Primary Forest Degradation and Secondary Re-growth Dynamics in the Brazilian Amazon

Yunxia Wang

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

School of Geography

September 2019

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 2

Wang, Y., Ziv, G., Adami, M., Almeida, C.A., Antunes, J.F.G., Coutinho, A.C., Esquerdo, G.C.D.M., Gomes, A.R., and Galbraith, D., 2019. Upturn in secondary forest clearing buffering primary forest loss in the Brazilian Amazon. *Nature Sustainability* (Accepted).

Y.W., D.G. and G.Z. developed the concept and methodological work plan. Y.W. performed the data analysis with support from G.Z. and D.G.. M.A., C.A.A., J.F.G.A., A.C.C., J.C.D.M.E. and A.R.G. coordinated the development of the TERRACLASS products. M.A. contributed visual interpretation for sampling-based analysis. Y.W., D.G. and G.Z. wrote the paper with contributions from M.A.. All authors commented on the manuscript.

Chapter 3

Wang, Y., Ziv, G., Adami, M., Mitchard, E., Batterman, S.A., Buermann, W., Marimon, B.S., Junior, B.H.M., Reis, S.M., Rodrigues, D. and Galbraith, D., 2019. Mapping tropical disturbed forests using multi-decadal 30 m optical satellite imagery. *Remote sensing of environment*, 221, pp.474-488.

Y.W., D.G. and G.Z. developed the concept and methodological work plan. Y.W. performed the data analysis with support from G.Z. and D.G.. M.A provided TERRACLASS products. E.M. contributed to radar data analysis. B.S.M., B.H.M.J., S.M..R and D.R. provided ground-truth validation data. Y.W., D.G. and

G.Z. wrote the paper with contributions from M.A., S.A.B and W.B.. All authors commented on the manuscript.

Chapter 4

Wang, Y., Ziv, G., Adami, M., and Galbraith, D., 2019. Historical Degradation of the Brazilian Amazon. This chapter is being prepared for submission to either *Geophysical Research Letters* or *Environmental Research Letters*.

Y.W., D.G. and G.Z. developed the concept and methodological work plan. Y.W. performed the data analysis with support from G.Z. and D.G.. M.A provided TERRACLASS, DEGRAD, DETEX products. Y.W., D.G. and G.Z. wrote the manuscript with contributions from M.A.. All authors commented on the manuscript.

Acknowledgments

First and foremost, I would like to thank my two brilliant supervisors, Dr. David Galbraith and Dr. Guy Ziv, for all of their training, guidance and fully support during the course of this project. I greatly appreciate being sent to the U.S.A. for inspiring training at the GEE (Google Earth Engine) Summit in 2016, and to the stunning Amazon jungle, both of which laid a solid foundation for this new journey. I am really grateful for David's infinite patience on guiding my writing skills, and Guy's timely responses on debugging my GEE programming scripts.

I also would like to thank Dr. Marcos Adami, for providing TERRACLASS and continued support through the data analysis, as well as for his guidance and support during the Rondônia fieldtrip and my stay in Belém. I thank Dr. Edward Mitchard, for his guidance and support for radar data analysis, and for his inspiration on remote sensing.

I thank all the members of my research support group, Dr. Tim Baker, Dr. Sarah Batterman and Prof. Wolfgang Buermann, for their useful feedback and ideas for this project. I also thank all members of Ecology and Global Change cluster, for their inspirition of science as a teacher, and for their support and encouragement as a friend.

I would especially like to thank my boyfriend, Dr. James Scott, for his love and encouragement whenever I needed, and for continuously making my favourite coffee. I also thank all my friends from the School of Geography: In particular, Julia Tavares and Alex Chambers-Ostler for being great company and keeping me sane at times when things weren't all going to plan, Amy Bennett for helping me with R analysis, Will Baker, Bart Crezee, Dr. Marta Giannichi, Dr. Michelle Kalamandeen, Dr. Adriane Muelbert, Dr. Fernanda Coelho, Dr. Jess Baker, Dr. Jiren Xu and Dr. Lili Xiang, for their wonderful support and encouragement. I thank all my friends from the badminton group for keeping me fit and helping me release pressure when I was down.

I would like to thank the China Scholarship Council, School of Geography, and Google Earth Engine, for their financial support throughout my studies.

Last but not the least, I would like to thank my parents, Cunchao and Shijun, my brothers, Tonghui and Binghui, and my sisters in law, for all of their love and support throughout my oversea studies.

Abstract

The Amazon rainforest is a vital biome that is of central importance for the provision of significant ecosystem services locally, regionally and globally. Brazil contains two-thirds of remaining Amazonian rainforests and is responsible for the majority of Amazonian forest loss. Over 0.7 million km^2 of primary forest area in the Brazilian Amazon has been deforested, of which ~20% are under secondary forest regeneration. However, the fate of secondary forests and the extent of degradation of the remaining primary forests (referred to as old growth forests in this thesis) are still unclear. In this thesis, I present: (1) the first large-scale analysis of secondary forest loss over 14 years (2000-2014) using recently released high resolution (30 m) post-deforestation land use datasets (TERRACLASS); (2) a novel machine learning classification method to map tropical forest disturbances using multi-decadal Landsat time-series imagery; and (3) first estimates of the historical degradation of remaining old growth forests using this newly-developed classification method. Our results show an accelerated loss of secondary forests across the entire Brazilian Amazon over our study period, in contrast to primary forest loss. Over 2000-2014, the proportion of total forest loss accounted for by secondary forests rose from (37 $\pm$ 3) % in 2000 to (72 $\pm$ 5) % in 2014. We developed a multi-decadal Landsat time-series imagery and machine learning random forest classification algorithm, which we found to be an efficient and accurate approach to map tropical disturbed forests. This approach allows me to map the historical degradation of old growth forests from 1984 to 2014. Until 2014, over 246,845 km^2 area of old-growth forests in the Brazilian Amazon (moist forest ecoregion) were degraded, accounted for approximately 10% of total area of old growth forests in the region. However, this approach may have underestimated the actual degradation of old growth forests as it did not detect the low intensity selective logging. In conclusion, the accelerated loss of secondary forests and extensive degradation of old growth forests in the Brazilian Amazon which we report have provided new insights into land use change dynamics in Amazonia. Both of these processes have important

implications for carbon storage and biodiversity and sustainable management of forest resources in the Brazilian Amazon.

Table of Contents

List of Tables

List of Figures

Chapter 1
Introduction

1.1 Background

Although tropical forests only cover 7% of Earth's land surface, they are hotspots for global biodiversity and carbon storage and play an important role in supplying multiple ecosystem services (Foley et al., 2005). Several decades of research have revealed the significance of tropical forests in the global carbon cycle. There are 247 petagrams (Pg) of live biomass carbon stocks in tropical forests, with 78% stored in aboveground biomass (Saatchi et al., 2011). It is estimated that tropical forests were a carbon sink of 2.83 Pg C yr^{-1} from 1990 to 2007 (Pan et al., 2011).

However, tropical forests have experienced significant degradation and deforestation (Chazdon, 2003), resulting in considerable carbon emissions. According to Hansen et al. (2013), 2.3 million km^2 of forest have been lost globally over 2000-2012, with great forest loss in the Tropics, where annual forest loss area increased by 2101 km^2 yr^{-1} between 2000 and 2012. These large areas of tropical forest deforestation resulted in 2.94 ± 0.47 Pg C yr^{-1} of emissions during 1990-2007, accounting for 40% of the global fossil fuel emissions (Pan et al., 2011). However, the effect of deforestation emissions is offset somewhat by the regrowth fluxes of secondary forests. When these are considered, the *net* carbon fluxes from land use changes have been estimated to account for 12.5% of anthropogenic CO_2 emissions from 1990 to 2010 (Houghton et al., 2012). Carbon emissions from deforestation and degradation are the second largest anthropogenic source of carbon dioxide to the atmosphere (Van der Werf et al., 2009).

Currently, a significant proportion of previously deforested tropical areas is under some form of secondary re-growth, although the magnitude of this effect remains unclear. For example, the tropical secondary forest sink was recently estimated to be 1.72 Pg C yr^{-1}, up to 70% greater than that of intact tropical forests (Pan et

al., 2011). These numbers take into account both estimates of secondary forest area and their assumed carbon accumulation potential, which can be very high. For example, Poorter et al. (2016) estimated a net carbon uptake of 3.05 Mg C ha^{-1} yr^{-1} for Neotropical secondary forests, approximately 11 times the mean uptake rate reported for old-growth forests.

On the other hand, secondary forests may also face repeated clearing and burning, which decreases their capability to absorb carbon (Zarin et al., 2005). Repeated cutting and burning of secondary forests has been found to reduce carbon accumulation rates of secondary forests by up to 50% (Zarin et al., 2005). However, this re-cutting frequency of secondary forests is not still not clear and studies of carbon fluxes in tropical forests rarely consider re-cutting of secondary forest.

Additionally, our understanding of pan-tropical secondary forest change dynamics is hampered by the absence of precise information on their extent and successional status. Although there have been a number of studies attempting to use various remote sensing metrics to classify tropical forests into different age or successional stage classes, these have typically been at a local scale and have only considered one or two remote sensing metrics at a time. There are currently no existing products that provide accurate information on the pan-tropical occurrence of secondary forests and their change dynamics. In the Brazilian Amazon, the TERRACLASS product (Almeida et al., 2016) developed by INPE (the Brazilian Space Institute), provides spatial maps of secondary forests with 30 meter resolution and these are now available for multiple years. Thus, it is possible to assess the large-scale temporal change dynamics of secondary forests in the Amazon for the first time.

Spanning an area of ~7 million km^2, the Amazon basin contains the world's largest area of remaining tropical forests and provides a host of ecosystem services of importance globally, regionally and locally. Approximately two-thirds of remaining Amazonian forests are found in Brazil (Gloor et al., 2012; RAISG, 2012; Grace, 2016), making the Brazilian Amazon the largest tropical forest area within a single country. Thus, accurate estimates of forest

deforestation/degradation and the fate of deforested areas (secondary forest dynamics) in the Brazilian Amazon are of paramount importance for understanding the broader Neo-tropical and global carbon cycle.

Given the significant importance of the Brazilian Amazon, the National Institute for Space Research (INPE) in Brazil has developed several projects to monitor forest change in the region (Diniz et al., 2015; Almeida et al., 2016): (1) PRODES for monitoring deforestation since 1988; (2) DEGRAD for mapping degradation (2007-2016); (3) DETEX for selective logging detection (2009-2015); (4) TERRACLASS for mapping the fate of deforested areas (1991, 2000, 2004, 2008-2014 bi-annually); (5) DETER-B for near real time deforestation and degradation monitoring. While there has been much analysis and discussion of PRODES data, the other more recent products (e.g. TERRACLASS) have been little explored to address questions of forest change dynamics.

Brazilian Amazon rainforests play an important but poorly quantified role in the global carbon cycle due. Key gaps in our understanding include 1) the change dynamics of secondary forests and 2) historical degradation of the old growth forests. With the support of Google Earth Engine (GEE), we undertake the first large-scale assessment of the spatio-temporal dynamics of secondary forests in the Amazonia, and develop a novel machine learning method to map historical degradation over 30 years (1984-2014). These will enhance understanding of the spatio-temporal dynamics of tropical secondary forests and degradation of the old growth forests, and support Brazil reporting on land use, land use change and forestry (LULUCF) and provide baseline information important for Brazil to meet its Nationally Determined Contribution (NDC) as mandated by the Paris Agreement.

1.2 Deforestation in the Brazilian Amazon

Land use and land cover in the Amazonia have undergone considerable changes over the last few decades. These changes are believed to have had a strong impact on regional and global carbon and water balance, contributed to global climate change. Accounting for ~60% of the Amazon rainforests, the Brazilian

Amazon is responsible for majority of forest loss in the region (Kalamandeen et al., 2018).

1.2.1 PRODES estimates of deforestation

Deforestation in the Brazilian Amazon has been well mapped annually by the National Institute for Space Research (INPE) through the PRODES project since 1988 (PRODES, 2018; Valeriano et al., 2004). PRODES uses satellite images from Landsat, CBERS (China-Brazil Earth Resources Satellite program) and IRS (the Indian Remote Sensing satellite program) to detect annual deforestation larger than 6.25 hectares. The detection process mainly depends on the visual interpretation of satellite composite image and the vegetation, soil, and shade fraction images generated from linear spectral mixing models (Valeriano et al., 2004).

Data from PRODES (Figure 1.1) show that annual deforestation area across the Brazilian Amazon fluctuated during the period of 1988-2018, with two peaks in 1995 and 2004. The large deforestation rates reported in 1995 almost have been linked to the 1994 Real Plan for economic stabilization and to large-scale forest fires (Lele 2000). After 1995, deforestation experienced a period of decease until 1997. For the period of 2001-2004, the significant increase of deforestation is primarily attributed to conversion of forest to cropland due to the high price of soybean, resulting in the increase of deforestation area from 18165 km^2 in 2001 to 27772 km^2 in 2004 (Morton et al., 2006) . From 2004 to 2014, there is a marked deforestation drop of 82% for the entire Brazilian Amazon and 91% for Mato Grosso. This is attributed to the implementation of a significant government program to combat illegal logging and forest fires in Amazonia, namely the PPCDAm program (The Action Plan for the Prevention and Control of Deforestation in the Legal Amazon) (Maia et al., 2011) and the Sustainable Amazon Plan PAS (2008) (CASTELO, 2015). PPCDAm was launched in 2004 to reduce deforestation rates and support sustainable development in Amazonia. In this period, the Brazilian government invested large sums of money to implement strategic actions related to land reform policy, improvement of deforestation

monitoring systems and the promotion of sustainable use of already deforested areas. All of these efforts together with international mechanisms such as the soybean (Gibbs et al., 2015; Rudorff et al., 2011) and beef moratoria (Gibbs et al., 2016; Massoca et al., 2017) made a considerable contribution to controlling deforestation in the Amazon.

The deforestation rates in Brazilian Amazon have increased in recent years. The area of deforestation increased by 50% from 5,012 km^2 in 2014 to 7,536 km^2 in 2018 (Figure 1.1). Furthermore, recent political changes in Brazil which focus less on the protection of the Amazon and more on its exploration have raised concerns that such increases may continue, at least in the near future.

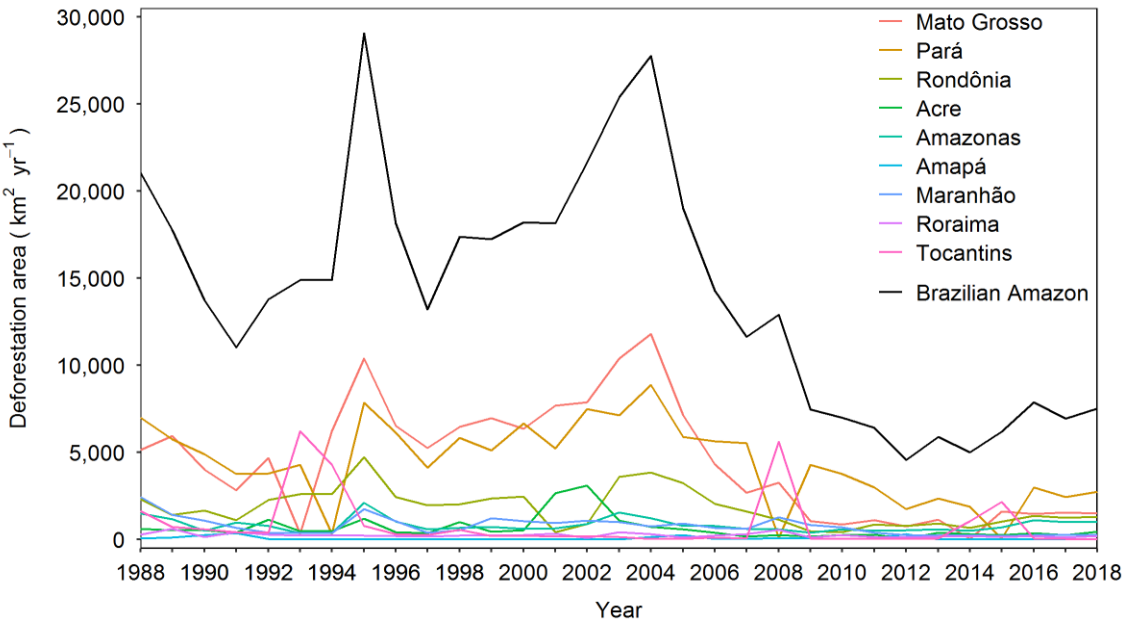


Figure 1.1 The area of annual deforestation of the Brazilian Amazon from PRODES project in INPE (National Institute for Space Research).

1.2.2 Consequences of deforestation

Continued deforestation in Amazonia has caused considerable impacts on the local and global climate (Houghton and Nassikas, 2017; Spracklen and Garcia-Carreras, 2015). In the last 60 years, the temperature in the Amazon has

increased by 1°C with deforestation reaching 20% of its forested areas (Nobre et al., 2016). During the dry season, local warming from deforestation can reach up to 1.5°C (Baker and Spracklen, 2019). Also, deforestation could result in the significant reduction of rainfall. Deforestation up to 2010 has reduced annual mean rainfall across the Amazon basin by 1.8 ± 0.3%, with a potential reduction of 8.1 ± 1.4 % by 2050 (Spracklen and Garcia-Carreras, 2015). Globally, deforestation in the Amazon was estimated to release carbon emissions of 0.18 ± 0.07 Pg C yr^{-1} between 2000 and 2010, 79% of which were accounted for by the Brazilian Amazon (Song et al., 2015). Deforestation in the Brazilian Amazon contributed 17% of global land use carbon emissions in the 1990s and early 2000s, reducing to 9% by 2010 as a result of the slow-down in deforestation (Aragao et al., 2014).

1.3 Secondary forest regrowth

A considerable fraction (~20%) of previously deforested land in the Brazilian Amazon is now under secondary regrowth (Almeida et al., 2016). Although they cannot fully compensate for the services provided by primary forests, these secondary forests play increasingly important roles in carbon sequestration and biodiversity maintenance in Amazonia.

1.3.1 The definition of secondary forest

In order to better quantify the extent and change dynamics of secondary forests, it is essential to first clearly define what is meant by a secondary forest. Although there are a large number of studies referring to secondary forests in the literature, there is considerable ambiguity regarding the use of the term “secondary forest”. Some regard secondary forests as regrowth after *natural or human* disturbances of original forests, while others only consider secondary forests as those that regrow following clear-felling for agriculture, pasture or *other human activities* (Guariguata and Ostertag, 2001). There is also disagreement with regards to the intensity of disturbance necessary for classification of secondary forests with some authors considering all disturbed forests to secondary, irrespective of the

degree of disturbance and others considering secondary forests to be vegetation regrowth on land which was totally cleared or where 90% of forest cover has been cleared (Chokkalingam and de Jong, 2001).

By synthesizing current definitions of secondary forests, Chokkalingam and de Jong (2001) considered secondary forest as “forests regenerating largely through natural processes after significant human and/or natural disturbance of the original forest vegetation at a single point in time or over an extended period, and displaying a major difference in forest structure and/or canopy species composition with respect to nearby primary forests on similar sites”, with minimum criteria that “land ≥ 0.5 ha in area and width of more than 20 m, with > 10% crown cover of trees = 5 m in height”. This definition is based on clear and objective criteria and is quite flexible, making it widely applicable.

Another accepted definition of secondary forests is from the International Tropical Timber Organisation (ITTO). They define secondary forests as “woody vegetation re-growing on land that was largely cleared of its original forest cover (i.e. carried less than 10% of the original forest cover). Secondary forests commonly develop naturally on land abandoned after shifting cultivation, settled agriculture, pasture, or failed tree plantations.” Also, secondary forests may be the result of natural forest regeneration after large natural disturbances such as wildfires, storms, landslides and floods (ITTO, 2002). From these two definitions, it is clear that causes of disturbance, intensity of these disturbances, vegetation development process should be considered when defining secondary forests. Both of these definitions emphasize severe natural disturbances and human-induced disturbances of original forests, as well as natural regeneration after disturbances.

In this thesis, secondary forests refer to areas that have been previously deforested according to PRODES (regrowth on previously clear-cut areas >6.25 ha) and converted to other land uses (e.g. pasture, agriculture and mining) but which have subsequently undergone a recovery process following abandonment.

1.3.2 Benefits of secondary forest

Secondary forests have considerable carbon mitigation potential due to their rapid biomass recovery, although these rates can vary considerably. Across the first 20 years of secondary regrowth in the Eastern Amazon, biomass was found to recover at 1.2% per year, equivalent to a carbon uptake of 2.25 Mg ha^{-1} yr^{-1} (Lennox et al., 2018). In a review of biomass recovery across the Neotropics, Poorter et al. (2016) estimated the annual net carbon uptake of secondary forests to be 3.05 Mg C ha^{-1} yr^{-1}, 11 times uptake rates of old-growth forests and 2.3 times uptake rates of selectively logged Amazonian forests.

The carbon accumulation rates of secondary forests differ across successional stages, being highest in the early stages (Brown and Lugo, 1990; Xaud et al., 2013). Based on a secondary forest growth model in central Amazonia, Neeff and dos Santos (2005) found that aboveground biomass (AGB) of secondary forests grows rapidly in the first 20 years, with the greatest growth rate at around 13 years. This result is quite similar to that from a recent study from Bonner et al. (2013), who estimated that for secondary forests younger than 18 years old, the mean aboveground carbon accumulation rate is 3.5 Mg ha^{-1} yr^{-1}, which on average was found to be 40% higher than secondary forests older than 18 years. In the Brazilian Amazon, it has been estimated that forest biomass recovers to approximately 70% of original biomass within 25 years and then takes another 50 years to recover the remaining 30% (Houghton et al., 2000).

Secondary forest carbon accumulation rates also vary substantially across various sites. For example, Poorter et al. (2016) found that after 20 years of recovery, secondary forest AGB accumulation differs from 20 to 225 Mg ha^{-1} across 44 different study sites, equivalent to 25%-85% of old-growth forest AGB in the same site. A key factor explaining these differential recovery rates is background climate, as forests on drier sites were found to recover less quickly than those on wetter sites.

Compared to the rapid biomass recovery of secondary forests, the biodiversity in tropical secondary forests recovers slower (Martin et al., 2013). While carbon

pools may take approximately 80 years to recover to similar levels as primary forests, plant biodiversity and faunal biodiversity recovery take more than 100 years and 150 years, respectively (Martin et al., 2013). However, secondary forests are still valuable for biodiversity conservation (Dent and Wright, 2009) and can support important ecosystem functions (Sayer et al., 2017). The avian phylogenetic diversity in secondary forests recovered to old-growth forest levels after 30 years (Edwards et al., 2017). A recent study (Lennox et al., 2018) showed a high degree of biodiversity resilience for secondary forests in the Brazilian Amazon. After over 40 years of regeneration, the secondary forests recovered 88% of species richness and 85% species composition relative to the undisturbed primary forests.

