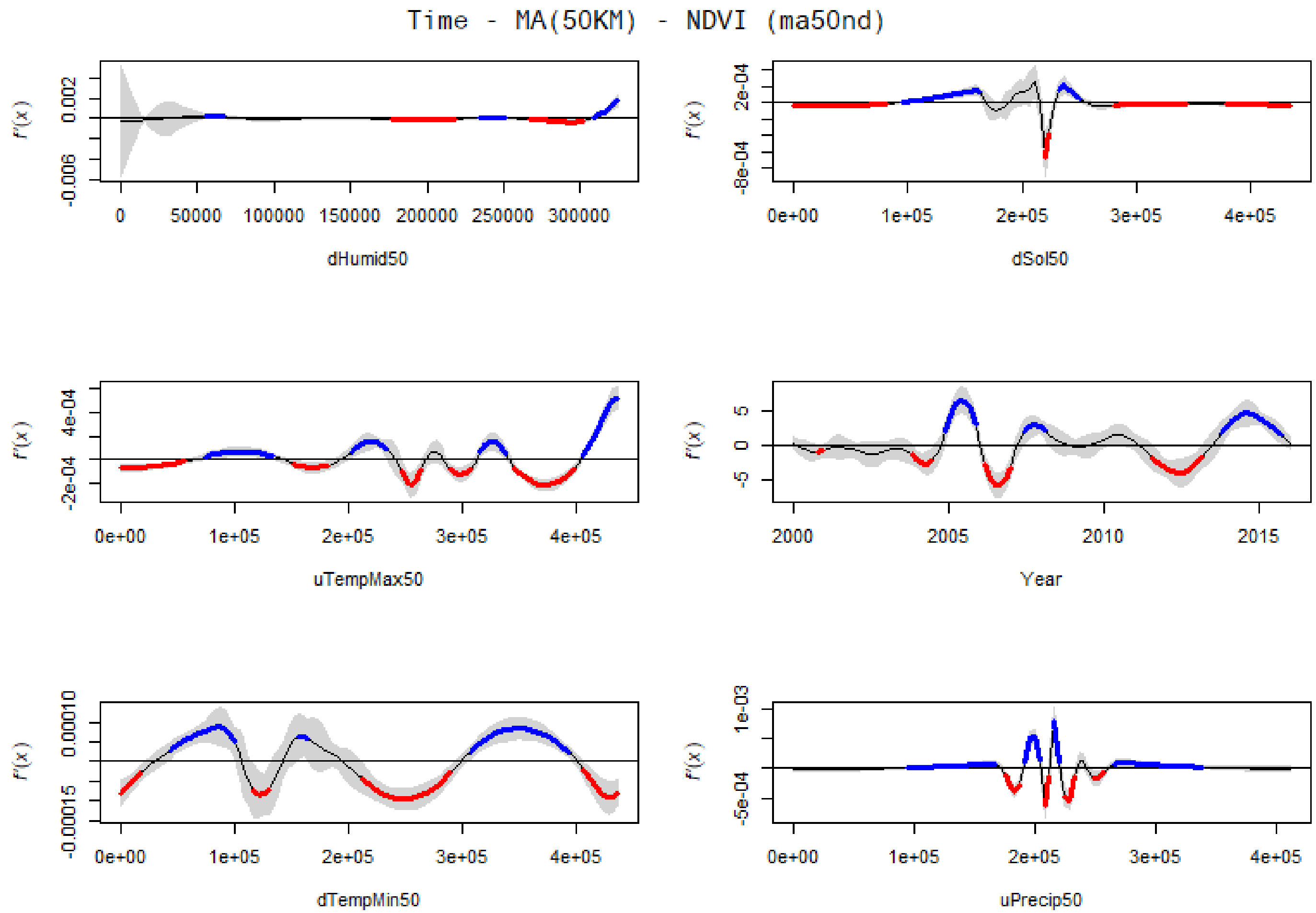

Figure A.2.41 Model ma50nd. Model Maranhão Cerrado for 50km bandwidth using NDVI values for dry season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

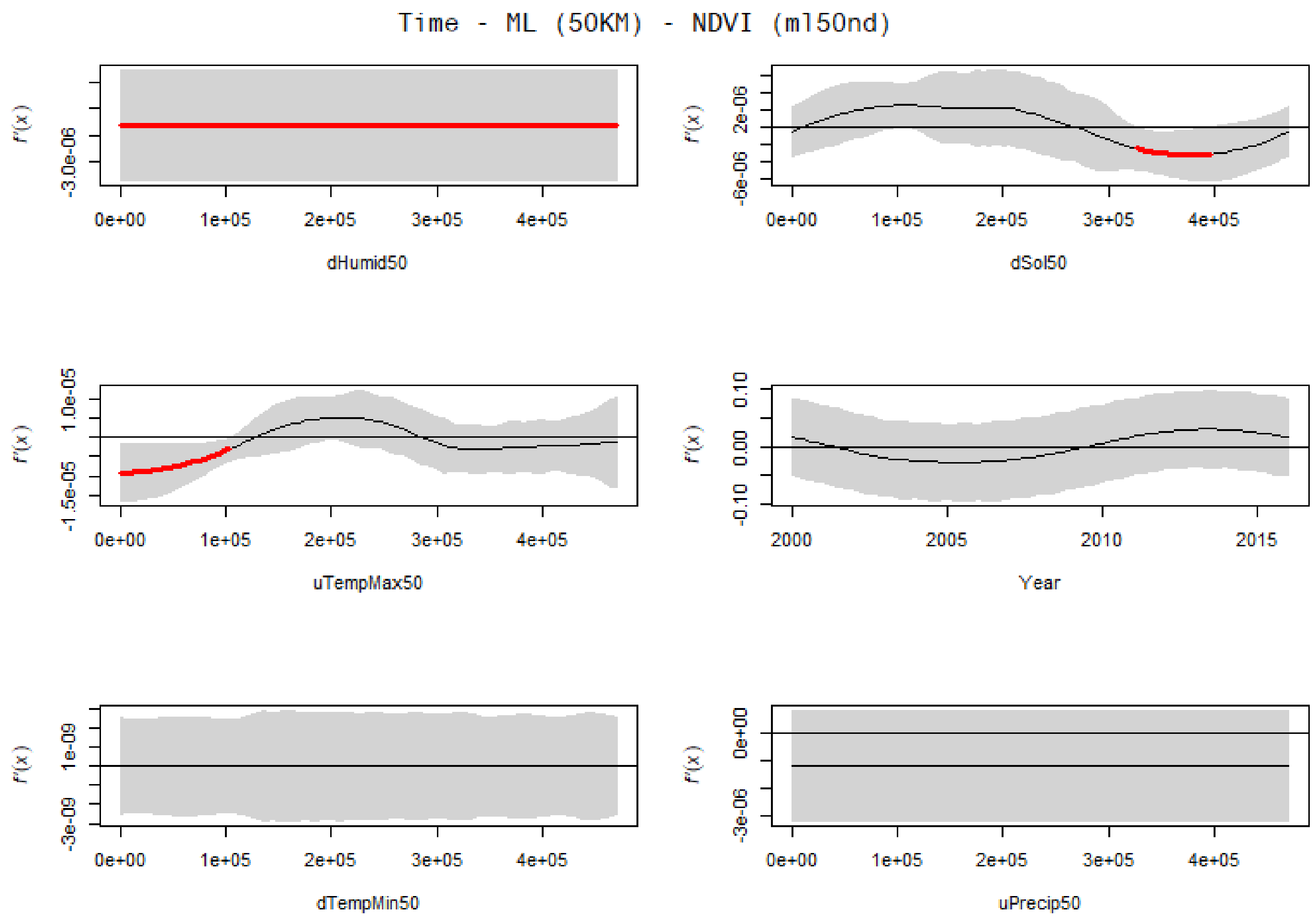

Figure A.2.42 Model lm50nd. Model Legal Maranhão for 50km bandwidth using NDVI values for dry season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

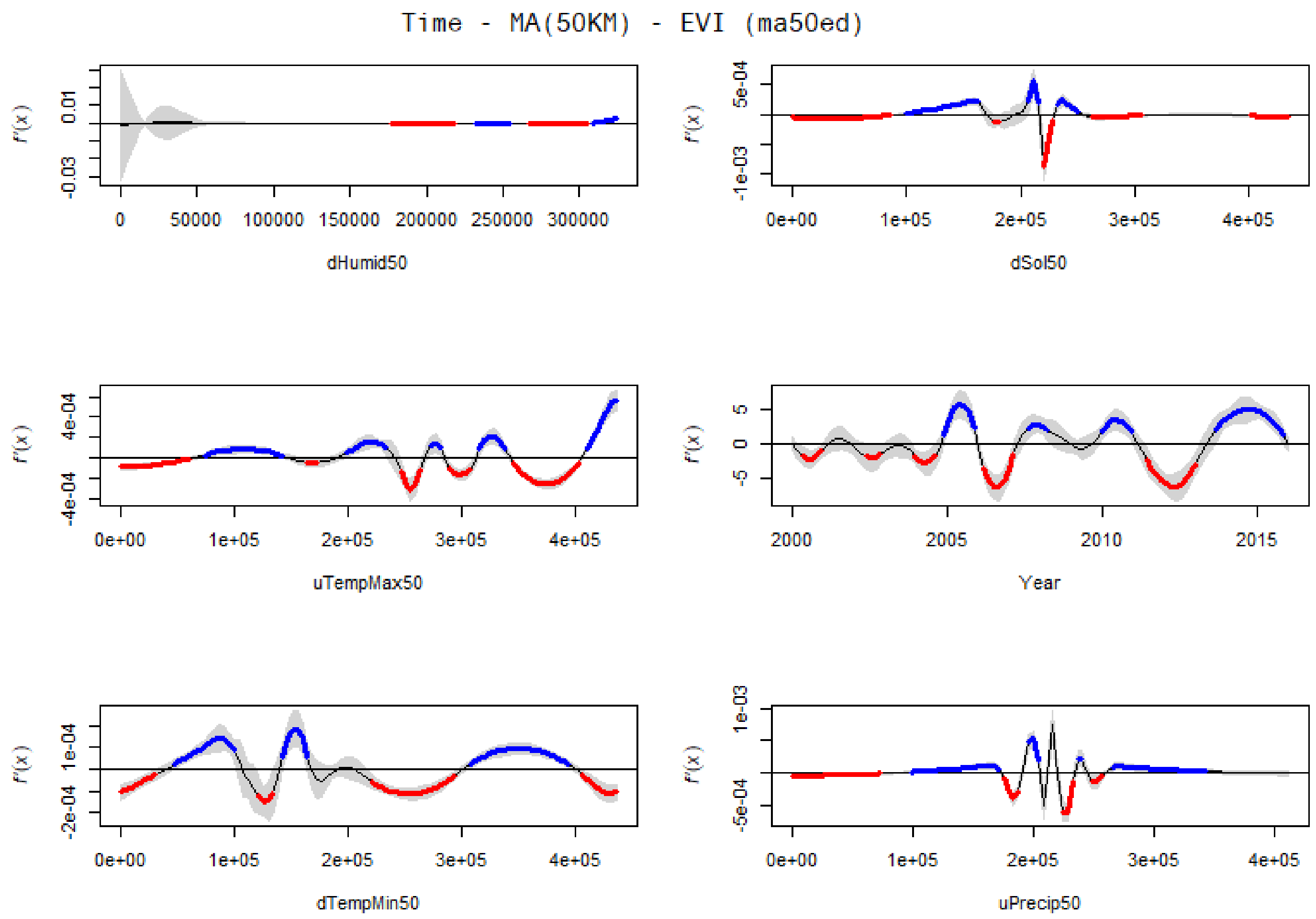

Figure A.2.43 Model ma50ed. Model Maranhão Cerrado for 50km bandwidth using EVI values for dry season.First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

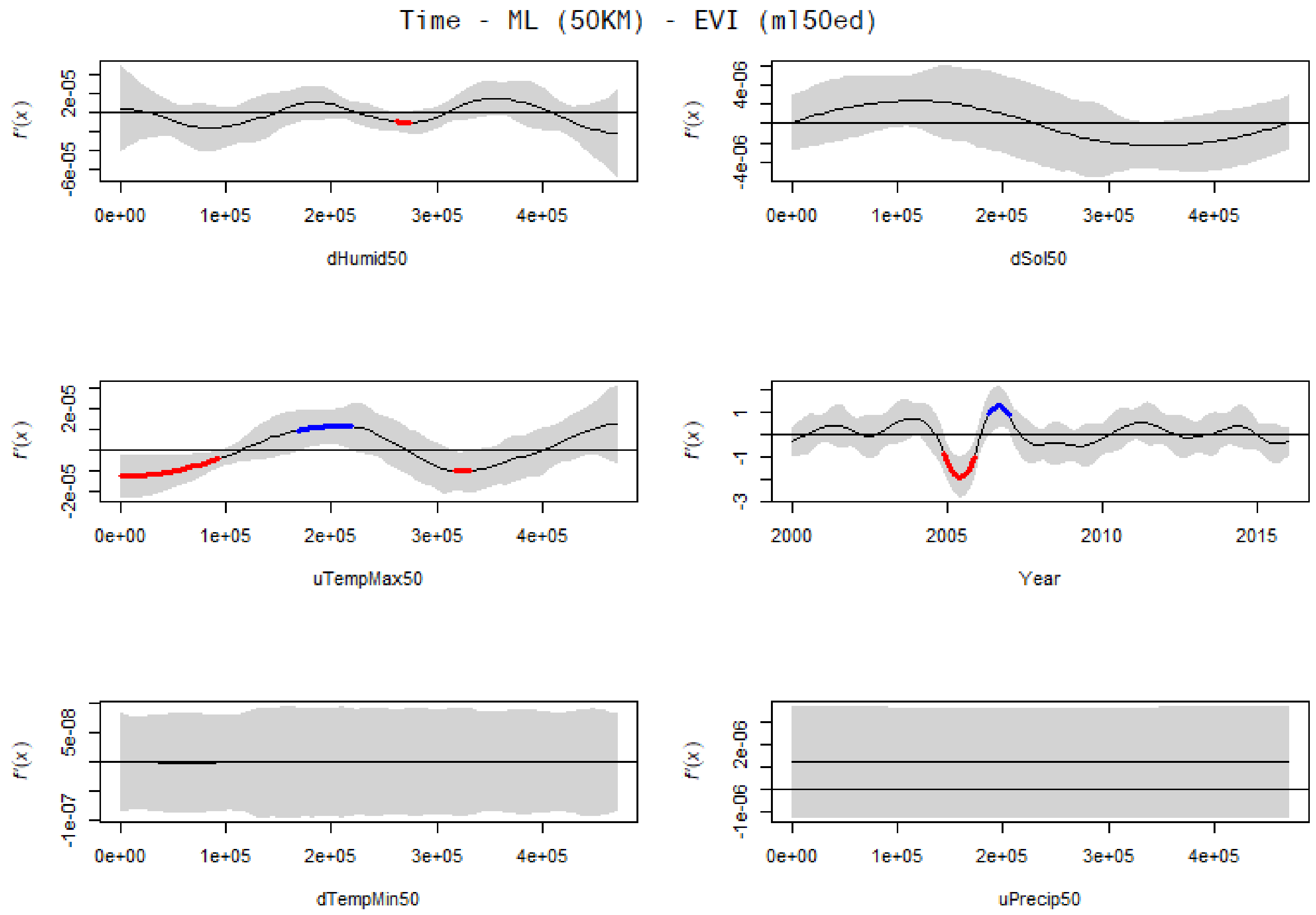

Figure A.2.44 Model lm50ed. Model Legal Maranhão for 50km bandwidth using EVI values for dry season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

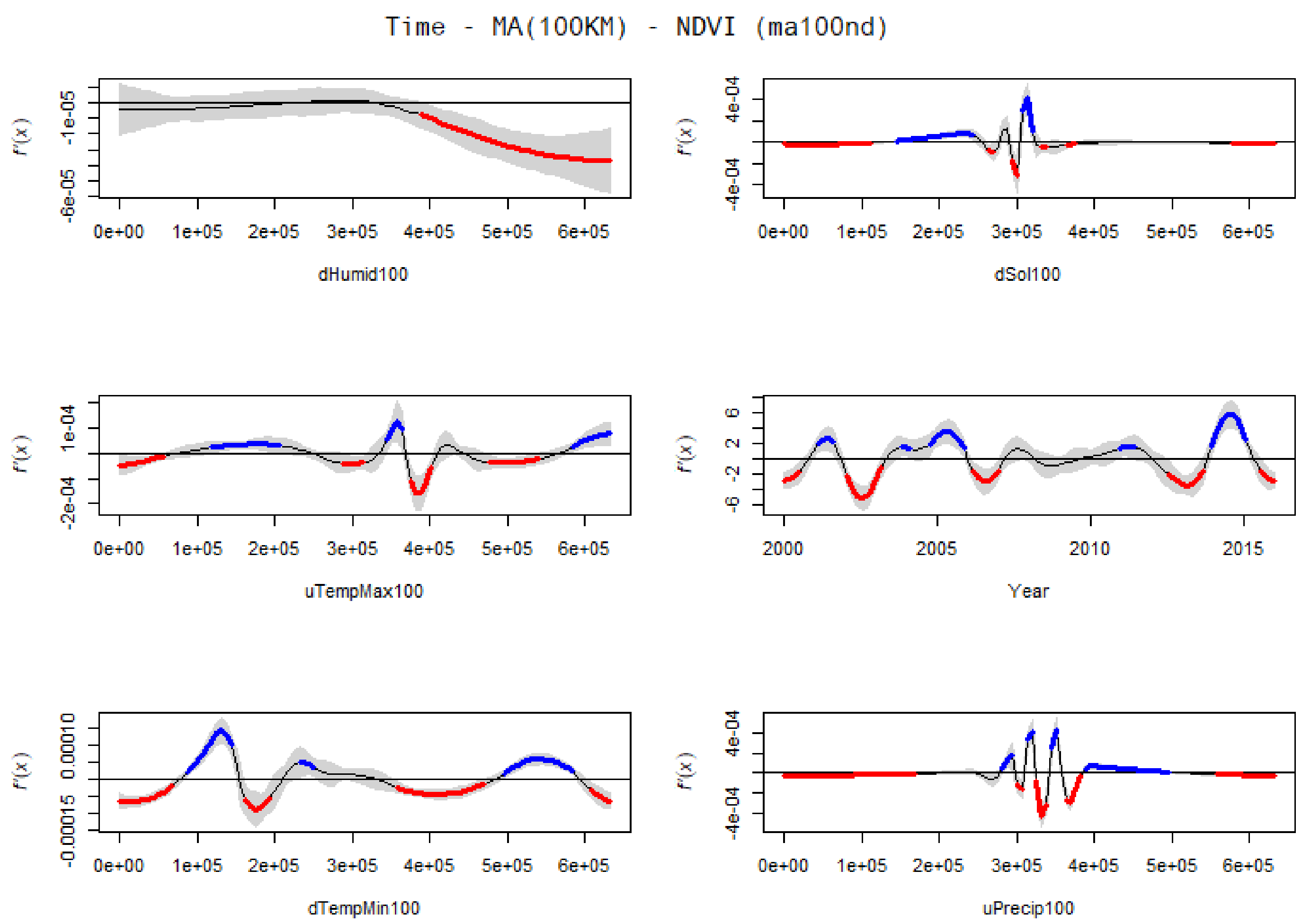

Figure A.2.45 Model ma100nd. Model Maranhão Cerrado for 100km bandwidth using NDVI values for dry season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

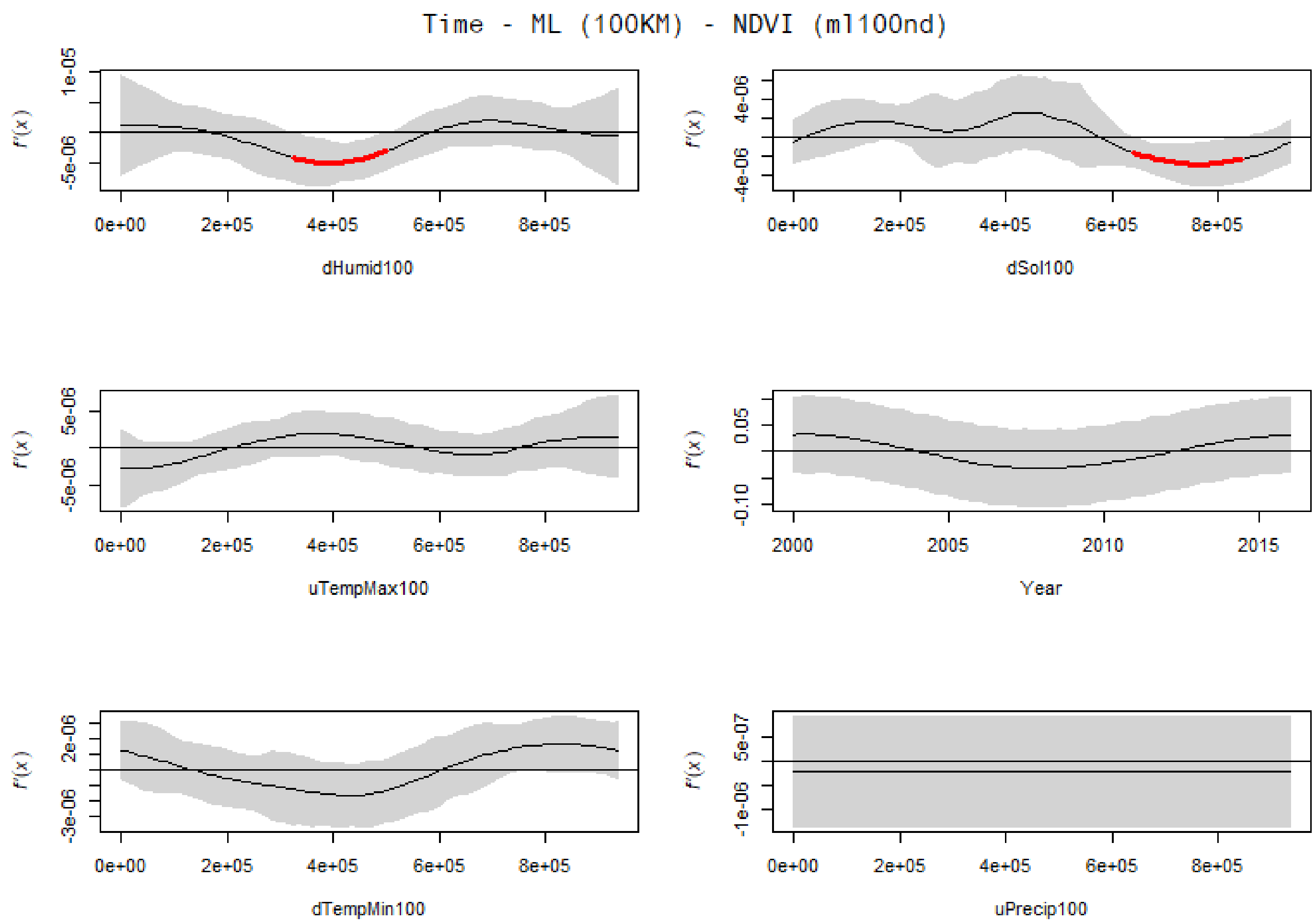

Figure A.2.46 Model lm100nd. Model Legal Maranhão for 100km bandwidth using NDVI values for dry season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

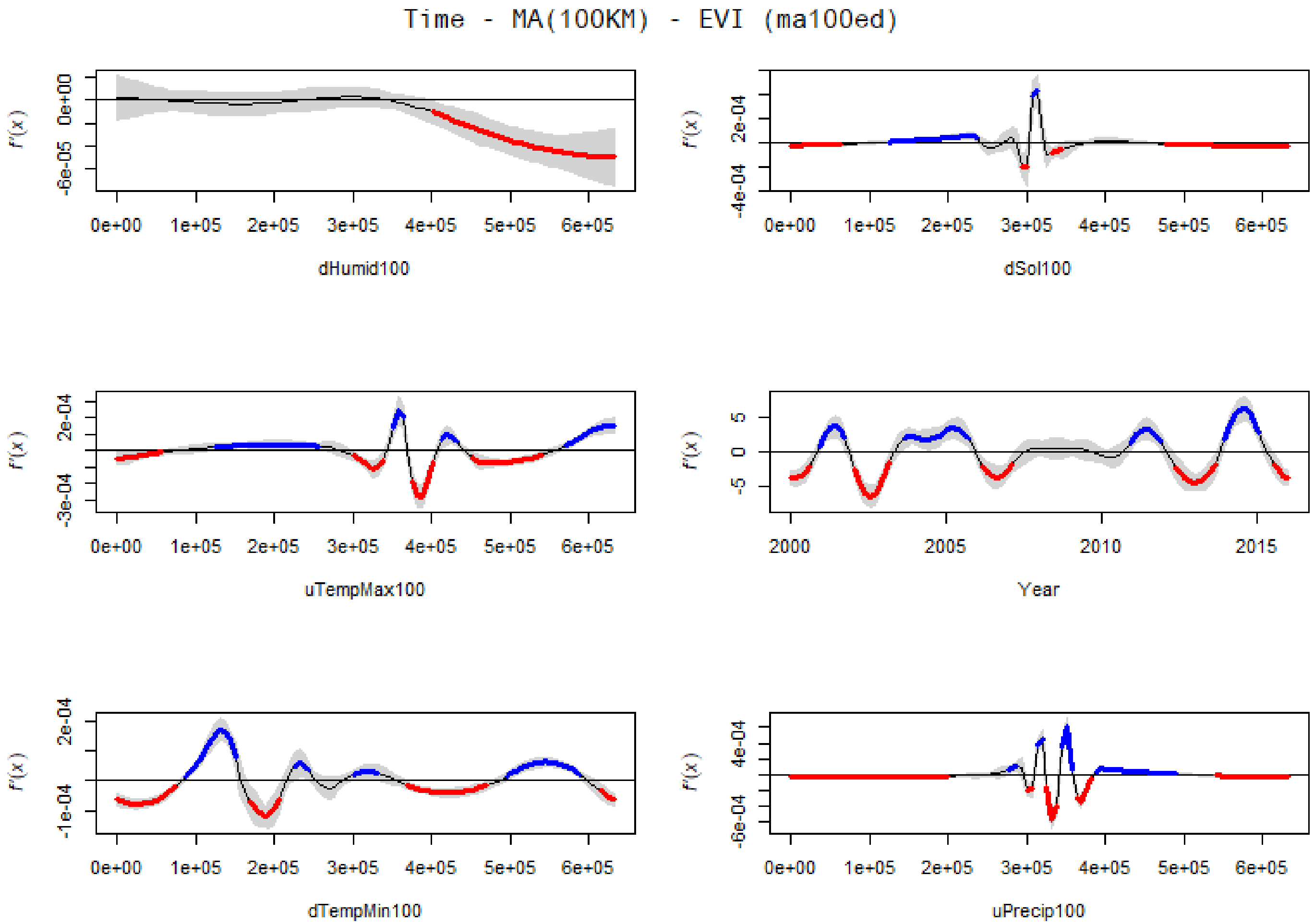

Figure A.2.47 Model ma100ed. Model Maranhão Cerrado for 100km bandwidth using EVI values for dry season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