1.3.3 The current status of secondary forest

During the last 100 years, tropical forests have been deforested and replaced by pastures, agriculture and regenerating forests at unprecedented rates. From 1990 to 2000, 8.7 million hectares were deforested annually in tropics, including humid and dry forests, with a further 2.3 million hectares of humid forest being apparently degraded annually through fragmentation, logging or fires (Mayaux et al., 2005). During this period, secondary forests become increasingly important in tropical landscapes, through re-growth on abandoned land. Over 30 years ago, Brown and Lugo (1990) estimated that approximately 40% of total tropical forest area was secondary forest with a formation rate of 9 million hectares per year. Based on 23 local, national and regional studies, Asner et al. (2009) provide the first pantropical estimate of natural forest regeneration. They concluded that at least 1.2% of humid tropical forests were undergoing secondary succession for at least 10 years. However, this figure was corrected by Wright (2010) to be 11.8% because the Asner et al. (2009) study did not correct for the total area surveyed in the 23 studies. In 2015, the FAO global forest resources assessment reported that 76%, 59% and 51% of forest area in Africa, Asia and South America was secondary forest (FAO, 2015), but such estimates depend on bottom-up national statistics which bear very high levels of uncertainty and are based on unreported methodologies. Nonetheless, the vast discrepancies in total secondary forest

area reported across these studies, even on a year-to-year basis, highlight our very limiting understanding of this topic. Furthermore, analysis on spatio-temporal changes of the tropical secondary forests over significant areas simply do not exist.

The recent development of the TERRACLASS products, however, allows for considerably greater analysis of the spatio-temporal change dynamics of secondary forests than has been possible thus far, at least for the Brazilian Amazon. TERRACLASS tracks the fate of deforested patches in Amazonia through visual interpretation of Landsat images, producing a post-deforestation land use classification of the Brazilian Amazon which at the time of our study was available over a 14-year time period (2000, 2004 and 2008-2014 biannually, *https://www.terraclass.gov.br*). TERRACLASS classifies previously deforested areas into one of twelve land use categories including pasture, annual crops, secondary vegetation and urban areas. It is extensively validated via field campaigns to determine the accuracy of classification. According to TERRACLASS, secondary forests comprised approximately 21% of previously deforested areas in the Brazilian Amazon in 2008 (Almeida et al., 2016).

1.3.4 Re-cutting of secondary forest

Although secondary forests have a substantial potential for offsetting carbon emissions, they can be repeatedly cleared and burned. It has been proposed that secondary forests in Brazilian Amazon are cut and burned on average every 5 years (Aguiar et al., 2016), but Almeida (2009) points out that the half-life of the secondary forests varies across the basin, ranging from 3 to 21 years. Two recent studies (Tyukavina et al., 2017; Carvalho et al., 2019) have illustrated the fate/clearance of secondary forests, but neither of them have delivered a comprehensive analysis of the spatio-temporal change dynamics of secondary forests. Tyukavina et al. (2017) is a sampling-based analysis of deforestation (including secondary forest clearance) in the Brazilian Amazon, and Carvalho et al. (2019) is a localised study of nature dynamics of secondary forests in the state of Pará.

1.3.5 Summary of knowledge gaps

Despite the importance of secondary forests for carbon sequestration and biodiversity conservation in Amazonia, a historical lack of spatio-temporal data on secondary forest area has precluded evaluation of their large-scale dynamics. Although prior studies have pointed out the frequent re-cutting of secondary forests in the Brazilian Amazon, a comprehensive analysis of secondary forest loss and its evolution over time does not exist.

1.4 Forest degradation in the Brazilian Amazon

Degradation of old growth forests is a fundamentally important, but poorly characterised, component of Amazonian ecosystems. Unlike deforestation, degradation in the Amazon does not entail a change in land use but rather, impoverishment of the state of old-growth forests. However, the tree cover losses and canopy gaps resulting from degradation can also affect carbon storage and local microenvironment of ecosystems, as well as often paving the way for clear-felling.

1.4.1 The definition of forest degradation

Forest degradation, commonly defined as the reduction of the capacity of a forest to provide goods and services (Simula, 2009; Thompson et al., 2013), is characterised by the partial destruction of forest canopy. In the Brazilian Amazon, national estimates of deforestation based on PRODES (INPE) consider only patches of forest that have been completely clear-felled following land use transitions to pasture or other non-forest land uses, and that have reached a minimum size threshold of 6.25 ha. Forests loss below that threshold or forest areas that have been degraded (e.g. burned) but not completely clear-felled is not included in PRODES deforestation mapping.

Forest degradation in this thesis therefore encompasses degradation arising from multiple drivers of disturbances with a gradient of disturbance intensity. It includes small-scale clearings less than 6.25 ha, heavy disturbances associated with

mining and road constructions, as well as fire damage which can span a broad range of intensity from mega-fires to small-scale burnings. Forest degradation is not necessarily restricted to anthropogenic drivers but also include natural disturbances associated with river flooding and wind-throws, for example (Foley et al., 2007; Espírito-Santo et al., 2014; Espírito-Santo et al., 2010). Following definition of forest degradation from IPCC (Intergovernmental Panel on Climate Change) (de Cambio Climático, 2003), normal forest management such as thinning and harvest (selective logging) were not considered as degradation in this thesis.

1.4.2 Impacts of forest degradation

Although degradation only involves removal of partial forest canopy, it can cause considerable damage. Degradation (e.g. selective logging, fire) can alter the structure and functioning of the old-growth forests (Struebig et al., 2013), leading to declines in biomass, losses of biodiversity (Barlow et al., 2016) and the reduction of productivity (Xaud et al., 2013).

A large filed study across eastern Amazonia has found that, on average, forests experienced both logging and understory fires lost 40% of their aboveground carbon compared with the undisturbed forests (Berenguer et al., 2014). Even for the low-impact logging without any fire occurrence could also cause the loss of 4-21% of initial forest aboveground carbon (Longo et al., 2016). Forest areas that burned by multiple times can lead to losses of their aboveground carbon by up to 94% (Longo et al., 2016). A recent remote-sensing based pan-tropical study concluded that tropical forest was a significant carbon source between 2003 and 2014, while 69% of carbon losses were due to forest degradation (Baccini et al., 2017). Forest degradation also affects the biodiversity (Barlow and Peres, 2004a; Struebig et al., 2013). A study in Brazilian state of Pará showed that the anthropogenic disturbances (e.g. logging, fire) could double the biodiversity loss arising from deforestation and more negatively affect species with the higher conservation importance (Barlow et al., 2016)

Degradation may also change the local microclimate and fire regimes, increasing the probability of deforestation in the region. For example, selective logging increases fire occurrence by providing abundant fuel loads and forest edges which are more vulnerable to desiccation during prolonged periods of dry weather (Nepstad et al., 1999; Laurance et al., 1998; Barlow and Peres, 2004b; Alencar et al., 2004). Logging has increased the probability of deforestation in the region by up to 400% (Asner et al., 2006). Approximately 19 ± 11% of previously logged forests were subsequently deforested within 3 years (Asner et al., 2005). Accidental fires have affected nearly 50% of remaining forests and caused more deforestation than has intentional clearings (Cochrane et al., 1999).

Forest degradation induced by natural disturbances such as wind-throws (Negrón-Juárez et al., 2017) could also change the forest structure (Marra et al., 2014), leading to an increase of tree mortality and significant loss of aboveground carbon (Espírito-Santo et al., 2014). The occurrence of blow- downs was found to be highly correlated with the frequency of heavy rainfall (Espírito-Santo et al., 2010), and the affected forest area by these natural disturbances could range from 0.01 ha up to 2,651 ha (Espírito-Santo et al., 2014). Negrón-Juárez et al. (2010) demonstrated that a single squall line aligned with storms in 2005 over Manaus caused the tree mortality that was equivalent to 30% of the observed annual deforestation in the region.

1.4.3 The current status of forest degradation

Compared to the long-term deforestation monitoring (PRODES) of Brazilian Amazon, the detection of forest degradation is relatively new and less developed. The DEGRAD satellite monitoring programme developed by INPE (the Brazilian National Institute for Space Research), provided degradation data from 2007 to 2016. The other programme – DETEX was for detecting selective logging but only last for 7 years (2009-2015).

Over 2007-2016 period on average, 23,181 km^2 yr^{-1} of forest were degraded due to either selective logging (DETEX) or other disturbance events (e.g. fire damage, DEGRAD), which is approximately the triple size area of the deforestation over

same period (7,502 km^2 yr^{-1}, PRODES). However, according to another ten-year report of degradation and deforestation in the Brazilian Amazon, the estimated area of degradation due to selective logging and forest fires during 2000-2010 was 5,082 km^2 yr^{-1}, only equivalent to 30% of the area of deforestation over same time period (Souza Jr et al., 2013). From 2009 to 2010, the combined annual estimated areas of degradation from DEGRAD and DETEX were up to 382% higher than the estimates from Souza Jr et al. (2013) (Figure 1.2**).** Such divergent estimates of degradation may be due to the definition of degradation and the different mapping methods. The polygon-based visual interpretation (DEGRAD) approach is more likely to overestimate the degradation than the pixel-based decision tree classification (Souza Jr et al., 2013). This highlights the existing large uncertainties of area estimates of degradation for the region.

To improve the degradation detection system in the Brazilian Amazon, the Amazon Regional Centre of INPE (INPE-CRA) has developed a project called DETER-B, the near real-time degradation and deforestation detection system in the Brazilian Amazon (Diniz et al., 2015). DETER-B is mainly based on the visual interpretation of AWFIS imagery (Advanced Wide Field Sensor, Indian Earth Observation satellite) at 56 m spatial resolution and 5 days of temporal resolution, complying with historical time-series images from Landsat, LISS (Linear Imaging Self Scanning Sensor, Indian Earth Observation satellite), and DMC (Disaster Monitoring Constellation satellite). DETER-B could effectively detect the clear-cut deforestation and the degradation resulting from selective logging, fire, and other moderate degradation events (Diniz et al., 2015). According to DETER-B, the degradation in the Brazilian Amazon had declined dramatically, from 28,798 km^2 yr^{-1} in 2015 to 4,953 km^2 yr^{-1} in 2019. However, the accuracy of such estimates is unknown.

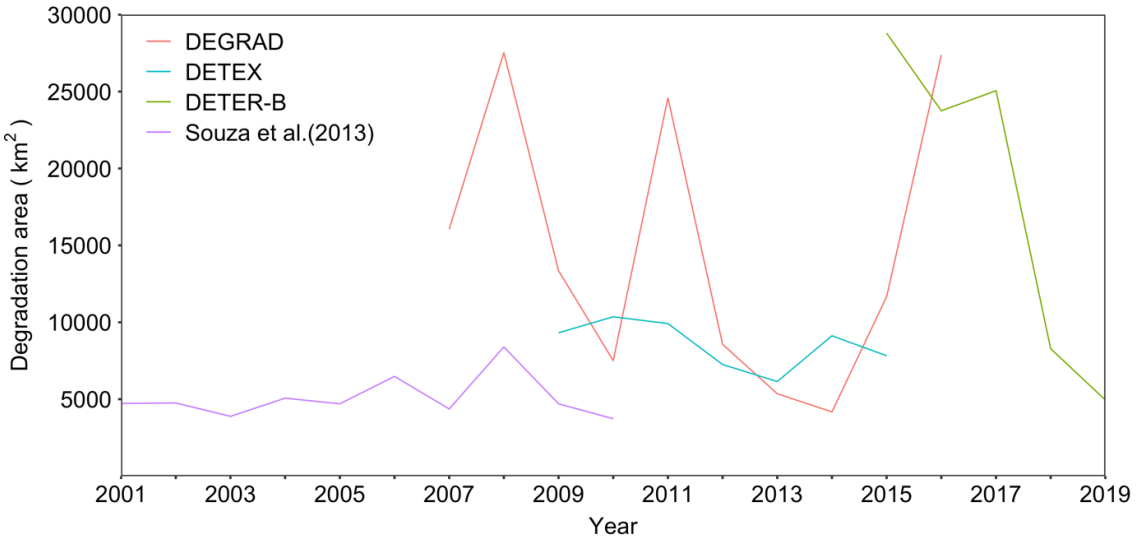


Figure 1.2 Degradation in the Brazilian Amazon.
DEGRAD (2007-2016): degradation without selective logging. DETEX (2009-2015): selective logging. DETER-B (post 2015): including degradation, selective logging, burned areas. Souza Jr et al. (2013): degradation from selective logging and fires.

1.4.4 Summary of knowledge gaps

Decades of studies have sought to quantify forest degradation in Amazonia. However, these have been incomplete as they have focused on studying one driver of degradation in isolation (e.g. logging) or have been restricted to short timeframes. Thus, a fully comprehensive evaluation of the extent of historical forest degradation in Amazonia, still remains elusive. In this thesis, I attempt to address this critical gap by explicitly estimating the area of old-growth forest in the Brazilian Amazon degraded between 1984 and 2014.

1.5 Mapping secondary forest and degradation from space

Although TERRACLASS (INPE project) provides detailed information on the spatial distribution of secondary forests in the Brazilian Amazon, it involves a huge effort based largely on visual interpretation and does not map the degradation of old growth forests. There is thus an important need to develop approaches for detecting forest disturbance that can be applied over large areas which consider both regrowth dynamics and degradation of old-growth forests.

1.5.1 Classification of land use/land cover

Remote sensing has played a key role in identifying land use/land cover changes (Table 1.1). Visual interpretation is the earliest and simplest method, but it is subjective and susceptible to interpreters' experiences and knowledge level (Mohammady et al., 2015). Automatic classification methods, divided into supervised classification and unsupervised classification, are much more widely used (Long and Sriharan, 2004; Xia et al., 2014; Bayoudh et al., 2015). Both supervised and unsupervised classification are pixel-based classification techniques. The main difference between them relates to the use of training data. In supervised classification, a set pixels whose class is known (training sites) are required to generate representative parameters for each class of interest, that are used to train classifiers to classify all image pixels into relevant classes. Unsupervised classification, on the other hand, is an approach that groups image pixels into several classes through clustering algorithms without any foreknowledge of these classes (Liu, 2005). Besides pixel-based classification, object-based classification is also an efficient approach to extract land use/land cover information (Mallinis et al., 2008).

Visual interpretation is a traditional approach that aims to interpret land cover information from satellite images on the basis of visual elements, such as shape, size, location, tone and texture, highly depending on interpreters' knowledge and their interpretation experience (Ali, 1989). Tone refers to the relative brightness or colour of objects in an image. Texture refers to the arrangement and frequency of tonal variation in particular areas of an image. Although the interpretation result is time consuming and subjective, this method is widely used because it is flexible and easy to operate (Prasad et al., 2002; Stuart et al., 2006). Van den Broek et al. (2004) found that visual interpretation classification was more effective than automatic classification when using polarimetric SAR data.

Supervised classification includes two phases: 1) the training phase, identifying a classification scheme based on spectral signatures of different bands obtained from training sites with known class labels; 2) the prediction phase, where the

classification scheme is applied to other sites without known class labels (Samaniego and Schulz, 2009). Several algorithms have been developed for supervised classification, such as Maximum Likelihood Classification (MLC) (Otukei and Blaschke, 2010), K Nearest Neighbourhoods (K-NN) (Samaniego and Schulz, 2009), Support Vector Machines (SVM) (Longepe et al., 2011), Logistic Discrimination (LD) classifier, decision trees (DTs) (Fu et al., 2010), fuzzy classifier (Sonmez and Onur, 2012). Marcal et al. (2005) found that SVM and LD were much better than MLC and K-NN in land use/land cover classification by comparing their classification accuracies based on a multispectral image from ASTER sensor. Shiraishi et al. (2014) compared five supervised classifiers in tropical land use/land cover classification, including Naïve Bayes, AdaBoost, multi-layer perceptron, random forest (RF), and support vector machine, and concluded that RF was a useful classifier for analysing ALOS-PALSAR mosaic data. Meanwhile others have concluded that decision tree-based algorithms provide higher classification accuracy than MLC and SVM (Otukei and Blaschke, 2010).

Compared with supervised classification, unsupervised classification is more time and cost efficient because for this method there is no need to choose specific training sites before classifying land covers (Mohammady et al., 2015). ISODATA is the most widely used variant of unsupervised classification, grouping pixels with similar spatial and spectral characteristics into relevant classes (Mohammady et al., 2015; Bakr et al., 2010).

In contrast to pixel-based classification approaches, object-based classification first aggregates image pixels into spectrally homogenous image objects through an image segmentation algorithm and then classifies the individual objects into various classes.

Because of various individual characteristics of these classification methods, many researchers choose more than one approach to produce more accurate classifications (Mohammady et al., 2015; Bakr et al., 2010; Long and Sriharan, 2004). Mohammady et al. (2015) suggested unsupervised classification is

suitable for obtaining agriculture land use due to significant changes of spectral signatures of agriculture pixels among different agriculture types and seasons.

Table 1.1 Main land use/land cover classification methods

Methods		Descriptions	Characteristics
Visual Interpretation		Images are analysed by experienced interpreters, who base their analyses on the objects' colours, spectral signatures, shapes, sizes and positions in relation to each other in the picture.	High accuracy, reliable; Time consuming
Spectra-based Classification	Supervised Classification (classes defined by user)	Two steps: 1) training--use available known pixels to generate representative parameters for each class; 2) the classifier is then used to attach labels to all the image pixels according to the trained parameters Classifier--Maximum likelihood, K-nearest neighbour, Support vector machine, Principal Component Analysis, Artificial neural network	Based on each pixels
	Unsupervised Classification (classes defined by computer)	No training samples, set the number of classes and maximum iterations change threshold, then classify automatically Classifier---ISOSEG classifier and ISODATA classifier	
Decision Tree Classification		Split a dataset into homogenous subgroups based on measured attributes. Based on attributes, create a decision tree which includes numbers of nodes and leaves. Each node represents a decision and each leaf represents a unique class.	Could add external dataset (DEM data)
Object Based Classification		1) Create objects through segmentation and merging of images 2) Create rules for each object based on spectral characteristics, texture, area, etc. 3) According to these rules and objects, generate a classification map Including rule-based classification and example based classification	Suitable for high resolution images

1.5.2 Secondary forest mapping

Many studies have been dedicated to classifying secondary forests according to age or successional stage (Kimes et al., 1999; Nelson et al., 2000; Neeff et al., 2006; Aguiar et al., 2016). In general, secondary forests are grouped into three stages: initial, intermediate and advanced successional stages (ITTO, 2002; Lu, 2005; Lu et al., 2014a). In the initial stage, herbs, shrubs and climbers establish quickly after disturbances, becoming the dominant structural elements, as well as seedlings and saplings. Several years later, pioneer tree species emerge gradually which develop a canopy very quickly and dominate for 10-20 years in what is referred to as an intermediate successional stage. This stage is a mix of dense saplings and young trees with higher canopy than initial stage. Eventually, pioneer species are replaced by other already-established light-demanding species that take advantage of improved growth conditions and gradually become dominant. This stage is referred to as an advanced successional stage and may last for 75-100 years. In this stage, there is a clear stratification of forest stand structure (Lu, 2005). With the availability of remote sensing technologies, various methodologies of distinguishing the successions of secondary forest have been developed.

1.5.2.1 Single-date data based approaches

Scientists have made great efforts in mapping forest regenerating stages based on the single-date various types of remote sensing data, including Landsat imagery (Lu, 2005; Vieira et al., 2003; Carreiras et al., 2014), Synthetic Aperture Radar (SAR) data (Kuplich, 2006), SPOT HRV (Kimes et al., 1999), and CHRIS/PROBA Hyperspectral Images (Millan et al., 2015). Although using one type of remote sensing data (e.g. Landsat TM data or SAR data) is efficient in discriminating pasture and mature forest, data from more than one sensor is often combined to classify successional stages of secondary forest (Kuplich, 2006). For example, different bands from SAR and optical TM have been combined to increase the classification accuracy of forest succession (Kuplich, 2006). There are two ways to combine these multi-sensor data: one is based on incorporation

of one sensor's data as extra bands added to multispectral imagery, the other is using data fusion techniques (Lu et al., 2011). The data fusion approach is an effective way for combining multi-sensor or multi-resolution data to enhance visual interpretation or quantitative analysis, generating a new dataset that contains improved information from both original datasets. The main data fusion methods consist of principal component analysis (PCA), wavelet-merging techniques (Wavelet), high pass filter resolution merging (HPF), and normalized multiplication (NMM). It has been found Wavelet multi-sensor fusion and HPF could increase 3.3%-5.7% classification accuracy by integrating Landsat TM and Radar Data, but PCA and NMM could decrease the classification accuracy (Lu et al., 2011). Carreiras et al. (2017) further demonstrated the use of combined Landsat spectral bands with ALOS PALSAR backscatter intensity to distinguish secondary regrowth forest and mature forest in three landscapes in Brazilian Amazon.

The majority of classification algorithms widely used to classify secondary forests are based on the above land use and land cover classification methods, including parametric algorithm like maximum likelihood classifier-MLC, and non-parametric approaches such as artificial neural network-ANN (Kuplich, 2006), linear analysis (Kimes et al., 1999; Lu, 2005), K-nearest neighbour-KNN, support vector machine-SVM, classification tree analysis-CTA (Millan et al., 2015). Lu et al. (2014b) compared a range of classification methods and concluded that MLC and CTA were suitable for Landsat data and fusion images and KNN was the best choice for the combination of Landsat and PALSAR data as extra bands. For moist tropical regions, MLC based on fusion images was suggested for vegetation classification (Lu et al., 2014b). Using a decision tree approach and CHRIS/PROBA Hyperspectral Images, Millan et al. (2015) found that images from the dry season were generally better for mapping successional forests, but for early stages of succession, using wet season images provided higher map accuracy. Thus, the selection of classification algorithms depends on both study areas and available remote sensing data.

1.5.2.2 Classification based on vegetation indices

Biophysical properties, such as tree height, diameter at breast height (DBH), tree density, leaf area index (LAI), vary within different secondary forest successions. A number of vegetation indices have been created based on different band combinations of remotely sensed data to measure differences in these biophysical properties, and have been used to identify secondary forests. The Normalised Difference Vegetation Index (NDVI) which uses a ratio between red and near-infrared bands is one of the most commonly used vegetation indices (VI) (Boyd et al., 1996; Zhao et al., 2009; Sader et al., 1989). NDVI values range from -1.0 to 1.0. In general, barren areas, sand, or snow show very low NDVI values (0.1 or less). Sparse vegetation such as shrubs and grasslands or senescing crops may result in moderate NDVI values (approximately 0.2 to 0.5). High NDVI values (approximately 0.6 to 0.9) correspond to dense vegetation which is at peak growth stage. Besides NDVI, several other vegetation indices are useful for mapping disturbed forests (Table 1.2, Table 1.3).

Using vegetation indices is efficient for identifying younger secondary forests. Study from (Sader et al., 1989) indicated NDVI difference is not detectable for secondary forests older than 15–20 years. Also, most studies are achieved by the combination of several indices instead of using one index. For example, by assessing Landsat TM radiance data, Boyd et al. (1996) found middle and thermal infrared wavebands contained significant information for detecting re-growing secondary forests in the Amazonia and demonstrated that secondary forests were more separable using a vegetation index acquired in the middle and thermal infrared wavebands than NDVI. A later study, which used both middle infrared band and NDVI to distinguish secondary forests, showed secondary forests at initial (3-6 years), intermediate (10-20 years), advanced (40-70 years) stages can be separated visually in a plot of NDVI and band 5 (middle infrared) reflectance (Vieira et al., 2003).

Besides using vegetation indices, extracting image texture information based on these vegetation indices or various reflectance bands, has been shown to be a promising tool to discriminate secondary forests (da Silva et al., 2014). Image

texture refers to spatial arrangement and variation of image pixels. A large number of variables have been used to measure image texture, such as calculating the range, skewness of pixel values, and mean, variance, correlation, entropy variables derived from the grey-level co-occurrence matrix (GLCM) which consider spatial relations between groups of two neighbouring pixels. By calculating 40 texture variables based on red, near infrared bands, EVI and NDVI derived from a high-resolution satellite image (Quickbird), Gallardo-Cruz et al. (2012) demonstrated image texture could well predict vegetation attributes (e.g. basal area, canopy cover) and reflect the internal heterogeneity of successional vegetation at the proper scale. Early work by Kimes et al. (1999), based on the time-series SPOT HRV data from 1986 to 1994, found that the use of texture information would highly increase secondary forest discrimination accuracy.

Table 1.2 Main Vegetation Indices used to identify secondary forest succession

Vegetation Index Type	Formulation
NDVI---Normalised Difference Vegetation Index	$NDVI = (NIR - Red)/(NIR + Red)$
EVI--- Enhanced Vegetation Index	$EVI = G * (NIR - Red)/(NIR + C1 * Red - C2 * Blue + L)$ (L is the canopy background adjustment, C1, C2 are the coefficients of the aerosol resistance term. For Landsat-EVI algorithm: L=1, C1 = 6, C2 = 7.5, G (gain factor) = 2.5)
SAVI --- Soil Adjusted Vegetation Index	$SAVI = ((1 + L) * (NIR - Red))/(NIR + Red + L)$ (L is a canopy background adjustment factor, L=0.5)
SR---Simple Ratio	$SR = Red/NIR$
RB---Ratio based	$RB = TIR/(Red + MIR)$
CD---Complex division	$CD = MIR/(TIR * Red * MIR)$
NBI---Normalized based index	$NBI = (MIR - (NIR + Red))/(MIR + (NIR + Red))$
IRI---Infrared Index ($NDWI_{1640}$)	$IRI = (NIR - MIR)/(NIR + MIR)$
MIRI---Mid-infrared index ($NDWI_{2130}$, also known as Normalized Burn Ratio)	$MIRI = (MIR - MIR2)/(MIR + MIR2)$

Table 1.3 Landsat 5 TM and Landsat 7 ETM+ Spectral bands

Bands	Wavebands(microns)	Useful for mapping
Band 1_Blue	0.45-0.52	Bathymetric mapping, distinguishing soil from vegetation and deciduous from coniferous vegetation
Band 2_Green	0.52-0.60	Emphasizes peak vegetation, which is useful for assessing plant vigor
Band 3_Red	0.63-0.69	Discriminate vegetation slopes
Band 4_Near Infrared (NIR)	0.77-0.90	Emphasizes biomass content of soil and shorelines
Band 5_Middle Infrared (MIR)	1.55-1.75	Discriminates moisture content of soil and vegetation; penetrates thin clouds
Band 6_Thermal Infrared (TIR)	10.40-12.5	Thermal mapping and estimated soil moisture
Band 7_ Middle Infrared (MIR2)	2.09-2.35	Hydrothermally altered rocks associated with mineral deposits
Band 8_Panchromatic (Landsat 7 only)	0.52-0.90	15 m resolution, sharper image definition

1.5.2.3 Old growth forest degradation mapping

Compared to deforestation monitoring, measuring partial reduction of forest cover (i.e. degradation) from remote sensing is far more difficult. Prior studies in the Amazon have generally focused on mapping one or two types of degraded forests (Souza Jr et al., 2005a; Costa et al., 2019), mainly logged forests (Monteiro et al., 2003; Souza Jr et al., 2003) or burned forests (COCHRANE, 1998).

The combination of linear spectral mixture models (Monteiro et al., 2003; Souza Jr et al., 2005b; Shimabukuro et al., 2019) and the decision tree classification was found to be one of the most efficient approaches to map forest degradation (Souza Jr et al., 2003). The soil fraction derived from the spectral mixture analysis were recognized as the spatial signature of mechanized logging activities, enhancing the detection of the log landings and logging roads (de Wasseige and

Defourny, 2004; Souza Jr and Barreto, 2000). And the non-photosynthetic vegetation fraction were useful for identifying the forest degradation caused by burn (COCHRANE, 1998; Shimabukuro et al., 2009). However, such an approach is better in detecting highly degraded forests that have been subject to recurrent logging and burning. For degraded forests with low canopy damage, Souza Jr et al. (2005a) proposed a Normalized Difference Fraction Index (NDFI) computed using the fraction images obtained from spectral mixture model (fraction of soil, fraction of non-photosynthetic vegetation and fraction of shade of green vegetation), which was subsequently used by other studies (Souza Jr et al., 2013; Daldegan et al., 2019). Yet, these enhanced techniques may still underestimate very small-scale logging disturbances because of the rapid recovery of logging gaps. Although Landsat could detect logging-induced canopy damage up to 3.5 years later, the majority of damaged canopies recovered within one year (Asner et al., 2004). Even for the very high resolution Ikonos images (1 m pixel size), log landings opened by selective logging became ‘cryptic’ after two years (Souza and Roberts, 2005). Moreover, the expensive cost of acquiring these very high-resolution images has limited their usage.

1.5.3 Summary of knowledge gaps

Decades of research have sought to quantify secondary forests regrowth and forest degradation using remote sensing in Amazonia. Whether through visual interpretation, statistical models or machine learning algorithm, they have been limited as they have generally focused on local-to-regional scales or have been restricted to the short timespans. The majority of mapping approaches used thus far are based on single date satellite imagery, which are limited in the discriminatory power they can provide as they make no use of temporal disturbance/recovery signals which characterise secondary forests and old growth forest degradation. Thus, a temporal analysis may prove more useful in this regard. There is also a glaring need for a historical perspective on forest disturbance, which provides insights of the cumulative extent of degradation has affected on the Amazonian old growth forests over multi-decadal timescales.

1.6 Summary

The deforestation of old growth forests in the Brazilian Amazon has been well mapped by PRODES (INPE project), but there are two forest change processes which have received much less attention, namely the spatiotemporal dynamics of secondary forests and the degradation of remaining old growth forests. The recently released TERRACLASS dataset makes it possible to evaluate the large-scale spatio-temporal change dynamics of secondary forests at pan-Brazilian Amazon level. However, there is no available spatial distribution of degraded old growth forests in the Brazilian Amazon. The extent that the old growth forests have been historically degraded in the region is still unknown.

Remote sensing is therefore needed to map the historical cumulative degradation of old growth forests. Prior approaches for degradation mapping have generally been based on the spectral information extracted from a single date satellite image, and have neglected the temporal features of the degradation process. Time-series based classification algorithm has been found to be very useful for mapping forest disturbances (White et al., 2017; Hirschmugl et al., 2017; Hermosilla et al., 2015; Kayastha et al., 2012; Huang et al., 2010; Kennedy et al., 2010; Kennedy et al., 2007), but the majority of these time-series based approaches are based on a single time-series trajectory and have mainly been implemented at local scales in extratropical regions (e.g. Canada, U.S.).

Mapping disturbed forest in this thesis therefore includes both secondary forests and degraded old growth forests. While the temporal spectral responses of secondary forests and degraded old-growth forests may be not exactly the same in terms of the timescales involved, they should exhibit similar overarching features (i.e. disturbance-induced declines in key metrics followed by recovery) which set them apart from intact forests. Based on this expectation, I provide a novel methodology that using long time-series Landsat images to map tropical secondary forests and degraded old growth forests, especially in the Brazilian Amazon.

1.7 Research aims and objectives

The overall aim of the thesis is to evaluate the large-scale spatio-temporal change dynamics of secondary forests in the Brazilian Amazon over 14 years (2000-2014), develop a novel approach to map tropical disturbed forests (secondary forests & degraded forests), and estimate the extent of old growth forest degradation for the Brazilian Amazon. Specifically:

1) To understand spatio-temporal change dynamics of secondary forests across entire Brazilian Amazon from 2000 to 2014.
2) To develop a novel remote-sensing methodology to map the extent of disturbed forests in Tropics, including secondary forests and degradation of old growth forests.
3) To apply the above remote-sensing methodology to the entire Brazilian Amazon moist forest, provide the extent of historical degradation of old growth forests for the region.

1.8 Thesis structure and publication status

This thesis consists five chapters, including the general introduction, three data chapters presented as paper manuscripts, and the final discussion and conclusion.

Chapter 1: Introduction. Literature review of the status of secondary forests and forest degradation in the Amazonia, and remote sensing approaches used for mapping the secondary forests and degradation. Summary of the knowledge gaps and the structure of the thesis.

Chapter 2 (Paper I): Upturn in secondary forest clearing buffering primary forest loss in the Brazilian Amazon. Analysis of spatio-temporal change dynamics of secondary forests in the Brazilian Amazon over 14 years (2000-2014) using high-resolution land use dataset (TERRACLASS, 30 m pixel size), providing the first estimates of secondary forest loss for the region and the lost

carbon sequestration opportunity due to the clearance of secondary forests. This chapter is accepted by *Nature Sustainability*.

Chapter 3 (Paper II): Mapping tropical disturbed forests using multi-decadal 30 m optical satellite imagery. Developing a novel methodology to map the tropical disturbed forests (i.e. secondary forests and degraded old growth forests) using Landsat multiple time-series trajectories, and test this method in three different ecoregions (moist forest, seasonal forest and dry forest ecoregions) in the state of Mato Grosso (Brazil). This chapter has been published in *Remote Sensing of the Environment*.

Chapter 4 (Paper III): Historical degradation of the Brazilian Amazon. Scaling up the developed approach as in Chapter 3, using the same multi-decadal 30 m Landsat time-series images (1984-2014) and the same algorithm to classify moist primary forests (i.e. those which have not been deforested according to PRODES, referred as old growth forests in the thesis) into *intact vs. degraded* in 2014. This chapter is being prepared for submission to either *Geophysical Research Letters* or *Environmental Research Letters*.

Chapter 5: Discussion and Conclusion. The contributions of this work (thesis) to the relevant scientific research fields, the challenges of monitoring secondary forests and old growth forest degradation, and the potential future directions that could lead to a better understanding of such dynamic land uses in other tropical/global areas.

References

Aguiar, A.P.D. et al. 2016. Land use change emission scenarios: anticipating a forest transition process in the Brazilian Amazon. *Global change biology.* **22**(5), pp.1821-1840.

Alencar, A.A. et al. 2004. Modeling forest understory fires in an eastern Amazonian landscape. *Ecological Applications.* **14**(sp4), pp.139-149.

Ali, A.E. 1989. Assessment of Visual Interpretation of Thematic Mapper (Tm) Images for Land-Use and Cover in Sudan. *Earth Surface Processes and Landforms.* **14**(5), pp.399-405.

Almeida, C.a.A.d. 2009. ESTIMATIVA DA ÁREA E DO TEMPO DE PERMANÊNCIA DA VEGETAÇÃO SECUNDÁRIA NA AMAZÔNIA LEGAL POR MEIO DE IMAGENS LANDSAT/TM.

Almeida, C.A.d. et al. 2016. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. *Acta Amazonica.* **46**(3), pp.291-302.

Aragao, L.E. et al. 2014. Environmental change and the carbon balance of A mazonian forests. *Biological Reviews.* **89**(4), pp.913-931.

Asner, G.P. et al. 2006. Condition and fate of logged forests in the Brazilian Amazon. *Proceedings of the National Academy of Sciences.* **103**(34), pp.12947-12950.

Asner, G.P. et al. 2004. Canopy damage and recovery after selective logging in Amazonia: field and satellite studies. *Ecological Applications.* **14**(sp4), pp.280-298.

Asner, G.P. et al. 2005. Selective logging in the Brazilian Amazon. *science.* **310**(5747), pp.480-482.

Asner, G.P. et al. 2009. A contemporary assessment of change in humid tropical forests. *Conservation Biology.* **23**(6), pp.1386-1395.

Baccini, A. et al. 2017. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science.* **358**(6360), pp.230-234.

Baker, J. and Spracklen, D. 2019. Climate benefits of intact Amazon forests and the biophysical consequences of disturbance. *Frontiers in Forests and Global Change.* **2**, p47.

Bakr, N. et al. 2010. Monitoring land cover changes in a newly reclaimed area of Egypt using multi-temporal Landsat data. *Applied Geography.* **30**(4), pp.592-605.

Barlow, J. et al. 2016. Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. *Nature.* **535**(7610), p144.

Barlow, J. and Peres, C.A. 2004a. Avifaunal responses to single and recurrent wildfires in Amazonian forests. *Ecological Applications.* **14**(5), pp.1358-1373.

Barlow, J. and Peres, C.A. 2004b. Ecological responses to El Niño–induced surface fires in central Brazilian Amazonia: management implications for flammable tropical forests. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences.* **359**(1443), pp.367-380.

Bayoudh, M. et al. 2015. Structural knowledge learning from maps for supervised land cover/use classification: Application to the monitoring of land cover/use maps in French Guiana. *Computers & Geosciences.* **76**, pp.31-40.

Berenguer, E. et al. 2014. A large-scale field assessment of carbon stocks in human-modified tropical forests. *Global change biology.* **20**(12), pp.3713-3726.

Bonner, M.T. et al. 2013. A meta-analytical global comparison of aboveground biomass accumulation between tropical secondary forests and monoculture plantations. *Forest Ecology and Management.* **291**, pp.73-86.

Boyd, D.S. et al. 1996. An assessment of radiance in Landsat TM middle and thermal infrared wavebands for the detection of tropical forest regeneration. *International Journal of Remote Sensing.* **17**(2), pp.249-261.

Brown, S. and Lugo, A.E. 1990. Tropical secondary forests. *Journal of tropical ecology.* **6**(1), pp.1-32.

Carreiras, J.M. et al. 2014. Land use and land cover change dynamics across the Brazilian Amazon: insights from extensive time-series analysis of remote sensing data. *PloS one.* **9**(8), pe104144.

Carreiras, J.M. et al. 2017. Mapping major land cover types and retrieving the age of secondary forests in the Brazilian Amazon by combining single-date optical and radar remote sensing data. *Remote Sensing of Environment.* **194**, pp.16-32.

Carvalho, R. et al. 2019. Changes in secondary vegetation dynamics in a context of decreasing deforestation rates in Pará, Brazilian Amazon. *Applied Geography.* **106**, pp.40-49.

CASTELO, T.B. 2015. BRAZILIAN FORESTRY LEGISLATION AND TO COMBAT DEFORESTATION GOVERNMENT POLICIES IN THE AMAZON (BRAZILIAN AMAZON). *Ambiente & Sociedade.* **18**(4), pp.221-242.

Chazdon, R.L. 2003. Tropical forest recovery: legacies of human impact and natural disturbances. *Perspectives in Plant Ecology, evolution and systematics.* **6**(1), pp.51-71.

Chokkalingam, U. and de Jong, W. 2001. Secondary forest: a working definition and typology. *The International Forestry Review.* pp.19-26.

COCHRANE, M.A. 1998. Linear mixture model classification of burned forests in the eastern Amazon. *International Journal of Remote Sensing.* **19**(17), pp.3433-3440.

Cochrane, M.A. et al. 1999. Positive feedbacks in the fire dynamic of closed canopy tropical forests. *Science.* **284**(5421), pp.1832-1835.

Costa, O.B.d. et al. 2019. Selective Logging Detection in the Brazilian Amazon. *Floresta e Ambiente.* **26**(2).

da Silva, R.D. et al. 2014. Spectral/textural attributes from ALI/EO-1 for mapping primary and secondary tropical forests and studying the relationships with biophysical parameters. *GIScience & Remote Sensing.* **51**(6), pp.677-694.

Daldegan, G.A. et al. 2019. Spectral mixture analysis in Google Earth Engine to model and delineate fire scars over a large extent and a long time-series in a rainforest-savanna transition zone. *Remote Sensing of Environment.* **232**, p111340.

de Cambio Climático, P.I. 2003. Definitions and Methodological Options to Inventory Emissions from Direct Human-Induced Degradation of Forests and Devegetation of Other Vegetation Types [Penman, J. y colaboradores (directores de la publicación)]. *The Institutefor Global Environmental Strategies (IGES), Japón.*

de Wasseige, C. and Defourny, P. 2004. Remote sensing of selective logging impact for tropical forest management. *Forest Ecology and Management.* **188**(1-3), pp.161-173.

Dent, D.H. and Wright, S.J. 2009. The future of tropical species in secondary forests: a quantitative review. *Biological conservation.* **142**(12), pp.2833-2843.

Diniz, C.G. et al. 2015. DETER-B: The new Amazon near real-time deforestation detection system. *Ieee journal of selected topics in applied earth observations and remote sensing.* **8**(7), pp.3619-3628.

Edwards, D.P. et al. 2017. Tropical secondary forest regeneration conserves high levels of avian phylogenetic diversity. *Biological Conservation.* **209**, pp.432-439.

Espírito-Santo, F.D. et al. 2014. Size and frequency of natural forest disturbances and the Amazon forest carbon balance. *Nature communications.* **5**, p3434.

Espírito-Santo, F.D. et al. 2010. Storm intensity and old-growth forest disturbances in the Amazon region. *Geophysical Research Letters.* **37**(11).

FAO. 2015. Global forest resources assessment 2015.

Foley, J.A. et al. 2007. Amazonia revealed: forest degradation and loss of ecosystem goods and services in the Amazon Basin. *Frontiers in Ecology and the Environment.* **5**(1), pp.25-32.

Foley, J.A. et al. 2005. Global consequences of land use. *Science.* **309**(5734), pp.570-574.

Fu, A.M. et al. 2010. Forest Cover Classification With MODIS Images in Northeastern Asia. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing.* **3**(2), pp.178-189.

Gallardo-Cruz, J.A. et al. 2012. Predicting tropical dry forest successional attributes from space: is the key hidden in image texture? *PLoS One.* **7**(2), pe30506.

Gibbs, H.K. et al. 2016. Did ranchers and slaughterhouses respond to zero-deforestation agreements in the Brazilian Amazon? *Conservation Letters.* **9**(1), pp.32-42.

Gibbs, H.K. et al. 2015. Brazil's soy moratorium. *Science.* **347**(6220), pp.377-378.

Gloor, M. et al. 2012. The carbon balance of South America: a review of the status, decadal trends and main determinants. *Biogeosciences.* **9**(12), pp.5407-5430.

Grace, J. 2016. The Amazon carbon balance: an evaluation of methods and results. *Interactions Between Biosphere, Atmosphere and Human Land Use in the Amazon Basin.* Springer, pp.79-100.

Guariguata, M.R. and Ostertag, R. 2001. Neotropical secondary forest succession: changes in structural and functional characteristics. *Forest ecology and management.* **148**(1), pp.185-206.

Hansen, M.C. et al. 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science.* **342**(6160), pp.850-853.

Hermosilla, T. et al. 2015. Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. *Remote Sensing of Environment.* **170**, pp.121-132.

Hirschmugl, M. et al. 2017. Methods for Mapping Forest Disturbance and Degradation from Optical Earth Observation Data: a Review. *Current Forestry Reports.* **3**(1), pp.32-45.

Houghton, R. and Nassikas, A.A. 2017. Global and regional fluxes of carbon from land use and land cover change 1850–2015. *Global Biogeochemical Cycles.* **31**(3), pp.456-472.

Houghton, R. et al. 2000. Annual fluxes of carbon from deforestation and regrowth in the Brazilian Amazon. *Nature.* **403**(6767), pp.301-304.

Houghton, R.A. et al. 2012. Carbon emissions from land use and land-cover change. *Biogeosciences.* **9**(12), pp.5125-5142.

Huang, C. et al. 2010. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment.* **114**(1), pp.183-198.

ITTO. 2002. Guidelines for the Restoration, Management and Rehabilitation of Degraded and Secondary Tropical Forests. *ITTO Policy Development Series No.13.* pp.ITTO, Yokohama, Japan.

Kalamandeen, M. et al. 2018. Pervasive rise of small-scale deforestation in Amazonia. *Scientific reports.* **8**(1), p1600.

Kayastha, N. et al. 2012. Monitoring wetland change using inter-annual landsat time-series data. *Wetlands.* **32**(6), pp.1149-1162.

Kennedy, R.E. et al. 2007. Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sensing of Environment.* **110**(3), pp.370-386.

Kennedy, R.E. et al. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. *Remote Sensing of Environment.* **114**(12), pp.2897-2910.

Kimes, D.S. et al. 1999. Mapping secondary tropical forest and forest age from SPOT HRV data. *International Journal of Remote Sensing.* **20**(18), pp.3625-3640.

Kuplich, T.M. 2006. Classifying regenerating forest stages in Amazonia using remotely sensed images and a neural network. *Forest Ecology and Management.* **234**(1-3), pp.1-9.

Laurance, W.F. et al. 1998. Rain forest fragmentation and the dynamics of Amazonian tree communities. *Ecology.* **79**(6), pp.2032-2040.

Lennox, G.D. et al. 2018. Second rate or a second chance? Assessing biomass and biodiversity recovery in regenerating Amazonian forests. *Global change biology.* **24**(12), pp.5680-5694.

Liu, X. 2005. *Supervised Classification and Unsupervised Classification*. Last.

Long, W. and Sriharan, S. 2004. Land cover classification of SSC image: Unsupervised and supervised classification using ERDAS imagine. *Igarss 2004: Ieee International Geoscience and Remote Sensing Symposium Proceedings, Vols 1-7.* pp.2707-2712.

Longepe, N. et al. 2011. Assessment of ALOS PALSAR 50 m Orthorectified FBD Data for Regional Land Cover Classification by Support Vector Machines. *Ieee Transactions on Geoscience and Remote Sensing.* **49**(6), pp.2135-2150.

Longo, M. et al. 2016. Aboveground biomass variability across intact and degraded forests in the Brazilian Amazon. *Global Biogeochemical Cycles.* **30**(11), pp.1639-1660.

Lu, D. 2005. Integration of vegetation inventory data and Landsat TM image for vegetation classification in the western Brazilian Amazon. *Forest Ecology and Management.* **213**(1-3), pp.369-383.

Lu, D. et al. 2014a. A comparative analysis of approaches for successional vegetation classification in the Brazilian Amazon. *GIScience & Remote Sensing.* **51**(6), pp.695-709.