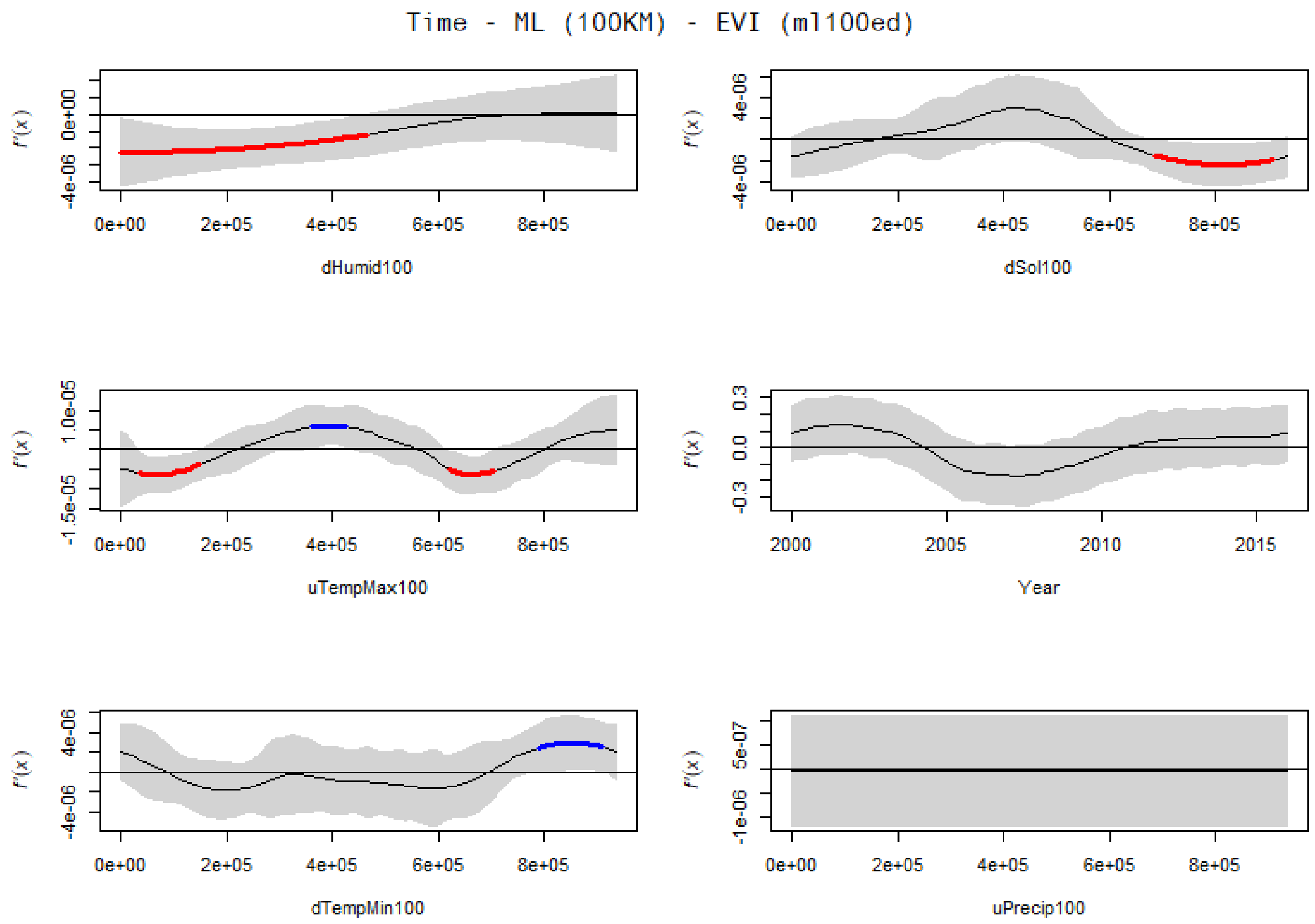

Figure A.2.48 Model lm100ed. Model Legal Maranhão for 100km bandwidth using EVI values for dry season. First derivative of the trends splines from the deforestation data Generalized Additive Model (GAM). The grey band is a 99% simultaneous point-wise confidence interval. Sections of the spline where the confidence interval does not include zero are indicated by coloured sections. Blue colour means positive values and red colour means negative values. This graph can show when the level of deforestation was statistically significantly increasing (blue) or decreasing (red). The approach taken is to compute the first derivatives of the fitted trend using the method of finite differences to compute them. To produce derivatives via finite differences, we compute the values of the fitted trend at a grid of points over the entire data. We then shift the grid by a tiny amount and recompute the values of the trend at the new locations. The differences between the two sets of fitted values are the first differences of the trend and give a measure of the slope of the trend at any point in time. We evaluate the trend at 200 equally spaced points.

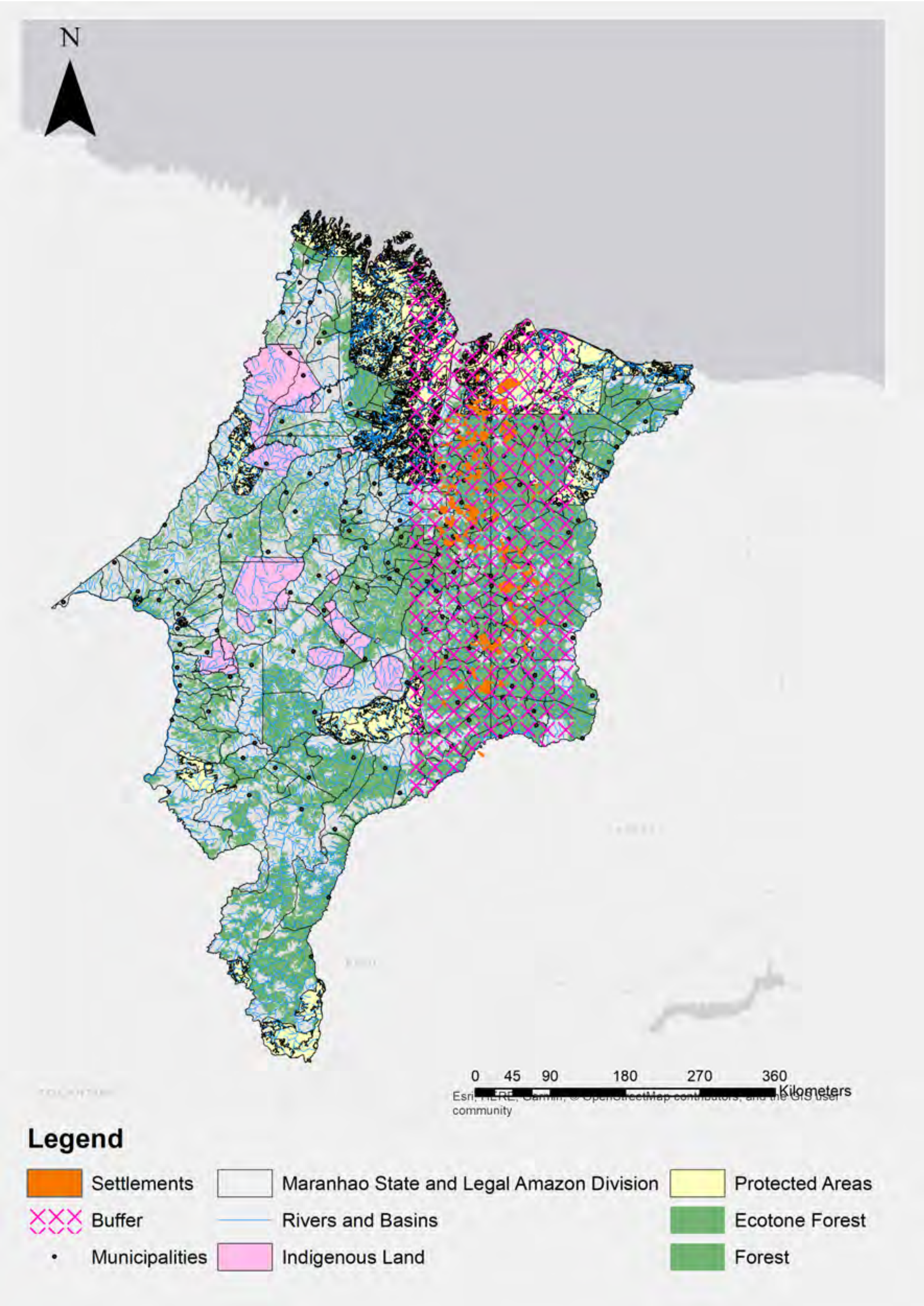

Figure A.2.49 Maranhão State and the Legal Amazon delimitation with buffers of 100km and the presence of settlements within the buffer. The map includes municipalities centre, rivers and basins, protected areas and indigenous land. Source: (EMBRAPA, 2018; INCRA, 2015; MMA, 2018b; Núcleo Geoambiental - NUGEO, 2018).

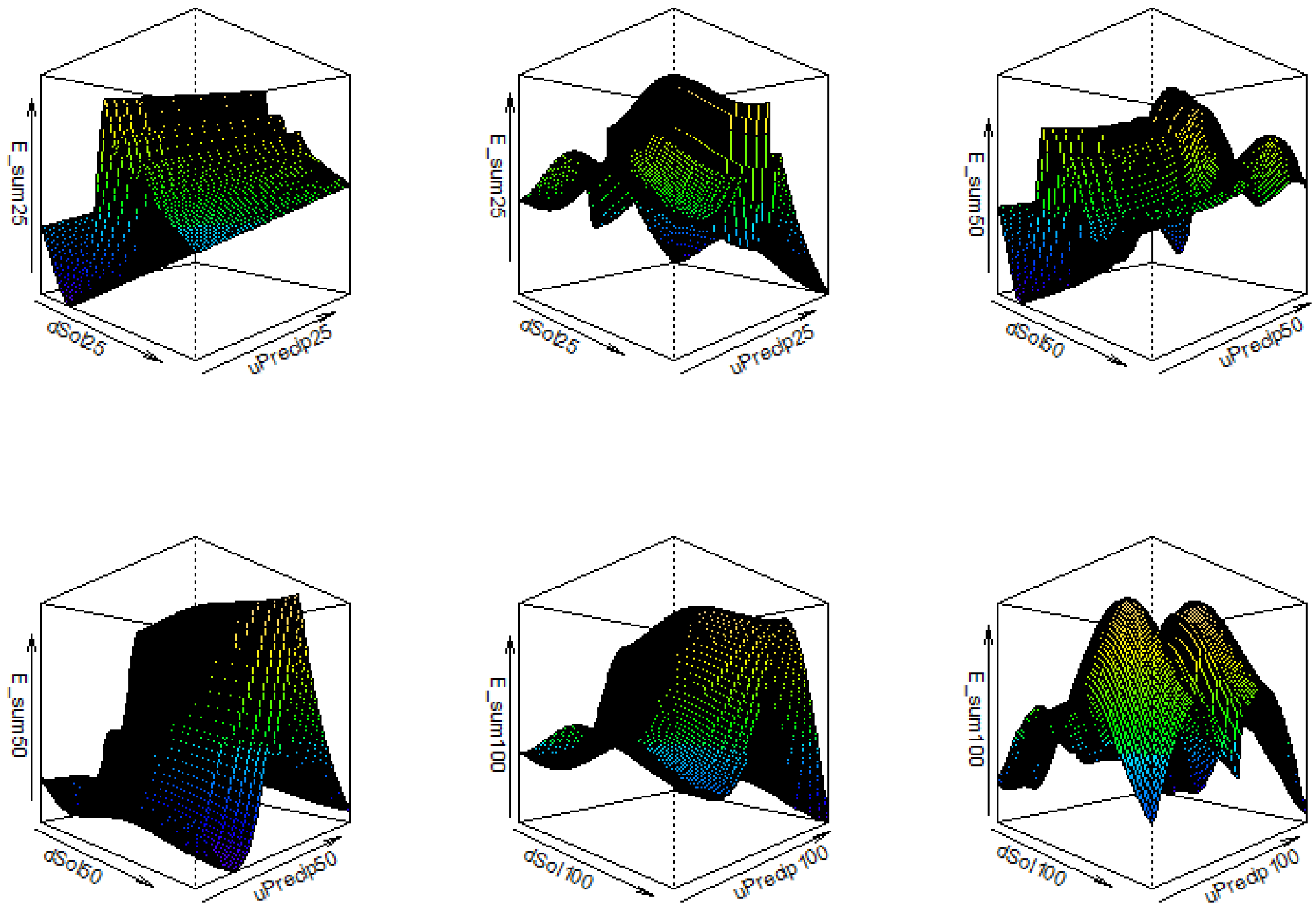

Figure A.2.50 Interactions of variables and the Baseline model for bandwidth of 25km, 50km and, 100km.

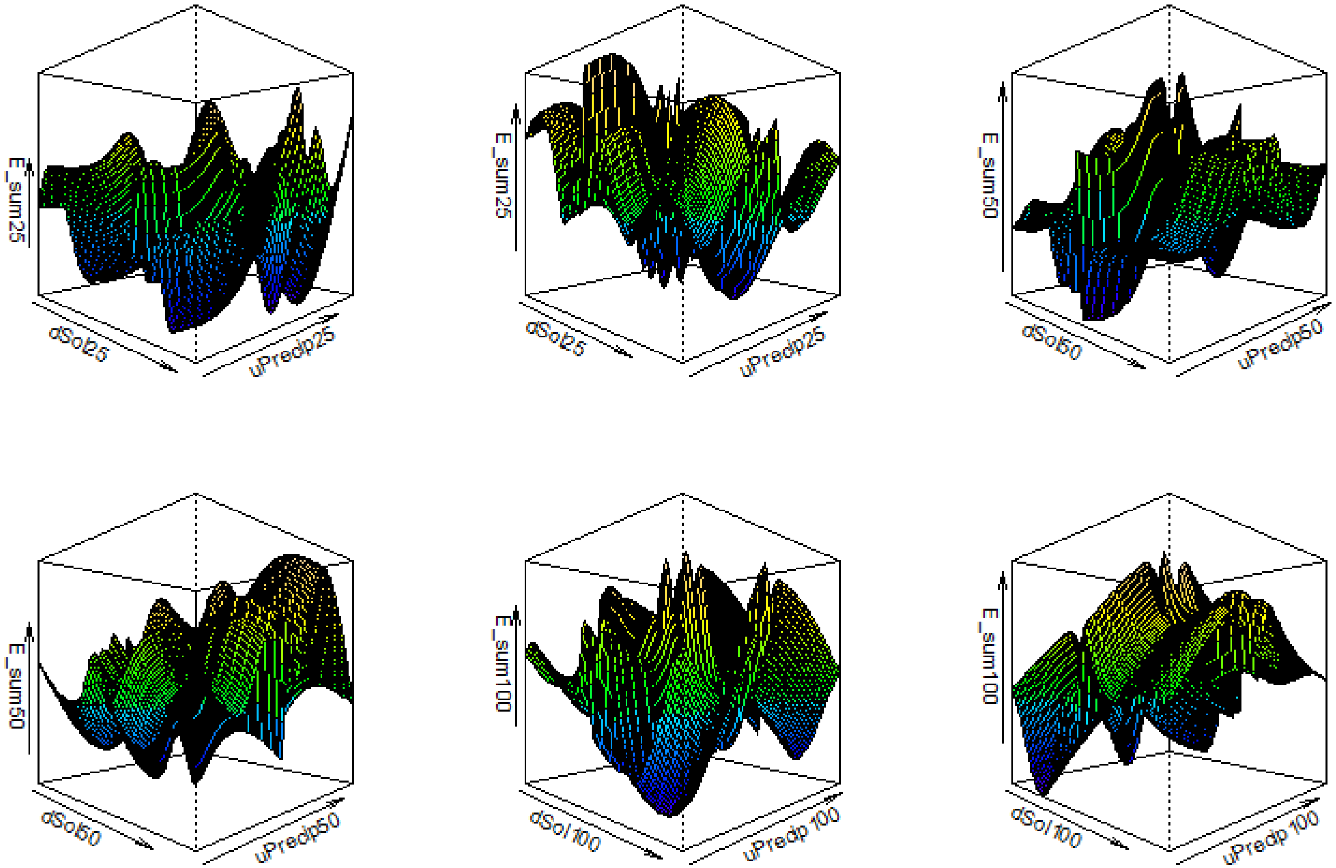

Figure A.2.51 Interactions of variables and the Baseline model for bandwidth of 25km, 50km and, 100km in the raining season.

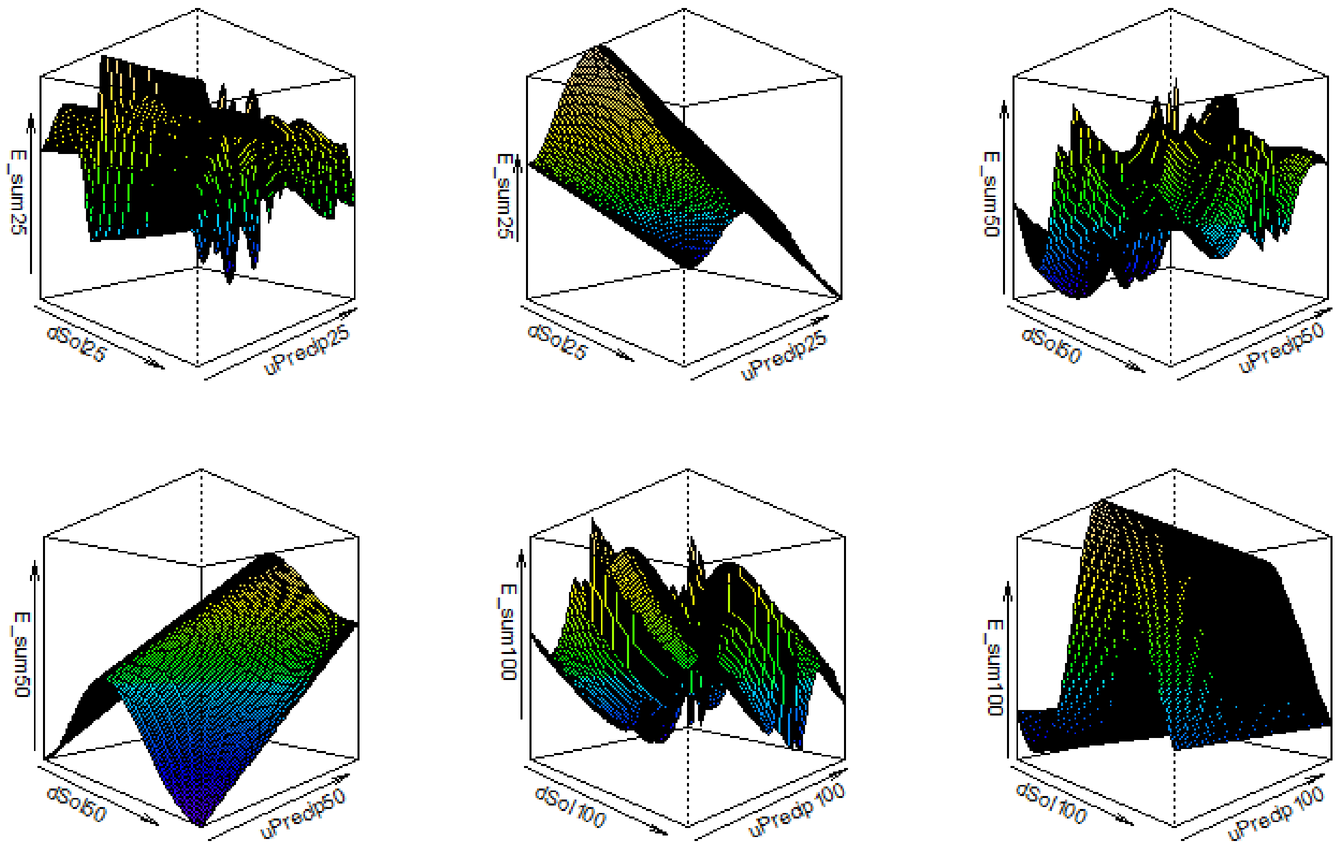

Figure A.2.52 Interactions of variables and the Baseline model for bandwidth of 25km, 50km and, 100km in the dry season.

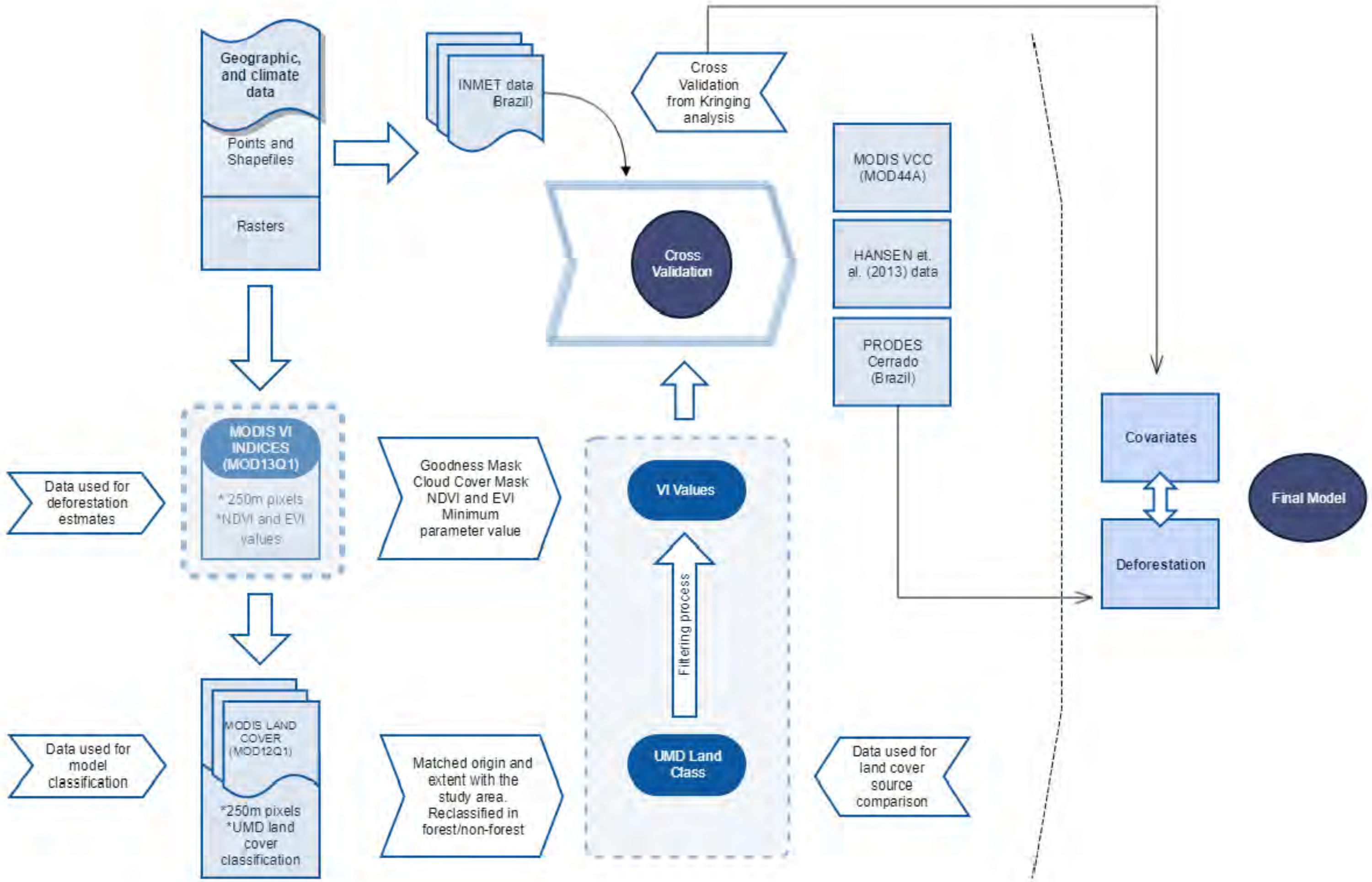

Figure A.2.53 Flowchart of method applied to data . Moving from left to right indicates increasing levels of data processing during the study.

Chapter 3

Satellite Monitoring of Deforestation and the Role of Clouds in Maranhão

Abstract

Deforestation rates in Brazil have declined noticeably over the past decade and it is believed that environmental policies used as instruments to encourage forest preservation have played a crucial role in this trend. In particular, the satellite monitoring program has enabled authorities to identify and react to deforestation in a much more timely manner than local field detection. However, importantly, cloud coverage, by delaying detection until skies are clear, has potentially acted as an important impediment to the policy's success. To investigate this we use satellite data within a survival analysis. Focusing on the ecological tension zone of Maranhão that is separated into two parts by an artificial line-one that was covered by environmental deforestation policy and the other that is not subject to this - we estimate how the probability of transition between intact forest to disturbed forest, given risk factors and conditional on the time elapsed until the occurrence of the transition, is affected by cloud coverage. The results suggest that the presence of clouds has increased deforestation in the region covered by the satellite detection program, and thus is likely an active barrier to legal compliance.