Lu, D.S. et al. 2011. A Comparison of Multisensor Integration Methods for Land Cover Classification in the Brazilian Amazon. *Giscience & Remote Sensing.* **48**(3), pp.345-370.

Lu, D.S. et al. 2014b. A comparative analysis of approaches for successional vegetation classification in the Brazilian Amazon. *Giscience & Remote Sensing.* **51**(6), pp.695-709.

Maia, H. et al. 2011. Avaliação do Plano de Ação para Prevenção e Controle do Desmatamento na Amazônia Legal: PPCDAm: 2007-2010.

Mallinis, G. et al. 2008. Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. *ISPRS Journal of Photogrammetry and Remote Sensing.* **63**(2), pp.237-250.

Marcal, A.R.S. et al. 2005. Land cover update by supervised classification of segmented ASTER images. *International Journal of Remote Sensing.* **26**(7), pp.1347-1362.

Marra, D.M. et al. 2014. Large-scale wind disturbances promote tree diversity in a Central Amazon forest. *PLoS One.* **9**(8), pe103711.

Martin, P.A. et al. 2013. Carbon pools recover more quickly than plant biodiversity in tropical secondary forests. *Proc. R. Soc. B.* **280**(1773), p20132236.

Massoca, P.E.D.S. et al. 2017. 4.6 Lessons from the soy and beef moratoria in Brazil. *Zero deforestation: A commitment to change.* p151.

Mayaux, P. et al. 2005. Tropical forest cover change in the 1990s and options for future monitoring. *Philosophical Transactions of the Royal Society of London B: Biological Sciences.* **360**(1454), pp.373-384.

Millan, V.E.G. et al. 2015. Mapping Tropical Dry Forest Succession With CHRIS/PROBA Hyperspectral Images Using Nonparametric Decision Trees. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing.* **8**(6), pp.3081-3094.

Mohammady, M. et al. 2015. A comparison of supervised, unsupervised and synthetic land use classification methods in the north of Iran. *International Journal of Environmental Science and Technology.* **12**(5), pp.1515-1526.

Monteiro, A. et al. 2003. Detection of logging in Amazonian transition forests using spectral mixture models. *International Journal of Remote Sensing.* **24**(1), pp.151-159.

Morton, D.C. et al. 2006. Cropland expansion changes deforestation dynamics in the southern Brazilian Amazon. *Proceedings of the National Academy of Sciences.* **103**(39), pp.14637-14641.

Neeff, T. and dos Santos, J.R. 2005. A growth model for secondary forest in Central Amazonia. *Forest Ecology and Management.* **216**(1), pp.270-282.

Neeff, T. et al. 2006. Area and age of secondary forests in Brazilian Amazonia 1978-2002: An empirical estimate. *Ecosystems.* **9**(4), pp.609-623.

Negrón-Juárez, R. et al. 2017. Windthrow variability in central Amazonia. *Atmosphere.* **8**(2), p28.

Negrón-Juárez, R.I. et al. 2010. Widespread Amazon forest tree mortality from a single cross-basin squall line event. *Geophysical Research Letters.* **37**(16).

Nelson, R.F. et al. 2000. Secondary forest age and tropical forest biomass estimation using thematic mapper imagery. *Bioscience.* **50**(5), pp.419-431.

Nepstad, D.C. et al. 1999. Large-scale impoverishment of Amazonian forests by logging and fire. *Nature.* **398**(6727), p505.

Nobre, C.A. et al. 2016. Land-use and climate change risks in the Amazon and the need of a novel sustainable development paradigm. *Proceedings of the National Academy of Sciences.* **113**(39), pp.10759-10768.

Otukei, J.R. and Blaschke, T. 2010. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation.* **12**, pp.S27-S31.

Pan, Y. et al. 2011. A large and persistent carbon sink in the world's forests. *Science.* **333**(6045), pp.988-93.

Poorter, L. et al. 2016. Biomass resilience of Neotropical secondary forests. *Nature.* **530**(7589), p211.

Prasad, R. et al. 2002. Visual interpretation of FCC image for land use & land cover mapping: An expert system approach. *Sice 2002: Proceedings of the 41st Sice Annual Conference, Vols 1-5.* pp.2093-2098.

PRODES, p. 2018. Projecto Prodes: Monitoramento de Floresta Amazonica Brasileira por satelite. http://www.obt.inpe.br/prodes/index.php. *Instituto Nacional de Pesquisas Espaciais*

RAISG. 2012. Amazonia under Pressure. p10.

Rudorff, B.F.T. et al. 2011. The soy moratorium in the Amazon biome monitored by remote sensing images. *Remote Sensing.* **3**(1), pp.185-202.

Saatchi, S.S. et al. 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences.* **108**(24), pp.9899-9904.

Sader, S.A. et al. 1989. Tropical forest biomass and successional age class relationships to a vegetation index derived from Landsat TM data. *Remote Sensing of Environment.* **28**, pp.143IN1159-156IN2198.

Samaniego, L. and Schulz, K. 2009. Supervised Classification of Agricultural Land Cover Using a Modified k-NN Technique (MNN) and Landsat Remote Sensing Imagery. *Remote Sensing.* **1**(4), pp.875-895.

Sayer, C. et al. 2017. Dynamics of avian species and functional diversity in secondary tropical forests. *Biological Conservation.* **211**, pp.1-9.

Shimabukuro, Y.E. et al. 2019. Monitoring deforestation and forest degradation using multi-temporal fraction images derived from Landsat sensor data in the Brazilian Amazon. *International Journal of Remote Sensing.* **40**(14), pp.5475-5496.

Shimabukuro, Y.E. et al. 2009. Fraction images derived from Terra Modis data for mapping burnt areas in Brazilian Amazonia. *International Journal of Remote Sensing.* **30**(6), pp.1537-1546.

Shiraishi, T. et al. 2014. Comparative Assessment of Supervised Classifiers for Land Use-Land Cover Classification in a Tropical Region Using Time-Series PALSAR Mosaic Data. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing.* **7**(4), pp.1186-1199.

Simula, M. 2009. Towards defining forest degradation: comparative analysis of existing definitions. *Forest Resources Assessment Working Paper.* **154**.

Song, X.-P. et al. 2015. Annual carbon emissions from deforestation in the Amazon Basin between 2000 and 2010. *PloS one.* **10**(5), pe0126754.

Sonmez, N.K. and Onur, I. 2012. Monitoring of land use and land cover changes by using fuzzy supervised classification method: A case study of Antalya, Turkey. *Journal of Food Agriculture & Environment.* **10**(3-4), pp.963-967.

Souza, C. and Roberts, D. 2005. Mapping forest degradation in the Amazon region with Ikonos images. *International journal of remote sensing.* **26**(3), pp.425-429.

Souza Jr, C. and Barreto, P. 2000. An alternative approach for detecting and monitoring selectively logged forests in the Amazon. *International Journal of Remote Sensing.* **21**(1), pp.173-179.

Souza Jr, C. et al. 2003. Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models. *Remote Sensing of Environment.* **87**(4), pp.494-506.

Souza Jr, C. et al. 2013. Ten-year Landsat classification of deforestation and forest degradation in the Brazilian Amazon. *Remote Sensing.* **5**(11), pp.5493-5513.

Souza Jr, C.M. et al. 2005a. Combining spectral and spatial information to map canopy damage from selective logging and forest fires. *Remote Sensing of Environment.* **98**(2-3), pp.329-343.

Souza Jr, C.M. et al. 2005b. Multitemporal analysis of degraded forests in the southern Brazilian Amazon. *Earth Interactions.* **9**(19), pp.1-25.

Spracklen, D. and Garcia-Carreras, L. 2015. The impact of Amazonian deforestation on Amazon basin rainfall. *Geophysical Research Letters.* **42**(21), pp.9546-9552.

Struebig, M.J. et al. 2013. Quantifying the biodiversity value of repeatedly logged rainforests: gradient and comparative approaches from Borneo. *Advances in ecological research.* Elsevier, pp.183-224.

Stuart, N. et al. 2006. Visual interpretation of synthetic aperture radar data for assessing land cover in tropical savannas - art. no. 64190L. *Geoinformatics 2006: Remotely Sensed Data and Information.* **6419**, pp.L4190-L4190.

Thompson, I.D. et al. 2013. An operational framework for defining and monitoring forest degradation. *Ecology & Society.* **18**(2).

Tyukavina, A. et al. 2017. Types and rates of forest disturbance in Brazilian Legal Amazon, 2000–2013. *Science advances.* **3**(4), pe1601047.

Valeriano, D.M. et al. 2004. Monitoring tropical forest from space: the PRODES digital project. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences.* **35**, pp.272-274.

Van den Broek, A.C. et al. 2004. Land use classification of polarimetric SAR data by visual interpretation and comparison with an automatic procedure. *International Journal of Remote Sensing.* **25**(18), pp.3573-3591.

Van der Werf, G.R. et al. 2009. CO 2 emissions from forest loss. *Nature geoscience.* **2**(11), p737.

Vieira, I.C.G. et al. 2003. Classifying successional forests using Landsat spectral properties and ecological characteristics in eastern Amazonia. *Remote Sensing of Environment.* **87**(4), pp.470-481.

White, J.C. et al. 2017. A nationwide annual characterization of 25 years of forest disturbance and recovery for Canada using Landsat time series. *Remote Sensing of Environment.* **194**, pp.303-321.

Wright, S.J. 2010. The future of tropical forests. *Annals of the New York Academy of Sciences.* **1195**(1), pp.1-27.

Xaud, H.A.M. et al. 2013. Tropical forest degradation by mega-fires in the northern Brazilian Amazon. *Forest Ecology and Management.* **294**, pp.97-106.

Xia, L.G. et al. 2014. An Automated Approach for Land Cover Classification Based on a Fuzzy Supervised Learning Framework. *Journal of the Indian Society of Remote Sensing.* **42**(3), pp.505-515.

Zarin, D.J. et al. 2005. Legacy of fire slows carbon accumulation in Amazonian forest regrowth. *Frontiers in Ecology and the Environment.* **3**(7), pp.365-369.

Zhao, B. et al. 2009. Monitoring rapid vegetation succession in estuarine wetland using time series MODIS-based indicators: An application in the Yangtze River Delta area. *Ecological Indicators.* **9**(2), pp.346-356.

Chapter 2
Upturn in secondary forest clearing buffering primary forest loss in the Brazilian Amazon (Paper Ⅰ)

Wang, Y., Ziv, G., Adami, M., Almeida, C.A., Antunes, J.F.G., Coutinho, A.C., Esquerdo, G.C.D.M., Gomes, A.R., and Galbraith, D., 2019. Upturn in secondary forest clearing buffering primary forest loss in the Brazilian Amazon. *Nature Sustainability* (Accepted).

Abstract

Primary forest deforestation in the Brazilian Amazon has declined considerably since 2004, but secondary forest loss has never been quantified. We use a recently-developed high-resolution land use dataset (TERRACLASS) to track secondary forests in the Brazilian Amazon over 14 years, providing the first estimates of secondary forest loss for the region. Secondary forest loss increased by (187 ± 48) % from 2008 to 2014. Moreover, the proportion of total forest loss accounted for by secondary forests rose from (37 ± 3) % in 2000 to (72 ± 5) % in 2014. Recent acceleration in secondary forests loss occurred across the entire region and was not driven simply by increasing secondary forest area but likely a conscious preferential shift towards clearance of a little-protected forest resource. Our results suggest that secondary forests have eased deforestation pressure on primary forests. However, this has been at the expense of a lost carbon sequestration opportunity of 2.59-2.66 Pg C over our study period.

2.1 Introduction

The Amazon rainforest provides significant ecosystem services locally, regionally and globally. The biome's forests are home to one-quarter of global biodiversity (Dirzo and Raven, 2003; Malhi et al., 2009), store in excess of 100 billion tonnes of carbon in their biomass (Baccini et al., 2012; Avitabile et al., 2016) and play a crucial role in the provision of rainfall in South America (Spracklen et al., 2012). Deforestation control is essential for maintaining the functional integrity of Amazon rainforests. In the Brazilian Amazon, which accounts for over two-thirds of Amazonian forests (RAISG, 2012), deforestation of primary forests fell by 82% from peak rates in 2004 to 2014 (PRODES, 2018). This substantial decline reflects the efficacy of Brazil's PPCDAm Program (Maia et al., 2011) (The Action Plan for the Prevention and Control of Deforestation in the Legal Amazon), which was launched in 2004 to reduce deforestation rates and support sustainable development in Amazonia. This program resulted in the implementation of new policies, enhanced detection frameworks (Assunção et al., 2013) and control measures to curtail deforestation in the Brazilian Amazon, and international mechanisms such as the soybean (Gibbs et al., 2015; Rudorff et al., 2011) and beef moratoria (Gibbs et al., 2016; Massoca et al., 2017). However, these mechanisms do not protect secondary forests, defined here as re-growing forests on previously deforested land.

Currently, secondary forests comprise approximately 21% of previously deforested areas in the Brazilian Amazon (Almeida et al., 2016). They can accumulate carbon very rapidly (Poorter et al., 2016), thereby providing a key pathway for Brazil to reduce net carbon emissions and mitigate climate change (Chazdon et al., 2016). At the same time, secondary forests are an important component of land management systems in the Brazilian Amazon, as their regrowth restores soil functioning, ensuring productivity of pastures and small-scale agriculture (Kato et al., 2004). Despite the importance of secondary forests for conservation planning, environmental policy and land management in Amazonia, a historical lack of spatio-temporal data on secondary forest area has precluded evaluation of their large-scale dynamics. Although a recent localised

study (Carvalho et al., 2019) for the state of Pará illustrates the dynamic nature of secondary forests, a comprehensive analysis of secondary forest loss in Amazonia does not exist.

Here we use a recently-developed 30 m land cover dataset for the Brazilian Amazon (TERRACLASS) (2018; Almeida et al., 2016), which provides unprecedented information on secondary forest occurrence over a 14-year period (2000-2014), to undertake the first large-scale assessment of the spatio-temporal dynamics of secondary forests in Amazonia. TERRACLASS takes the deforested areas from PRODES(PRODES, 2018) as an input layer and classifies each deforested patch into one of twelve different land covers (Table 2.1), including secondary forest. From TERRACLASS, we computed the areas of secondary and primary forest cleared annually, generated secondary forest loss by age structure and evaluated the fate (land cover type) of cleared secondary forests. To account for classification error in the TERRACLASS base map, we use a sampling-based approach combined with expert validation, following best practice in the field (Olofsson et al., 2014; Tyukavina et al., 2017). The summary forest loss estimates presented in the main text of this manuscript refer to sampling-based and not map-based estimates. A comparison of sampling-based and map-based estimates is provided in the supplementary information (Appendix Table A.7).

2.2 Results

2.2.1 Spatio-temporal distribution of secondary forest loss

Our results reveal two distinct phases of secondary forest loss in Amazonia. Between 2000-2008, we find a marked decline in secondary forest loss, mirroring the declines in primary forest loss seen over this period. However, we find that secondary forest loss between 2008-2014 increased sharply from approximately 6,040 ± 1,417 km^2 yr^{-1} to 10,757 ± 1,486 km^2 yr^{-1}, despite an apparent levelling off of primary forest loss over this period (Figure 2.1). This second period, therefore, was marked by an increase pressure on forest ecosystems, which was largely absorbed by intensified secondary forest loss. These large increases in

secondary forest loss translate into considerable overall increases (123 ± 21 %) in total (primary and secondary) forest loss between 2008-2010 and 2012-2014, reversing the downward trend in total forest loss up to 2008 (Figure 2.1). Over our study period, the proportion of total forest loss due to secondary forest clearance increased from 37 ± 3 % in 2000-2004 to 72 ± 5 % in 2012-2014 (Figure 2.1). Map-based estimates of forest loss were very consistent with those derived from our sampling-based analysis and exhibited the same temporal pattern (Appendix Figure A.2).

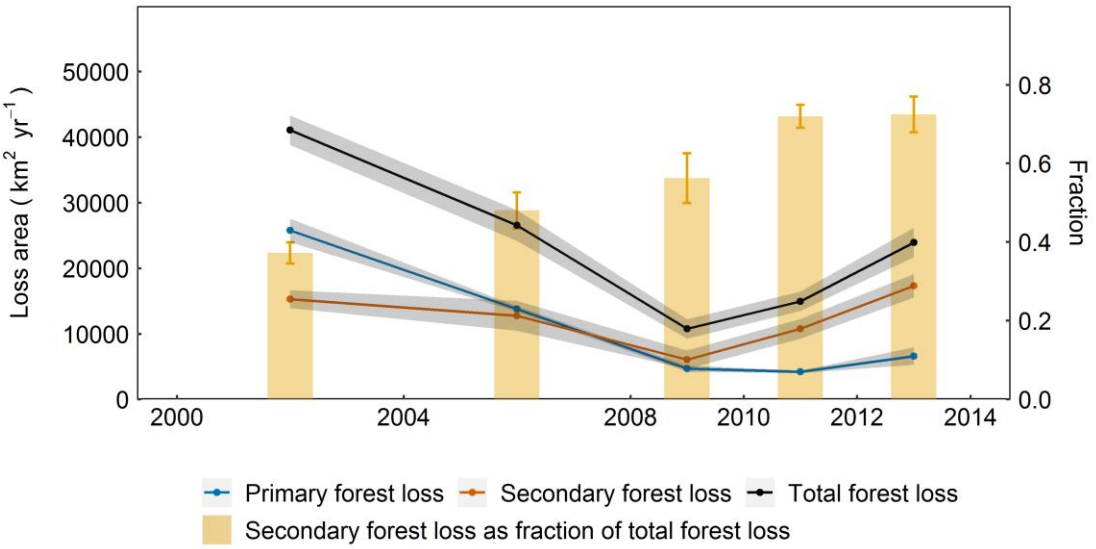


Figure 2.1 Sample-based estimates of annual primary and secondary forest loss in the Brazilian Amazon from 2000-2014.

Total forest loss is the sum of primary and secondary forest loss. The uncertainties (grey shaded areas) denote standard errors from our sample-based validation (all intervals) as well as time-interval corrections which account for missed secondary forest loss in 4-year intervals (2000-2004 and 2004-2008 only). See Appendix Table A.7 for numerical values and comparison to map-based estimates.

The preferential cutting of secondary forests was found to be geographically widespread. In 2000-2004, secondary forest loss mainly outstripped primary forest loss in the far northeast of the Brazilian Amazon (Figure 2.2) which has historically been subject to high primary forest deforestation, with little remaining primary forest (Appendix Figure A.4). By 2012-2014, however, secondary forest loss exceeded primary forest loss across almost all of the Brazilian Amazon (Figure 2.2).

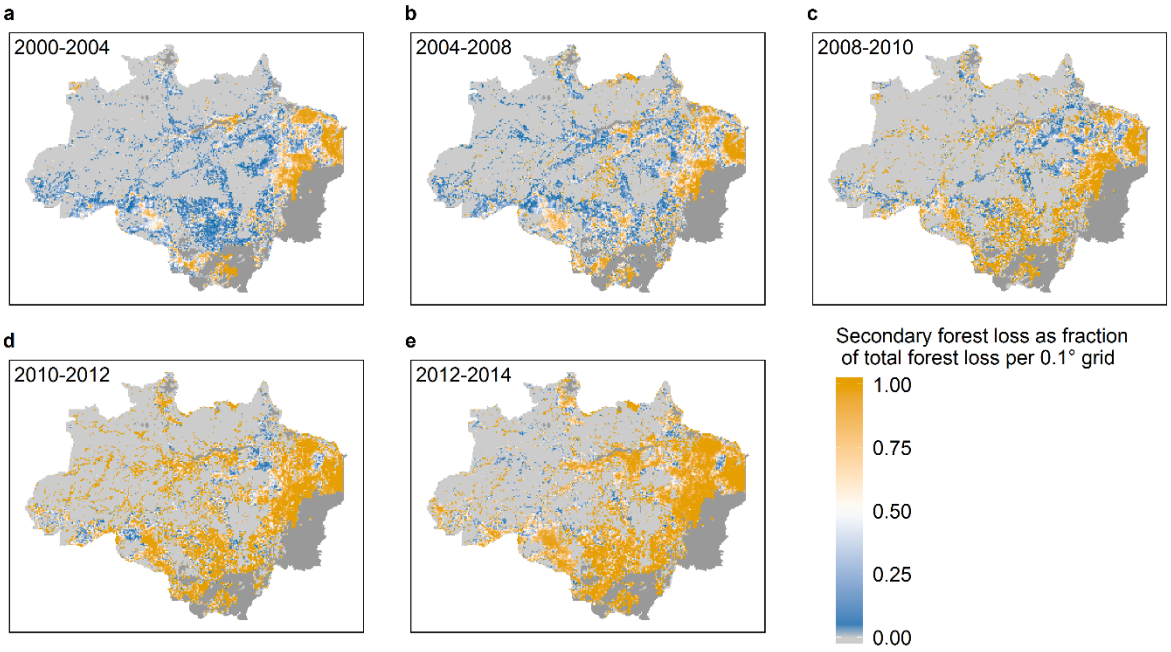


Figure 2.2 Spatio-temporal variation of secondary forest loss as fraction of total forest loss in the Brazilian Amazon.
Darker blue (warmer orange) colours indicate areas where majority of forest loss occurred in primary (secondary) forests. The lighter grey colours represent areas with no recorded forest loss. Darker grey colours represent non-forest areas (e.g. savannas). Time interval corrections were applied in the first two intervals (i.e. 2000-2004, 2004-2008). See Appendix Figure A.3 for the spatial distribution of the absolute area of secondary forest loss. Analysis of spatial patterns was undertaken directly on the TERRACLASS wall-to-wall maps.

We further examined the age structure of secondary forest loss. Within any given interval, we find that the percentage loss rate of secondary forests declines progressively with increasing secondary forest age (Appendix Table A.8). In the 2012-2014 interval, for example, the percentage loss rate of the youngest secondary forest age category (0-2 years) was over five times greater than that

of the oldest age category (>12 years old). Between 2008-2014, increases in secondary forest loss were observed across all age categories (Figure 2.3) but were particularly marked for young (0-4 years) secondary forests (Figure 2.3a). Over this time period, the annual percentage loss rates of young secondary forests increased by 250% from 6% in 2008-2010 to 21% in 2012-2014 (Figure 2.3a, mean), compared to increases of 192% and 106% for intermediate (4-8 years) and old (>8 years) secondary forests respectively (Figure 2.3 b-c).

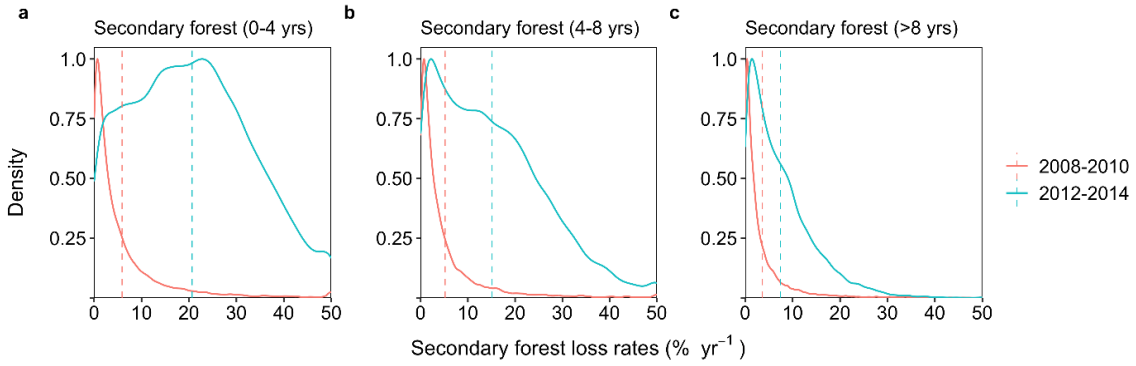


Figure 2.3 Distribution of percentage loss rate of secondary forests by age group (0-4 years, 4-8 years and over 8 years).
Annual percentage loss rates of secondary forests were computed for individual 0.1° grid cells, based on TERRACLASS maps. Grid cells without secondary forest loss were excluded. Panel a, 10539 valid grid cells, 87% of which showed an increase in secondary forest loss rates; Panel b, 10915 valid grid cells, 81% of which showed an increase in secondary forest loss rates; Panel c, 11248 valid grid cells, 76% of which showed an increase in secondary forest loss rates. Solid lines depict density distributions of secondary forest loss rates across all valid grids. Dashed vertical lines denote mean values. Density, computed through R-'stat_density', was the kernel density estimate which is a non-parametric way to estimate the probability density function of a random variable. It's a smoothed version of the histogram.

2.2.2 Fate of secondary forest loss

The vast majority (91%) of cleared secondary forests (almost identical for young, intermediate and old secondary forests) in the Brazilian Amazon over our study period became pastureland (Appendix Table A.9 Table A.10), mirroring the fate of deforested primary forests (Tyukavina et al., 2017). Pasture expansion from primary forest deforestation in Amazonia slowed considerably following the

establishment of the 2008 beef moratorium (Massoca et al., 2017), in which retailers pledged to stop purchasing meat produced on illegally deforested land. Since these measures were introduced, secondary forests have absorbed much more of the pasture expansion in the Brazilian Amazon, with conversion of secondary forest to pastureland increasing by 282% between 2008-2010 and 2012-2014 (Appendix Table A.9). Conversely, about 90% of new secondary forests observed in TERRACLASS between 2008 and 2014 were previously identified as pasture (Appendix Table A.9). Although conversion of secondary forest to agricultural land increased by 106% between 2008-2010 and 2012-2014, the absolute area of secondary forest converted to agricultural land in 2012-2014 was >40 times lower than that converted to pastureland and only accounted for approximately 2% of the total cleared secondary forest area (Appendix Table A.9).

Overall, our results point to an acceleration of the pasture-forest-pasture management system since the introduction of the beef moratorium. Post-deforestation landscapes in the Brazilian Amazon are highly dynamic in nature. In these landscapes, secondary forests are often cut and usually burned, as part of the pasture cycle. Their regrowth on pasturelands improves soil integrity by replenishing nutrients, enhancing organic matter storage and improving the physical structure of soils, which can become heavily degraded following sustained pasture activity (Cordeiro et al., 2017). Our results suggest that the permanence time of secondary forests in these cycles has decreased substantially over time, as cutting rates have accelerated greatly but with no underlying trend over time in the fate of secondary forests. Whereas only 2.86 ± 0.67 % of total secondary forest area was cut annually between 2008 and 2010, this increased to 7.43 ± 0.81 % in 2012-2014 (Appendix Table A.4Table A.6).

2.2.3 Area of secondary forests

The upturn in overall forest loss, including both primary and secondary forests, since 2008 indicates an enhanced demand for new pasture and agricultural lands. This enhanced demand has increasingly been met by secondary forests,

thus providing a buffer that has stalled deforestation of primary forests. Ultimately, however, the strength of this buffer depends on the area of secondary forest available. Between 2000-2010, the sampling-derived area of secondary forest increased by 34,183 ± 12,392 km^2 (an overall change of 0.87 ± 0.29% in agreement with Aguiar et al. (2016)), but did not change significantly over the last two intervals (Appendix Table A.11). Moreover, the area of stable secondary forest (secondary forests which persisted over an entire TERRACLASS interval) increased progressively over time up to the last interval, when it declined for the first time (Appendix Table A.2-A.6). Future depletions in secondary forest area would likely lead to increasing pressure on primary forests as the available pool of easily accessible secondary forests for cutting is diminished.

2.3 Discussion and Conclusion

While primary forests have benefited from strong legal protection in the Brazilian Amazon, secondary forests have little protection status in Brazilian law. This partially stems from the lack of clear definitions for secondary forests themselves - e.g. the point in the recovery process where they effectively become 'forests'. Pará is currently the only Brazilian state to adopt an explicit definition of secondary forests, where secondary forests are defined as those that have regenerated from previously cleared land and that can no longer be considered as fallow(Vieira et al., 2014). The right to cut secondary forests in Pará is directly related to forest age, as state law(Pará State Government. INSTRUÇÃO NORMATIVA SEMA Nº 08, DE 28-10-2015, 2015) dictates that areas younger than five years can be cleared irrespective of their physical structure, whilst areas older than 20 years must be conserved. Clearance of forests in intermediate stages of succession (5-20 years) follows basal area thresholds which vary according to background forest cover status. While such legislation is beneficial for ensuring the recovery of older forests, it encourages the cutting of secondary forests before they reach the age or basal area thresholds that would render their cutting illegal. In other Brazilian Amazon states, legislation governing the cutting of secondary forests has yet to be developed. This limited legal protection means that secondary forest loss is largely unregulated.

To formally test whether the increase in secondary forest loss over time can be explained purely by increasing availability of secondary forests relative to primary forests, we compared the observed secondary forest cutting to a null model which assumes a time-invariant preference for secondary forest clearance relative to primary forest clearance. We find that across the Brazilian Amazon, this null model predicted secondary forests losses well up to 2008-2010. In the last two intervals, however, the null model greatly underestimated secondary forest loss and its relative contribution to total forest loss (Figure 2.4). This recent rise in secondary forest clearance may reflect a conscious behavioural shift towards preferential cutting of secondary forests over primary forests - i.e. the increase in secondary forest loss in our statistical model would only be captured if the preference (bias) for cutting secondary forest relative to primary forest was allowed to increase over time.

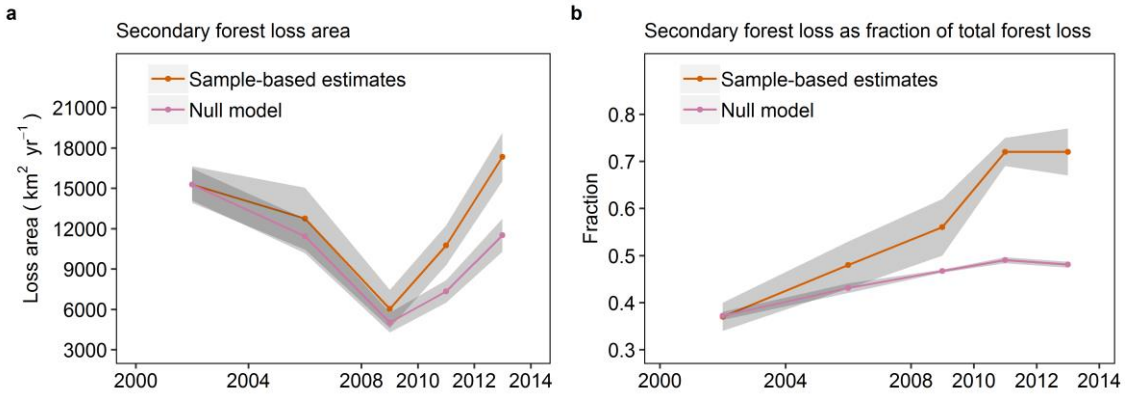


Figure 2.4 Comparison of secondary forest loss between actual estimates from TERRACLASS and null model predictions.

The null model predicts secondary forest loss by sampling without replacement based on Fisher's non-central hypergeometric distribution, given known available areas of primary and secondary forests in each interval and assuming a bias (odds ratio, estimated to be 13.69) for cutting secondary forests relative to primary forest computed for the first interval (2000-2004) and subsequently maintained across all intervals. Points on the null model curves are based on mean values from Fisher's non-central hypergeometric distribution. See Appendix Table A.11 for numerical values.

The large losses of secondary forests observed in this study have significant implications. On the one hand, their accelerated cutting has been important for

curbing losses of primary forests whose biodiversity value is irreplaceable (Gibson et al., 2011). The enhanced preference for cutting secondary forests instead of primary forests also reinforces the effectiveness of measures in place to inhibit primary forest loss. On the other hand, secondary forests are themselves an important biodiversity reservoir in an increasingly fragmented landscape (Brockerhoff et al., 2008; Lennox et al., 2018), and if left to regrow, can act as substantial carbon sinks (Martin et al., 2013). Brazil has committed to restore 120,000 km^2 of forest land by 2030 as part of its Nationally Determined Contribution (NDC) for the Paris Agreement(UNFCCC, 2015). A cost-effective way to do this would be to allow part of its existing Amazonian secondary forest area to recover naturally. Over the 14-year period of our study, over 180,329 $\pm$ 11,760 km^2 of secondary forests were cut, exceeding its total NDC commitment by over 60,329 $\pm$ 11,760 km^2. Applying a simple biomass accumulation model (see Methods), we estimate that this loss of secondary forests prevented the potential accumulation of 2.59-2.66 billion tonnes of carbon. This represents approximately 18 years of Brazil's fossil fuel emissions, based on 2014 emissions (World Bank, 2014).

Despite the accelerated rate of secondary forest loss, the Brazilian Amazon still has in excess of 235,718 ± 7,773 km^2 of secondary forests. Managing this resource sustainably so as to maximise the conservation value of these forests, while not intensifying pressure on primary forests, requires an integrated strategy that includes active monitoring of secondary forests in Amazonia and strengthening of their governance.

2.4 Methods

2.4.1 TERRACLASS

We used the land use/land cover classification maps produced by the TERRACLASS Project(2018) (https://www.terraclass.gov.br) as the basis for all analyses of secondary forest dynamics. TERRACLASS maps post-deforestation land cover at 2 to 4 year intervals across the Brazilian Legal Amazon. We used all TERRACLASS maps available at the time of the study (2000, 2004, 2008, 2010, 2012 and 2014). TERRACLASS assigns areas designated as deforested by PRODES (primary forest deforestation monitoring program for the Brazilian Amazon) into one of twelve different land cover types (Table 2.1). In this study, we combined shrubby pasture and herbaceous pasture categories into a single pasture category and further combined perennial agriculture, semi-perennial agriculture and temporal agriculture into a single agriculture category. For areas not observed in a specific TERRACLASS year due to persistent cloud cover, we assume the same land use categories as for the preceding TERRACLASS map. Non-forest and hydrology categories were excluded from the study. TERRACLASS 2004-2014 products inherited historical PRODES misalignment issues(PRODES, 2017) which were subsequently corrected for TERRACLASS-2000. To ensure consistency across all TERRACLASS products, we aligned the TERRACLASS-2000 to other TERRACLASS products using an image displacement algorithm in Google Earth Engine (See Appendix Figure A.1 for the example image for diaplacement correction).

Table 2.1 TERRACLASS land use/cover classification categories.

Land use categories	Description	Notes
1. Primary forest	Areas that have never been deforested, as mapped by PRODES.	
2. Secondary forest	Areas that were clear-cut and are at an advanced stage of regeneration with trees and shrubs.	
3. Shrubby pasture	Areas of productive pasture with grass coverage between 50% and 80% and 20-50% coverage of shrubby vegetation	Combined as pasture
4. Herbaceous pasture	Productive pasture with 90-100% grass coverage.	
5. Perennial agriculture	e.g. palm oil	
6. Semi-perennial agriculture	e.g. sugarcane	Combined as agriculture
7. Temporal agriculture	e.g. soybean	
8. Mining	Areas of mineral extraction with the presence of bare soil and deforestation in the proximity of water bodies.	
9. Urban	Population concentrations forming small inhabited places, villages and cities.	
10. Others	Areas not encompassed by other categories such as rocky or mountain outcrops, river shores and sand banks.	
11. Deforestation	Areas recently deforested with no defined land use at this stage. Mapped by PRODES as deforested in that year.	
12. Reforestation	Human-cultivated forest (plantation).	
13. Non-observed area	Areas not possible to be interpreted due to clouds or cloud shadows at the moment of the satellite overpass or recently burned areas.	Assume same land use categories as previous year
14. Non-forest	e.g. Amazon savannas.	Excluded from this study
15. Hydrology	Water bodies.	

2.4.2 Estimates of forest loss

We computed forest loss estimates from TERRACLASS in two ways: 1) simple wall-to-wall calculations based directly on the TERRACLASS map and 2) a sampling-based approach in which classification accuracy and the map areas of different land cover categories are used to construct forest loss estimates with appropriate error quantification[19,31].

We estimated annual primary and secondary forest loss as well as secondary forest gain for five individual time intervals: 2000-2004, 2004-2008, 2008-2010, 2010-2012, 2012-2014. Primary forest loss was considered as the land use change from primary forest to any non-primary forest categories (i.e. pasture, agriculture, secondary forest, urban, mining, others, and reforestation). Secondary forest loss was regarded as the land use change from secondary forest to other non-forest categories. Secondary forest was defined as being represented only by the 'secondary forest' class from TERRACLASS. No post-hoc re-classification of any other land classes (e.g. shrubby pasture) as secondary forest was applied. Thus, all estimates of secondary forest area and loss reported in this study refer specifically to the 'secondary forest' category from TERRACLASS. Total forest loss was computed as the sum of primary forest loss and secondary forest loss. Secondary forest gain was defined as the regrowth of secondary forests following abandonment from other non-forest categories. Wall-to-wall primary/secondary forest loss rates were constructed by summing the pixel areas of all pixels that were defined as primary/secondary forest at the beginning of a TERRACLASS interval but not these classes at the end of the interval.

We used the map-based calculations to evaluate spatial patterns of secondary/primary forest loss. To do this, we applied a 0.1 degree grid over the Brazilian Amazon and computed the fraction of total forest loss accounted for by secondary forests within each grid cell.

2.4.3 Sampling-based estimates.

Our wall-to-wall calculations may be subject to biases related to TERRACLASS classification errors (Almeida et al., 2016). To account for this, we estimated forest loss by applying an unbiased estimator to a stratified sample of reference observations following best practice recommendations (Olofsson et al., 2014; Olofsson et al., 2013). For each TERRACLASS time interval (i.e. 2000-2004, 2004-2008, 2008-2010, 2010-2012, 2012-2014), we used stratified random sampling to generate an independent set of samples, for subsequent visual

assessment by three experts. Sampling was stratified according to six land cover change categories: 1) stable primary forest, 2) primary forest loss, 3) stable secondary forest, 4) secondary forest loss, 5) secondary forest gain, and 6) stable others (e.g. pasture, agriculture, mining). The stable primary forest stratum occupied >70% of the study area (Appendix Table A.1). Given the very large area of this stratum, stable forest samples interpreted as change categories in the reference classification will carry a disproportional area weight and may considerably reduce the accuracy of estimates of change categories (Arévalo et al., 2019). To account for this, we introduced a buffer stratum (1 km) for stable primary forest areas surrounding areas of primary forest loss and partitioned our stable forest sample to account for stable forests inside and outside of this buffer (Arévalo et al., 2019). We calculated the total sample size *n* following Olofsson et al. (2014), as follows:

$$n = \left(\frac{\sum(w_i S_i)}{S(\hat{O})}\right)^2 \qquad \text{(eq. 2.1)}$$

where w_i is the mapped proportion of area of stratum *i,* $S(\hat{O})$ is the standard error of the estimated overall accuracy that we would like to achieve (0.015), S_i is the standard deviation of stratum *i,* $S_i = \sqrt{U_i(1-U_i)}$ where U_i is the anticipated user's accuracy of stratum *i* (0.70 for all strata in this study). This yielded a total of 933 sampled pixels for each time interval with 50-100 pixels allocated to the smaller strata and the remaining pixels proportionally allocated to other strata based on their area weights (w_i)(Olofsson et al., 2014; Arévalo et al., 2019) (Appendix Table A.1). All pixels were sampled so that they were at least 200 m away from the edge of an individual stratum to avoid potential misalignments between TERRACLASS and the reference images(PRODES, 2017).

Reference classification for each sampled pixel was conducted through Collect Earth Online(Saah et al., 2019) by three experts through visual interpretation of annual Landsat composite images acquired during 1st July – 31st August and, when available, very high resolution imagery from Digital Globe and Google Earth. Information from time-series trajectories of Landsat spectral bands (red and short-wave infrared bands) and vegetation indices (NDVI-normalized

difference vegetation index, NDWI- normalized difference water index) were also utilized by the experts for the reference classification. Each sampled pixel was classified by the experts as stable forest, forest loss, forest gain or stable others, and flagged if no clear Landsat image was available. Experts did not distinguish between stable primary forest and stable secondary forest or between primary forest loss or secondary forest loss as TERRACLASS only classifies land use/cover on historically deforested areas, so that misclassifications between primary and secondary forests are not technically possible in TERRACLASS. Initially, each expert assessed each reference pixel independently. Pixels with disagreement between experts were subsequently revisited until agreement was reached. Flagged pixels (with no clear Landsat imagery between 1st July and 31st August) were re-interpreted using Landsat composite imagery for the entire year or excluded if no clear reference image was available for that year.

Area estimates of each individual reference class were based on the above reference data and sample classification protocol. Following Olofsson *et al.* (2014)(Olofsson et al., 2014), the estimated area of reference class *k* was computed as:

$$\hat{A}_k = A \times \hat{p}_{\cdot k} \qquad \text{(eq. 2.2)}$$

where *A* is the total area of the entire domain considered (3,924,375.63 km^2), and $\hat{p}_{\cdot k}$ is the proportion of area of class k as determined from the reference classification, which was computed as:

$$\hat{p}_{\cdot k} = \sum_{i=1}^{q} w_i \frac{n_{ik}}{n_i} \qquad \text{(eq. 2.3)}$$

where *q* represents the number of mapped strata (*i*), w_i is the proportion of area of each mapped stratum *i*, n_{ik} is the number of samples from mapped stratum *i* interpreted as reference class *k*, and n_i is the total number of samples for mapped stratum *i*.

The standard error for the proportion of area of reference class k was computed as(Olofsson et al., 2014):

$$S(\hat{p}_{\cdot k}) = \sqrt{\sum_i \frac{w_i \hat{p}_{ik} - \hat{p}_{ik}^2}{n_i - 1}} \qquad \text{(eq. 2.4)}$$

where $\hat{p}_{ik}$ is the proportion of area from mapped stratum *i* interpreted as reference class *k*, $\hat{p}_{ik} = w_i \frac{n_{ik}}{n_i}$ (refer to the above eq. (3)).

The standard error of the estimated areas was then computed as:

$$S(\hat{A}_k) = A \times S(\hat{p}_{\cdot k}) \qquad \text{(eq. 2.5)}$$

The summary forest loss estimates reported in the main text of this manuscript denote the sampling-based estimates $\hat{A}_k \pm S(\hat{A}_k)$.

2.4.4 Correcting for varying interval lengths

The time structure of TERRACLASS products (2000, 2004, 2008, 2010 , 2012, 2014), is such that the first two intervals used to compute forest loss span four years while the remaining intervals span two years. These differences in interval length do not affect calculation of primary forest loss but do have implications for secondary forest loss and gain due to the much more transient nature of secondary forests, which are often cleared within 2 years of regrowth. Thus, 4-year intervals can miss the clearance of secondary forests that have established and been cut again within the interval. To account for this, we derived a correction factor α, where secondary forest loss/gain estimates for 4-year intervals were rescaled as:

$$A_{corrected} = A_{uncorrected} \times \alpha \qquad \text{(eq. 2.6)}$$

where $A_{uncorrected}$ is the original, uncorrected loss/gain over 4-year TERRACLASS intervals (2000-2004, 2004-2008). We calculated α as follows, based on available 2-year TERRACLASS intervals (2008-2014), which we then regrouped into 4-year intervals (e.g. 2008-2012, 2010-2014):

$$\alpha = \left(B_{2yr(i)} + B_{2yr(ii)}\right)/B_{4yr} \qquad \text{(eq. 2.7)}$$

where B_{4yr} is the secondary forest loss/gain over the regrouped 4-year interval and $B_{2yr(i)}$ and $B_{2yr(ii)}$ are secondary forest loss/gain for 1st and 2nd 2-year intervals respectively. We found that on average, 4-year intervals underestimated secondary forest loss by 16.84-26.52% and underestimated secondary forest gain by 10.31-24.61% relative to 2-year intervals.

We applied the above underestimates of secondary forest loss/gain to provide revised best estimates (based on mean underestimates) of secondary forest loss/gain for 4-year intervals and used the full range of underestimates (minimum and maximum values) to provide uncertainty bounds on our re-scaled values.

The interval length corrections were applied to both our map-based and sampling-based estimates for the 4-year intervals (i.e. 2000-2004, 2004-2008). For sample-based estimates, the total errors for the loss rates were computed by adding the sampling-derived errors in quadrature with the interval correction errors (only relevant for 4-year intervals).

2.4.5 Determining secondary forest loss from different forest ages

To calculate secondary forest loss for different forest age groups, we generated four age category maps for 2004, 2008, 2010 and 2012 by tracking individual secondary pixels in time back to their year of first emergence in the dataset (Appendix Table A.8). To account for the differences in forest area among different age groups, we report secondary forest losses as proportional loss rates whereby the annual secondary forest loss for individual age categories were divided by the corresponding total secondary forest area for that age category (Appendix Table A.8). The number of age categories that we considered increased over time for each map. For example, the secondary forest age map for 2004 only has two age categories (0-4, >4 years), while the secondary forest age map for 2012 contains five age categories (0-2, 2-4, 4-8, 8-12, >12 years). As it was not possible to compare the same age category across all intervals, we restricted our analysis of changes in forest loss by age category to two intervals (2008-2010 and 2012-2014) for which it was possible to compare identical age categories (0-4, 4-8, >8 years). For these two intervals, we computed the

percentage of secondary forest loss annually for each age categories (i.e. 0-4, 4-8 and >8 years) within individual 0.1° × 0.1° grid cells and compared the pixel-level forest loss distributions between both intervals.