Keywords: Remote Sensing, Survival Analysis, Environmental Policies, Deforestation

3.1 Introduction

The clear-cutting of forests plays a central role in many environmental threats of our time, including global climate change, habitat degradation, and species extinction. Reassuringly, deforestation rates in Brazil have declined over the past decade, and it is generally believed that an important reason for this reduction has been environmental policies used as instruments to encourage forest preservation (Celentano et al., 2017; Nepstad et al., 2014; Richards, 2015; Richards and VanWey, 2015). The best example in this regard is the satellite monitoring program that has allowed the Brazilian environmental police to considerably speed up their response to punish clear-cutting agents by detecting local deforestation remotely and much quicker than through local inspection. More specifically, in 2004, the Brazilian government created the Action Plan for the Prevention and Control of Deforestation in the Legal Amazon (PPCDAm in Portuguese), where the purpose of this program was to plan development, control land use and ensure compliance with the environmental law and promote sustainable practices. In order to control land use and prevent further deforestation, the PPCDAm importantly included two satellite-based monitoring programme PRODES (Programa de Cálculo do Desflorestamento da Amazônia in Portuguese) (INPE, 2017) that recorded the annual rate of deforestation within the policy area using fine resolution, and the DETER (Sistema de Detecção de Desmatamento em Tempo Real in Portuguese) programme, which is a system to support the supervision and control of deforestation in the Amazon in a coarse resolution throughout the year within the environmental policy area. DETER reports deforestation alerts on areas, greater than 250m, in the process of deforestation by degradation forest management to fully deforested areas (shallow cut).[1]

[1] In order to control for degradation of the forest by selective logging and forest fires, the government uses the DETER program. In addition, two other systems were introduced in 2007: DEGRAD (Mapeamento da Degradação Florestal na Amazônia Brasileira in Portuguese), for mapping forest degradation in the Legal Amazon, and DETEX (Mapeamento da Cobertura Florestal na Amazônia Brasileira in Portuguese), for detecting logging operations in the Legal Amazon region (Pinheiro et al., 2016). In order to control land use and prevent further deforestation, the PPCDAm also includes the satellite-based monitoring programme

While DETER has certainly allowed much quicker detection of local deforestation, an important impediment to its success has been local climate. More specifically, because the satellite used is incapable of detecting land cover changes when its view of land is obscured by clouds, detection will be delayed until skies are clear again. As a matter of fact, Assunçao et al. (2017) show that cloud coverage is an important predictor of the extent of deforestation fines issued within municipalities in the Brazilian amazon. This may not be surprising since, according to Assunçao et al. (2017); Pinheiro et al. (2016), Brazil's institutional setup is such that law enforcement agents can more easily punish offenders for illegal forest clearings when catching them red-handed, since offenders can thereby be held directly accountable for the crime. That is, although in principle Brazilian past deforestation acts can be legally punished, this is difficult to implement ex post in the Amazon, where land and production property rights are unclear. Moreover, once deforestation has already occurred, the absence of well-defined property rights makes it very difficult or even impossible to identify who owns the cleared land, so that sanctions or punishment becomes pointless if no one can be held responsible for the crime. Not to mention that accessing these cleared areas requires substantial efforts since illegal roads are built under precarious way (Pfaff et al., 2007). Additionally, there is considerable evidence of corruption among local government officials who monitor and fine loggers. Furthermore, some firms can exert political influence to obtain favourable terms and often bribe government officials during the concession process to overlook noncompliance. The reason for this ranges from underpaid inspectors, potentially high rents to resource stocks, and low penalties and lack of enforcement by centrally located governments in open access distant native forests (Amacher et al., 2012).

PRODES (Projeto de Estimativa do Desflorestameno da Amazônia in Portuguese) (INPE, 2017) that attempts to record incidents of deforestation throughout the year within the policy area. All the data gathered by the action plan are used to enforce the PPPCDAm plan, which includes the issuing of fines for agents who clear or damage the forest, embargoes of areas in the process of being cleared with the confiscation of equipment, and restrictions on access to subsidised credit (Aubertin, 2015).

Despite the potentially crucial importance of cloud coverage in deterring the detection of deforestation through the DETER satellite program, there is no study that explicitly examines this.[1] The current paper thus sets out to explicitly quantify the effect of cloud coverage on local levels of deforestation in the Brazilian Amazon. To this end we focus on the ecological tension zone of Maranhão because it is divided by an artificial line that separates it in two parts: the Legal Amazon Maranhão and the Cerrado Maranhão.[2] This division, occurring approximately 44° west of the meridian, allows the state to be divided in terms of environmental policies, providing us with a spatial division for which the effect of cloud coverage is likely to be very different. More precisely, unlike forests in the Legal Amazon Maranhão, Cerrado Maranhão was not under the satellite monitoring program and thus clouds should play no direct role in deforestation other than for climatic reasons. Importantly, the area on both sides of the border is homogeneous in several aspects such as biota, institutions, and, climate, with the only difference being the surveillance environmental program that is observed in only one part of the region. This provides us with a unique context within which to examine the role of cloud coverage in deforestation, in that we can compare regions that biologically homogeneous but heterogeneous in terms of deforestation detection policy. To identify local deforestation and cloud coverage within our two regions we use remote sensing sources (MODIS Vegetation Indices (MOD13Q1) product) to construct a time event dataset at the individual pixel level (250m). We then use survival estimation methods to quantify the role of cloud coverage in local deforestation across these two region.

[1] Assunçao et al. (2017) do use cloud coverage as an instrumental variable to identify how deforestation fines have affected deforestation across municipalities in the Brazilian Amazon. However, their implied estimate of the impact of fines on deforestation through cloud coverage can only be viewed as capturing the effect of clouds partially and very roughly. Firstly, the cloud coverage will have affected other aspects of deforestation through other legal coercion such confiscation of assets or access to credit and commercialisation channels; see (Börner et al., 2015). Moreover, since their analysis undertaken at the municipality level, where as satellite detection is at the 250m level, i.e., at the resolution of the satellite images.

[2] Importantly, in July 2018, the National Institute of Spatial Research (INPE in Portuguese) together with several other institutions published a dataset covering annual and biannual of Cerrado biome deforestation at state and national level. From the dataset, it is shown that deforestation in great Cerrado Maranhão, which includes the transitional forest, has been two times higher than in the Amazon region of Maranhão.

There are already a number of other studies that use satellite data to examine the factors driving deforestation. For example, Vance and Geoghegan (2002) estimated a spatially explicit model of forest clearance process in the southern Mexico implementing a time event analysis to identify the effect of households on the probability of deforestation. The results showed a non-linear probability of forest clearance. Greenberg et al. (2005) assessed the impact of spatial, cultural and economic factors on deforestation using survival analysis on satellite data from eastern Ecuador. Their results showed that deforestation prediction was higher when there was proximity to roads. Until now, however, studies combining satellite data in areas of ecotone forest to detect time event (survival) analysis of the risk of deforestation process remains scarce in the Brazilian literature.[1]

According to our survival analysis results, forests inside the specific surveillance policy area had a lower probability of survival compared to the area not covered by the environmental policy. Forested pixels close to protected areas, which include conservation units and indigenous land, had a higher chance of being cleared comparing to forested pixels far from these special zones. More importantly, the presence of clouds increased the time to local deforestation in the monitoring area, but not such effect was found for the non-monitored counterpart. Thus clouds arguably were an active barrier to legal compliance to deforestation legislation which is corroborated by the results found in Assunçao et al. (2017) that showed that environmental police is systematically less present in municipalities with greater cloud cover in any given year, and that these municipalities exhibit higher deforestation the following year.

[1]Monitoring the transitional biome is difficult due to the high heterogeneity of the forests (open and dense forest, for example) which are substantially influenced by the climatic seasonality (Bayma and Sano, 2015). Also, there is no environmental policy in place to prevent rampant deforestation. However, in the context of the Amazon region which is under a specific environmental policy, it is arguably crucial to understand the dynamic of the transitional forest borders and its potential to influence adjacent Amazon forests since it provides a valuable endpoint from which climate and anthropogenic related aspects in the Amazon forest my be better understood.

3.2 Material and Methods

3.2.1 Study Context

In the central part of Maranhão there is a contact area between the Amazonian and Cerrado biomes, where it is possible to observe a mosaic of savanna vegetation *Cerrado* and open and dense forest formation, which configures as an Ecological Tension Zone (ETZ). Various authors discuss the difficulty of delimiting forest areas into transitional and / or ecological tension regions, mainly Cerrado-Amazon Forest, due to the innumerable indentations and interpenetrations of savanna formations in the territory of the Legal Amazon. Areas with these characteristics can be found in the states of Amazonas, Mato Grosso, Pará, Tocantins and, especially, in Maranhão. Garcia et al. (2017) studied part of the Maranhão central region and defined forest as a combination of riparian forest, transitional forest, and Cerrado woodland which definition is adopted in this study.

Site

The studied area comprises a total of 29,307 km^2 and corresponds to 21 municipalities. All the municipalities are crossed by the artificial division that occurs approximately 44° west of the meridian and was established in 1953 in order to plan the economic development of the region comprised of the tropical forest areas of the Maranhão state called Legal Maranhão (LM). We depict this delineation in Figure 3.1. More specifically, areas to both sides of the artificial line is homogeneous in several aspects, such as biota, institutional framework, and, climate with the only difference being the surveillance environmental program that is observed in only one part of the municipality. To ensure that we are comparing homogeneous areas we reduce our study region to a 0.2 degree buffer zone around the artificial line.

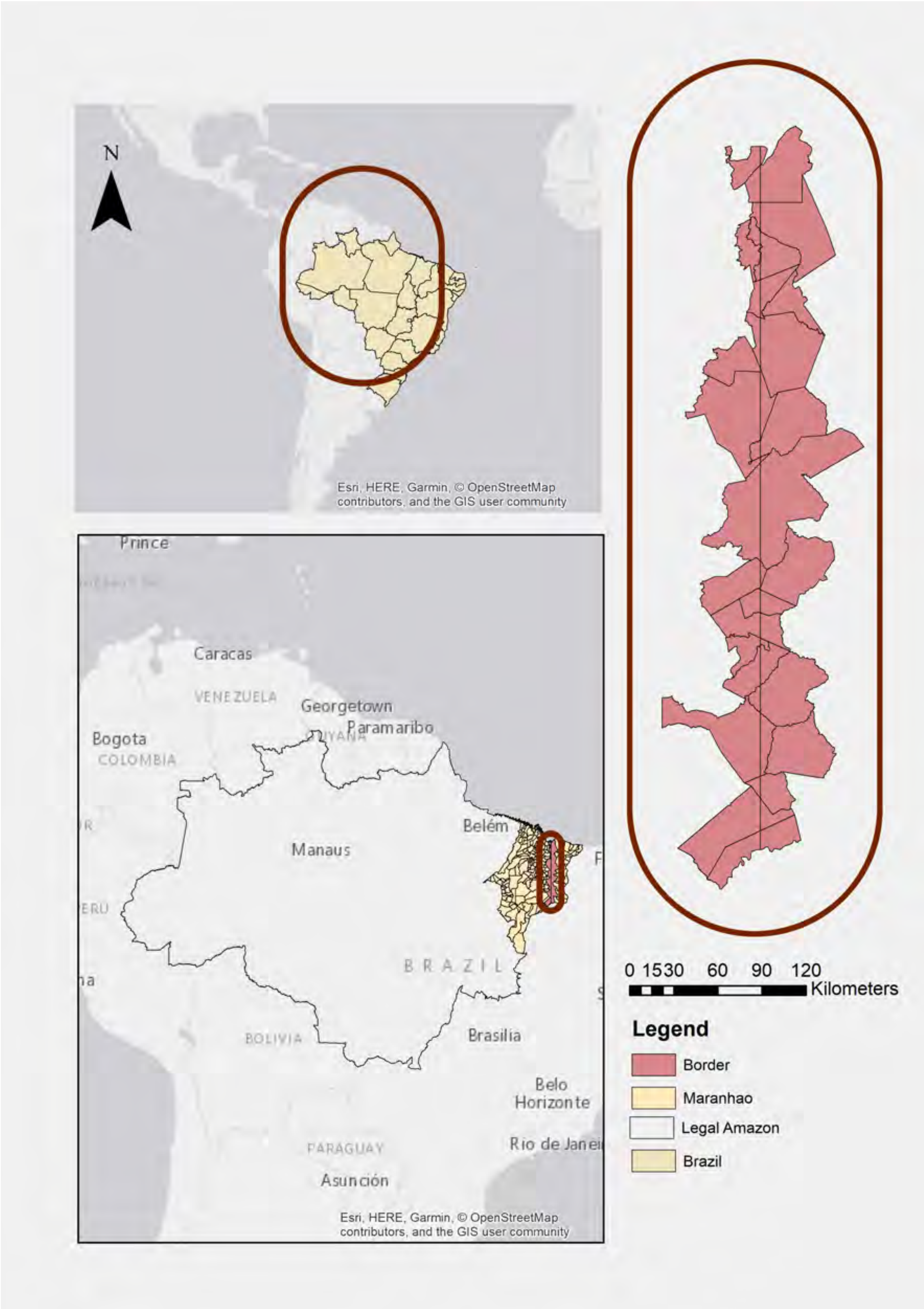

Figure 3.1 Map of Maranhão Studied Area. The vertical line in the Border refers to the artificial line that divides Legal Maranhão and Cerrado Maranhão. Source: (MMA, 2018b; Núcleo Geoambiental - NUGEO, 2018).

Environmental Policy: DETER monitoring system

As noted previously, the PPCDAm policy has at its core two satellite-monitoring operating systems, the PRODES and DETER, which were designed to meet different objectives at different times. PRODES system measures annual clearing rates since 1988, for increments of over 6.25 hectares. Because it uses a fine resolution, it is more detailed and depends on the climatic conditions of the dry season for acquisition of cloud-free images. These images are obtained between May and September, and the deforestation rate is calculated once a year.

In contrast, DETER alerts are conducted for the purposes of support to inspection. The alerts are issued bi-weekly and sent to the environmental police (IBAMA) and to environmental agencies of the states of the Legal Amazon to plan their field actions and operations to combat illegal deforestation. For the general public, the alerts are combined into monthly reports for the period between May and October, when the cloud cover decreases and it becomes possible to observe the region relatively cloud free. During the period from November to April, when there is greater cloud cover, DETER's public reports are quarterly.

The real-time detection system identifies and maps deforested areas in forest formations in the Amazon. This system uses images from MODIS sensors, aboard NASA's TERRA satellite and WFI images aboard the Brazilian CBERS-2B satellite (BRASIL, 2018). These sensors cover the Amazon with high temporal frequency, of two and five days, respectively, but with limited spatial resolution of 250 meters and 260 meters (WFI). With this spatial resolution, the images allow the detection of deforestation the areas of which are greater than 250m (or 25 hectares). Since the satellite used in DETER is incapable of detecting land cover patterns in areas covered by clouds, no forest clearing activity is identified in these areas, and thus no alerts pinpointing the location of recent deforestation activity are issued for the region and time covered by clouds (Assunçao et al., 2017).

The high frequency of observation is one way to compensate for the limitation of the spatial resolution. Most importantly, the high frequency reduces the problems imposed by the frequent cloud cover in the Amazon region. In the Amazon forest, the cloud formation depends on the topography of the place. Usually at the north of Amazon forest there is more cloud persistence, and distance to water basin and the ocean matters, as normally cloud forms far from water reservoirs. Chagnon et al. (2004); Heiblum et al. (2014); Koren et al. (2004); Pinto et al. (2009); Wang et al. (2009) discovered that there is a relationship between cloud cover and land cover in Amazon forest. The evapotranspiration properties of the land cover vegetation are tightly linked to the dynamics of the boundary layer and the formation of clouds, which commonly cap the boundary layer. Deforested areas in the Amazon (either pasture or cropland) usually display higher sensible heat and lower latent heat fluxes in comparison with the forested areas, which in turn can enhance the growth of the boundary layer during the day and favour the formation of larger clouds, i.e. deforested areas enact the presence of clouds and it is difficult to disentangle whether cloud favours deforestation or deforestation favours the formation of clouds.

In the present study this phenomenon does not seems to happen because there is no presence of dense forest our study area that could favour cloud formation over the surrounding forested areas or that could influence the clouds to other areas (deforested ones). The ecological tension zone allows the clouds to spread uniformly within the area which excludes the causal effect found by several studies. We hypothesise that cloud coverage limiting satellite visibility in the DETER monitoring system affects the enforcement of the environmental policy in the forest transition area of Maranhão.

3.2.2 Image preparation

When considering vegetation mapping for forest cover loss and forest regrowth path, traditional methods such as field surveys are time consuming, date acquisition lagged, and generally too expensive. Over the past four decades, a feasible alternative considered by researchers and specialists is to apply remote sense technology and subsequent image analysis. More precisely, the use of satellite time series along with statistical analysis can be helpful in understanding the characteristics of vegetation dynamics. In this paper we use images derived from the MODIS sensors.[1]

Freely available and with high temporal resolution, the MODIS sensor has two instruments. The Terra satellite is on an AM overpass, whereas the Aqua platform provides complementary observations in the afternoon. The Terra orbital configuration and MODIS viewing geometry produce full global coverage every one to two days, except for the equatorial zone, where the repeat frequency is approximately 1.2 days (Setiawan et al., 2014; Zhan et al., 2002a). The high temporal resolution of MODIS is a determining factor in phenological studies and spectral discrimination, and can be used to obtain detailed knowledge about the seasonal cycles of vegetation in biomes with strong seasonal contrast, such as the Cerrado biome and the Ecotone forest. An additional advantage of the MODIS data is that is in a ready-to-use format.

To derive our dataset we use two products created by MODIS sensor: MCD12Q1 and MOD13Q1. These products were retrieved from the online Application for Extracting and Exploring Analysis Ready Samples (AppEEARS) tool courtesy of the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, (Didan, 2015; Didan et al., 2015; Sulla-Menashe and Friedl, 2015, 2018). The MODIS Land Cover Type

[1]A detailed summary of satellites, sensors and databases relevant to vegetation can be found in Horning (2010).

Product (MCD12Q1) provides 13 science data sets (SDSs) that map global land cover at the 500m spatial resolution in annual time steps for six different land cover legends from 2001-2016 and includes 5 legacy classification schemes, such as the University of Maryland classification (UMD) (see Table 3.1), which recognises 17 classes, covering natural vegetation (11 classes), mosaic lands (2 classes), and non-vegetated lands (4 classes) (Setiawan et al., 2014; Sulla-Menashe and Friedl, 2018).

Table 3.1 University of Maryland (UMD) legend and class definitions

Name	Class	Description
Water	0	At least 60% of area is covered by permanent water bodies
Evergreen Needleleaf forest	1	Needleleaf Forests 1 Dominated by evergreen conifer trees (canopy >2m). Tree cover >60%.
Evergreen Broadleaf forest	2	Dominated by evergreen broadleaf and palmate trees (canopy >2m). Tree cover >60%.
Deciduous Needleleaf forest	3	Dominated by deciduous needleleaf (larch) trees (canopy >2m). Tree cover >60%.
Deciduous Broadleaf forest	4	Dominated by deciduous broadleaf trees (canopy >2m). Tree cover >60%.
Mixed forest	5	Dominated by neither deciduous nor evergreen (40-60% of each) tree type (canopy >2m). Tree cover >60%.
Closed shrublands	6	Dominated by woody perennials (1-2m height) >60% cover.
Open shrublands	7	Dominated by woody perennials (1-2m height) 10-60% cover.
Woody savannas	8	Tree cover 30-60% (canopy >2m).
Savannas	9	Tree cover 10-30% (canopy >2m).
Grasslands	10	Dominated by herbaceous annuals (<2m).
Permanent Wetlands	11	Permanently inundated lands with 30-60% water cover and >10% vegetated cover.
Croplands	12	At least 60% of area is cultivated cropland.
Urban and built-up	13	At least 30% impervious surface area including building materials, asphalt, and vehicles.
Cropland/Natural Vegetation Mosaics	14	Mosaics of small-scale cultivation 40-60% with natural tree, shrub, or herbaceous vegetation
Non-Vegetated Land	15	At least 60% of area is non-vegetated barren (sand, rock, soil) or permanent snow and ice with less than 10% vegetation.
Unclassified	255	Has not received a map label because of missing inputs

Note: Source: **?**.

The MODIS Vegetation Indices (VI) (MOD13Q1) product consists of time series comparisons of global vegetation conditions that can be used to monitor the Earth's terrestrial change detection. The two vegetation indices that we derive from these are the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI). The NDVI is a normalized transformation of the NIR (Near Infrared) and red reflectance ratio standardized to range from -1 to 1. The EVI is an optimisation of the vegetational signal that minimises noise, and has been reported to be more responsive to canopy structural variations, including canopy type. The EVI formula is written as:

$$EVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + C_{1\rho_{red}} - C_{2\rho_{blue}} + L}(G) \tag{3.1}$$

where ρ_{red} and ρ_{NIR} and ρ_{blue} are the reflectance in MODIS bands 1,2 and 3 (459-479nm) and C_1 and C_2 are the atmospheric resistance coefficients. L and G are the canopy background adjustment and the gain factor, respectively. The coefficients adopted for the MODIS EVI algorithm are L=1, C_1 =6, C_2 =7.5, and G=2.5. The EVI differs from NDVI by attempting to correct for atmospheric and background effects. In addition, EVI is superior in discriminating subtle differences in areas of high vegetation density than NDVI because the latter tends to saturate (Didan et al., 2015; Ratana et al., 2005). The NDVI takes the form of:

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \tag{3.2}$$

where ρ_{red} and ρ_{NIR} are the surface bidirectional reflectance factors for MODIS bands 1 (620-670nm) and 2 (841-876nm).

To construct our forest survival dataset we collected the 16-days-250m image from the first period and compared it to the 16-days image from the second period of the same month. First, we used two masks in the analysis, namely the Goodness of Fit mask and

Land Cover mask. The Goodness mask was used to filter pixels flagged with bad quality. For the Land Mask, we resampled the rasters to 250m and applied the mask to differentiate pixels from land, forest, built-in and water. Through comparison it was possible to detect if the pixel with NDVI and EVI values survived from period one to period two within the month. This approach was undertaken for all the images corresponding to NDVI and EVI values for each 2 periods of the month of each year, which corresponds to 776 image analyses. For the leap years process we stopped at the land cover mask filtering process and we used the first period of 16-days as the main period. To compose the dataset we created a Boolean list of every image acknowledging if the pixel survived during that period or not, then we aggregated the list to year periods and computed the corresponding year of deforestation. To determine if the pixel survived, we used the rules outlined in Table 3.2. At the end of the process pixels were selected within the area studied (see Figure 3.1), and measured in distance from the artificial Legal Amazon line to the west and east portions of the municipalities. When the pixel had a variation greater than 0.1 within a month, it was labelled as a disturbance. Any disturbances below 0.1 are considered measurement error. Disturbances greater than 0.1 that result in negative changes in vegetation indices are considered as at least some deforestation having taken place in that pixel. We do not re-examine the pixel once it has been identified as having been deforested. With the identification of the time point of deforestation of a pixel we calculate the time of survival as the time since the beginning of our sample period. As a demonstrative example, we show the remaining pixels in the Legal Maranhão (LM) and Cerrado Maranhão (see Figures A.3.1 and A.3.2).

Table 3.2 Algorithm Assumption for NDVI and EVI values

$NDVI_1 > NDVI_2$	-> $NDVI_1 - NDVI_2$	Numbers (1) and (2) refer to the order of the period of the month
$NDVI_1 <= NDVI_2$	-> $NDVI_1 = NDVI_2$	Numbers (1) and (2) refer to the order of the period of the month. The second equation assumes that values did not change within the month and then the value assigned is from the last observation
$EVI_1 > EVI_2$	-> $EVI_1 - EVI_2$	Numbers (1) and (2) refer to the order of the period of the month
$EVI_1 <= EVI_2$	-> $EVI_1 = EVI_2$	Numbers (1) and (2) refer to the order of the period of the month. The second equation assumes that values did not change within the month and then the value assigned is from the last observation

To obtain the cloud cover dataset we first used the Goodness mask to create an image from the 16-days first and second periods containing only pixels flagged with clouds. Flagged pixels considered either period to create the final image. After creating the image for a month, we summed all month images in order to have the number of months whiting the period year that a pixel with a cloud upon it. With these we performed a Kernel Regression of the share of deforested pixels on the number of months with cloud cover to identify the threshold at which cloud coverage causes deforestation. To this end we used the optimal bandwidth suggested by Bowman and Azzalini (1997) with 1000 replications with cross-validation. This suggested a threshold of about 9 months for the whole sample period of seventeen years. With the identified threshold, we applied a third mask called

Cloud Mask to all the images processed of each year, considering only clouded pixels with 9 months persistence within the year. If within a year the pixel was equal to or exceeded the threshold, then it was flagged as a cloud persistence pixel; see algorithm implementation in A.3.3 for details. Finally, to integrate this information with the survival dataset framework, we create a Boolean list to transform the clouded pixels in the binary form. Considering the time period a forested pixel survived, we indicate as clouded if within the whole sample period the pixel had cloud at least in one year period. While this decision rule will inevitable identify some pixels as clouded after their deforestation, excluding them does not change our results in any noticeable manner. [1]

The risk factor (covariate) variables were acquired in shapefile format from different sources (see Table A.3.4). To get the distance from each pixel to the covariates we transformed the files into raster format and performed an Euclidean distance calculation from each pixel to the variable source within the studied area. The source variables were roads, protected areas, indigenous land, markets, municipalities centre, and, mining/mineral resources existent before or from 2000. For rivers, latitude, longitude, and elevation we simply used the distance to them. Time dependent variables were partitioned into five year averages in order to make their computation feasible. More precisely we calculated 5-year measures of neighbouring forested pixels and average levels of rainfall and temperature for the years 2001, 2006, 2011 and, 2016.