2.4.6 Null model analysis.

To test whether the accelerated loss of secondary forest was driven simply by increases in secondary forest area relative to primary forest area over time, we compared TERRACLASS secondary forest loss estimates to predictions from a statistical null model based on Fisher's non-central hypergeometric distribution, a modification of the hypergeometric distribution which allows the sampling probabilities of two binomially distributed variables to be adjusted according to an odds ratio. The odds ratio for cutting of secondary forests relative to primary forests was computed for the first TERRACLASS interval (2000-2004), based on the known total areas of both secondary and primary forest at the beginning of the interval from sample-based estimates (stable secondary forest + secondary forest loss within the interval) and the known secondary and total forest loss during the interval. For the first interval (2000-2004), this odds ratio was found to be 13.69 (i.e. secondary forests were >13 times more likely to be cut than primary forests). We applied the null model to each TERRACLASS interval, considering interval-specific total forest loss and available primary and secondary forest areas but maintaining the same odds ratio for preferential cutting of secondary forests as in the first interval. The null model analysis was conducted in R using the 'BiasedUrn' package.

2.4.7 Calculating carbon sequestration forgone due to the clearance of secondary forest.

To estimate the lost carbon sequestration potential arising from secondary forest cutting, we applied a Michaelis-Menten model commonly used in assessments of secondary biomass recovery (Batterman et al., 2013; Galbraith et al., 2019; Poorter et al., 2016). In this model, the amount of carbon sequestered in secondary forests at age t is given by: $C(t) = (C_{max} * t)/(\alpha_{50} + t)$, where C_{max} is average old-growth carbon storage for Amazon forests (170.60 Mg C ha^{-1})[30],

α_{50} is the half-saturation content denoting the time taken to reach half of the maximum carbon sequestration (35 years) (Galbraith et al., 2019), and age t is the average age of secondary forest when cleared. We estimated *t* as the area-weighted mean age of secondary forest loss in the last time interval (Appendix Table A.8, 2012-2014 time interval), taking the midpoint of each age category to represent the actual age of the secondary forest when cut. For the oldest age category, we conducted a sensitivity analysis where the mean age varied from 12-20 years. The final value of *t* used in the calculation above ranged from 5.50-6.57 years, once the uncertainty associated with the midpoint of the oldest age category was accounted for. The lost carbon sequestration opportunity due to secondary forest cutting was calculated by subtracting the secondary forest carbon sequestration at average cutting age t ($C(t)$) from its potential maximum carbon sequestration (C_{max}) and scaling this by the total area of lost secondary forest over our study period (180,329 $\pm$ 11,760 km^2 from sample-based estimates).

References

Aguiar, A.P.D. et al. 2016. Land use change emission scenarios: anticipating a forest transition process in the Brazilian Amazon. *Global change biology.* **22**(5), pp.1821-1840.

Almeida, C.A.d. et al. 2016. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. *Acta Amazonica.* **46**(3), pp.291-302.

Arévalo, P. et al. 2019. Continuous monitoring of land change activities and post-disturbance dynamics from Landsat time series: A test methodology for REDD+ reporting. *Remote Sensing of Environment.* p111051.

Assunção, J. et al. 2013. DETERring deforestation in the Brazilian Amazon: environmental monitoring and law enforcement. *Climate Policy Initiative.* pp.1-36.

Avitabile, V. et al. 2016. An integrated pan-tropical biomass map using multiple reference datasets. *Global change biology.*

Baccini, A. et al. 2012. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change.* **2**(3), pp.182-185.

Bank, W. 2014. CO^2 emissions (metric tons per capita).

Batterman, S.A. et al. 2013. Key role of symbiotic dinitrogen fixation in tropical forest secondary succession. *Nature.* **502**(7470), p224.

Brockerhoff, E.G. et al. 2008. Plantation forests and biodiversity: oxymoron or opportunity? *Biodiversity and Conservation.* **17**(5), pp.925-951.

Carvalho, R. et al. 2019. Changes in secondary vegetation dynamics in a context of decreasing deforestation rates in Pará, Brazilian Amazon. *Applied Geography.* **106**, pp.40-49.

Chazdon, R.L. et al. 2016. Carbon sequestration potential of second-growth forest regeneration in the Latin American tropics. *Science Advances.* **2**(5), pe1501639.

Cordeiro, I.M.C.C. et al. 2017. O manejo da floresta secundária na Amazônia Oriental. *Embrapa Amazônia Oriental-Capítulo em livro científico (ALICE).*

Dirzo, R. and Raven, P.H. 2003. Global state of biodiversity and loss. *Annual review of Environment and Resources.* **28**(1), pp.137-167.

Galbraith, D. et al. 2019. High-resolution Assessment of the Net Biomass Balance of the Brazilian Amazon: 2004-2014. *Global Change Biology.* **In review**.

Gibbs, H.K. et al. 2016. Did ranchers and slaughterhouses respond to zero-deforestation agreements in the Brazilian Amazon? *Conservation Letters.* **9**(1), pp.32-42.

Gibbs, H.K. et al. 2015. Brazil's soy moratorium. *Science.* **347**(6220), pp.377-378.

Gibson, L. et al. 2011. Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature.* **478**(7369), p378.

Kato, O. et al. 2004. Plantio direto na capoeira. *Embrapa Amazônia Oriental-Artigo em periódico indexado (ALICE).*

Lennox, G.D. et al. 2018. Second rate or a second chance? Assessing biomass and biodiversity recovery in regenerating Amazonian forests. *Global change biology.* **24**(12), pp.5680-5694.

Maia, H. et al. 2011. Avaliação do Plano de Ação para Prevenção e Controle do Desmatamento na Amazônia Legal: PPCDAm: 2007-2010.

Malhi, Y. et al. 2009. Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest. *Proceedings of the National Academy of Sciences.* **106**(49), pp.20610-20615.

Martin, P.A. et al. 2013. Carbon pools recover more quickly than plant biodiversity in tropical secondary forests. *Proc. R. Soc. B.* **280**(1773), p20132236.

Massoca, P.E.D.S. et al. 2017. 4.6 Lessons from the soy and beef moratoria in Brazil. *Zero deforestation: A commitment to change.* p151.

Olofsson, P. et al. 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment.* **148**, pp.42-57.

Olofsson, P. et al. 2013. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment.* **129**, pp.122-131.

Pará State Government. INSTRUÇÃO NORMATIVA SEMA Nº 08, DE 28-10-2015. 2015. *Available at: https://sogi8.sogi.com.br/Arquivo/Modulo113.MRID109/Registro1230833/instrucao%20normativa%20no%2008.pdf.* pp.3-4.

Poorter, L. et al. 2016. Biomass resilience of Neotropical secondary forests. *Nature.* **530**(7589), p211.

PRODES. 2017. Correcting the PRODES Mask. http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes/pdfs/nt_deslocamentomascara.pdf. *National Institute for Space Research (INPE).*

PRODES, p. 2018. Projecto Prodes: Monitoramento de Floresta Amazonica Brasileira por satelite. http://www.obt.inpe.br/prodes/index.php. *Instituto Nacional de Pesquisas Espaciais*

RAISG. 2012. Amazonia under Pressure. p10.

Rudorff, B.F.T. et al. 2011. The soy moratorium in the Amazon biome monitored by remote sensing images. *Remote Sensing.* **3**(1), pp.185-202.

Saah, D. et al. 2019. Collect Earth: An online tool for systematic reference data collection in land cover and use applications. *Environmental Modelling & Software.* **118**, pp.166-171.

Spracklen, D.V. et al. 2012. Observations of increased tropical rainfall preceded by air passage over forests. *Nature.* **489**(7415), p282.

TerraClass, p. 2018. TERRACLASS: mapping land uses/land covers for all previous deforested areas in the Brazilian Amazon. https://www.terraclass.gov.br. *National Institute for Space Research (INPE) and Brazilian Agricultural Research Corporation (EMBRAPA).*

Tyukavina, A. et al. 2017. Types and rates of forest disturbance in Brazilian Legal Amazon, 2000–2013. *Science advances.* **3**(4), pe1601047.

UNFCCC. 2015. Nationally Determined Contributions (NDCs).

Vieira, I.C.G. et al. 2014. Challenges of governing second-growth forests: A case study from the Brazilian Amazonian State of Pará. *Forests.* **5**(7), pp.1737-1752.

Chapter 3
Mapping tropical disturbed forests using multi-decadal 30 m optical satellite imagery (Paper Ⅱ)

Wang, Y., Ziv, G., Adami, M., Mitchard, E., Batterman, S.A., Buermann, W., Marimon, B.S., Junior, B.H.M., Reis, S.M., Rodrigues, D. and Galbraith, D., 2019. Mapping tropical disturbed forests using multi-decadal 30 m optical satellite imagery. *Remote sensing of environment*, 221, pp.474-488.

Abstract

Tropical disturbed forests play an important role in global carbon sequestration due to their rapid post-disturbance biomass accumulation rates. However, the accurate estimation of the carbon sequestration capacity of disturbed forests is still challenging due to large uncertainties in their spatial distribution. Using Google Earth Engine (GEE), we developed a novel approach to map cumulative disturbed forest areas based on the 27-year time-series of Landsat surface reflectance imagery. This approach integrates single date features with temporal characteristics from six time-series trajectories (two Landsat shortwave infrared bands and four vegetation indices) using a random forest machine learning classification algorithm. We demonstrated the feasibility of this method to map disturbed forests in three different forest ecoregions (seasonal, moist and dry forest) in Mato Grosso, Brazil, and found that the overall mapping accuracy was high, ranging from 81.3% for moist forest to 86.1% for seasonal forest. According to our classification, dry forest ecoregion experienced the most severe disturbances with 41% of forests being disturbed by 2010, followed by seasonal forest and moist forest ecoregions. We further separated disturbed forests into degraded old-growth forests and post-deforestation regrowth forests based on an existing post-deforestation land use map (TerraClass) and found that the area of degraded old-growth forests was up to 62% larger than the extent of post-

deforestation regrowth forests, with 18% of old-growth forests actually being degraded. Application of this new classification approach to other tropical areas will provide a better constraint on the spatial extent of disturbed forest areas in Tropics and ultimately towards a better understanding of their importance in the global carbon cycle.

3.1 Introduction

As hotspots of global biodiversity and carbon storage, tropical forests play an important role in biodiversity conservation, climate change mitigation and the provision of multiple other ecosystem services (Foley et al., 2005). However, millions of hectares of tropical forests have been lost due to deforestation and degradation disturbances, resulting in estimated net carbon emissions of 1.4 ± 0.5 Pg yr^1 from 1990-2010 (Houghton, 2012). These emissions represent the second largest anthropogenic source of carbon dioxide to the atmosphere after burning of fossil fuels (van der Werf et al., 2009). In contrast, a significant proportion of previously disturbed tropical forests are regrowing, trapping some of the carbon we are adding to the atmosphere, and with the potential to sequester more in the future. The carbon sink due to tropical forest recovering from deforestation and logging has been reported to be up to 70% greater than that of intact tropical forests (Pan et al., 2011). However, our ability to accurately assess tropical carbon sources or sinks is hampered by the lack of precise information on the extent of disturbed forests in the tropics (Baccini et al., 2017).

Remote sensing has played a key role in identifying forest disturbances and recovery, especially with the recent proliferation of high-resolution satellite data (Hansen et al., 2013). Several approaches have previously been used to map disturbed forests in tropical regions, including optical approaches based on moderate resolution MODIS imagery (Langner et al., 2007), high-resolution Landsat imagery (Vieira et al., 2003; Lu, 2005) and very high-resolution SPOT data (Kimes et al., 1999; Souza et al., 2003; Carreiras et al., 2014) , as well as Synthetic Aperture Radar (SAR) (Kuplich, 2006; Trisasongko, 2010) and Lidar-based approaches (Andersen et al., 2014). However, the majority of these studies have focused on local scales and have been based on single date images. For example, Vieira et al. (2003) classified forests into young, intermediate, advanced and mature forests for one municipality in the state of Pará, using Landsat spectral information and vegetation indices, and found that combining Landsat shortwave infrared band (1.55-1.75 μm) with NDVI generated a better classification than using any individual band/index. Carreiras et al. (2017) further

demonstrated the use of combined Landsat spectral bands with ALOS PALSAR backscatter intensity to distinguish secondary regrowth forest and mature forest in three landscapes in Brazilian Amazon. Such multiple multi-sensor fusion approaches have yet to be applied over regional scales.

Several regional satellite-based land cover classifications that include secondary regrowth and forest degradation have become available for Neotropical regions. Two prominent examples are the TerraClass post-deforestation land use/land cover classification (Almeida et al., 2016) and the DEGRAD forest degradation product (INPE, 2007) , both of which were developed by Brazilian National Institute for Space Research (INPE) specifically for the Brazilian Amazon. In TerraClass, available since 2004, secondary regrowth forest is mapped on previously deforested areas larger than 6.25 ha using a semi-manual approach (Almeida et al., 2016). The DEGRAD product is produced mainly by visual interpretation of Landsat and CBERS satellite images from a single year and is annually available between 2007 and 2013 (INPE, 2007). Recently, another product, MapBiomas, has become available that provides annual national-level land cover and land use maps for Brazil (MapBiomas, 2015). MapBiomas, available from 2000 to 2016, classifies forest land cover as dense forest, open forest, secondary forest, degraded forest, flooded forest or mangrove, using an empirical decision tree classification algorithm based on single date spectral mixture analysis. All of those single date imagery based approaches are limited in the discriminatory power they can provide as they make no use of temporal degradation/recovery signals which characterise disturbed forests. Thus, none of the existing products fully exploits the potential of existing Landsat time-series data spanning multiple decades to provide reliable maps of both forest regrowth and degradation. Furthermore, none of these products captures historical (pre-2000) disturbances. There is therefore a clear need for a product that provides a more comprehensive picture of historical disturbances over tropical regions.

Methods that exploit temporal information in satellite data (e.g. threshold approaches, trajectory fitting or segmentation) have been found to be very useful for mapping forest disturbances (White et al., 2017; Hirschmugl et al., 2017;

Hermosilla et al., 2015; Kayastha et al., 2012; Huang et al., 2010; Kennedy et al., 2010; Kennedy et al., 2007). However, majority of these time-series based approaches are based on a single time-series trajectory and have mainly been implemented at local scales in extratropical regions (e.g. Canada, U.S.). For example, the recently developed LandTrendr (Kennedy et al., 2010), Vegetation Change Tracker (Huang et al., 2010) and patch-based VeRDET (Vegetation Regeneration and Disturbance Estimates through Time) (Hughes et al., 2017) algorithms have all only been extensively tested in the United States. A recent inter-comparison of disturbance detection algorithms for US forests found that different time-series analysis algorithms are sensitive to different disturbance patterns, with little agreement among these disturbance detection results (Cohen et al., 2017). Thus, when applying these algorithms elsewhere, local calibration and further secondary classification are needed to improve the algorithm's classification performance (Cohen et al., 2018). Machine learning approaches (i.e. random forest) offer the potential to harness the differential sensitivities of different time-series once provided with an appropriate training dataset, but have rarely been coupled with multiple time-series trajectories in Tropics.

In this study, we develop a novel Landsat multiple time-series based classification methodology to map cumulative disturbed forest areas in Tropics, which exploits the power of 1) time-series images relative to single date images, 2) an ensemble of reflectance bands/indices trajectories relative to single trajectories, and 3) machine learning algorithms which enhances classification power by harnessing the differential sensitivities of different time-series. The 'disturbed forests' in this study include both degraded old growth forests and post-deforestation regrowth forests. The former are characterised by a reduction of forest canopy cover (e.g. selective logging, windfall, fire) but have not been clearfelled and thus have not been included in deforestation estimates. The latter refer to areas that have been previously deforested (clearfelled) and converted to other land uses (e.g. pasture, agriculture and mining) but which have subsequently undergone a recovery process following abandonment. Our approach integrates information from six different time-series trajectories (Landsat 5/7 short-wave infrared band 5, band 7, NDVI, SAVI, $NDWI_{2130}$, $NDWI_{1640}$), extracting both statistical and temporal

characteristics from each trajectory which then serve as inputs for random forest classification. It not only captures disturbances occurring within study period (1984-2010), but also areas disturbed prior to 1984 which thereafter have exhibited clear recovery patterns. Here, we apply this method to three forest ecoregions (seasonal, moist and dry forests) in the Brazilian state of Mato Grosso.

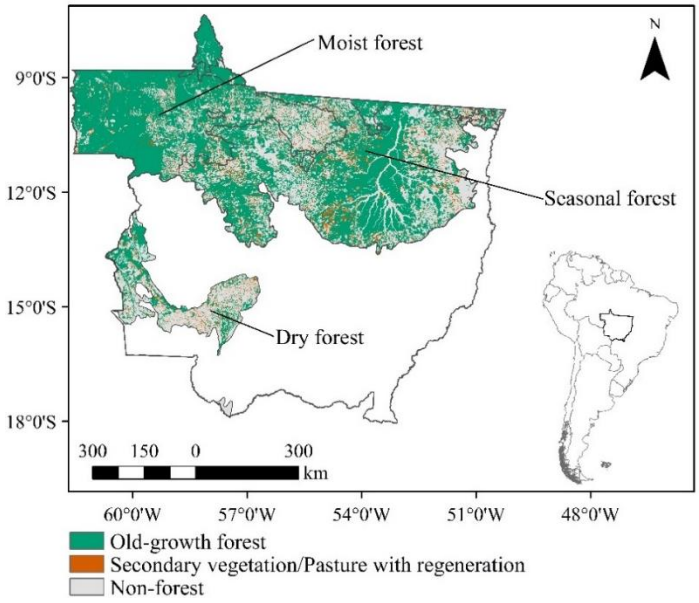
3.2 Study Area

Our study area (Figure 3.1), the state of Mato Grosso, is located in the southern edge of Brazilian Legal Amazon. Mato Grosso is the third largest state in Brazil, covering a total area of 903,357 km^2. According to the Terrestrial Ecoregions of the World (TEOW) from World Wildlife Fund (WWF), 43% of Mato Grosso area is covered by *Cerrado* (tropical savanna), 27% by seasonal forest, 18% by moist forest, 6% by dry forest and 6% by *Pantanal* (tropical wetlands) (Olson et al., 2001). In Mato Grosso, 139,917 km^2 have been deforested since 1988 (PRODES, 2018) amounting to 26.5 % of the state's intact forest in that year (Skole and Tucker, 1993), most of which has been converted into pasture and agricultural land use due to demand for beef and soy beans (Barona et al., 2010). According to TerraClass (Almeida et al., 2016), herbaceous pasture and shrubby pasture cover 61.4% of the total deforested areas in Mato Grosso while 19.2% of deforested areas are under secondary regrowth (including secondary vegetation and regeneration with pasture). The combination of extensive disturbances and significant amount of remaining intact forest makes Mato Grosso an ideal testbed for the application of our newly developed disturbed forests mapping approach (see section 3).

As indicated, TerraClass is a project that maps land use/land cover on previous deforested areas provided by PRODES (Program for Deforestation Monitoring, PRODES, 2018) at approximately bi-annual intervals across the Brazilian Legal Amazon (Almeida et al., 2016). TerraClass classifies previously deforested areas into 12 land use categories including pasture, annual crops, secondary vegetation and urban areas. It is extensively validated via field campaigns to determine the accuracy of classification. These have been conducted across

different Amazonian regions, including the state of Mato Grosso. This is the best available information on the distribution of secondary forests in any region of the Tropics. However, TerraClass involves a huge effort based largely on visual interpretation and does not map degradation.

The aim of this study is to propose a Landsat multiple time-series based approach in Tropics to 1) improve the efficiency/cost-effectiveness of mapping disturbed forests vs. intact forests, facilitating future TerraClass efforts, 2) map degraded old-growth forests (outside of TerraClass), and 3) eventually enable mapping of disturbed forests over domains for which no reliable data on forest disturbance exist. Only forest areas are considered in this study. To make sure all non-forest areas are excluded, we created a forest cover mask by merging TerraClass-2010 old-growth forest, secondary vegetation and pasture with regeneration categories (Figure 3.1). The latter category effectively captures the beginning of the regenerative process containing shrubs and early successional vegetation (Almeida et al., 2016).



Old-growth forest

Secondary vegetation

Pasture with regeneration

Figure 3.1 TerraClass classification map for 2010.
Pasture with regeration in TerraClass is treated as young secondary vegetation. Later, we merged old-growth forest, secondary vegetation and pasture with regeneration into the forest cover mask as the forest boundary. The study area encompasses three WWF forest ecoregions (moist, seasonal and dry forest).

3.3 Methodology and dataset

The whole approach was developed in Google Earth Engine (GEE) (Gorelick et al., 2017). GEE is a cloud-based geospatial processing platform which consists of over 40 years of historical and current Earth observation imagery, making pixel-based land use and land cover classification feasible across large regions through its inbuilt machine learning algorithms. The overall methodology (Figure 3.2) consisted of building Landsat multiple (six) annual time-series trajectories,

calculating trajectory metrics (eleven metrics divided into four groups, Table 3.2), generating a training and validation database, applying a machine learning random forest classification algorithm and validating the disturbed forests vs. intact forests classification map, all of which were coded and processed in GEE. We subsequently used the post-deforestation regrowth forest mask generated from TerraClass-2010 to separate the disturbed forests identified through our classification map into post-deforestation regrowth forests and degraded forests (Table 3.1). Finally, we performed a relative important analysis of trajectories and trajectory metrics used in the random forest classification to evaluate the extent to which the full suite of all trajectories/metrics enhanced discriminatory power relative to a single trajectory or individual group of trajectory metrics. To do this, ten separate classifications were performed whereby our classification procedure was repeated for each individual trajectory separately (but using all four groups of trajectory metrics), or separately for individual groups of trajectory metrics (but using all six trajectories).