3.2.3 Survival Models

Survival analysis is a statistical method designed to study the amount of time an experimental unit survives. Originally, the event of interest was mortality and the analysis consisted of following the subjects until death. In recent years, the identification of risk

[1] We ran regressions taking into consideration pixels presenting cloud cover before deforestation only and the results don't change from what we are presenting in this chapter.

and/or prognostic factors related to survival have been applied much more widely. What distinguishes survival analysis from other areas in statistics is that survival data are usually censored (Cao, 2005; Lee and Wang, 2003). The defining feature of censored data is that the event time of interest is not always fully observed on all subjects under study. We consider three survival approaches in this study: the Kaplan Meier (KM) model, the Cox Proportional Hazard model (CPH), and the Extended Cox Proportional Hazard model (ECPH).

Let n be the total number of pixels of NDVI and EVI values whose survival times, censored or not, are available. Relabel the n survival times in order of increasing magnitude. Then

$$\hat{S}(t) = \prod_{t_{(r)} \leq t} \frac{n-r}{n-r+1} \tag{3.3}$$

where r runs through those positive integers for which $t_{(r)} \leq t$ and $t_{(r)}$ is not censored. The values of r are consecutive integers 1, 2,. . . , n if there are no censored observations; if there are censored observations, they are not. The estimated median survival time is the 50th percentile, which is the value of t at $\hat{S}(t)$ = 0.5. More precisely, it is a step function, and is the nonparametric maximum-likelihood estimator of the product limit estimates proposed by Kaplan and Meier (Lee and Wang, 2003).

Extending the analysis to the inclusion of risk factors and time dependent variables, the most common approach used is the Cox Proportional Hazard model which can handle censored and/or truncated observations (Cao, 2005; R., 1972).

Again, let n be the total number of pixels of NDVI and EVI values which consists of $t_{(j)}$, $\delta_{(j)}$, $z_{(j)}$, j = 1,2,. . . ,n, where $t_{(j)}$ is the time under study for the jth pixel, $\delta_{(j)}$ is the deforestation indicator ($\delta_{(j)}$ = 1 if the deforestation has occurred and $\delta_{(j)}$ = 0 if the pixel is right-censored), and $z_{(j)}$ is the vector of risk factor for the *j*th pixel that may affect the distribution of *X*, the time to deforest.

Let $h(t|z)$ be the hazard rate in the sub population with covariate value(s) z. The Cox proportional hazards regression model relates covariates to the hazard function as follows:

$$h(t|z) = h_0(t)c(\beta' z) \tag{3.4}$$

where $h_0(t)c(0)$ is the hazard function for the sub population with covariate value z = 0 and is called the baseline hazard function, β = (β_1,β_2,. . . , β_p) is a parameter vector of regression coefficients, $\beta' z = \sum_{i=1}^{p} \beta_k z_k$, and $c(.)$ is a fixed, known scalar function. This model is semi-parametric because the baseline hazard model is estimated non parametrically, while the risk factors are constrained by the parametric representation $c(\beta' z)$. The parametric function usually assumes the exponential form

$$c(\beta' z) = exp(\beta' z) = exp(\sum_{k=1}^{p} = \beta_k z_k) = e^{\sum_{k=1}^{p} = \beta_k z_k} \tag{3.5}$$

where

$$h(t|z) = h_0(t)c(\beta' z) = h_0(t)exp(\sum_{k=1}^{p} = \beta_k z_k) \tag{3.6}$$

The Cox model is often called a proportional hazards model because, if we look at two pixels with covariate values z_1 and z_2, the ratio of their hazard functions at time *t* does not depend on *t* and the hazard rates are proportional (Cao, 2005).

$$\frac{h(t|z_1) = h_0(t)exp(\beta' z_1)}{h(t|z_2) = h_0(t)exp(\beta' z_2)} = exp[(\beta' (z_1 - z_2))] \tag{3.7}$$

An issue might arise from the possibility of unobserved heterogeneity (spatial and temporal dimensions) that would result from a misspecification of the survival model. To control for these two types, we include the 5-year averages of the share of neighbouring pixels with forest remaining, rainfall, and temperature. Extending the Cox Proportional Hazard model by adding a time dependent variable, the basic model is done by replacing z by z(t) ($T_{(j)}$, $\delta_{(j)}$, [$z_{(j)}(t), 0 \leq t \leq T_j$]), so that

$$h(t|z_s, 0 \leq s \leq t) = h_0(t)exp[\sum_{k=1}^{p} \beta_k z_k(t)] \quad (3.8)$$

where T_j is time on study for the *j*th pixel, and $\delta_{(j)}$ the deforestation event for the *j*th pixel ($\delta_{(j)}$ = 1 if deforestation has occurred and, $\delta_{(j)}$ = 0 if the pixel was not deforested during the period (right-censored)). In addition, $z_{(j)}(t)$ corresponds to the vector of risk factors for the *j*th pixel which includes measures of elevation, and distance to rivers, mining, roads, markets, municipality centres, protected areas and, cloud persistence. We interact some of the controls and risk factors with a region dummy (Legal Maranhão = 1 and Cerrado Maranhão = 0).

Data Analysis

An important assumption of our analysis is that the two buffer zones that we created around the artificial line to isolate and compare pixels, are homogeneous biologically. To find further support for this we first calculated an effect size index for the two areas divided. Following Cohen (1977), we take differences of means expressed in terms of the pooled within areas standard deviation. The Index is interpreted in terms of the average percentile standing of one area relative to the other area. The test result shows an index of 0.2 which indicates that the mean of one area is at the 58th percentile of the buffer zone, i.e the dissimilarities for the two areas are close to zero. Since the chosen buffer zone may seem somewhat arbitrary, we also examined whether forests between the two regions differ right outside the buffer zone. More precisely, we isolated the areas 0.2 degrees just outside buffer zone on each side of the two regions and similarly tested their difference. The corresponding Cohen index was 0.59, which is at the 69th percentile, and thus one can reject the null hypothesis that there is no difference between these two regions. As a matter of fact, the index value of 0.59 indicates a distinction of 33% in the two distributions (Cohen, 1977).

The non parametric and semi parametric models presented in this study are estimated using MODIS satellite derived data as the dependent variable and several biophysical, economic and environmental spatial data for the covariates. Table A.3.2 provides the summary statistics for our response and risk factors variables and Table A.3.4 provides the variables' sources. Our sample contains approximately 530,000 observations (for the combined MA and LM region) with the time each pixel was deforested as well as eight influencing risk factors and several controls.

Overall, the average time a pixel in our sample becomes deforested is 2.6 years for NDVI values and 2.4 years for EVI values. Note that nearly 71% of the pixels are covered by clouds and that the average distance to protected areas is about 58 km^2 for the studied area. Markets, municipalities and roads have an average distance of 55 , 13, and 4 km^2, respectively. The region is extensively surrounded by rivers to which the average distance is less than 2 km^2. Particularly in the region of Maranhão there is a high concentration of mineral resources, where the average distance between the pixels and mines is equivalent to approximately 6 km^2. In addition, 40% of the pixels belong to the LM region and almost 97% of the analysed pixels were at some time deforested.

To conduct the analysis, we extensively used Python and R language and specific packages and modules, such as lifelines from Davidson-Pilon et al. (2018) and survival from Terry M. Therneau and Patricia M. Grambsch (2000); Therneau (2015). In survival analysis with censorship it is not appropriate to use a loss function like mean-squared-error or mean-absolute-loss. Instead, we use the concordance-index, also known as the c-index. This measure evaluates the accuracy of the ordering of predicted time. It is in fact a generalization of the AUC (area under the curve), another common loss function, and is interpreted similarly (Davidson-Pilon et al., 2018). More precisely, if the c-index has a value around 0.5, then it corresponds to the expected result from random predictions, whereas 1.0 is perfect concordance and 0.0 is perfect anti-concordance. Usually, survival models range from 0.5 to 0.7.

To validate the results, we perform k-fold validations. This entails splitting a training set from the data into k smaller sets (k=3, as default). A model is trained using $k-1$ of the folds as training data. The resulting model is validated on the remainder of the data, i.e., it is used as a test set to compute a performance measure such as accuracy. The performance measure reported by k-fold cross-validation is then the average of the values computed in the loop and should be close to 0 (Pedregosa et al., 2011). The c-index and k-fold validation results are computed in Table A.3.1. As a further validation process we compute variable importance with p-values for high dimensional data. For this task we applied random forests classification. This testing approach is based on a modified version of the permutation variable importance, which is inspired by cross-validation procedures (Janitza et al., 2016). With an unbiased variable importance measure, the importance values of non-associated variables vary randomly around zero. Thus, all non-positive importance values are assumed to correspond to these non-associated variables and are used to construct a distribution of the importance under the null hypothesis of no association to the response. Since only the non-positive values of this distribution can be observed, the positive values are created by mirroring the negative distribution. We calculated less than 250 permutations and added 1 to the numerator and denominator to avoid zero p-values (Wright and Ziegler, 2017). The results indicated that all variables are important to the model and, hierarchically, clouds, rivers, and roads are the most important measure in the model.

Statistical validation included several tests on non parametric and semi parametric fitters. For the KM estimator we conducted two different logrank tests. For the Cox proportional Hazard method we conducted Likelihood ratio, Wald, and score tests. All models showed good tests results rejecting the null hypothesis decisively.

3.3 Results

We first examine the non parametric estimation of Kaplan and Meier (1958) of deforestation using NDVI and EVI values applied to the two sites (LM and MA) distinguished by the Legal Amazon line. We secondly implement the semi parametric proportional hazard model proposed by R. (1972) and, as an extension, the extended proportional hazard model with time dependent variables. The results from the semi parametric analysis uses NDVI values because these responded better to our validation process[1].

3.3.1 Non parametric Analysis

As can be seen from Figure 3.2, the chance of survival of the forests through the analysis of pixels was higher in the Cerrado Maranhão (MA) than in Legal Maranhão (LM). For the 529,680 pixel analysis, the rate of survival is lower for EVI values in all settings. In 2001 the rate of survival for EVI pixels in the whole Region was about 59% as compared to 61% for NDVI pixels, but by 2010, the rate of survival decreased to 2% and 3.2%, respectively (see Table 3.3). The Log Rank test and Weighted Log Rank test confirms that the stratified samples differ in terms of distributions.

[1]Results from EVI values do not differ from the results presented here. The suitability of the vegetational index is based on validation process.

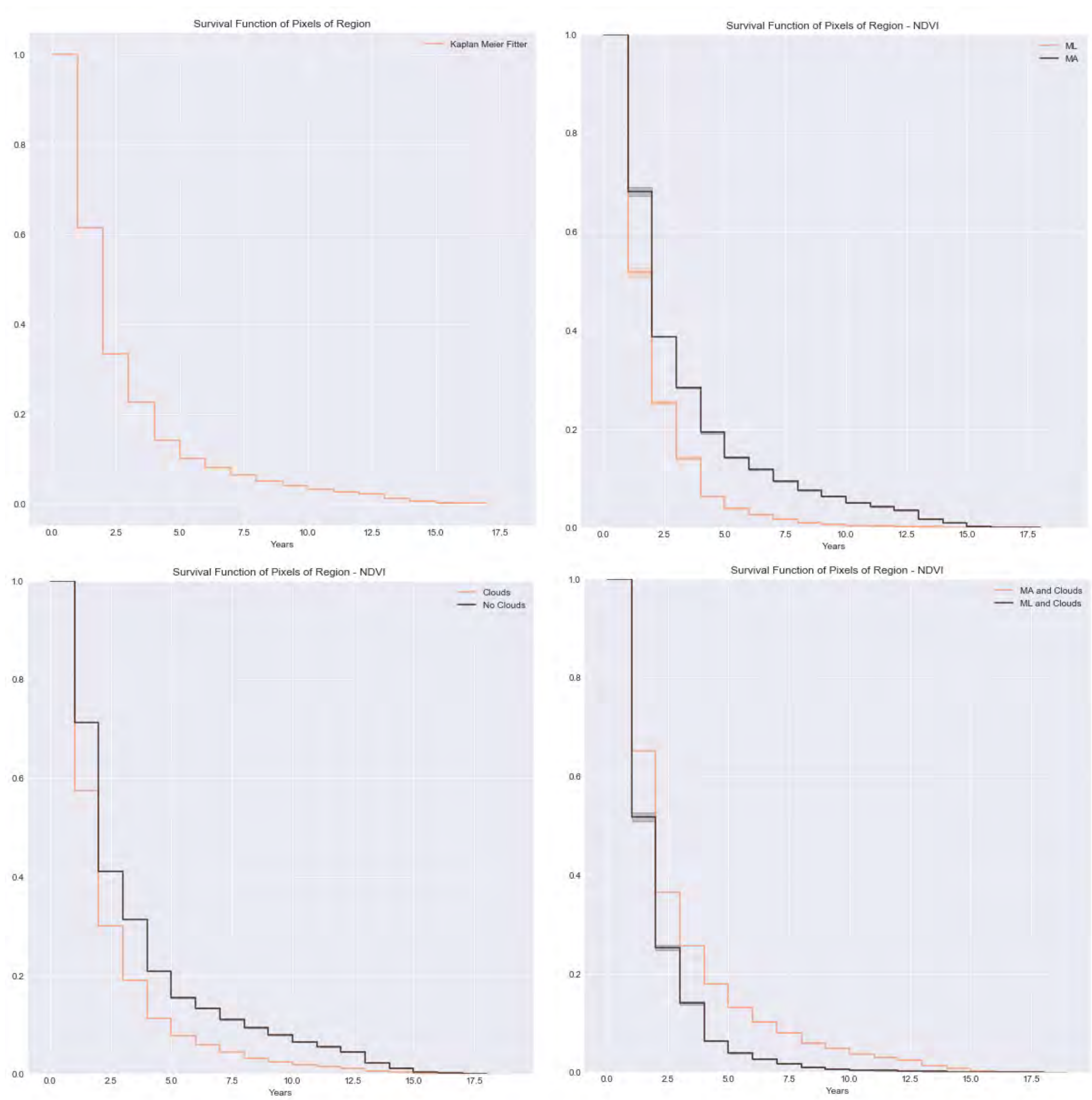

Figure 3.2 Kaplan Meier Fitter for NDVI values in the complete Region (LM+MA). Top left shows the KM fitter for the Region and, at the top right, the stratified KM fitter of the Region. Bottom left shows the KM fitter for the Region stratifying by cloud coverage. At the bottom right, it shows the KM fitter stratifying by Clouds and Region.

Table 3.3 Kaplan Meier Estimation

Time	Numbers at risk	Number of events	Survival	Standard Errors	lower 95% CI	upper 95% CI
Region						
2001	512017	197454	0.61436	6.80E-04	0.61303	0.6157
2005	71707	263281	0.10016	4.20E-04	0.09934	0.10098
2011	20594	34998	0.0318	2.45E-04	0.03133	0.03229
2015	3189	15180	0.00216	6.48E-05	0.00203	0.00229
Region MA						
2001	302075	96251	0.68137	0.000848	0.6797	0.683
2005	58271	162794	0.14245	0.000636	0.1412	0.1437
2011	19179	27662	0.05087	0.0004	0.0501	0.0517
2015	2946	14404	0.00319	0.000103	0.003	0.0034
Region LM						
2001	209942	101203	0.517948	1.09E-03	0.515815	0.52009
2005	13436	100487	0.039306	4.24E-04	0.038484	0.040146
2011	1415	7336	0.004363	1.44E-04	0.00409	0.004654
2015	243	776	0.000667	5.63E-05	0.000565	0.000787

The Table 3.3 shows the number of pixels (forested pixels) at risk of deforestation or disturbance. In 2001, for the total region, 512,017 pixels were at risk while 197,454 pixels had been deforested. The survival rate at 2001 was around 61%. For the Cerrado Maranhão (MA), the Kaplan Meier fitter estimation shows that from the 302,075 forested pixels in 2001, between 64% to 68% had a chance of survival considering both vegetation indices. By 2010 this decreased to a range of 3% to 5%. At the end of our sample period, the chance of survival for pixels within Cerrado Maranhão was around 0.3% (see upper left graph from Figure 3.3 and Figure 3.4) and Table 3.3. The median time of survival for the model was two years. This means that the half life of the sample pixels was around two years (the 50th percentile).

In calculating the survival curves by the presence of clouds one discovers that pixels with no clouds over the studied period had a higher rate of survival comparing to pixels with clouds. At the end of 2001, the rate of survival for pixels with no clouds was about 66.6% to 71.2% for both indices. In 2005, the rate lowered to around 15% and, in 2015, the chance of survival of the forests was approximately 0.4%. Looking at the survival curve of pixels covered by clouds, the chance of forest survival in 2001 was around 61% to 65%, then decreased to between 11% and 13% in 2005. The rate of survival for forests by 2015 was around 0.2% for both vegetation indices (see upper right graph from Figure 3.3 and Figure 3.4). The median time of survival for the survival curves stratified by clouds did not change from the original set.

We conducted a Log Rank test (Peto and Peto, 1972) to check whether the sub samples by cloud coverage were originated from the same distribution. The results showed that, in fact, these sub samples come from significantly different distributions. Since the majority of the pixels are deforested within two years, we also applied the generalized Wilcoxon test, also referred to as the Log Rank Weighted test because it gives more weight to earlier deforestation than later deforestation observations, while the Log Rank test gives equal weight to the whole deforestation process (Lee and Wang, 2003). The results reinforce the earlier findings that the two survival curves are different from each other.

The Legal Maranhão (LM) experienced a similar pattern in that the median time of survival of the pixels was two years. The Kaplan Meier fitter estimation shows that from the 209,942 forested pixels in 2001 at risk, 51% had a chance of survival considering both vegetation indices. By 2015, the number of pixels at risk dropped to 243 and the number of pixels deforested was about 776 giving a probability of survival of less than 0.07%. As apparent from the survival curve, the rate of survival was lower for the Legal Maranhão

(LM). Stratifying by cloud coverage, the survival curves show that pixels without the presence of clouds had a higher rate of survival for both indices (see lower right graph from Figure 3.3 and Figure 3.4). The rate of survival for pixels with no clouds was about 58.8% to 59.2% in 2001. then dropped to around 2.5% in 2010, and was reduced to a zero rate of survival of pixels by 2015. The chance of survival for pixels covered by clouds was around 51% in 2001 for both indices, i.e., almost 10% lower when compared to non-clouded pixels. In 2010, the rate of survival was less than 0.5% and in 2015 was equivalent to 0.065%.

We performed a Log Rank test (Peto and Peto, 1972) to check whether the two sub samples originated from the same distribution and the differences from these survival curves were due to other characteristics than the presence of clouds. The test result showed that the two sub samples are from different distributions. We was also applied the generalized Wilcoxon test (Lee and Wang, 2003) which rejected the null hypothesis that the two series came from the same distributional origin.

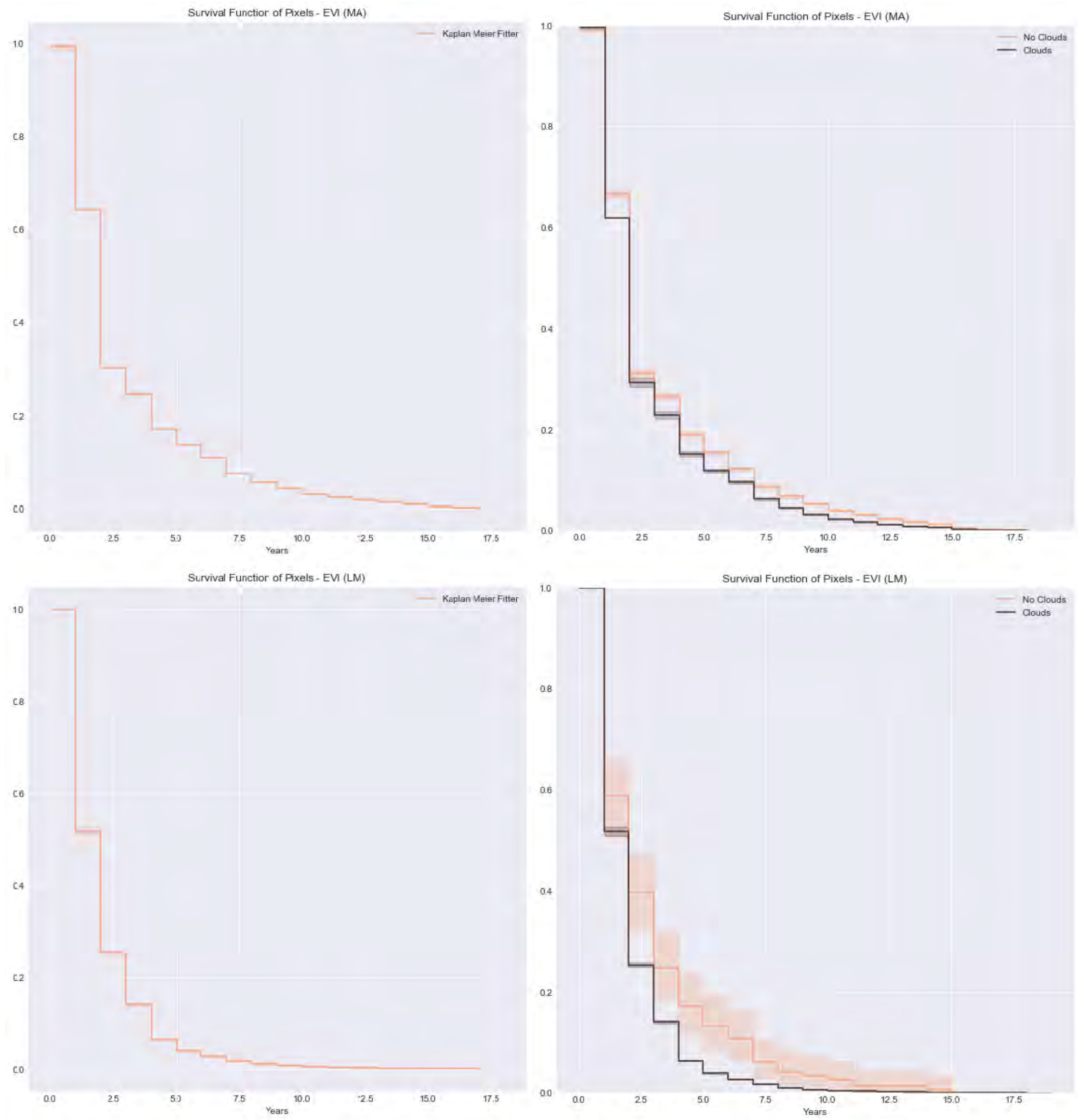

Figure 3.3 Kaplan Meier Fitter for EVI values in LM and MA region separately. Top left and right show the KM fitter for the Cerrado Maranhão and, at the bottom, for the Legal Maranhão . The top and bottom left show the sides KM fitter. The top right and bottom show the sample stratified by cloud coverage. The coloured band represents 99% confidence interval.

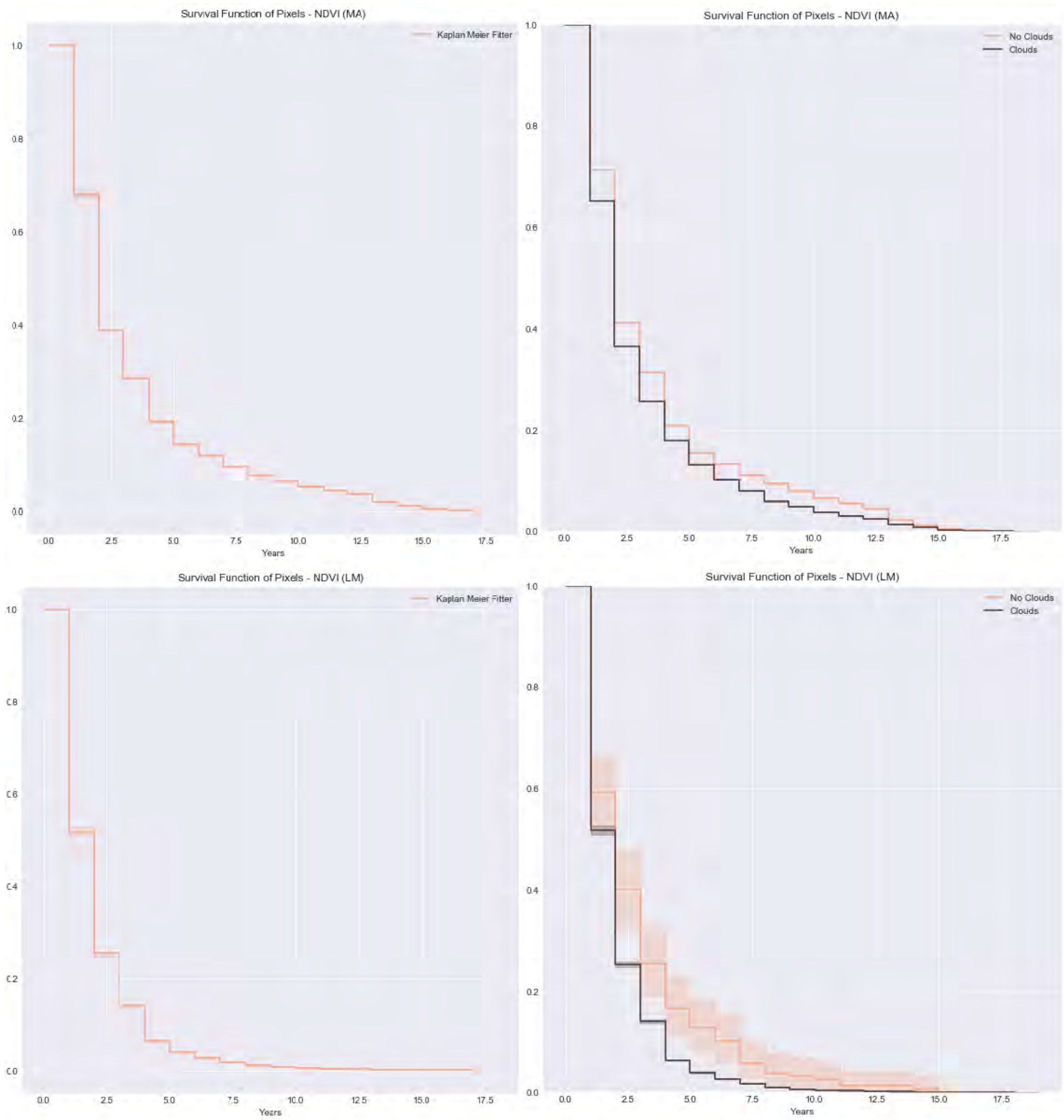

Figure 3.4 Kaplan Meier Fitter for NDVI values in LM and MA region separately. Top left and right show the KM fitter for the Cerrado Maranhão and, at the bottom, for the Legal Maranhão . The top and bottom left show the sides KM fitter. The top right and bottom show the sample stratified by cloud coverage. The coloured band represents 99% confidence interval.

3.3.2 Semi parametric Analysis

Studying forest loss and, in consequence, deforestation reveals that many factors potentially affect the process. In this sense, we expand the analysis to include risk factors (covariates) to evaluate the effect of these variables on forest survival (see 3.6).

We first start off by pooling the data across the two regions of interest. In this regard, the preferred model was the NDVI model considering the concordance index (0.755) and k-fold model validation (0.03). We also checked the proportional hazard assumption with plots of each estimated regression coefficient as a function of time through the smoothed scaled Schoenfeld residuals. The results showed that each risk factor curve presents a flat format suggesting that the proportional hazard assumption holds. To access the goodness of fit of this model, we plotted the graph of deviance residuals against the survival time. The results showed that the residuals were distributed about zero, indicating a good fit of the model.

Table 3.4 shows the estimated regression coefficients, their standard errors, the z values, and relative hazards. Accordingly, elevation has no effect on the relative hazard, while the presence of clouds per say increases the hazard by 1.5%. The interaction term between clouds and the Legal Maranhão region is not significant, however, on its own being on the Legal Maranhão side increases a cell's hazard by 29.4%. Thus while clouds and being on the Legal Maranhão side encourages deforestation, clouds appear to not encourage deforestation to any greater extent regionally. We also find that greater distance from protected areas and paved roads decreases hazard by 4.7% and 23.8%, respectively. On the other hand, greater distance from markets and municipalities centre increases hazard by 6.5% and 19.1%, respectively. This might reveal that urbanisation accounted for deforestation in these areas. On the other hand, distancing from markets increased the probability of a pixel being deforested. It comes to the knowledge that great part of the environmental and inspection police are located close to markets, specially those dealing with logger, grains and cereals products, so forests close to the municipalities had more visibility to be inspected and accessed and then preserved (Rural, 2018). Latitude and longitude show an interesting result, forests close to the coast have a higher chance of survival comparing to forests country-inside. The relative hazard decreases by 7.1% for forests close to the coastline. Distancing from the artificial line the relative hazard also increase, in this sense, forests far from the institutional line have a probability 1.29 higher of deforestation comparing to forests close to the border.

Table 3.4 Cox Proportional Hazard Model - Region (LM+MA)

Variables	Regression Coefficient	Relative Hazards	Standard Errors	z score	Pr (>|z|)
Clouds	0.015	1.015	0.004	3.634	3E-04 ***
Pas	-0.049	0.953	0.010	-4.912	9E-07 ***
Mining	0.340	1.405	0.022	15.633	< 2E-16 ***
Elevation	0.001	1.001	0.003	0.173	0.86
Markets	0.063	1.065	0.010	5.998	2E-09 ***
Municipalities	0.175	1.191	0.026	6.842	8E-12 ***
Rivers	0.252	1.286	0.114	2.199	0.03 *
Roads	-0.271	0.762	0.043	-6.250	4E-10 ***
Lat	-0.074	0.929	0.004	-19.147	< 2E-16 ***
Lon	0.254	1.290	0.013	19.407	< 2E-16 ***
ML	0.258	1.294	0.080	3.220	1E-03 **
Cloud*LM	0.089	1.093	0.080	1.107	0.27

Signs stand for '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 and denote hazard ratios that are significantly different from 1 at 99%, 95% and 90% confidence levels. The model contains controls for grouped year periods. PAs stand for Protected Areas (Indigenous Lands and Conservational Units).

We performed a likelihood ratio test to check if the model could be improved by splitting the sample into different regions as we did with the non parametric analysis. The results indicate that we can improve the model by dividing the sample area according to the region. Analysing the Cerrado Maranhão (MA) sample separately, the NDVI model was preferred in many settings when accounting for concordance index (0.767) and k-fold model validation (0.30) (see Table A.3.1. The diagnostics of the model corroborated with the assumption of proportional hazard during the studied period and the deviance residuals exhibit a good fit. Table 3.5 shows the results of the specified model. Pixels covered by clouds have a decreased risk of the pixel being deforested by 2% compared to those not covered by clouds. Pixels far from rivers increase the relative hazard by 55%, while greater distance from protected areas increases the risk of deforestation by 4%. Similarly, being further away from mines results in an increase of 32% in the relative hazard. Being one degree further away from markets and municipalities increases the hazard by 6% and 8%, while a greater distance to roads decreases the relative risk of the pixels being deforested in that area by 48%.

Table 3.5 Cox Proportional Hazard Model - Cerrado Maranhão (MA)

Variables	Regression Coefficient	Relative Hazards	Standard Errors	z score	Pr (>|z|)
Clouds	-0.015	0.98	0.004	-3.65	3E-04 ***
PAs	0.036	1.04	0.017	2.05	0.041 *
Mining	0.275	1.32	0.028	10.00	< 2E-16 ***
Elevation	0.020	1.02	0.005	4.18	3E-05 ***
Markets	0.061	1.06	0.017	3.48	0.001 ***
Municipalities	0.077	1.08	0.033	2.34	0.019 *
Rivers	0.436	1.55	0.150	2.90	0.004 **
Roads	-0.660	0.52	0.052	-12.61	< 2E-16 ***
Lat	-0.149	0.86	0.007	-21.59	< 2E-16 ***
Lon	0.258	1.29	0.017	15.21	< 2E-16 ***

Signs stand for '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 and denote hazard ratios that are significantly different from 1 at 99%, 95% and 90% confidence levels. The model contains controls for grouped year periods. PAs stand for Protected Areas. Indigenous Lands and Conservational Units.

Following the same selection procedure adopted for the previous model, in terms of the two vegetation indices models for Legal Maranhão (LM) the NDVI model was favoured considering the c-index and the k-fold validation. Table 3.6 shows the results of the proportional hazard model for the Maranhão region under environmental policy (LM). Clouds are not significant in this model and elevation has a constant effect on the relative hazard. Distancing one unit from protected areas, the risk of deforestation increases by almost 9%. One possible reason could be that PAs, specially conservational units, are established in response to previous deforestation pressure which is then displaced to neighbouring areas just outside the PAs and is therefore capturing the presence of active groups of loggers in a closed area. The second possible explanation regards to the fact that there are potentially limited public monitoring and enforcement in this transitional forest area, specially in the not satellite monitored side.[1]

Greater distance from river basins decreases the hazard by approximately 34%. Distancing from roads in the policy area increases hazard by 25%. Roads as a driver of deforestation is also discussed in the literature (Baynard et al., 2012; Cropper et al., 2001; Mani and Griffiths, 1997; Pfaff et al., 2007) which support this finding with regard to roads. The Legal Maranhão (LM) deviance residuals showed that the model presents a good fit with no pattern detected. The scaled Schoenfeld residuals also show a flat curve for all the risk factor variables. These results support the assumption of the proportional hazard model.

[1]These possibilities were observed in the Conservational Unit of *Morros Garapenses* on the Cerrado Maranhão (MA) side. The unit was created in 2008 with the aim to protect the diversity of representative regional ecosystems that functioned as habitat for native and migratory species, as well as wildlife refuges from areas already devastated by human activities (ISA, 2018).

Table 3.6 Cox Proportional Hazard Model - Legal Maranhão (LM)

Variables	Regression Coefficient	Relative Hazards	Standard Errors	z score	Pr (>lzl)
Clouds	0.108	1.113	0.080	1.344	0.179
PAs	0.081	1.084	0.016	5.137	3E-07 ***
Mining	-0.019	0.981	0.049	-0.387	0.699
Elevation	-0.011	0.989	0.004	-2.498	0.012 *
Markets	-0.023	0.977	0.015	-1.588	0.112
Municipalities	-0.070	0.932	0.045	-1.562	0.118
Rivers	-0.413	0.662	0.179	-2.307	0.021 *
Roads	0.224	1.251	0.085	2.645	0.008 **
Lat	-0.046	0.955	0.006	-8.218	2E-16 ***
Lon	0.282	1.325	0.025	11.396	< 2E-16 ***

Signs stand for '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 and denote hazard ratios that are significantly different from 1 at 99%, 95% and 90% confidence levels. The model contains controls for grouped year periods. PAs stand for Protected Areas. Indigenous Lands and Conservational Units.

Up until now the proportional hazards models were assumed to have risk factors independent of time. However, in practice, deforestation is a process that might have spatially time dependent factors. In this sense, we extend the Cox proportional Hazard model for the Region and sampled areas.

To incorporate the spatially and temporal dependence, we include the ratio of neighbouring forested pixels for each pixel in four different periods (2001, 2006, 2011, 2016). We also include average values of rainfall and temperature distributed within the specified years and interact them with policy region of Legal Maranhão. Table 3.7 shows the results from this setting. Clouds increase the chance of deforestation by 1.8%. A one unit distance from protected areas decreases the hazard by 4.4%. Distancing one unit from mines, increase hazard by 20% and elevated areas increased hazard by 3.5%. A degree distance from roads decrease hazard by almost 18%. Clouded pixels in the Legal Maranhão region do not have an effect on hazard. Observing the time dependent variables, having neighbouring pixels with deforestation increase the risk of being cleared by 31%. Higher rainfall and temperature in the whole region decreases the hazard by 0.1%. Finally, rainfall and temperature interacted with the Legal Maranhão dummy show that this region has a lower precipitation effect (0.1%) and higher temperature impact (1.4%) on the deforestation hazard. This suggests, as might be expected, that forested pixels in the Legal Maranhão had a higher chance of survival during dry season.

Table 3.7 Cox Proportional Hazard Model Time Dependent - Region (LM+MA)

Variables	Regression Coefficient	Relative Hazards	Standard Errors	z score	Pr (>|z|)
Clouds	0.018	1.018	0.01	3.20	0.001 **
Pas	-0.045	0.956	0.01	-3.07	0.002 **
Mining	0.187	1.205	0.03	6.40	2E-10 ***
Elevation	0.034	1.035	0.00	8.42	< 2E-16 ***
Markets	0.024	1.025	0.02	1.56	0.118
Municipalities	-0.057	0.945	0.03	-1.66	0.097 .
Rivers	-0.452	0.637	0.155	-2.915	0.003 **
Roads	-0.197	0.821	0.06	-3.36	0.001 ***
Lat	0.000	1.000	0.01	-0.06	0.952
Lon	-0.034	0.966	0.02	-2.05	0.041 *
ML	0.293	1.341	0.10	3.06	0.002 **
Cloud*LM	0.018	1.018	0.01	1.52	0.130
Neighbours	0.274	1.315	0.03	10.44	< 2E-16 ***
Rainfall	-0.001	0.999	0.00	-6.47	1E-10 ***
Rainfall*LM	0.001	1.001	0.00	6.61	4E-11 ***
Temperature	-0.006	0.994	0.00	-2.28	2E-02 *
Temperature*LM	-0.014	0.986	0.00	-5.36	8E-08 ***

Signs stand for '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 and denote hazard ratios that are significantly different from 1 at 99%, 95% and 90% confidence levels. The model contains controls for grouped and time-dependent year periods. PAs stand for Protected Areas. Indigenous Lands and Conservational Units.

We also checked whether we should split our sample into different regions. To this end we employed a likelihood ratio test comparing the base specification with that where all covariations are interacted with our region dummy. The test statistic clearly indicated that the covariates had different impacts across regions and we thus estimated our survival model separately for each region, where the Cerrado Maranhão (MA) results are presented in Table 3.8 and the Legal Maranhão (LM) results are shown in Table 3.9. The results from the area not monitored shows that the presence of clouds has no effect on the relative hazard. Forested pixels distant from protected areas decrease the risk of being deforested by 18%. Greater distance from roads and rivers also decreases the hazard by 23% and 33%, respectively. Examining the time dependent controls, one can see that forested pixels surrounded by deforested pixels double the relative hazard. Rainfall decreases the risk of the pixel being deforested by 0.1% and high temperature increases the chance of deforestation by 6.3%. The interpretation relies possibly in the fact that clear cutting process became costly when intensive rain occurred. When temperature increased, the risk of deforestation also increased, differing from the previous findings.

Table 3.8 Cox Proportional Hazard Model Time Dependent - Cerrado Maranhão (MA)

Variables	Regression Coefficient	Relative Hazards	Standard Errors	z score	Pr (>lzl)
Clouds	0.003	1.003	0.006	0.534	0.593
PAs	-0.207	0.814	0.024	-8.446	< 2E-16 ***
Mining	0.055	1.057	0.041	1.352	0.176
Elevation	0.066	1.069	0.006	10.581	< 2E-16 ***
Markets	0.177	1.193	0.026	6.881	6E-12 ***
Municipalities	-0.243	0.784	0.045	-5.373	8E-08 ***
Rivers	-0.455	0.635	0.204	-2.232	0.026 *
Roads	-0.269	0.764	0.072	-3.713	2E-04 ***
Lat	-0.282	0.755	0.011	-24.792	< 2E-16 ***
Lon	0.155	1.167	0.025	6.180	6E-10 ***
Neighbours	0.742	2.099	0.040	18.379	< 2E-16 ***
Rainfall	-0.009	0.991	0.000	-26.481	< 2E-16 ***
Temperature	0.062	1.063	0.003	19.476	< 2E-16 ***

Signs stand for ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 and denote hazard ratios that are significantly different from 1 at 99%, 95% and 90% confidence levels. The model contains controls for grouped and time-dependent year periods. PAs stand for Protected Areas. Indigenous Lands and Conservational Units.

In the surveilled area (Table 3.9) the pattern is different. More precisely, the presence of clouds increases the risk of a pixel being deforested by 3%. One unit distance from markets and mining projects increases the hazard by 9.7% and 1.6%, accordingly, being located to neighbouring deforested pixels increases the hazard by 33% and the increasing level of temperature decreases the relative risk of a pixel being deforested by 2.1%. Interestingly, roads is not significant in the monitored area when controlling for time and space. This aspect might relate to several actions implemented in the Legal Amazon to detain deforestation through roads surveillance which refrained deforestation process along major roads.[1] Moreover, increasing one unit of temperature decreases the relative hazard

[1] ’Arc of Fire’ special force was created in 2007 by the federal government to combat the increased pace of deforestation in the Amazon recorded from INPE results (BRASIL, 2018, 2012).

by 2.1% which corroborates for the fact that pixels in the Legal Maranhão had a higher chance of survival during dry season.

Table 3.9 Cox Proportional Hazard Model Time Dependent - Legal Maranhão (LM)

Variables	Regression Coefficient	Relative Hazards	Standard Errors	z score	Pr (>|z|)
Clouds	0.030	1.030	0.011	2.592	0.010 **
Pas	-0.060	0.941	0.020	-3.055	0.002 **
Mining	0.175	1.191	0.050	3.518	4E-04 ***
Elevation	0.016	1.016	0.006	2.724	0.006 **
Markets	0.093	1.097	0.021	4.497	7E-06 ***
Municipalities	-0.102	0.903	0.054	-1.901	0.057 .
Rivers	-0.090	0.914	0.241	-0.374	0.708
Roads	0.026	1.026	0.103	0.252	0.801
Lat	0.034	1.034	0.008	4.149	3E-05 ***
Lon	-0.145	0.865	0.027	-5.338	9E-08 ***
Neighbours	0.289	1.335	0.034	8.37	< 2E-16 ***
Rainfall	0.000	1.000	0.000	1.075	0.283
Temperature	-0.021	0.979	0.001	-20.965	< 2E-16 ***

Signs stand for '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 and denote hazard ratios that are significantly different from 1 at 99%, 95% and 90% confidence levels. The model contains controls for grouped and time-dependent year periods. PAs stand for Protected Areas. Indigenous Lands and Conservational Units.

Settlements

We have examined the Region (MA and LM), considering their spatial definition. To further investigate whether the effectiveness of environmental policy was undermined by the presence of clouds we also look specifically at the survival time of forested pixels within settlements. The reason for doing so is that these are known to be allocated and supervised by Brazil's Special Secretary of Agrarian Development, and that there is virtually no law enforcement in these areas. More specifically, usually any deforestation fines are sent to the INCRA and not to the settler. Since the justice system claims that this agrarian agency does not commit an environmental crime because it leaves part of the forest of the whole settlement intact, it cannot be obliged to detain forest clearing. While, under pressure by public opinion, INCRA established in 2012 the 'Green Settlement Program' to deal with this problem of environmental debt of settlements, as of date this policy has still not implemented Pacheco (2009); Schneider and Peres (2015). Under these circumstances one would thus not expect that clouds play a role in deterring deforestation, since in practise there is no penalty for doing so. To verify this we identified, within our region of study, settlements areas from both sides MA and ML (see figure 3.5) to check whether the time of survival of the pixels differed or was equal to the rest of the region and under cloud coverage.

The survival analysis in settlements follows the Kaplan Meier fitter presented in Section 3.3.1 and the Cox proportional hazard model with time dependent variable as demonstrated in Section 3.3.2.

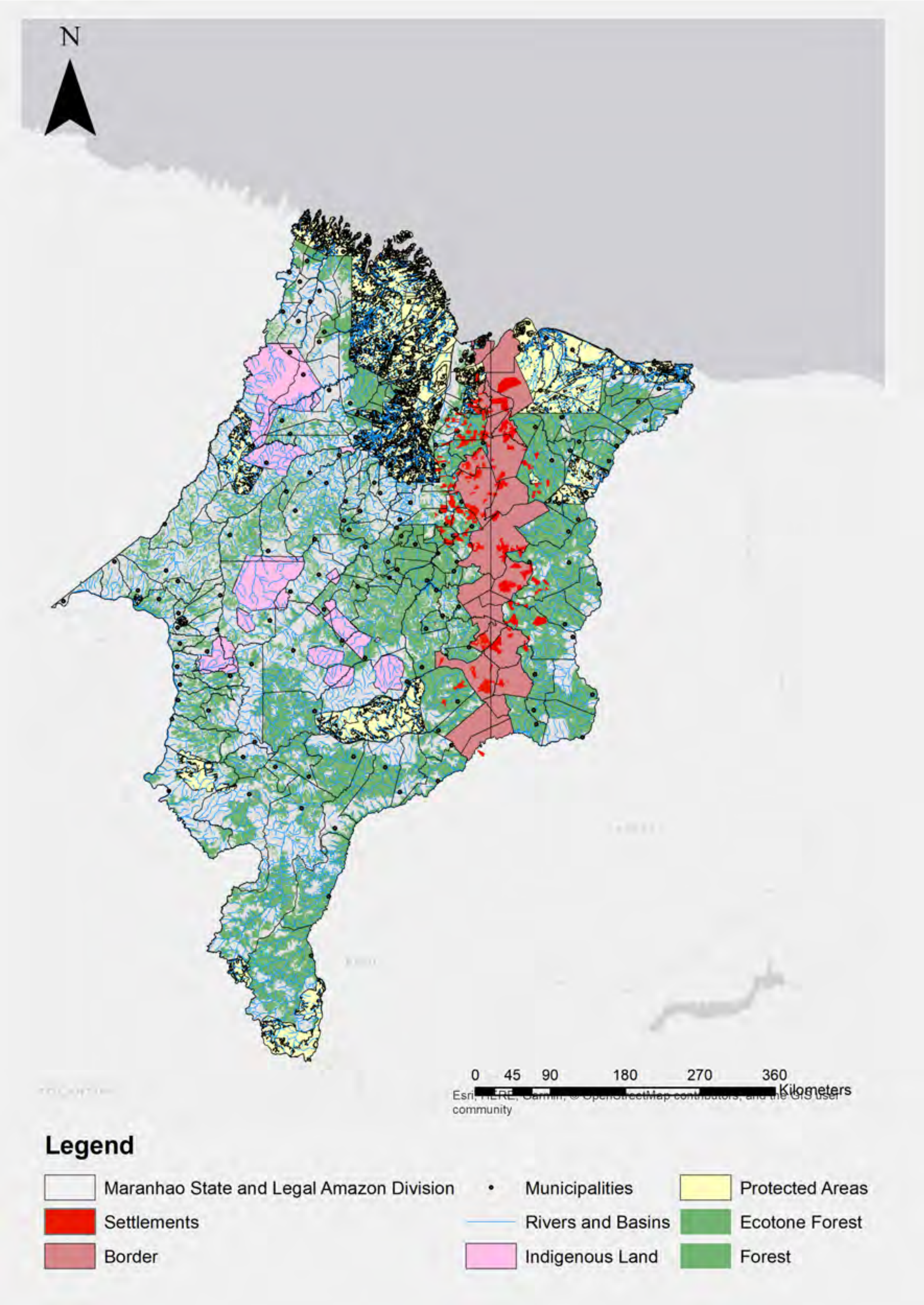

Figure 3.5 Maranhão state and settlements within the studied municipalities. Source: (MMA, 2018b; Núcleo Geoambiental - NUGEO, 2018).

The Kaplan Meier fitter shows that from the 37,532 pixels, in 2001, about 55% of pixels had a chance of survival considering both vegetation indices. For 2010, the chance of survival decreased to the range of 5.8% to 6.1%. At the end of the studied period, the chance of survival for pixels within settlements were around 0.4% and 0.6% (see upper and bottom left graph from Figure 3.6). The median time of survival for the model was two years. This means that the half life of the sample pixels was around two years (the the 50th percentile). Sub-setting the survival curves by region, it's possible to identify that pixels within Legal Maranhão (LM) had a higher rate of survival comparing to pixels within Cerrado Maranhão (MA). The presence of clouds in this subset still shows that the chance of survival was higher in the Legal Maranhão (not shown).

We performed a Log Rank test (Peto and Peto, 1972) to check whether the two sub samples were originated from the same distribution and the differences from these survival curves are produced from other characteristics than the region. The results suggested that these sub samples came from the same distribution which cannot reject the null hypothesis. However, given that the majority of the settlement pixels are deforested within two years, we also applied the generalized Wilcoxon test, which can be also referred to Log Rank Weighted test, because it gives more weight to early deforestation than later deforestation, while a Log Rank test gives equal weight to all deforestation process (Lee and Wang, 2003). From the Log Rank Weighted test, at 5% level, the two sub samples actually come from different distribution.

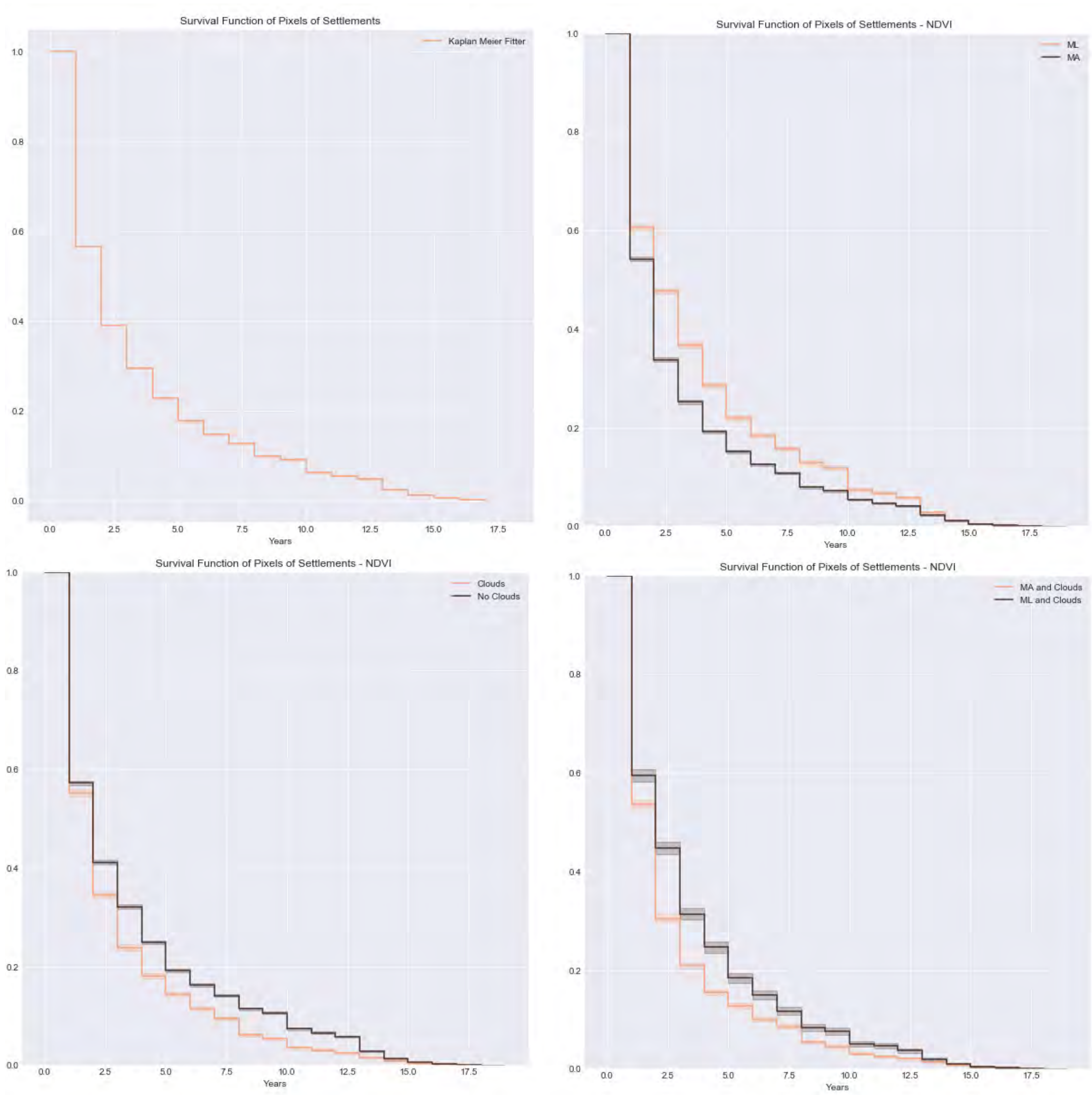

Figure 3.6 Kaplan Meier Fitter for NDVI values in the Settlements areas of Cerrado Maranhão (MA) and Legal Maranhão (LM). Top left shows the KM fitter for the Region and, at the top right, the stratified KM fitter of the Region. Bottom left shows the KM fitter for the Region stratifying by cloud coverage. At the bottom right, it shows the KM fitter stratifying by Clouds and Region.