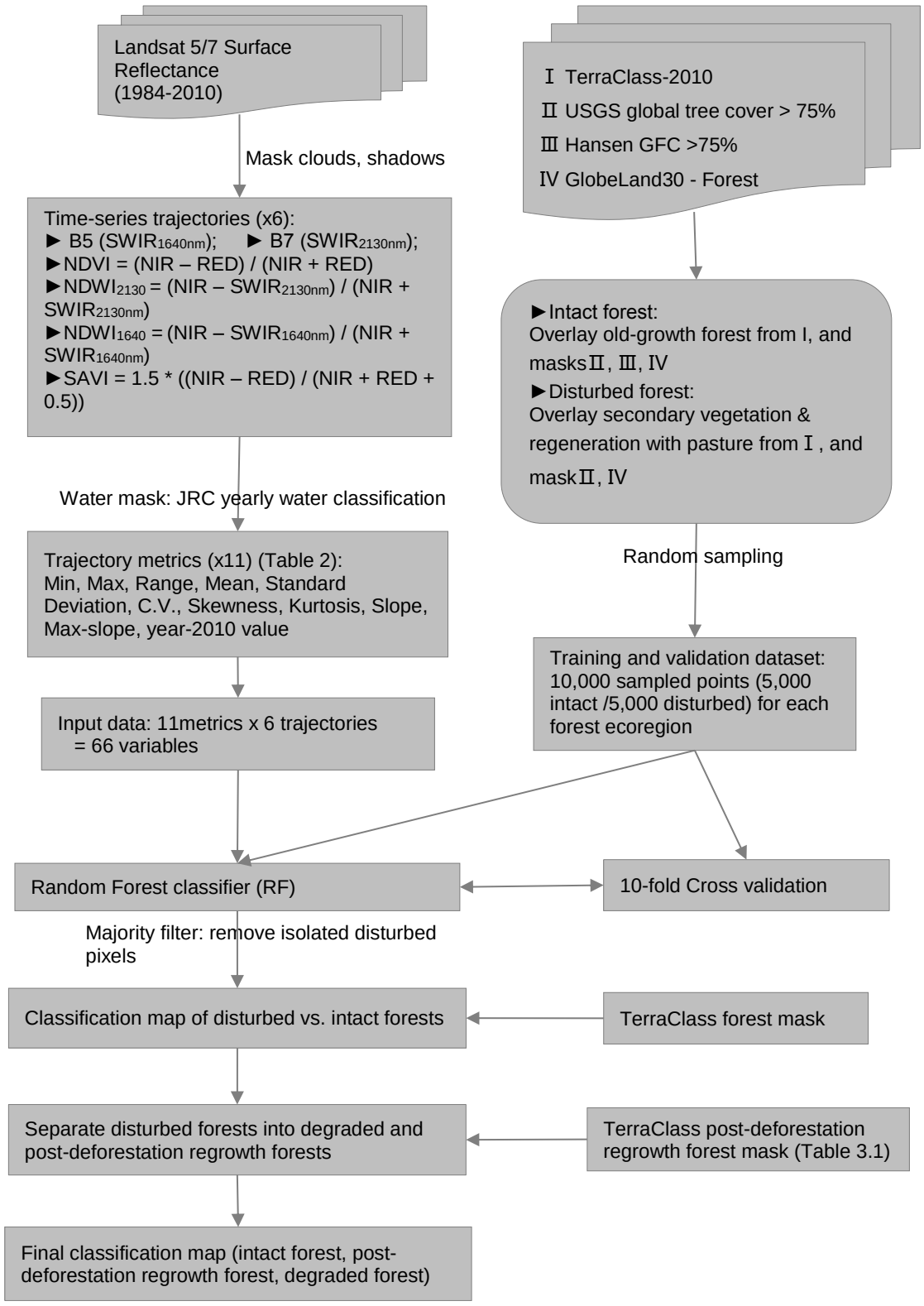


Figure 3.2 Classification Methodology for discrimination of disturbed forests and intact forests.

Table 3.1 Classification categories for forested land cover types used in this study.

Categories	Description
Total area	Total area of each ecoregion
Forest cover	Forest mask from TerraClass classification for the year of 2010, combining TerraClass categories of old-growth forest, secondary vegetation and regeneration with pasture.
Intact forest	Forests that have never been experienced any detectable disturbances during 1984-2010. Classified from this study.
Disturbed forest	Cumulative disturbed forest areas during 1984-2010. Classified from this study. Further separated into Post-deforestation regrowth forest & Degraded forest.
Post-deforestation regrowth forest	Areas that have been previously deforested (clearfelled) and converted to other land uses (e.g. pasture, agriculture and mining) but which have subsequently undergone a recovery process following abandonment. Secondary vegetaion or regeneration with pasture in TerraClass-2010.
Degraded forest	Degraded old-growth forests. Characterised by a reduction of forest canopy cover (e.g. selective logging, windfall, fire) but have not been clearfelled and thus have not been included in deforestation estimates.

3.3.1 Time-series trajectories

3.3.1.1 Landsat surface reflectance dataset

We used Landsat atmospherically corrected surface reflectance (SR) products (30 m resolution) (Masek et al., 2006; USGS, 2018) to generate annual time-series trajectories. All Landsat-5 Thematic Mapper (TM) surface reflectance images aquired during the period of 1984-2010 were used except for 2001 and 2002. In 2001, most images had striping artifacts limiting their use, while in 2002, images from Landsat 5 only covered 61% of our study area. For these reasons, we used Landsat-7 Enhanced Thematic Mapper Plus (ETM+) images, which are compatible in their spectral characteristics (Home et al., 2013; Claverie et al., 2015), for these two years. In terms of spectral bands, we chose spectral bands 3 (red, 0.52 - 0.60 µm) which is sensitive to the amount of chlorophyll, 4 (near-infrared, 0.76 - 0.90 µm) which is related to leaf cellular structure, 5 (shortwave-infrared, 1.55 - 1.75 µm) and 7 (shortwave-infrared, 2.08 - 2.35 µm) which relate to leaf water content (Nelson et al., 2000). To minimize the influence of variable

extent of rivers on the classification, we excluded water bodies in our analysis using the Joint Research Center (JRC) Yearly Water Classification History v1.0 product. This dataset contains maps of the location and temporal distribution of surface water from 1984 to 2015 at annual resolution, generated using more than three million scenes from Landsat 5, 7 and 8 (Pekel et al., 2016).

3.3.1.2 Generating time-series trajectories

We processed 11,483 images in total for our entire study period (1984-2010), ranging from 257 to 715 annual images depending on data availablity, with annual spatial coverage of 99% of our study area (see Appendix Table B.1). Five steps were involved to process the Landsat SR data and produce time-series image stacks for 1984-2010. First, areas covered by clouds and cloud shadows were removed based on the pixel quality and radiometric saturation attributes of the Landsat surface reflectance product. Second, original surface reflectance (16-bit signed integer) values were converted to 0-1 range values by multiplying by the scale factor of 0.0001. Third, four vegetation indices (VIs) were calculated including the Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index ($NDWI_{2130}$, $NDWI_{1640}$) (Chen et al., 2005) and Soil-Adjusted Vegetation Index (SAVI) (Huete, 1988). Fourth, to minimise the influence of cloud contamination and improve the quality of input data, we selected the maximum value of individual VIs for each year (Maxwell and Sylvester, 2012). For time-series of reflectance from spectral bands 5 and 7, median values were calculated for each year. In the final step, we used the JRC Yearly Water Classification History v1.0 product to mask water areas (Pekel et al., 2016). After processing, annual time-series trajectories (1984-2010) of Landsat SR spectral band 5 (1.55 - 1.75 µm), band 7 (2.08 - 2.35 µm), NDVI, $NDWI_{2130}$, $NDWI_{1640}$ and SAVI were used for the classification of disturbed forests and intact forests.

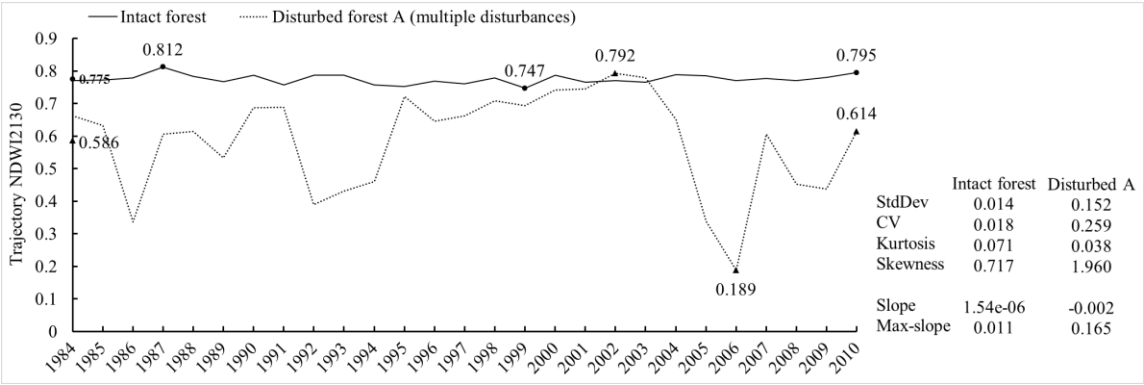
3.3.2 Trajectory metrics

We calculated eleven metrics divided into four groups (Table 3.2) for each of the six spectral trajectories to act as inputs for random forest algorithm (see section 3.4), based on *a priori* expectations of divergence between intact and disturbed forests. Each of these 11 metrics may capture information that is linked to a particular disturbance type. For example, the coefficient of variation (C.V.) shows the extent of variability in relation to the mean. Forests which have experienced large disturbances would be expected to have higher C.V. than undisturbed intact forests. We further hypothesized that time-series trajectories of intact forest would follow a normal distribution, while those of disturbed forest would tend not to and be much more likely to exhibit greater skewness and kurtosis. Finally, trends (based on linear regressions) were also estimated from the time-series trajectories. We hypothesized that disturbance events would likely result in either decreasing (deforestation/degradation) or increasing (regrowth) trends over time, and thus expected that the regression slopes of disturbed pixels would be much smaller/greater than undisturbed pixels where we expected that the slope value is close to zero. It has been found that regrowth secondary forests in Amazonia are cut and burned on average every 5 years (Aguiar et al., 2016). Thus, we also considered the maximum absolute regression slopes derived from individual 5-year windows within the 1984-2010 study period.

Figure 3.3 demonstrates differences in trajectories and trajectory metrics between intact and disturbed forest pixels. For intact forests (undisturbed during 1984-2010), we expected trajectories to fluctuate, but to follow a normal distribution pattern, while trajectories of disturbed forests were expected to exhibit more pronounced decrease and increase patterns. Trajectories of disturbed forest pixels' can follow various patterns, depending on whether they have been disturbed once (Figure 3.3 Disturbed B) or multiple times (Figure 3.3 Disturbed A) within the study period (1984-2010) or disturbed before 1984 but following a clear recovery pattern within study period (Figure 3.3 Disturbed C).

Table 3.2. Metrics for each time-series trajectory and related main GEE algorithms.
The metrics were divided into location, scale, temporal and single year groups which were further used for metric important analysis (see section 4.4).

Group	Name	Description	Main GEE algorithm
Location metrics	Min	Minimum of time-series	ee.Reducer.min()
	Max	Maximum of time-series	ee.Reducer.max()
	Range	The range between maximum and minimum of time-series	Code equation 'max-min'
	Mean	The mean of time-series	ee.Reducer.mean()
Scale metrics	StdDev	Standard deviation of time-series	ee.Reducer.stdDev()
	C.V.	Coefficient of variation of time-series	Code equation 'stdDev/mean'
	Kurtosis	Dispersion measure related to the tails of Normality distribution test (D'Agostino, 1970, see methods)	Code equations based on the reference
	Skewness	Symmetry measure related to Normality distribution test (D'Agostino, 1970, see methods)	Code equations based on the reference
Temporal metrics	Slope	Linear regression slope of total time-series	ee.Reducer.linearFit()
	Max-slope	Maximum linear regression slope of every 5-year window	Function of 5-year window; ee.Reducer.linearFit(); ee.Reducer.max()
Single year	Year-2010	Time-series trajectory value at year 2010	'FilterMetadata' equals 2010



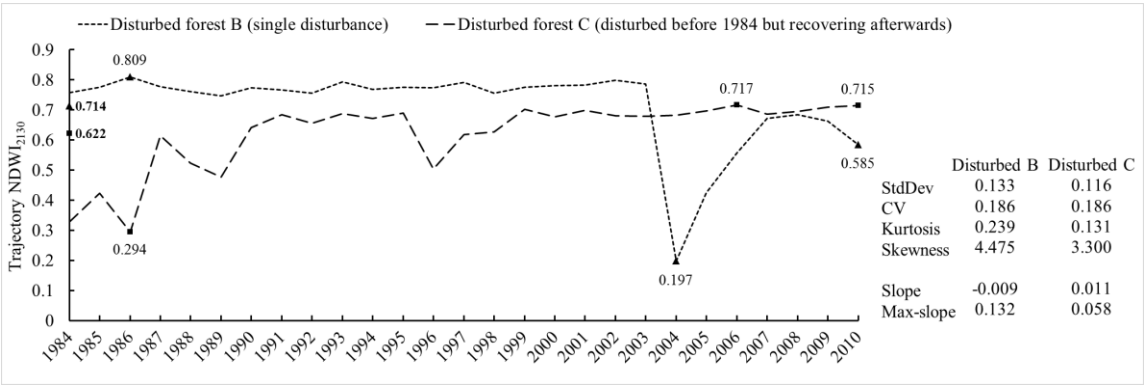


Figure 3.3 Examples ($NDWI_{2130}$) of time-series trajectories for illustrative intact forest pixel and disturbed forest pixels.

Values of trajectory scale and temporal metrics extracted from each trajectory (Table 3.2) are shown to the right of the graph. Metrics of max, min and year-2010 value are shown on the trajectory with the mean marked on y axis.

3.3.3 Sampling design

We used GEE random sampling to generate a set of spatially representative points of disturbed and intact forests for classification training and validation based on TerraClass-2010 map of old-growth forest, secondary vegetation and pasture with regeneration, USGS (United States Geological Survey) 30 m Global Tree Cover 2010 (Hansen et al., 2013), the Hansen Global Forest Change (GFC) product (Hansen et al., 2013), and 30 m Global Land Cover 2010 (GlobeLand30-2010) produced by National Geomatics Centre of China (Chen et al., 2015). Since TerraClass uses deforestation vector data from PRODES (PRODES, 2018) as input data to map subsequent land use/covers (Almeida et al., 2016), it inherited

PRODES historical misalignment issues. To better align TerraClass with GFC products, we registered the TerraClass-2010 classification map using the GEE image displacement algorithm by calculating the displacement between TerraClass-2010 forest mask and GFC forest mask (Hansen et al., 2013).

For intact forests, points were randomly sampled from areas that met the following conditions: i) classified as old-growth forest in TerraClass-2010; ii) tree canopy cover > 75% in GFC in 2000 and no forest loss during 2000-2010; iii) tree cover >75% in USGS 30 m Global Tree Cover 2010; and, iv) classified as forest in GlobeLand30-2010. Similarly, disturbed forest pixels were sampled from areas that satisfied the following conditions: i) classified as secondary vegetation or regeneration with pasture in TerraClass-2010; ii) tree cover > 75% in USGS 30 m Global Tree Cover 2010; iii) classified as forest in GlobeLand30-2010. To reduce the influence of unwanted positional errors among these land cover products and avoid edge effects, we required that both intact forest and disturbed forest sampled points were located at least 100m away from the patch boundary. For each forest ecoregion (moist/seasonal/dry forest), 10000 points (5000 intact and 5000 disturbed) were randomly sampled, respectively. In total, we sampled 30000 intact and disturbed points across the study area as the training and validation database.

3.3.4 Random forest classifier

Mapping of disturbed forests was performed by using the GEE Random Forest classifier algorithm, which has been recently successfully applied to cropland mapping (Xiong et al., 2017; Shelestov et al., 2017), oil palm plantation detection (Lee et al., 2016), mapping urban settlement and population (Patel et al., 2015) and soil mapping (Padarian et al., 2015). Random Forest (RF) classification is a relatively well-known supervised machine leaning algorithm that iteratively produces an ensemble of decision tree classifications by using corresponding randomly selected subsets of the training dataset (Breiman, 2001). It grows classification trees by splitting each node using a random selection subset of input variables, which reduces overfitting and yields a more robust classification

compared to other classifiers (Breiman, 2001). RF uses a voting system to classify data and the final classification category for each pixel is determined by the plurality vote of all trees generated to build the forest.

We used 66 variables comprising 11 metrics (Table 3.2) for each of the six time-series trajectories as input predictors for the RF classification. RF classifications were applied in moist, seasonal and dry forest ecoregions, respectively. All classifications were based on the outputs of 500 decision trees (See Appendix Figure B.1). Each tree split was based on eight variables randomly selected from all 66 input variables, which was the default configuration for the GEE random forest classifier. After constructing our disturbed forest classification, we performed a post-classification filtering to reduce noise and remove spurious classification artefacts by applying a 90m x 90m majority filter.

3.3.5 Classification validation

To evaluate how well our classification performed, we used ten-fold cross-validation (Schaffer, 1993; Kohavi, 1995) based on above randomly sampled database (See section 3.3, i.e. 10000 points for each forest ecoregion), which randomly partitions our sampled database into ten equal sized subsets. Of the ten subsets, a single subset (1000 points) was retained as the validation data for testing the classification algorithm, and the remaining nine subsets (9000 points) were used as training data for RF classifier. The cross-validation process was repeated ten times. The final accuracy estimation was determined by the average of ten-fold results. The accuracy matrix included overall accuracy (OA), producer's accuracy (PA), user's accuracy (UA) and Kappa statistic (Kohavi, 1995).

For an additional independent confirmation for our Landsat optical sensor based classification of disturbed forests vs. intact forests, we used another microwave radar based satellite product, ALOS/PALSAR 25 m spatial resolution mosaic imagery, as visual interpretation. ALOS PALSAR imagery consists of dual polarization HH (transmission of horizontal wave and reception of horizontal component) and HV (horizontal transmission and vertical reception), but it has

been shown that the polarization mode HV is more effective in deforestation detection than HH polarization (Motohka et al., 2014), which corresponds with findings of close relations between HV backscatter and vegetation structural properties (e.g. forest height, forest cover) (Joshi et al., 2015). Thus, we visually compared the 2007-2010 ALOS/PALSAR HV backscatter change with our final classification results.

SAR data are stored as digital number (DN) in unsigned 16 bit and typified by a high degree of speckles in the image (random 'salt and pepper' noise). To reduce noise and improve image interpretability, a multi-temporal speckle filter (7×7) (Lee, 1980; Lopes et al., 1990) was implemented in GEE and applied to 2007-2010 PALSAR images, without significant loss of spatial resolution. Filtered ALOS/PALSAR HV backscatter DN values were converted to sigma-naught (σ^0) in decibel (dB) units using the following equation:

$$\sigma^0 = 10 * log_{10}(DN^2) - 83 \qquad \text{(eq. 3.1)}$$

$\sigma^0$is generally negative and can vary from -35 dB in very low backscatter areas (degraded/deforested area), up to 0 dB for extremely high backscatter (dense forest area). For visual interpretation, we expected a decrease or an increase in σ^0 in forest areas that have been recently disturbed or are recovering from past disturbances (Joshi et al., 2015). However, we also expected that many disturbed areas in our classification would not be captured by PALSAR due to its short time period (2007-2010).

3.4 Results

3.4.1 Classification results

As represented in Figure 3.2, the new developed disturbed forests vs. intact forests classification approach was applied to three different ecoregions in Mato Grosso. The final classification map (Figure 3.4) was generated by training the random forest classifier individually for each ecoregion on the entire sampled database. Our classification results representative of the year 2010 show that disturbed forests (both post-deforestation regrowth forests and degraded forests) were widely spread across Mato Grosso, but were most prevalent along rivers and next to non-forest areas (Figure 3.4). Forests in Mato Grosso covered a total area of 295,383 km^2 in 2010 (Table 3.3), accounting for about 63% of the total study area. Our results show that, until 2010, 25% of the total forested area was disturbed (Table 3.3). Forest cover percentage varied considerably across ecoregions, ranging from 37% in dry forest to 74% in moist forest (Table 3.3). Dry forest experienced the most severe disturbances with 41% of forest cover classified as disturbed, followed by seasonal forest and moist forest where disturbed forests accounted for 28% and 20% of forest cover, respectively (Table 3.3).

We further separated disturbed forests identified through our classification map into post-deforestation regrowth forests and degraded forests. It shows that the area of degraded forests was up to 62% larger than the area of post-deforestation regrowth forests across ecoregions, with degraded forests and post-deforestation regrowth forests covering a total area of 47,039 km^2 and 28,246 km^2, respectively (Table 3.4). By comparing degraded forests and old-growth forests classified in TerraClass for the year of 2010, we found that 18% of areas identified as old-growth forests in TerraClass were actually degraded forests, ranging from 15% to 27% across various ecoregions (Table 3.4).

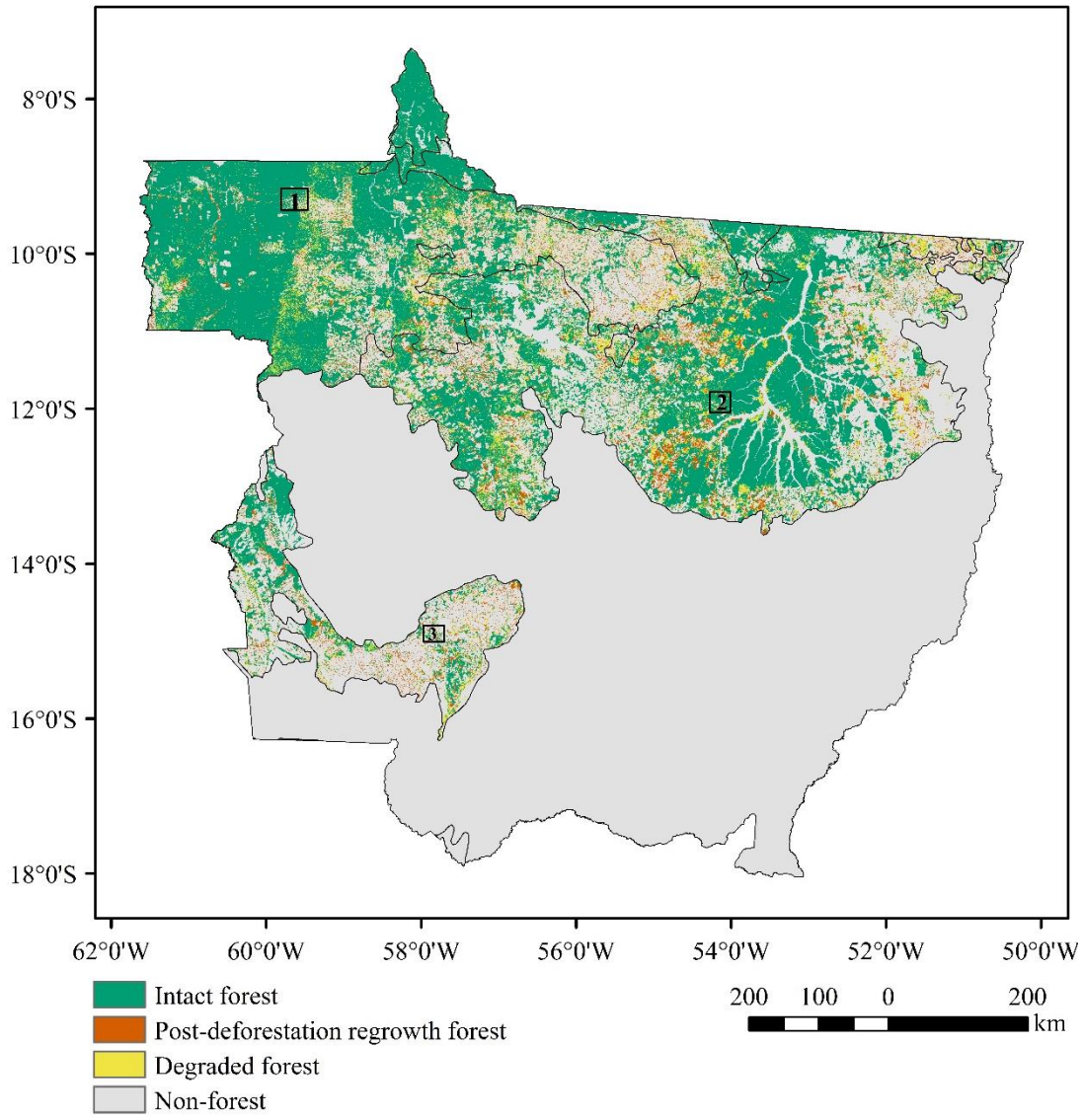


Figure 3.4 Classification map of intact forest, post-deforestation regrowth and degraded forest representative of the year 2010.