The results from the proportional hazard model is shown in Table 3.10. The presence of clouds decreases hazard by 5%. Elevation has a constant relative hazard. Distancing from municipalities centre decreases hazard by almost 27%. Forested pixels within settlements and under the monitored policy area has a decreasing risk of being deforested comparing to non monitored areas. However, the presence of clouds in settlements located in the Legal Maranhão side increases the hazard by 4.2%.

The results from the proportional hazard model are shown in Table 3.10. The presence of clouds decreases the hazard by 5%. Elevation has a constant relative hazard. Distancing from municipalities centre decreases hazard by almost 27%. Forested pixels within settlements and under the monitored policy area hves a decreasing risk of being deforested comparing to non monitored areas. However, the presence of clouds in settlements located in the Legal Maranhão side increases the hazard by 4.2%, as can be seen from the coefficient on the interaction term between our regional dummy and cloud variable.

Table 3.10 Cox Proportional Hazard Model - Settlements

Variables	Regression Coefficient	Relative Hazards	Standard Errors	z score	Pr (>|z|)
Clouds	-0.052	0.950	0.014	-3.778	2E-04 ***
Pas	-0.011	0.989	0.045	-0.241	0.809
Mining	-0.419	0.658	0.085	-4.933	8E-07 ***
Elevation	0.000	1.000	0.000	3.309	0.001 ***
Markets	0.026	1.027	0.029	0.903	0.366
Municipalities	-0.306	0.737	0.084	-3.654	3E-04 ***
Rivers	1.431	4.183	0.345	4.148	3E-05 ***
Roads	0.234	1.264	0.139	1.685	0.092 .
Lat	-0.154	0.857	0.018	-8.76	< 2E-16 ***
Lon	0.332	1.393	0.048	6.944	4E-12 ***
ML	-0.079	0.924	0.022	-3.525	4E-04 ***
Clouds*LM	0.041	1.042	0.021	1.995	0.046 *

Signs stand for '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 and denote hazard ratios that are significantly different from 1 at 99%, 95% and 90% confidence levels. The model contains controls for grouped year periods. PAs stand for Protected Areas. Indigenous Lands and Conservational Units.

A likelihood ratio test again suggested that it was better to divide the sample area according to the region. The results are presented in Tables 3.11 and 3.12. For the non surveilled area, the presence of clouds decreased hazard by almost 4%. Also increasing the distance to protected areas and mining centres by one degree decreases the hazard by 56% and 50%, respectively. For the Legal Maranhão side, the presence of clouds has no effect on the relative hazard. Differently from the neighbouring area, distancing from protected areas and mining centres increase hazard by 38% and 76%. Distancing from urban and market areas has a negative impact on the hazard. For markets the relative hazard decreases by 10% and for municipalities centre the amount decreased is almost 56%.

Table 3.11 Cox Proportional Hazard Model - Settlements (MA)

Variables	Regression Coefficient	Relative Hazards	Standard Errors	z score	Pr (>|z|)
Clouds	-0.038	0.963	0.014	-2.665	0.008 **
PAs	-0.626	0.535	0.081	-7.69	0.000 ***
Mining	-0.684	0.504	0.100	-6.855	0.000 ***
Elevation	0.000	1.000	0.000	-0.915	0.360
Markets	0.535	1.708	0.057	9.469	< 2E-16 ***
Municipalities	-0.358	0.699	0.108	-3.328	0.001 ***
Rivers	0.997	2.711	0.448	2.228	0.026 *
Roads	0.266	1.305	0.156	1.702	0.089 .
Lat	0.043	1.044	0.030	1.452	0.147
Lon	0.446	1.562	0.060	7.407	0.000 ***

Signs stand for '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 and denote hazard ratios that are significantly different from 1 at 99%, 95% and 90% confidence levels. The model contains controls for grouped year periods. PAs stand for Protected Areas. Indigenous Lands and Conservational Units.

Table 3.12 Cox Proportional Hazard Model - Settlements (LM)

Variables	Regression Coefficient	Relative Hazards	Standard Errors	z score	Pr (>lzl)
Clouds	0.006	1.006	0.017	0.340	0.734
PAs	0.323	1.382	0.084	3.849	0.000 ***
Mining	0.567	1.762	0.222	2.558	0.011 *
Elevation	0.001	1.001	0.000	3.616	0.000 ***
Markets	-0.097	0.907	0.049	-1.976	0.048 *
Municipalities	-0.615	0.541	0.170	-3.614	0.000 ***
Rivers	1.377	3.963	0.581	2.372	0.018 *
Roads	-0.543	0.581	0.345	-1.573	0.116
Lat	-0.236	0.790	0.028	-8.442	< 2E-16 ***
Lon	0.471	1.602	0.110	4.300	2E-05 ***

Signs stand for '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 and denote hazard ratios that are significantly different from 1 at 99%, 95% and 90% confidence levels. The model contains controls for grouped year periods. PAs stand for Protected Areas. Indigenous Lands and Conservational Units.

The extended Cox model presented in equation 3.8 is also applied to the settlements. Accordingly, the presence of clouds has no effect in the presence of the relative hazard. In contrast, a one unit greater distance from protected areas and mines increase the relative hazard by 12% and 22% , as can be seen from Table 3.13. Municipalities centre are not significant in the model. The relative hazard decreases 80% if the pixels are far from markets. Rivers and roads have no impact on relative hazard. The incidence of clouds in settlement pixels within the Legal Maranhão also has no impact on the relative hazard. Increasing the level of rain decreases the hazard by 0.01%. Even though an increasing the level of temperature increases the relative risk of a pixel being deforested, increasing the temperature in the Legal Maranhão decreases the hazard by almost 3%.

The analysis for the two regions separated suggests similar patterns. Clouds have no effect on the relative hazard. In contrast, neighbouring forested areas increase the hazard by 15% in the Legal Maranhão, and twice the relative hazard in the Cerrado Maranhão.

Increasing rainfall levels will decrease the hazard in both sides by approximately 2%, while higher levels of temperature increase hazard in both regions ranging from 1 to 5%.

Table 3.13 Cox Proportional Hazard Model Time Dependent - Settlements

Variables	Regression Coefficient	Relative Hazards	Standard Errors	z score	Pr (>lzl)
Clouds	0.031	1.032	0.020	1.535	0.125
PAs	0.116	1.123	0.066	1.766	0.077 .
Mining	0.573	1.774	0.111	5.162	2E-07 ***
Elevation	0.000	1.000	0.000	2.376	0.018 *
Markets	-0.218	0.804	0.042	-5.232	2E-07 ***
Municipalities	0.067	1.069	0.115	0.582	0.561
Rivers	0.873	2.393	0.735	1.188	0.235
Roads	-0.293	0.746	0.191	-1.53	0.126
Lat	-0.106	0.900	0.026	-4.074	5E-05 ***
Lon	0.016	1.016	0.057	0.288	0.773
ML	0.573	1.773	0.111	5.162	2E-07 ***
Clouds*LM	-0.028	0.972	0.028	-0.99	0.322
Neighbours	0.317	1.373	0.111	2.863	0.004 **
Rainfall	-0.003	0.997	0.001	-5.331	1E-07 ***
Rainfall*LM	0.001	1.001	0.001	0.859	0.390
Temperature	0.025	1.025	0.004	6.855	7E-12 ***
Temperature*LM	-0.021	0.979	0.004	-5.476	4E-08 ***

Signs stand for '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 and denote hazard ratios that are significantly different from 1 at 99%, 95% and 90% confidence levels. The model contains controls for grouped and time-dependent year periods. PAs stand for Protected Areas. Indigenous Lands and Conservational Units.

Finally, with the settlements analysis, it is likely that the environmental policy could not detain deforestation during the last two decades in the ecotonic region of Maranhão. According to the results, the forests inside the surveillance area had a lower probability of survival comparing to the area not covered by the policy. The presence of clouds along with climatic variables was an active barrier to the legal compliance and, since the studied area has no systematic differences, we can agree on the lack of effectiveness due to cloud barrier. Since these barrier indicate different form of seasonality, the probability of a pixel being deforested in the rainy season is higher comparing to dry season which in turn evidence the behaviour change. We deduce that our results justify the undermining effect of the environmental policy caused by the changing behaviour and the presence of the artificial line.

3.4 Conclusions

In this study, we quantified the effect of cloud coverage, which inhibits satellite detection of disturbances to local forest, on local levels of deforestation in the Brazilian Amazon. To this end we focused on the ecological tension zone of Maranhão because it is divided by an artificial line that separates it in two parts, one for which there is satellite monitoring (the Legal Amazon Maranhão) and one (Cerrado Maranhão) for which no such monitoring was in place. Thus, arguably the role of cloud coverage in local forest losses should be different across this synthetic border. Identification of local deforestation was done through an algorithm to capture Vegetation Indices changes over time using satellite images. Similarly, satellite images were used to detect corresponding local cloud coverage. To estimate the impact of the latter on the former we employed a survival analysis on homogeneously forested regions near the border. Our results showed that clouds indeed played an important role in encouraging deforestation on the side that is under the satellite-monitoring program, while they were not determinants of forest loss in cells across the border. This was further supported by examining settlement areas on both sides of the border where there was also no satellite detection in place. This suggests that clouds inhibited the detection process of the satellite monitoring program in place even though we could not capture significance when interacting the environmentally controlled region with cloud cover.

We acknowledge a number of limitations to the resulted analysis. Firstly, the model implicitly assumes that in the first year of the study all pixels were fully occupied by forest, which in reality might not be true. It is also possible that the spatial distribution of the vegetation indices may exhibit dynamic behaviour over time, so that a potential area may or may not be sparsely vegetated for a certain period (e.g., during sampling) due to progressive succession of attacks of pests and diseases, which could not be distinguished from true permanent deforestation by the vegetation indices. Secondly, the study was based on coarse image resolution which could neglect local changes (< 250m) in the sample area. Thirdly, our results may not be generalised to other areas, such as dense tropical forests and sparse open fields. Fourthly, since our measure of forest disturbance is based on the remainder after controlling for seasonality, it is inevitably just a proxy and thus subject to measurement error. Our fifth limitation exposes the computational constraints existent for this analysis. For instance, it would actually improve the model if we considered yearly climatic control variables for our analysis. However, computing yearly levels of rainfall and temperature were beyond the processing time available. In addition, we only considered for the analysis covariates existent before or from 2000, which thus excluded roads, protected areas, indigenous land, markets, municipalities centre, and, mining/mineral resources created or discovered after 2000. Finally, we acknowledge that the models were derived from NDVI values and that one could alternatively have used EVI values, which in some instances might be more suited for ecotone forests (Bayma and Sano, 2015; Didan et al., 2015; Ratana et al., 2005).

Appendix A.3

Table A.3.1 Model Selection and Validation

Models	Concordance index	Std Deviation	K-fold (k=3)	Std Deviation
NDVI MA	0.767	0.001	0.30	0.0004
NDVI ML	0.691	0.001	0.69	0.0008
NDVI Region	0.755	0.001	0.03	0.0004
EVI MA	0.735	0.001	0.39	0.0011
EVI ML	0.691	0.001	0.69	0.0008
EVI Region	0.732	0.001	0.08	0.0008
NDVI Sett MA	0.780	0.003	0.77	0.0040
NDVI Sett ML	0.798	0.003	0.79	0.0019
NDVI Sett Region	0.785	0.002	0.78	0.0013
NDVI Buffer MA	0.582	0.005	0.58	0.0066
NDVI Buffer ML	0.689	0.003	0.68	0.0052
NDVI Buffer Region	0.668	0.003	0.66	0.0003

Concordance Index is computed along with the statistical results. K-fold validation is computed using the results from the statistical analysis as input for the calculation.

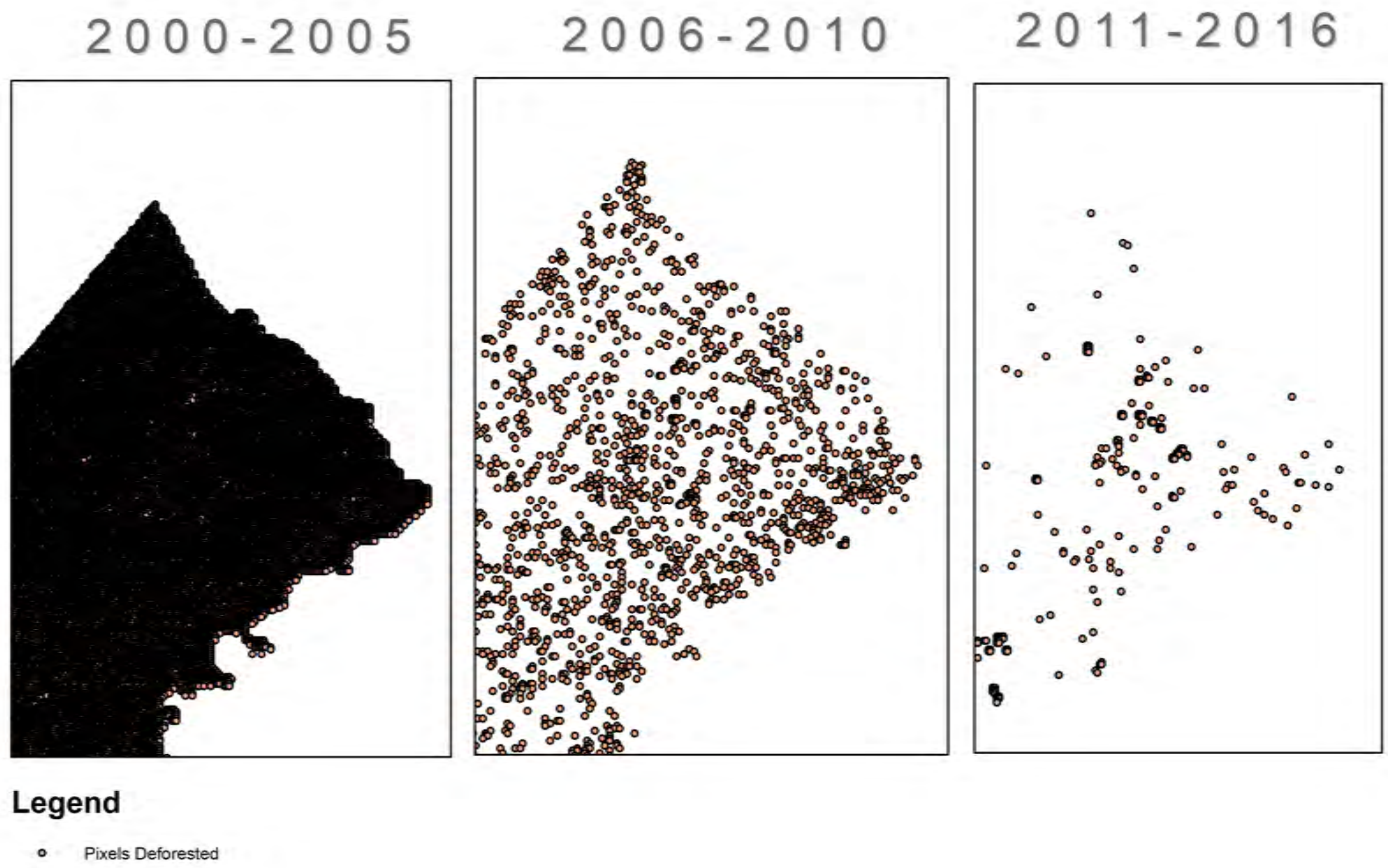

Figure A.3.1 Remaining Pixels in Legal Maranhão side for three different year periods. Pixels are only counted once. Area was randomly chosen.

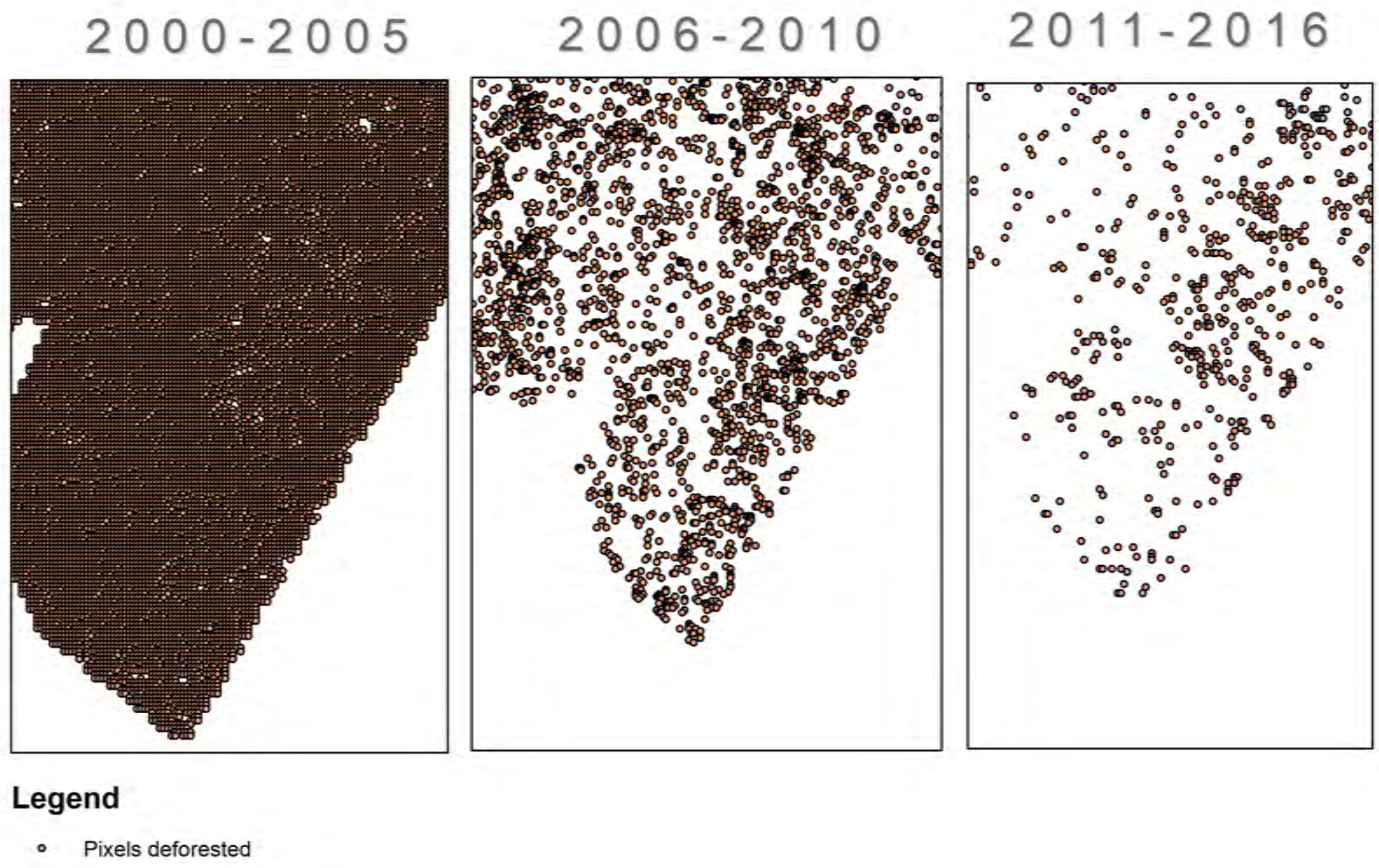

Figure A.3.2 Remaining Pixels in Cerrado Maranhão side for three different year periods. Pixels are only counted once. Area was randomly chosen.

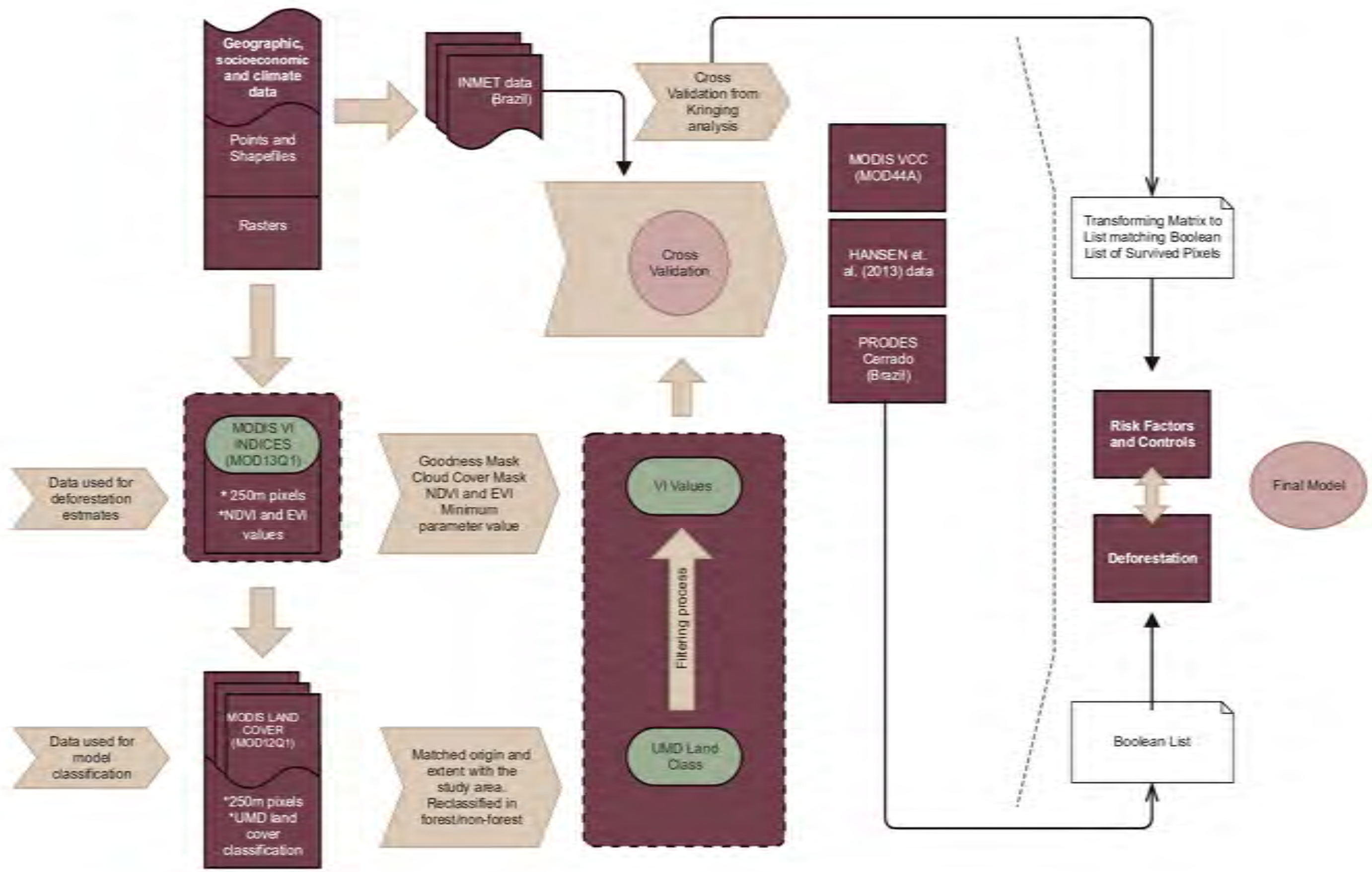

Figure A.3.3 Flowchart of method applied to data. Moving from left to right indicates increasing levels of data processing during the study.

Table A.3.2 Descriptive Statistics - Vegetation Indices (Region)

Variables	Mean	Std Deviation	Min	Max
Period (N)	2.653	2.628	0.000	17.000
Censored (N)	0.967	0.180	0.000	1.000
Cloud (N)	0.711	0.453	0.000	1.000
Period (E)	2.459	2.344	0.000	17.000
Censored (E)	0.967	0.180	0.000	1.000
Cloud (E)	0.711	0.453	0.000	1.000
Lat	-2.683	0.223	-3.064	-2.347
Lon	-41.892	1.171	-44.601	-39.584
PAs	0.539	0.355	0.000	1.228
Mining	0.053	0.085	0.000	0.461
Elevation	137.934	100.009	-26	993
Markets	0.502	0.235	0.000	1.095
Municipalities	0.128	0.058	0.000	0.341
Rivers	0.017	0.015	0.000	0.211
Roads	0.037	0.037	0.000	0.218
Region	0.404	0.491	0.000	1.000
Neighbours	0.992	0.161	0.000	1.000
Rainfall	130	35.695	0.000	214
Temperature	31.7	6.241	29	34

Statistics refer to N=529,680 pixels observations. (N) refers to NDVI and (E) refers to EVI. All distancing values are in decimal degrees. The conversion assumes 0.1 degree to 11km^2.

Table A.3.3 Descriptive Statistics - Vegetation Indices (MA and LM)

Variables (MA)	Mean	Std Deviation	Min	Max
Period (N)	3.059	3.053	0.000	17.000
Censoring (N)	0.956	0.205	0.000	1.000
Clouds (N)	0.517	0.500	0.000	1.000
Lat	-2.527	0.119	-2.974	-2.347
Lon	-41.820	1.174	-43.980	-39.584
Pas	0.512	0.388	0.000	1.228
Mining	0.065	0.098	0.000	0.461
Elevation	123.893	87.687	-26.000	993.000
Markets	0.499	0.172	0.111	0.968
Municipalities	0.130	0.060	0.000	0.341
Rivers	0.017	0.016	0.000	0.211
Roads	0.042	0.040	0.000	0.218
Neighbours	0.995	0.049	0.000	1.000
Rainfall	138.9	24.877	0.000	203.000
Temperature	32.845	1.448	29.000	34.000
Variables (LM)	**Mean**	**Std Deviation**	**Min**	**Max**
Period (N)	2.091	1.643	0.000	17.000
Censoring (N)	0.486	0.500	0.000	1.000
Clouds (N)	0.941	0.235	0.000	1.000
Lat	-2.896	0.139	-3.064	-2.347
Lon	-41.987	1.134	-44.599	-39.603
Pas	0.580	0.310	0.000	1.227
Mining	0.037	0.065	0.000	0.461
Elevation	1.542	1.094	-0.030	4.940
Markets	0.509	0.291	0.000	1.095
Municipalities	0.129	0.058	0.000	0.341
Rivers	0.016	0.013	0.000	0.079
Roads	0.032	0.031	0.000	0.211
Neighbours	0.991	0.066	0.000	1.000
Rainfall	116.867	44.567	0.000	214.000
Temperature	30.011	9.522	29.000	34.000

Statistics refer to N=529,680 pixels observations. (N) refers to NDVI and (E) refers to EVI. All distancing values are in decimal degrees. The conversion assumes 0.1 degree to 11km^2.

Table A.3.4 Data Description - Sources

Variable	Source	Description	Source Data Type	Source Resolution
Vegetation Indices	USGS/NASA	Vegetation Index (VI) value at a per pixel basis.	Raster	250m
Land Cover	USGS/NASA	Global land cover types at yearly intervals derived from six different classification schemes.	Raster	500m
Pas	MMA	Euclidean distance to nearest protected area in decimal degrees	Polygon	-
Mines	EMBRAPA	Euclidean distance to nearest mineral resource/mining in decimal degrees	Point	-
Elevation	EMBRAPA	Digital elevated map of Maranhão	Raster	30m
Markets	CONAB	Euclidean distance to nearest market in decimal degrees	Point	-
Municipalities	IBGE	Euclidean distance to nearest municipality centre in decimal degrees	Point	-
Rivers	IBGE	Euclidean distance to nearest river/basin in decimal degrees	Polyline	-
Roads	IBGE	Euclidean distance to nearest road in decimal degrees	Polyline	-
Settlements	INCRA	Polygons of the settlement areas	Polygon	-
Latitude and Longitude	IBGE	The latitude and longitude of each pixel in the study. This is used to account for spatial autocorrelation across large distances.	Raster	250m

Conclusions

The natural process of societal evolution has brought many changes in the political, social, economic and environmental contexts. These changes were mainly a result of technological progress that spurred an improvement in the use of the environment and consequently led to its degradation. Considered one of major the causes of climate change, deforestation releases billions of tonnes of carbon dioxide and other greenhouse gases into the atmosphere and causes the biodiversity loss in the tropical regions and is damaging the environmental system in the planet.

The purpose of this thesis has been to examine the economic determinants of deforestation in Brazil and the effectiveness of environmental policies taking place in the country using innovative and interdisciplinaries techniques. We analysed first the institutional environmental framework (IEF) in the economic delimitation of Brazilian Amazon (Legal Amazon) controlling for market expansion which was characterised the main driver of deforestation at the time of study. The results from the first chapter suggest that the Institutional Environmental Framework conditioned on policies and prices curbed deforestation within municipalities. Since deforestation has a spatial dimension we expand the study to include spatial analysis. We observe that the Institutional Environmental Framework when established in a municipality tend to reduce deforestation in neighbouring municipalities. An anecdotal counter-factual simulation indicated that the existence of institutional environmental framework avoided forest clearing that would have occurred had the institutional framework not been implemented.

Following the results from the first chapter, we focused our analysis on the deforestation trends observed in the Ecological Tension Zone of Maranhao, which provided us with unique natural experiment in that there were spatially heterogeneous environmental policies to combat deforestation. To understand the deforestation trends in that region we used non-linear modelling for the task since it is recognised that most ecological and climatic data represent complex relationships and thus non-linear models, such as Generalized Additive Models (GAMs), may be particularly suited to capture confounding effects in trends. Our findings suggest that deforestation is related to year and several climatic covariates, but also revealed that there are substantially differences in trends between seasons and regions. For the region under a surveillance system most of the deforestation happened during the rainy season and, for the region not under the monitoring policy, there was a well-defined deforestation trend for both seasons.

Finally, in chapter 3 we combine the findings from the previous chapters to elaborate on the motivation. Deforestation rates have declined in Brazil over the past two decades and it is believed that environmental policies conditioned on the institutional framework have played a crucial role. Moreover, the satellite monitoring program has enabled authorities to identify and react to deforestation in a much more timely manner than local field detection. Given that the trends of deforestation in two regions, under different environmental policies, of an ecological tension zone (ecotone forest) in Maranhão showed diverged results. We assumed that cloud coverage, by delaying detection until skies are clear, has acted as an important impediment to the policy's success. Focusing on the ecological tension zone of Maranhão that is separated into two parts by an artificial line- one that was covered by environmental deforestation policy and the other that is not subject to this - we use satellite data within a survival analysis framework to estimate how the probability of transition between intact forest to disturbed forest, given risk factors and conditional on the time elapsed until the occurrence of the transition, is affected by cloud coverage. Our findings suggest that the presence of clouds has increased deforestation in the region covered by the satellite detection program, and thus was likely an active barrier to legal compliance.

Overall, our results has policy implications for environmental policies in Brazil. We've seen that the institutional environmental framework is important for the protection of the tropical forests when combined with environmental enforcement. The actual institutional framework needs to be tightened up in terms of implementation. Most important, it is pertinent that the established efforts proceed in the Brazilian Amazon, in spite of any political changes in Brazil and, strengthening the institutional framework must be detached from any transient actions. We then, observed the past trend of deforestation in two areas of great importance for the Brazilian biome, Amazon and Cerrado and, the results indicate that the environmental policies in the Legal Amazon should be expanded to ecotonic/transition forests along the Amazon Forest because these forests represent the first indication of anthropic intervention. Also, we believe that the deforestation monitoring system should be improved by the use of satellites that are not constrained by climatic events such as cloud cover. Finally, it is important to acknowledge that significant deforestation is happening in areas of transition between Amazon and Cerrado and, the implementation or expansion of satellite monitoring program must be applied to further biomes like Cerrado, the second most degraded biome in the country.

LIMITATIONS Although our results presented in this thesis are in line with previous studies and our empirical evidence has been corroborated by robustness checks and model validations, our analysis still suffers from a number of weaknesses.

Our main results in the first chapter might suffer from the issue of omitted variable bias. We tried to include all possible variables that could potentially affect deforestation but data limitations prevented us to capture for all possible determinants. To deal with this issue we looked for proxies that could minimise this problem. Another possible issue is the potential endogeneity of many of the explanatory variables, and hence their interpretation in terms of causality. Since it would be difficult to find plausible instruments for many, if not for all, of our independent variables, we instead tried to control for municipality fixed effects, allowing us to purge all time invariant unobservables from the specifications and, we lagged all control variables by one period, so that under assumption that, after controlling for fixed effects all confounding shocks are only contemporaneous in nature, we are left with solely exogenous variation.

In the second and third chapters, we a have a number of limitations that must be taken into account. First, the model implicitly assumes that the predicted range or potential space is fully occupied by forest, which in reality might not be true. Secondly, the spatial distribution of the vegetation indices may exhibit dynamic behaviour over time, so that a potential area may or may not be sparsely vegetated for a certain period (e.g., during sampling) due to progressive succession of forest. Or a temporary absence could be due to natural causes, such as, an attack of pests or diseases or inter-species competition. Thirdly, the study was based on coarse image resolution which could neglect local changes in the sample area. Finally, our results may not be generalised to other areas, such as dense tropical forest and open fields. The same issues for chapter two apply to chapter three since we use the same dataset. In addition, for the third chapter, we only considered for the analysis covariates before or from 2000, which thus excluded roads, protected areas, indigenous land, markets, municipalities centre, and, mining/mineral resources created or discovered after 2000. Finally, we acknowledge that the models were derived from NDVI values and that one could alternatively have used EVI values, which in some instances might be more suited for ecotone forests.

FUTURE RESEARCH This thesis represent a starting point for a research agenda which can be extended in the future. First, departing from the analysis of deforestation in Brazil, it would be interesting to look at other tropical countries investigating the role of environmental policies taking into account the different settings of the institutional framework. There are data available on the proxy of deforestation for many tropical countries and different policies approach have been undertaken for different countries, such as Bolivia, Colombia and, Venezuela. For this reason, it would be interesting to conjecture the differences existing with different institutional apparatus. Secondly, the estimation of the deforestation trends used coarse resolution to the analysis. However, a more in-depth analysis using fine resolution might be needed given the increasing action of selective logging that cannot be captured observing at a coarse resolution. Finally, our analysis of chapter 3 has shown the relevance of climatic events as an impediment of satellite monitoring program effectiveness. We could corroborate the findings by applying a fine resolution satellite which is not affected by cloud cover to reassure our findings. In addition, we could incorporate the yearly climatic event controls to the model that were not able to be included in this study due to computational limitations.

List of References

Achard, F., Stibig, H.-J., Eva, H. D., Lindquist, E. J., Bouvet, A., Arino, O., and Mayaux, P. (2010). Estimating tropical deforestation from Earth observation data. *Carbon Management*, 1(2):271–287.

Albuquerque, B. and Ramos, F. (2006). Análise teórica e empírica dos determinantes de corrupção na gestão pública municipal. Anais do XXXIV encontro nacional de economia [proceedings of the 34th brazilian economics meeting], ANPEC - Associação Nacional dos Centros de Pós-Graduação em Economia [Brazilian Association of Graduate Programs in Economics].

Aleixandre-Benavent, R., Aleixandre-Tudó, J. L., Castelló-Cogollos, L., and Aleixandre, J. L. (2018). Trends in global research in deforestation. a bibliometric analysis. *Land Use Policy*, 72:293–302.

Almeida, E. (2012). *Econometria Espacial Aplicada*. Alínea, First edition.

Almeyda Zambrano, A. M., Broadbent, E. N., Schmink, M., Perz, S. G., and Asner, G. P. (2010). Deforestation drivers in Southwest Amazonia: Comparing smallholder farmers in Iapari, Peru, and Assis Brasil, brazil. *Conservation and Society*, 8(3):157.

Amacher, G. S., Ollikainen, M., and Koskela, E. (2012). Corruption and forest concessions. *Journal of Environmental Economics and Management*, 63(1):92–104.

Anselin, L. (1988). *Spatial econometrics*. Kluwer Academic Publishers.

Antunez, P., Hernandez-Diaz, J., Wehenkel, C., and Clark-Tapia, R. (2017). Generalized models: An application to identify environmental variables that significantly affect the abundance of three tree species. *Forests*, 8(3):59.

Araújo, E., Barreto, P., Baima, S., and Gomes, M. (2017). *Most deforested Conservation Units in the Legal Amazon 2012-2015*. Imazon.

Arima, E. Y., Barreto, P., Araujo, E., and Soares-Filho, B. (2014). Public policies can reduce tropical deforestation: Lessons and challenges from Brazil. *Land Use Policy*, 41:465–473.

Assunçao, J., Gandour, C., and Rocha, R. (2015). Deforestation slowdown in the Brazilian Amazon: prices or policies? *Environment and Development Economics*, 20(06):697–722.

Assunçao, J., Gandour, C., and Rocha, R. (2017). Detering deforestation in the amazon: Environmental monitoring and law enforcement. *Climate Policy Initiative*.

Aubertin, C. (2015). Deforestation control policies in Brazil: sovereignty versus the market. *Forests, Trees and Livelihoods*, 24(3):147–162.

Auderset Joye, D. and Rey-Boissezon, A. (2015). Will charophyte species increase or decrease their distribution in a changing climate? *Aquatic Botany*, 120:73–83.

Barbier, E. B. (2004). Explaining agricultural land expansion and deforestation in developing countries. *American Journal of Agricultural Economics*, 86(5):1347–1353.

Barni, P. E., Pereira, V. B., Manzi, A. O., and Barbosa, R. I. (2015). Deforestation and forest fires in Roraima and their relationship with phytoclimatic regions in the Northern Brazilian Amazon. *Environmental Management*, 55(5):1124–1138.

Barretto, A. G. O. P., Berndes, G., Sparovek, G., and Wirsenius, S. (2013). Agricultural intensification in Brazil and its effects on land-use patterns: an analysis of the 1975-2006 period. *Glob Change Biol*, 19(6):1804–1815.

BASA (2017). Banco da Amazônia. Website: http://www.bancoamazonia.com.br/index.php/acesso-informacao-menusup. Last checked on Oct 27, 2017.

Batistella, M., Bolfe, E. L., Vicente, L. E., Victoria, D. d. C., and Spinelli Araújo, L. (2013a). *Relatorio do Diagnóstico do Macrozoneamento Ecólogico-Econômico do Estado do Maranhão. Relatório Técnico, v. 2.* Embrapa Monitoramento por Satélite.

Batistella, M., Luis Bolfe, E., Vicente, L. E., de Castro Victoria, D., and Spinelli Araujo, L. (2013b). *Relatorio final do Macrozoneamento Ecológico -Econômico do Estado do Maranhão*. EMBRAPA.

Batistella, M., Luis Bolfe, E., Vicente, L. E., de Castro Victoria, D., and Spinelli Araujo, L. (2014). *Macrozoneamento Ecológico-Econômico: potencialidades e fragilidades do estado do Maranhão*, pages 449–453. EMBRAPA.

Bayma, A. P. and Sano, E. E. (2015). Séries temporais de indices de vegetacao (NDVI e EVI) do sensor MODIS para detecção de desmatamentos no bioma cerrado. *Boletim de Ciências Geodésicas*, 21(4):797–813.

Baynard, C. W., Ellis, J. M., and Davis, H. (2012). Roads, petroleum and accessibility: the case of Eastern Ecuador. *GeoJournal*, 78(4):675–695.

BCB (2017a). Banco Central do Brasil. Website http://www.bcb.gov.br. Last checked on Oct 01, 2017.

BCB (2017b). Programa Nacional de Fortalecimento da Agricultura Familiar - Pronaf. Website https://www.bcb.gov.br/pre/bc_atende/port/PRONAF.asp. Last checked on Oct 01, 2017.

BDMEP (2018). BDMEP - Banco de Dados Meteorologicos para Ensino e Pesquisa do INMET. Website http://www.inmet.gov.br/projetos/rede/pesquisa/. Last checked on July 01, 2018.

Bebber, D. P. and Butt, N. (2017). Tropical protected areas reduced deforestation carbon emissions by one third from 2000. *Scientific Reports*, 7(1).

Bell, T. W., Cavanaugh, K. C., Reed, D. C., and Siegel, D. A. (2015). Geographical variability in the controls of giant kelp biomass dynamics. *Journal of Biogeography*, 42(10):2010–2021.

BenYishay, A., Heuser, S., Runfola, D., and Trichler, R. (2017). Indigenous land rights and deforestation: Evidence from the Brazilian Amazon. *Journal of Environmental Economics and Management*, 86(Supplement C):29 – 47. Special issue on environmental economics in developing countries.

Bhattarai, M. and Hammig, M. (2001). Institutions and the environmental kuznets curve for deforestation: A crosscountry analysis for Latin America, Africa and Asia. *World Development*, 29(6):995–1010.

Bio, A., Alkemade, R., and Barendregt, A. (1998). Determining alternative models for vegetation response analysis: a non-parametric approach. *Journal of Vegetation Science*, 9(1):5–16.

BNB (2017). Banco do Nordeste do Brasil. Website https://www.bnb.gov.br/. Last checked on Oct 27, 2017.

Borges de Lima, I. and Buszynski, L. (2011). Local environmental governance, public policies and deforestation in Amazonia. *Management of Env Quality*, 22(3):292–316.

Börner, J., Kis-Katos, K., Hargrave, J., and König, K. (2015). Post-crackdown effectiveness of field-based forest law enforcement in the Brazilian Amazon. *PLoS One*, 10(4):1–19.

Bowman, A. W. and Azzalini, A. (1997). *Applied Smoothing Techniques for Data Analysis: The Kernel Approach With S-Plus Illustrations (Oxford science publications)*. Oxford University Press.

BRASIL, C. d. D. (2017). Câmara, dos Deputados - Frente Parlamentar Mista da Agropecuaria – FPA. Website http://www.camara.leg.br/internet/deputado/frenteDetalhe.asp?id=53476. Last Checked on Aug 01, 2017.

BRASIL, I.-D. (2018). DETER — Coordenação-Geral de Observação da Terra. Website http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/deter. Last Checked on Dez 10, 2018.

BRASIL, M. d. M. A. (2016). *Plano Operativo do PPCDAM (2016-2020)*. MMA.

BRASIL, P. F. (2012). Website http://www.pf.gov.br/agencia/noticias/2012/setembro/pf-combate-extracao-ilegal-de-madeira-no-maranhao. Published on September 2012. Last Checked on Nov 10, 2018.

BRASIL, Serviço Florestal Brasileiro (2016). Cadastro Nacional de Florestas Públicas. Distribuição por Estados. Website http://www.florestal.gov.br/documentos/informacoes-florestais/cadastro-nacional-de-florestas-publicas/2665-tabela-de-distribuicao-por-estado-cnfp-2016. Published in 2016. Last Checked on Aug 30, 2017.

BRASIL, Tribunal Superior Eleitoral (2018). Tribunal superior eleitoral (tse) - eleitor e eleições. Website http://www.tse.jus.br/eleitor-e-eleicoes/eleicoes/eleicoes-anteriores/eleicoes-2004/candidaturas-votacao-e-resultados/reeleicao-eleicoes-2004. Last Checked on March 30, 2018.

Brito, A. d., Valeriano, D. d. M., Ferri, C., Scolastrici, A., and Sestini, M. (2018). Metodologia da detecção do desmatamento no bioma Cerrado. Website http://www.dpi.inpe.br/fipcerrado/report_funcate_metodologia_mapeamento_bioma_cerrado.pdf. Last checked on Nov 01, 2018.

Brito, B. (2017). IMAZON - nota Técnica sobre o impacto das novas regras de regularizaçao fundiaria na Amazônia. Technical report, IMAZON. Website http://imazon.org.br/publicacoes/nota-tecnica-sobre-o-impacto-das-novas-regras-de-regularizacao-fundiaria-na-amazonia/. Last Checked on Aug 01, 2017.

Brollo, F., Nannicini, T., Perotti, R., and Tabellini, G. (2013). The political resource curse. *American Economic Review*, 103(5):1759–1796.

Bugarin, M. and Meneguin, F. B. (2016). Incentivos à corrupção e à inação no serviçõ público: Uma análise de desenho de mecanismos.

Buizer, M., Humphreys, D., and de Jong, W. (2014). Climate change and deforestation: The evolution of an intersecting policy domain. *Environmental Science & Policy*, 35:1–11.

Cabral, C. d. S. R., Gurgel, A. C., and Paltsev, S. (2012). Economic analysis of deforestation reduction in Brazil. http://purl.umn.edu/211378 Online; Accessed 2016-2-26.

Calixto, B. (2016). Criminosos mudaram a metodologia do desmatamento ilegal da Amazônia – Revista Época (Blog do Planeta). Website http://epoca.globo.com/colunas-e-blogs/blog-do-planeta/noticia/2016/07/criminosos-mudaram-metodologia-do-desmatamento-ilegal-da-amazonia.html. Published on July 2016. Last Checked on Sep 15, 2017.

Cao, H. (2005). *A Comparison Between the Additive and Multiplicative Risk Models*. PhD thesis, Faculte des sciences et de genie - Universite Laval Quebec.

Caviglia-Harris, J. L. (2018). Agricultural innovation and climate change policy in the Brazilian Amazon: Intensification practices and the derived demand for pasture. *Journal of Environmental Economics and Management*, 90:232 – 248.

Celentano, D., Rousseau, G. X., Muniz, F. H., Varga, I. v. D., Martinez, C., Carneiro, M. S., Miranda, M. V., Barros, M. N., Freitas, L., and Narvaes, I. d. S. e. a. (2017). Towards zero deforestation and forest restoration in the Amazon region of Maranhão state, Brazil. *Land Use Policy*, 68:692–698.

CEPEA (2017). Centro de Estudos Avançados em Economia Aplicada - CEPEA/ESALQ. Website http://www.cepea.esalq.usp.br/br. Last Checked on Aug 01, 2017.

Chagnon, F. J. F., Bras, R. L., and Wang, J. (2004). Climatic shift in patterns of shallow clouds over the Amazon. *Geophysical Research Letters*, 31(24).

Chaves, L. F., Cohen, J. M., Pascual, M., and Wilson, M. L. (2008). Social exclusion modifies climate and deforestation impacts on a vector-borne disease. *PLOS Neglected Tropical Diseases*, 2(2):1–8.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences. In Cohen, J., editor, *Statistical Power Analysis for the Behavioral Sciences*, pages 1 – 17. Academic Press.

Cohen, S., Ianetz, A., and Stanhill, G. (2002). Evaporative climate changes at Bet Dagan, Israel, 1964–1998. *Agricultural and Forest Meteorology*, 111(2):83–91.

Combes Motel, P., Pirard, R., and Combes, J.-L. (2009). A methodology to estimate impacts of domestic policies on deforestation: Compensated Successful Efforts for avoided deforestation (REDD). *Ecological Economics*, 68(3):680–691.

Cornillon, P., Gallagher, J., and Sgouros, T. (2003). OPeNDAP: Accessing data in a distributed, heterogeneous environment. Website https://datascience.codata.org/articles/abstract/10.2481/dsj.2.164/.

Costa, D. (2018). O globo - com alta no preço do gás, mais brasileiros passam a usar lenha e carvão. Website https://oglobo.globo.com/economia/com-alta-no-preco-do-gas-mais-brasileiros-passam-usar-lenha-carvao-22629819. Published on April, 2018. Last Checked on Oct 01, 2018.

Cropper, M., Puri, J., and Griffiths, C. (2001). Predicting the location of deforestation: The role of roads and protected areas in north Thailand. *Land Economics*, 77(2):172.

Culas, R. J. (2014). Causes of Deforestation and Policies for Reduced Emissions (REDD+): A Cross-Country Analysis. *IUP Journal of Applied Economics*, 13(4):7–27.

da Silva, V. d. P. R. (2004). On climate variability in Northeast of Brazil. *Journal of Arid Environments*, 58(4):575–596.

Davalos, L. M., Holmes, J. S., Rodriguez, N., and Armenteras, D. (2014). Demand for beef is unrelated to pasture expansion in northwestern Amazonia. *Biological Conservation*, 170:64–73.

Davidson-Pilon, C., Kalderstam, J., Kuhn, B., Zivich, P., Fiore-Gartland, A., Moneda, L., Parij, A., Stark, K., Anton, S., Besson, L., Jona, Gadgil, H., Golland, D., Hussey, S., Noorbakhsh, J., Klintberg, A., Evans, N., Braymer-Hayes, M., Lukasz, Séguin, J., Rose, J., Slavitt, I., Martin, E., Ochoa, E., Albrecht, D., dhuynh, Zgonjanin, D., Chen, D., Fournier, C., and Rendeiro, A. F. (2018). Camdavidsonpilon/lifelines: v0.14.6.

de Souza, J. B., Reisen, V. A., Franco, G. C., Ispány, M., Bondon, P., and Santos, J. M. (2017). Generalized additive models with principal component analysis: An application to time series of respiratory disease and air pollution data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 67(2):453–480.

Dias, L. F. O., Dias, D. V., and Magnusson, W. E. (2015). Influence of environmental governance on deforestation in municipalities of the Brazilian Amazon. *PLOS ONE*, 10(7):e0131425.

Didan, K. (2015). Mod13q1 MODIS/terra vegetation indices 16-day l3 global 250m sin grid v006 [data set]. *NASA EOSDIS LP DAAC*, V006.

Didan, K., Munoz, A. B., Solano, R., and Huete, A. (2015). MODIS vegetation index user's guide (mod13 series) - version 3.00 - collection 6. *NASA EOSDIS LP DAAC*.

Diprose, G. and McGregor, A. (2009). Dissolving the sugar fields: Land reform and resistance identities in the Philippines. *Singapore Journal of Tropical Geography*, 30(1):52–69.

Domingos, P. (2012). A few useful things to know about machine learning. *Commun. ACM*, 55(10):78–87.

Driscoll, J. C. and Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, 80(4):549–560.

Elhorst, J. P. (2012). Matlab software for spatial panels. *International Regional Science Review*, 37(3):389–405.

EMBRAPA (2018). Empresa Brasileira de Pesquisas Agropecuária - Portal Embrapa – Acesso a Informação. Website https://www.embrapa.br/.

ESRI (2016a). *ArcMap - Environmental Systems Research Institute*. Environmental Systems Research Institute.

ESRI (2016b). ArcPy - Environmental Systems Research Institute. Website http://resources.arcgis.com/en/help/main/10.4.1/index.html#//000v000000v7000000.

Ezzine-de Blas, D., Barner, J., Violato-Espada, A.-L., Nascimento, N., and Piketty, M.-G. (2011). Forest loss and management in land reform settlements: Implications for REDD governance in the Brazilian Amazon. *Environmental Science & Policy*, 14(2):188–200.

Fearnside, P. M. (2016). Brazil's Amazonian forest carbon: the key to Southern Amazonia's significance for global climate. *Regional Environmental Change*, 18(1):47–61.

Feitosa, A. C. (2006). *Relevo do Estado do Maranhão: uma nova proposta de classificaçao topomorfologica*, pages 1–11. VI Simposio Nacional De Geomorfologia/ Regional Conference on Geomorphology.

Ferreira, A. J. d. A. (2008). *Políticas territoriais e a reorganização do espaço Maranhense*. PhD thesis, Universidade de Sao Paulo.

Friedrich, C. J. (2002). *Corruption Concepts in Historical Perspective*, pages 15–24. Transaction Publishers, 3 edition.

FUNAI (2017). Fundação Nacional do índio - funai. Website http://www.funai.gov.br/. Last Checked on Dec 15, 2017.

Garcia, A. S., Sawakuchi, H. O., Ferreira, M. E., and 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. *Journal of Environmental Management*, 187:16 – 23.

Garcia, R. L. (2003). *A economia da corrupção - teoria e evidências. Uma aplicação ao setor de obras rodoviárias no Rio Grande do Sul*. PhD thesis, Universidade Federal do Rio Grande do Sul.

Geerken, R. A. (2009). An algorithm to classify and monitor seasonal variations in vegetation phenologies and their inter-annual change. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(4):422–431.

Geist, H. and Lambin, E. F. (2001). *What drives tropical deforestation?* LUCC International Project Office.

Geist, H. J. and Lambin, E. F. (2002). Proximate causes and underlying driving forces of tropical deforestation. *BioScience*, 52(2):143.

Gerold, G., Fremerey, M., and Guhardja, E. (2004). *Land use, nature conservation and the stability of rainforest margins in Southeast Asia*. Springer.

GFC, G. F. C. (2017). Global forest change.

Girardi, G. (2015). Revista galileu - sem dúvida. Índios, santos e geografia. Website http://revistagalileu.globo.com/Galileu/0,6993,ECT498531-1716-5,00.html. Last Checked on Sep 15, 2016.

Girardi, G. (2017). O Estado de S.Paulo. Desmatamento cresce em Unidades de Conservação no meio da Amazônia. *Sustentabilidade. O Estado de S.Paulo*. Last Checked on Sep 15, 2017.

GLCF, G. L. C. F. (2018). Glcf: MODIS vegetative cover conversion.

Gollnow, F. and Lakes, T. (2014). Policy change, land use, and agriculture: The case of soy production and cattle ranching in Brazil, 2001-2012. *Applied Geography*, 55:203–211.

Gordon, A. D., Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). Classification and regression trees. *Biometrics*, 40(3):874.

Green, J. M., Larrosa, C., Burgess, N. D., Balmford, A., Johnston, A., Mbilinyi, B. P., Platts, P. J., and Coad, L. (2013). Deforestation in an african biodiversity hotspot: Extent, variation and the effectiveness of protected areas. *Biological Conservation*, 164:62–72.

Greenberg, J. A., Kefauver, S. C., Stimson, H. C., Yeaton, C. J., and Ustin, S. L. (2005). Survival analysis of a neotropical rainforest using multitemporal satellite imagery. *Remote Sensing of Environment*, 96(2):202–211.

Griffin, O. (2012). *Modelling the Severity of Deforestation in Coastal Tanzania and a comparison of data sources*. PhD thesis, Imperial College London.

Halperin, J., LeMay, V., Coops, N., Verchot, L., Marshall, P., and Lochhead, K. (2016). Canopy cover estimation in miombo woodlands of Zambia: Comparison of LANDSAT 8 OLI versus RapidEye imagery using parametric, nonparametric, and semiparametric methods. *Remote Sensing of Environment*, 179:170–182.

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., and Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160):850–853.

Hargrave, J. and Kis-Katos, K. (2012). Economic causes of deforestation in the Brazilian Amazon: A panel data analysis for the 2000s. *Environmental and Resource Economics*, 54(4):471–494.

Heiblum, R. H., Koren, I., and Feingold, G. (2014). On the link between Amazonian forest properties and shallow cumulus cloud fields. *Atmospheric Chemistry and Physics*, 14(12):6063–6074.

Horning, N. (2010). *Remote Sensing for Ecology and Conservation: A Handbook of Techniques (Techniques in ecology and conservation series)*. Oxford University Press.

Hu, J. and Hu, P. (2009). On the performance of the cross-entropy method. In *Proceedings of the 2009 Winter Simulation Conference (WSC)*, pages 459–468.

Huang, X. and Friedl, M. A. (2014). Distance metric-based forest cover change detection using MODIS time series. *International Journal of Applied Earth Observation and Geoinformation*, 29:78–92.

IBAMA (2017). Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renovaveis - IBAMA. Website http://www.ibama.gov.br/. Last Checked on Aug 01, 2017.

IBGE (2017). Instituto Brasileiro de Geografia e Estatistica - IBGE. Website http://www.ibge.br/. Last Checked on Aug 30, 2017.

ICMBIO (2017). Instituto Chico Mendes de Conservação a Biodiversidade - ICMBIO. Website http://www.icmbio.gov.br/. Last Checked on Dec 15, 2017.

Imori, D. and Guilhoto, J. J. M. (2015). Tracing Brazilian regions CO2 emissions in domestic and global trade. Website http://www.usp.br/nereus/?txtdiscussao=tracing-Brazilian-regions%E2%80%99-co2-emissions-in-domestic-and-global-trade. Published on August 2015. Last Checked on Feb 26, 2016.

INCRA (2015). *Na luta pela Reforma Agrária: INCRA 45 anos.* Secretaria Especial de Agricultura Familiar e do Desenvolvimento Agrário (MDA) /Instituto Nacional de Colonização e Reforma Agrária (INCRA).

INCRA (2017). Instituto Nacional de Colonização e Reforma Agrária - INCRA . Website http://www.mda.gov.br/sitemda/tags/incra. Last Checked on Aug 01, 2017.

INMET (2018). Instituto Nacional de Meteorologia – INMET. Estações Automáticas. Website http://www.inmet.gov.br/portal/index.php?r=estacoes/estacoesautomaticas. Last Checked on July 30, 2018.

INPE (2017). Instituto Nacional de Pesquisas Espaciais - National Institute for Spatial Research - INPE. Website http://www.inpe.br/ingles/. Last Checked on Aug 01, 2017.

IPEA (2008). O que é Amazônia legal? instituto de pesquisa econômica aplicada (ipea). Website http://www.ipea.gov.br/desafios/index.php?option=com_content&id=2154:catid=28&Itemid=23. Published in June 2008. Last Checked on March 2, 2016.

ISA (2018). Instituto Socioambiental (ISA) – Unidade de Conservação APA dos Morros Garapenses. Website https://uc.socioambiental.org/uc/591527. Last Checked on Nov 20, 2018.

Janitza, S., Celik, E., and Boulesteix, A.-L. (2016). A computationally fast variable importance test for random forests for high-dimensional data. *Advances in Data Analysis and Classification*.

Kaimowitz, D. and Angelsen, A. (1998). *Economic models of tropical deforestation*. Center For International Forestry Research (CIFOR).

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481.

Kintisch, E. (2007). Carbon emissions: Improved monitoring of rainforests helps pierce haze of deforestation. *Science*, 316(5824):536–537.

Klintowitz, D. (2016). Por que o programa Minha Casa Minha Vida so poderia acontecer em um governo petista? *Cadernos Metropole*, 18:165 – 190.

Koren, I., Kaufman, Y. J., Remer, L. A., and Martins, J. V. (2004). Measurement of the effect of Amazon smoke on inhibition of cloud formation. *Science*, 303(5662):1342–1345.

Krause, C., Balbim, R., and Neto, V. C. L. (2013). Minha casa minha vida, nosso crescimento: Onde fica política habitacional? Texto para Discussão, Instituto de Pesquisa Econômica Aplicada (IPEA) 1853, IPEA, Brasília.

Kuik, O. (2013). REDD+ and international leakage via food and timber markets: a CGE analysis. *Mitig Adapt Strateg Glob Change*, 19(6):641–655.

Lambin, E. F. and Geist, H. (2006). *Land-use and land-cover change. Local Processes and Global Impacts.* Springer - Verlag.

Lambin, E. F. and Meyfroidt, P. (2011). Global land use change, economic globalization, and the looming land scarcity. *Proceedings of the National Academy of Sciences*, 108(9):3465–3472.

Larsen, K. (2015). Gam: The predictive modelling silver bullet. Unpublished.

Latawiec, A. E., Strassburg, B. B. N., Valentim, J. F., Ramos, F., and Alves-Pinto, H. N. (2014). Intensification of cattle ranching production systems: socioeconomic and environmental synergies and risks in Brazil. *Animal*, 8(08):1255–1263.

Lee, E. T. and Wang, J. W. (2003). Statistical methods for survival data analysis. *Wiley Series in Probability and Statistics*.

Leemans, R. (2009). *Why regional and spatial specificity is needed in environmental assessments? In: Integrated regional assessment of global climate change / Knight, G., Jäger, J.* Cambridge University Press.

LeSage, J. P. and Pace, R. K. (2009). *Introduction to spatial econometrics*. CRC Press.

Liu, X., Song, Y., Yi, W., Wang, X., and Zhu, J. (2018). Comparing the random forest with the generalized additive model to evaluate the impacts of outdoor ambient environmental factors on scaffolding construction productivity. *Journal of Construction Engineering and Management*, 144(6):04018037.

Lopez-Carr, D. and Burgdorfer, J. (2013). Deforestation drivers: Population, migration, and tropical land use. *Environment: Science and Policy for Sustainable Development*, 55(1):3–11.

Lusk, C. H., McGlone, M. S., and Overton, J. M. (2016). Climate predicts the proportion of divaricate plant species in New Zealand arborescent assemblages. *Journal of Biogeography*, 43(9):1881–1892.

Mani, M. and Griffiths, C. (1997). Roads, population pressures, and deforestation in Thailand, 1976-89. *Policy Research Working Papers*.

Marcellus Caldas, Robert Walker, E. A. S. P. S. A. and Simmons, C. (2007). Theorizing land cover and land use change: The peasant economy of Amazonian deforestation. *Annals of the Association of American Geographers*, 97(1):86–110.

MATLAB (2017). *Matrix Laboratory - MATLAB*. The MathWorks Inc.

Mello, N. and Artaxo, P. (2017). Evoluçao do Plano de Ação para Prevenção e Controle do Desmatamento na Amazônia Legal - PPCDAm. *Revista do Instituto de Estudos Brasileiros*, 0(66):108–129.

Mendes, C. M. and Junior, S. P. (2012). Deforestation, economic growth and corruption: a nonparametric analysis on the case of Amazon forest. *Applied Economics Letters*, 19(13):1285–1291.

Mendes, C. M. d. V. (2009). How does corruption drive illegal deforestation in Amazon forest? http://portalrevistas.ucb.br/index.php/rbee/article/viewFile/4212/2547 Online; Accessed 2016-2-19.

Mendes, N. (2017). Estrutura organizacional do Sistema Nacional do Meio Ambiente (SISNAMA). Website https://nathymendes.jusbrasil.com.br/noticias/315451463/estrutura-organizacional-do-sistema-nacional-do-meio-ambiente-sisnama?ref=topic_feed. Last Checked on Oct 12, 2017.

MMA
(2017). Ministerio do Meio Ambiente (MMA) - O que sao Unidades de Conservação? Website http://www.mma.gov.br/areas-protegidas/unidades-de-conservacao/o-que-sao. Last Checked on Aug 01, 2017.

MMA (2018a). Ministério do Meio Ambiente - Governo divulga desmatamento no Cerrado. Website http://www.mma.gov.br/index.php/comunicacao/agencia-informma?view=blog&id=3066. Last Checked on Dec 10, 2018.

MMA (2018b). Ministério do Meio Ambiente (MMA). Website http://www.mma.gov.br/. Last Checked on Aug 01, 2017.

MMA-CNUC, M. d. M. A. (2017). Cadastro Nacional das Unidades de Conservação. Tabela consolidada das Unidades de Conservação. Website http://www.mma.gov.br/cadastro_uc, Online; Accessed 2017-08-11. Last Checked on Aug 11, 2017.

Molina Vale, P. (2014). The conservation versus production trade-off: Does livestock intensification increase deforestation? the case of the Brazilian Amazon. *SSRN Electronic Journal*.

Moore, L., Hanley, J. A., Turgeon, A. F., and Lavoie, A. (2011). A comparison of generalized additive models to other common modeling strategies for continuous covariates: Implications for risk adjustment. *Journal of Biometrics 'I&' Biostatistics*, 02(01).

Mor, M. F. (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, 6(4):525–533.

Moraes, M. M. G. d., Villanueva, P., Belmont Figueira, M. A., Netto Lasmar, E., Affonso Swerts, L., Cavanha de Rezende Caminha, A. F., Moreira Angelim, D. C., and Vieira da Silva, S. (2016). *Nota Técnica MMA - Estrategias e mecanismos de articulacao do Sistema Nacional do Meio Ambiente - Sisnama, para a gestão descentralizada, democrática e eficiente*. Ministério do Meio Ambiente.

Moreno-Fernández, D., Augustin, N. H., Montes, F., Cañellas, I., and Sánchez-González, M. (2018). Modeling sapling distribution over time using a functional predictor in a generalized additive model. *Annals of Forest Science*, 75(1).

Murase, H., Nagashima, H., Yonezaki, S., Matsukura, R., and Kitakado, T. (2009). Application of a generalized additive model (GAM) to reveal relationships between environmental factors and distributions of pelagic fish and krill: a case study in Sendai Bay, Japan. *ICES Journal of Marine Science*, 66(6):1417–1424.

Mustin, K., Carvalho, W. D., Hilario, R. R., Costa-Neto, S. V., Silva, C., Vasconcelos, I. M., Castro, I. J., Eilers, V., Kauano, E. E., and Mendes-Junior, R. N. G. e. a. (2017). Biodiversity, threats and conservation challenges in the Cerrado of Amapa, an Amazonian savanna. *Nature Conservation*, 22:107–127.

Nabuurs, J., Masera, O., Andrasko, K., Benitez-Ponce, P., Boer, R., Dutschke, M., Elsiddig, E., Ford-Robertson, J., Frumhoff, P., and Karjalainen, T. e. a. (2007). Forestry. In

Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. *IPCC*.

Nepstad, D., McGrath, D., Stickler, C., Alencar, A., Azevedo, A., Swette, B., Bezerra, T., DiGiano, M., Shimada, J., and Seroa da Motta, R. e. a. (2014). Slowing Amazon deforestation through public policy and interventions in beef and soy supply chains. *Science*, 344(6188):1118–1123.

Nepstad, D. C., Boyd, W., Stickler, C. M., Bezerra, T., and Azevedo, A. A. (2013). Responding to climate change and the global land crisis: Redd+, market transformation and low-emissions rural development. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1619):20120167–20120167.

North, D. C. (1994). Economic performance through time. *American Econ. Rev*, pages 359–368.

Núcleo Geoambiental - NUGEO (2018). Website http://www.nugeo.uema.br/. Last Checked on Oct 10, 2018.

Oestreicher, J. S., Farella, N., Paquet, S., Davidson, R., Lucotte, M., Mertens, F., and Saint-Charles, J. (2014). Livelihood activities and land-use at a riparian frontier of the Brazilian Amazon: quantitative characterisation and qualitative insights into the influence of knowledge, values, and beliefs. *Hum Ecol*, 42(4):521–540.

Oliveira, A. E. S. (2011). Políticas socioambientais Brasileiras e o aprendizado de uma nova ação. Website http://ojs.c3sl.ufpr.br/ojs/index.php/made/article/viewFile/20891/14462. Last Checked on Feb 28, 2016.

Olson, S. H., Gangnon, R., Silveira, G. A., and Patz, J. A. (2010). Deforestation and Malaria in Mancio Lima county, Brazil. *Emerg. Infect. Dis.*, 16(7):1108–1115.

Pacheco, P. (2009). Agrarian reform in the Brazilian Amazon: Its implications for land distribution and deforestation. *World Development*, 37(8):1337–1347.

Pailler, S. (2018). Re-election incentives and deforestation cycles in the Brazilian Amazon. *Journal of Environmental Economics and Management*, 88:345–365.

Pascale, C. M., COMBES, J.-L., Catherine, A. B., ARAUJO, C., and REIS, E. J. (2010). Does land tenure insecurity drive deforestation in the Brazilian Amazon? Copyright - Copyright Federal Reserve Bank Of St Louis 2010; Última atualizaçao em - 2015-10-03.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Pereira, E. B., Martins, F. R., Gonçalves, A. R., Costa, R. S., de Lima, F. J. L., Rüther, R., de Abreu, S. L., Tiepolo, G. M., Pereira, S. V., and de Souza, J. G. (2017). *Atlas Brasileiro de Energia Solar*. INPE, 2 edition.

Peto, R. and Peto, J. (1972). Asymptotically efficient rank invariant test procedures. *Journal of the Royal Statistical Society. Series A (General)*, 135(2):185–207.

Pfaff, A., Robalino, J., Walker, R., Aldrich, S., Caldas, M., Reis, E., Perz, S., Bohrer, C., Arima, E., and Laurance, W. e. a. (2007). Road investments, spatial spillovers and deforestation in the Brazilian Amazon. *Journal of Regional Science*, 47(1):109–123.

Pfaff, A. S. (1999). What drives deforestation in the Brazilian Amazon? *Journal of Environmental Economics and Management*, 37(1):26–43.

Pfaff, A. S. P. (1997). What drives deforestation in the Brazilian Amazon? evidence from satellite and socioeconomic data. *Policy Research Working Papers*.

Pinheiro, T. F., Escada, M. I. S., Valeriano, D. M., Hostert, P., Gollnow, F., and Müller, H. (2016). Forest degradation associated with logging frontier expansion in the Amazon: The br-163 region in Southwestern Para, Brazil. *Earth Interactions*, 20(17):1–26.

Pinto, E., Shin, Y., Cowling, S. A., and Jones, C. D. (2009). Past, present and future vegetation-cloud feedbacks in the Amazon Basin. *Climate Dynamics*, 32(6):741–751.

Pourtaghi, Z. S., Pourghasemi, H. R., Aretano, R., and Semeraro, T. (2016). Investigation of general indicators influencing on forest fire and its susceptibility modeling using different data mining techniques. *Ecological Indicators*, 64:72–84.

R., C. D. (1972). Regression models and life tables. *Journal of the Royal Statistic Society*, B(34):187–202.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ratana, P., Huete, A. R., and Ferreira, L. (2005). Analysis of Cerrado Physiognomies and Conversion in the MODIS Seasonal-Temporal Domain. *Earth Interactions*, 9(3):1–22.

Reis, C. S. d. and Conceição, G. M. d. (2010). Aspectos Florísticos de um Fragmento de Vegetação, localizado no Município de Caxias, Maranhão, Brasil. *SCIENTIA PLENA*, 6(2):1–17.

Richards, P. (2015). What drives indirect land use change? How Brazil's agriculture sector influences frontier deforestation. *Annals of the Association of American Geographers*, 105(5):1026–1040.

Richards, P. and VanWey, L. (2015). Where deforestation leads to urbanization: How resource extraction is leading to urban growth in the Brazilian Amazon. *Annals of the Association of American Geographers*, 105(4):806–823.

Rico, J., Panlasigui, S., J. Loucks, C., Swenson, J., and Pfaff, A. (2017). Logging concessions, certification and protected areas in the peruvian amazon: forest impacts from combinations of development rights and land-use restrictions. FAERE Working Paper.

Rochedo, P. R. R., Soares-Filho, B., Schaeffer, R., Viola, E., Szklo, A., Lucena, A. F. P., Koberle, A., Davis, J. L., Rajao, R., and Rathmann, R. (2018). The threat of political bargaining to climate mitigation in Brazil. *Nature Climate Change*, 8(8):695–698.

Rossatto, D. R., Hoffmann, W. A., de Carvalho Ramos Silva, L., Haridasan, M., Sternberg, L. S. L., and Franco, A. C. (2013). Seasonal variation in leaf traits between congeneric savanna and forest trees in Central Brazil: implications for forest expansion into savanna. *Trees*, 27(4):1139–1150.

Rural, R. G. (2018). Revista Globo Rural - MPF e Ibama fazem operação contra o desmatamento ilegal no Cerrado. Website https://revistagloborural.globo.com/Noticias/Sustentabilidade/noticia/2018/05/mpf-e-ibama-fazem-operacao-contra-o-desmatamento-ilegal-no-cerrado.html Publised in May 2018. Last Checked on Nov 20, 2018.

Sanches, A. C., Caleman, S. M. D. Q., Melo, M. F. S., and Campos-Silva, W. (2017). Descentralização da gestão ambiental no Brasil: análise histórica dos principais momentos do processo. *Revista Gestão e Desenvolvimento*, 14(2):51.

Santos-Filho, F. S., Almeida Junior, E. B. d., and Soares, C. J. d. R. S. (2013). Cocais: Zona ecotonal natural ou artificial? *Revista Equador*, 1(1):2–12.

Scardua, F. P. and Bursztyn, M. A. A. (2003). Descentralização da política ambiental no Brasil. *Sociedade e Estado*, 18:291 – 314.

Schneider, M. and Peres, C. A. (2015). Environmental costs of government-sponsored agrarian settlements in Brazilian Amazonia. *PLOS ONE*, 10(8):e0134016.

Setiawan, Y., Yoshino, K., and Prasetyo, L. B. (2014). Characterizing the dynamics change of vegetation cover on tropical forestlands using 250m multi-temporal MODIS EVI. *International Journal of Applied Earth Observation and Geoinformation*, 26:132–144.

Silva Costa, L. G., Miranda, I. S., Grimaldi, M., Silva, M. L., Mitja, D., and Lima, T. T. S. (2012). Biomass in different types of land use in the Brazil's arc of deforestation. *Forest Ecology and Management*, 278:101–109.

Simpson, G. L. (2018). *gratia: Ggplot-based graphics and other useful functions for GAMs fitted using mgcv*. R package version 0.1-0.

Sluiter, R. (2009). Interpolation methods for the climate atlas. *KNMI*, pages 10–16.

Soler, L., Verburg, P., and Alves, D. (2014). Evolution of land use in the Brazilian Amazon: From frontier expansion to market chain dynamics. *Land*, 3(3):981–1014.

Stella, A. (2011). *Plano de prevenção e controle do desmatamento e queimadas do Maranhão*. Secretaria de Estado do Meio Ambiente e Recursos Naturais.

Stickler, C. M., Coe, M. T., Costa, M. H., Nepstad, D. C., McGrath, D. G., Dias, L. C. P., Rodrigues, H. O., and Soares-Filho, B. S. (2013). Dependence of hydropower energy generation on forests in the Amazon basin at local and regional scales. *Proceedings of the National Academy of Sciences*, 110(23):9601–9606.

Sulla-Menashe, D. and Friedl, M. (2015). Mcd12q1 MODIS/terra+aqua land cover type yearly l3 global 500m sin grid v006 [data set]. *NASA EOSDIS Land Processes DAAC*.

Sulla-Menashe, D. and Friedl, M. (2018). User guide to collection 6 MODIS land cover (mcd12q1 and mcd12c1) product. *NASA EOSDIS Land Processes DAAC*.

Tacconi, L., Boscolo, M., and Brack, D. (2003). *National and international policies to control illegal forest activities*. Center for International Forest Research.

Terry M. Therneau and Patricia M. Grambsch (2000). *Modeling Survival Data: Extending the Cox Model*. Springer, New York.

Therneau, T. M. (2015). *A Package for Survival Analysis in S*. version 2.38.

UNEP (2018). The UNEP Environmental Data Explorer, as compiled from Food and Agriculture Organization of the United Nations (FAO) - FAOStat . United Nations Environment Programme. Website http://ede.grid.unep.ch. Last Checked on Dec 10, 2018.

USGS, L. D. (2018). USGS - U.S. Geological Survey, LP DAAC - Land Processes Distributed Active Archive Center - AppEEARS. Website https://lpdaacsvc.cr.usgs.gov/appeears/help?section=area%252Freprojection.

van Wagtendonk, J. W., Root, R. R., and Key, C. H. (2004). Comparison of aviris and LANDSAT etm+ detection capabilities for burn severity. *Remote Sensing of Environment*, 92(3):397–408.

Vance, C. and Geoghegan, J. (2002). Temporal and spatial modelling of tropical deforestation: a survival analysis linking satellite and household survey data. *Agricultural Economics*, 27(3):317–332.

Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer, New York, fourth edition. ISBN 0-387-95457-0.

Vianello, R. L. (2011). *A Estação Meteorológica E Seu Observador. Uma parceria secular de bons serviços prestados à humanidade*. INMET.

Wang, J., Chagnon, F. J. F., Williams, E. R., Betts, A. K., Renno, N. O., Machado, L. A. T., Bisht, G., Knox, R., and Bras, R. L. (2009). Impact of deforestation in the Amazon basin on cloud climatology. *Proceedings of the National Academy of Sciences*, 106(10):3670–3674.

Wood, S. (2017). *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, 2 edition.

Wood, S. N. (2003). Thin-plate regression splines. *Journal of the Royal Statistical Society (B)*, 65(1):95–114.

Wood, S. N. (2004). Stable and efficient multiple smoothing parameter estimation for generalized additive models. *Journal of the American Statistical Association*, 99(467):673–686.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, 73(1):3–36.

Wright, M. N. and Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, 77(1):1–17.

Xie, Y., Sha, Z., and Yu, M. (2008). Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology*, 1(1):9–23.

Zhan, X., Sohlberg, R., Townshend, J., DiMiceli, C., Carroll, M., Eastman, J., Hansen, M., and DeFries, R. (2002a). Detection of land cover changes using MODIS 250 m data. *Remote Sensing of Environment*, 83(1-2):336–350.

Zhan, X., Sohlberg, R., Townshend, J., DiMiceli, C., Carroll, M., Eastman, J., Hansen, M., and DeFries, R. (2002b). Detection of land cover changes using MODIS 250 m data. *Remote Sensing of Environment*, 83(1-2):336–350.

Zuur, A. F. (2011). *Mixed Effects Models and Extensions in Ecology with R*. Springer.

Zuur, A. F., Saveliev, A. A., and Ieno, E. N. (2014). *A beginner's guide to generalised additive mixed models with R*. Highland Statistics Ltd.