Non-forest areas include areas under anthropogenic use or natural savannahs/wetlands. Small areas 1 to 3 represent three focal regions within individual ecoregions, for which subsequent fine-scale visual interpretation confirmation were performed (Figure 3.5, Figure 3.6, Figure 3.7).

Table 3.3 Areal extent (km^2) of intact forest and historically disturbed forest representative of 2010.

	Moist forest	Seasonal forest	Dry forest	Total
Total area	170,154	245,514	54,454	470,122
Forest cover (% of total area)	125,474 (73.74%)	149,571 (60.92%)	20,338 (37.35%)	295,383 (62.83%)
Intact forest (% of forest cover)	100,050 (79.74%)	107,991 (72.20%)	12,058 (59.29%)	220,099 (74.51%)
Disturbed forest (% of forest cover)	25,424 (20.26%)	41,581 (27.80%)	8,280 (40.71%)	75,285 (25.49%)

Table 3.4 Areal extent (km^2) of post-deforestation regrowth forest and degraded forest representative of 2010.

	Moist forest	Seasonal forest	Dry forest	Total
Post-deforestation regrowth (% of disturbed forest)	8,188 (32.21%)	15,950 (38.36%)	4,108 (49.62%)	28,246 (37.52%)
Degraded forest (% of disturbed forest)	17,236 (67.79%)	25,631 (61.64%)	4,171 (50.38%)	47,039 (62.48%)
TerraClass old-growth forest	116,226	131,703	15,622	263,551
% of degraded forest within TerraClass	14.83%	19.46%	26.70%	17.85%

3.4.2 Ten-fold cross validation

Ten-fold cross validation was used as the main validation of our disturbed forests and intact forests classification map, with accuracy matrices provided in Table 3.5. Overall, all the classification accuracies were above 80% with Kappa agreements above 62%. Across ecoregions, the overall accuracy was the highest in seasonal forest at 86.1%, with a producer's accuracy of 88.9% for intact forests

and 83.3% for disturbed forests. In moist forest and dry forest regions, the overall accuracies were lower at 81.3% and 82.6%, respectively.

Table 3.5 Ten-fold cross validation accuracy based on sampled points from our study.

Regions	Overall accuracy	Producer's accuracy		User's accuracy		Kappa statistic
		Intact forest	Disturbed forest	Intact forest	Disturbed forest	
Moist forest	0.813	0.888	0.737	0.772	0.867	0.625
Seasonal forest	0.861	0.889	0.833	0.842	0.882	0.722
Dry forest	0.826	0.856	0.797	0.809	0.846	0.653

3.4.3 High-resolution image interpretation

To further validate our classification, we consider in detail one landscape within each biome, comparing our results to radar and other very high-resolution data. Examples in Figure 3.5-3.7 allow for visual comparison of our classification in selected focal areas within each forest ecoregion with corresponding ALOS PALSAR HV backscatter (σ^0) temporal (2007-2010) change composite images and very high-resolution (5 m) RapidEye true-colour composite images (Team, 2017). Overall, this comparison at local scales shows a very good visual agreement between our classification and the PALSAR temporal change as well as with RapidEye images across ecoregions (Figure 3.5-3.7), especially those logging roads shown in Figure 3.6. As expected, there were some mismatches between our classification and the temporal change in PALSAR HV σ^0, such as several disturbed areas from our classification not appearing in PALSAR temporal change image. This is likely due to PALSAR images only being available from 2007 and thus not capturing much forests disturbed before 2007.

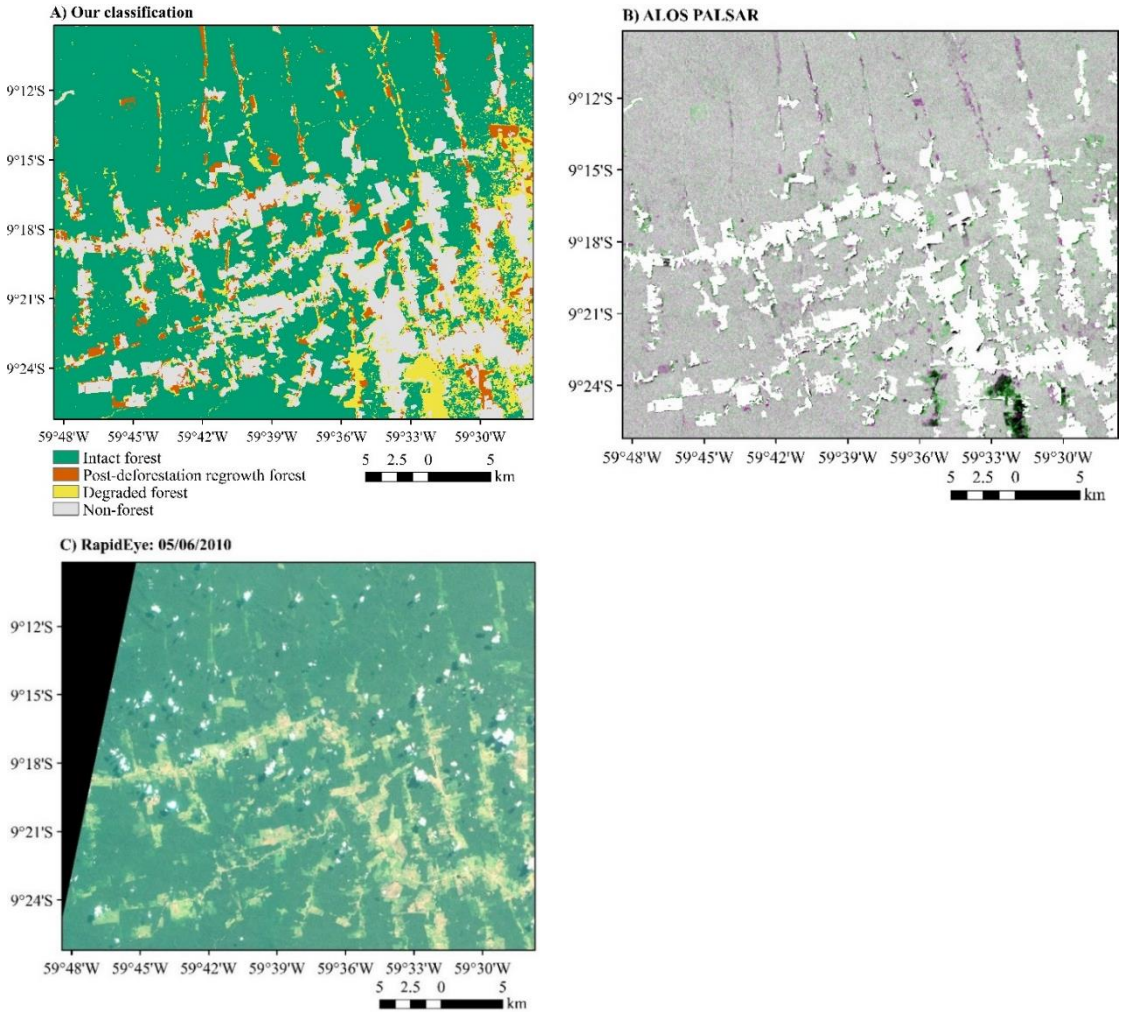


Figure 3.5 Moist forest focal region (area 1 in Figure 3.4).
A) Detailed classification map. B) Forest masked ALOS PALSAR HV σ0 temporal change, pink represents increase of σ0, green represents decrease of σ0 between 2007-2010, grey represents little/no change between 2007-2010, white areas are non-forest. C) RapidEye true-colour composite image (See Appendix Figure B.2 for better visualization).

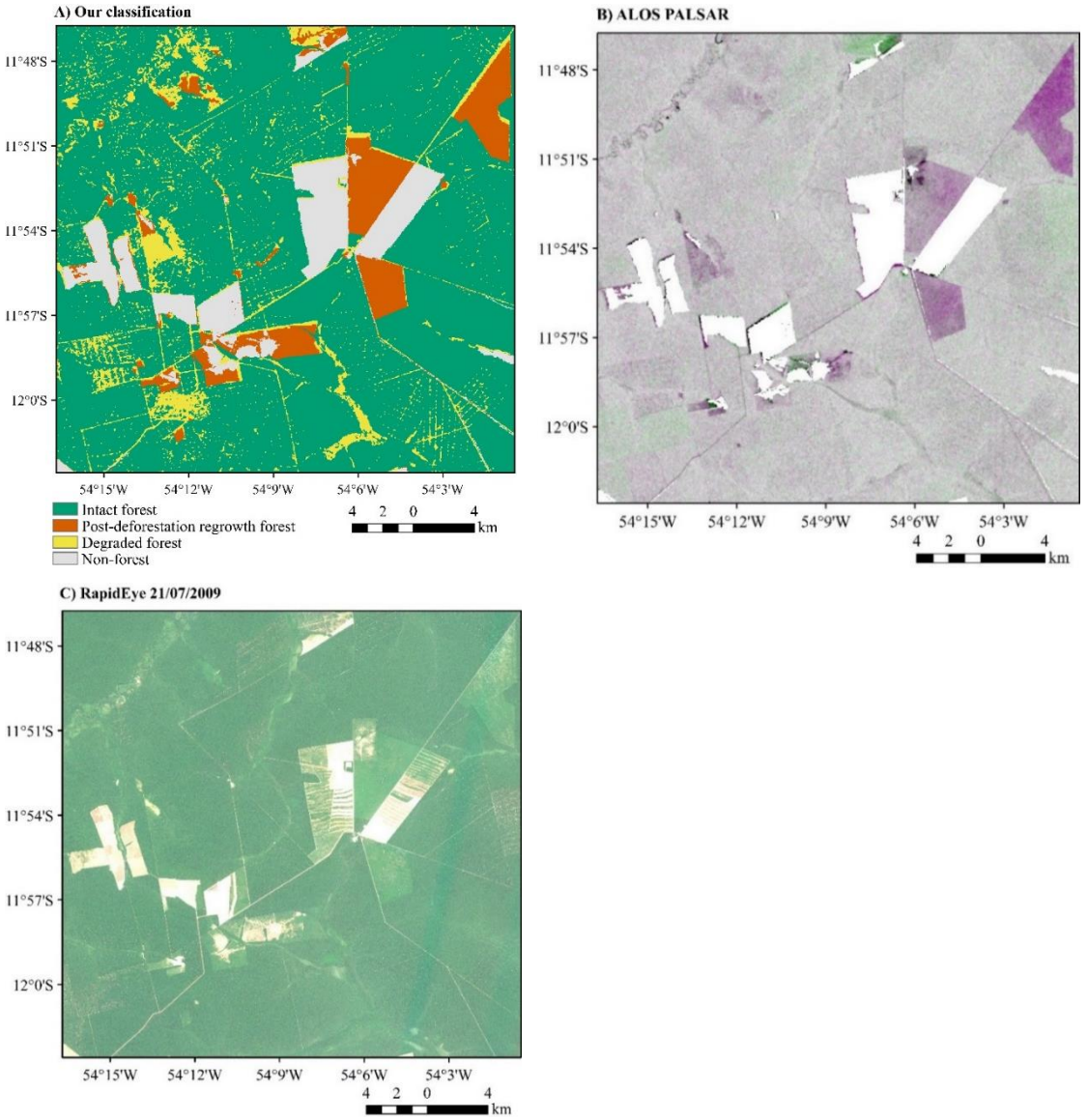


Figure 3.6 Seasonal forest focal region (area 2 in Figure 3.4).
A) Detailed classification map. B) Forest masked ALOS PALSAR HV σ0 temporal change, pink represents increase of σ0, green represents decrease of σ0 between 2007-2010, grey represents little/no change between 2007-2010, white areas are non-forest. C) RapidEye true-colour composite image (See Appendix Figure B.3 for better visualization).

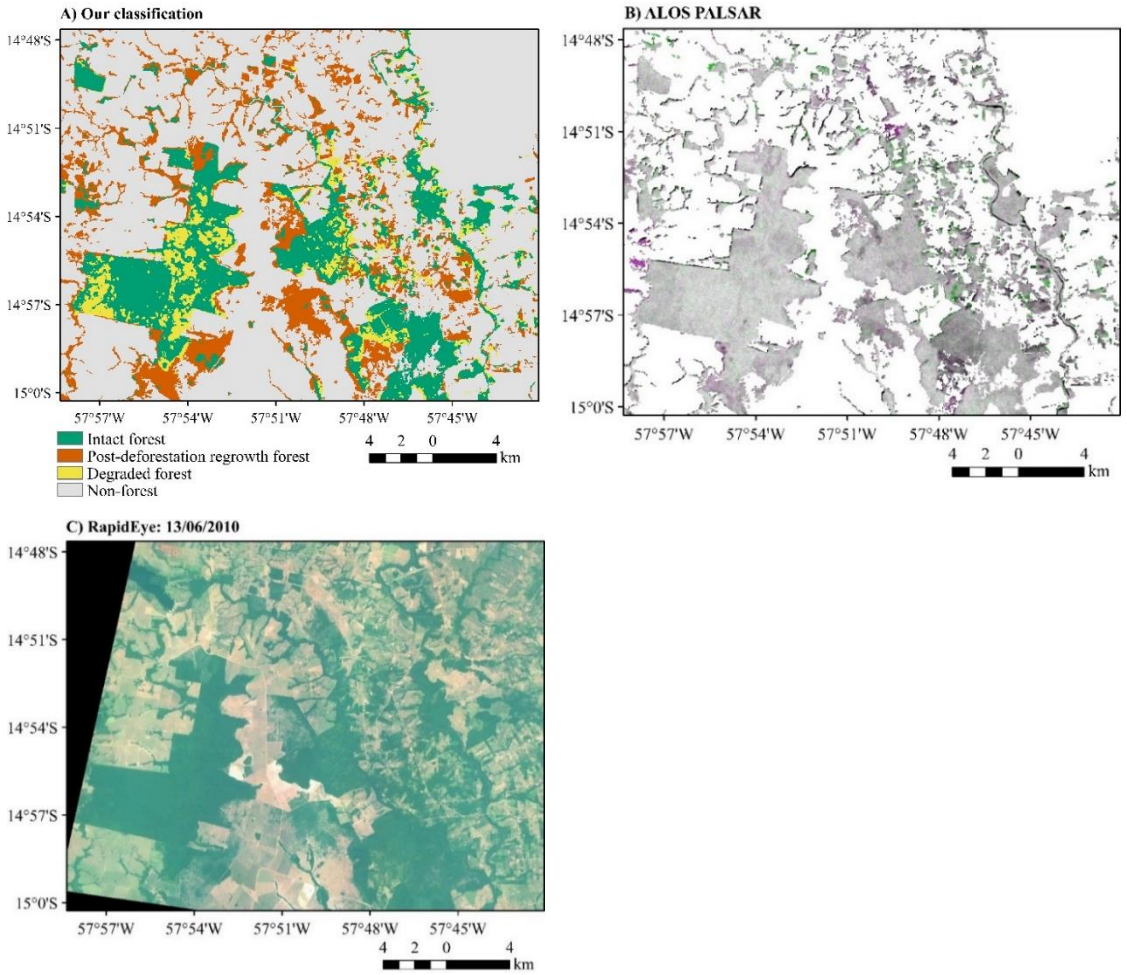


Figure 3.7 Dry forest focal region (area 3 in Figure 3.4).
A) Detailed classification map. B) ALOS PALSAR HV σ0 temporal change, pink represents increase of σ0, green represents decrease of σ0 between 2007-2010, grey represents little/no change between 2007-2010, white areas are non-forest. C) RapidEye true-colour composite image (See Appendix Figure B.4 in supplementary information for better visualization).

3.4.4 Importance of individual trajectories and metrics

The relative importance of individual trajectories in our classification was measured by the percentage of overall accuracy change (% OAC) when running our classification for a single trajectory (but using all four groups of trajectory metrics) relative to our full suite multi-trajectory classification (Table 3.5). The larger the overall accuracy change, the less important an individual trajectory is in distinguishing the differences between disturbed forests and intact forests. All of the single time-series trajectories based classifications had much lower (3-15% across ecoregions) overall classification accuracy than our full suite classification (Figure 3.8). In moist forest and dry forest ecoregions, Landsat shortwave spectral band 5 and 7 were the most important trajectories for distinguishing disturbed forests and intact forests, decreasing %OAC the least relative to our full suite classification. However, in the seasonal forest ecoregion, NDWI trajectories were the most important, decreasing the overall accuracy the least, followed by spectral band 7.

The important of specific groups of trajectory metrics (Table 3.2) was determined in an analogous manner to the importance of specific trajectories. Importance patterns for groups of metrics were similar across ecoregions (Figure 3.8B), with location metrics being the most important in distinguishing disturbed and intact forests, followed by temporal metrics, scale metrics and single year (2010) values. However, single year (2010) values alone were found to have much less discriminatory power than other metrics, resulting in much lower (up to 20%) classification accuracy relative to our full suite classification with all groups of metrics included (Figure 3.8B).

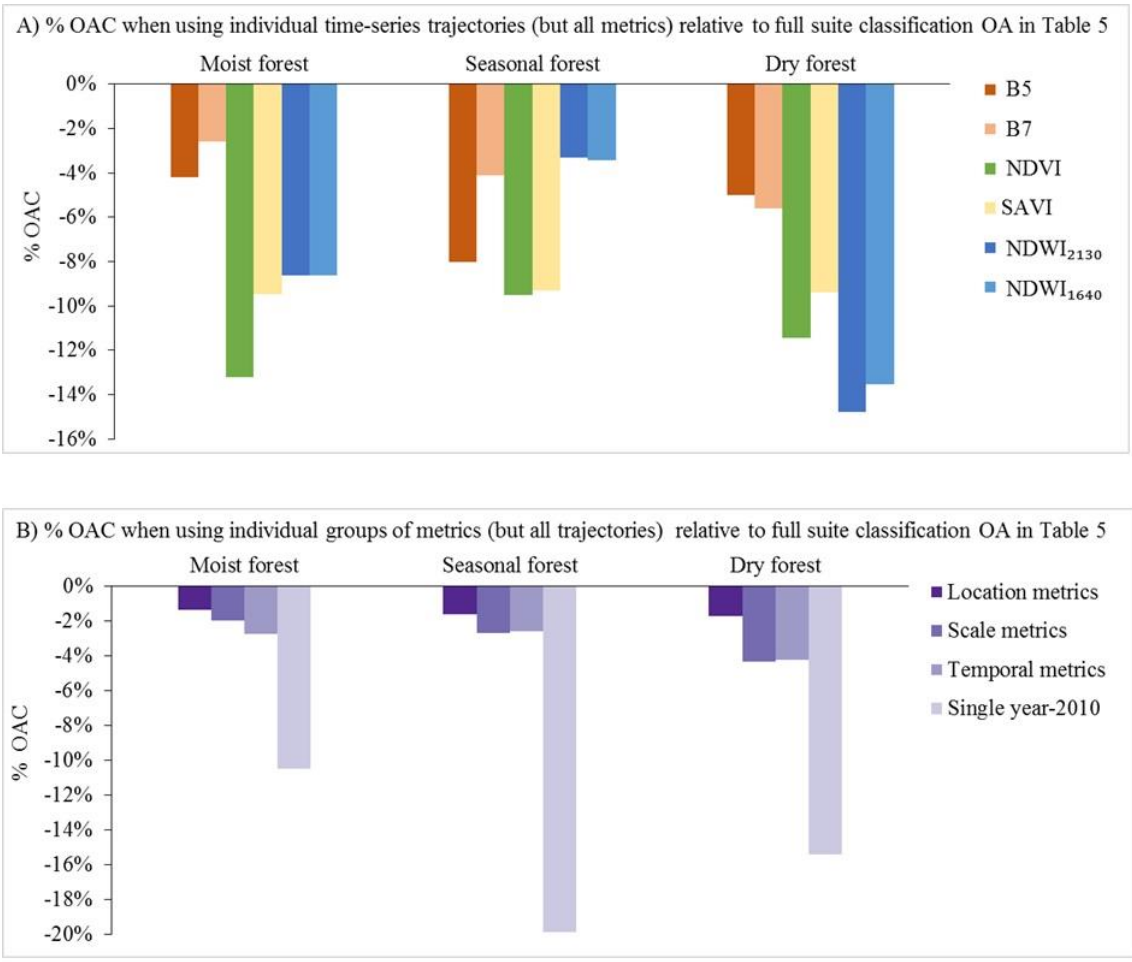


Figure 3.8 The percentage of overall accuracy change (% OAC) when running our classification procedure for individual trajectories separately (but using all four groups of trajectory metrics) or separately for individual groups of trajectory metrics (but using all six trajectories) relative to our full suite classification with all trajectories/metrics included (Table 3.3-3.5).

The larger the absolute % OAC, the less important the particular trajectory (or the group of trajectory metrics) is.

3.4.5 Comparing with other products

We compared our classification of disturbed forests in Mato Grosso with other relevant products which have recently become available (Figure 3.9). These include the MapBiomas land use/cover products (2000-2010) and the Latin American secondary forest map recently produced by Chazdon et al. (2016). The latter was derived from the map of Neotropical forest aboveground biomass of Baccini et al. (2012) for 2008. To ensure comparability in time, we only compared disturbed forests from our classification against the area of secondary forests <

24 years old from Chazdon et al. (2016). To compare against MapBiomas products (2000-2010), we reclassified open forest, degraded forest, secondary forest, and flooded forest categories from MapBiomas-2010 map into one disturbed forest class. Areas classified as non-dense forest in 2000-2009 MapBiomass products but classified as dense forest in 2010 were also considered as disturbed forests.

Our estimate of disturbed forest area in Mato Grosso was three times larger than disturbed forests from MapBiomas with corresponding spatial distribution shown in Figure 3.9 (A&B). The biggest classification differences was located in moist forest ecoregion, followed by seasonal forest and dry forest. The difference relative to MapBiomas may be due to the use of different classification methods (single date based classification) and the limited time period (2000-2010) for MapBiomas. However, secondary forest area estimates from Chazdon *et al.* (2016) were approximately three times greater than the disturbed area from our classification (Figure 3.9C), increasing to four times greater in the dry forest biome. This may be due to the coarse resolution (500 m) of forest age map, the misclassification of some anthropogenic land use areas as forest or to errors arising from interpreting the age from the forest biomass map (Chazdon et al., 2016).

The large discrepancies of estimated disturbed forests among those products highlight the importance of using high-resolution time-series images and the consideration of historical disturbances when mapping secondary forest regrowth and forest degradation. By excluding pre-2000 historical disturbances and ignoring time-series spectral characteristics, MapBiomas significantly underestimate the area of disturbed forests (Figure 3.9B), and correspondingly may underestimate the impacts of disturbance on tropical biodiversity and carbon cycles.

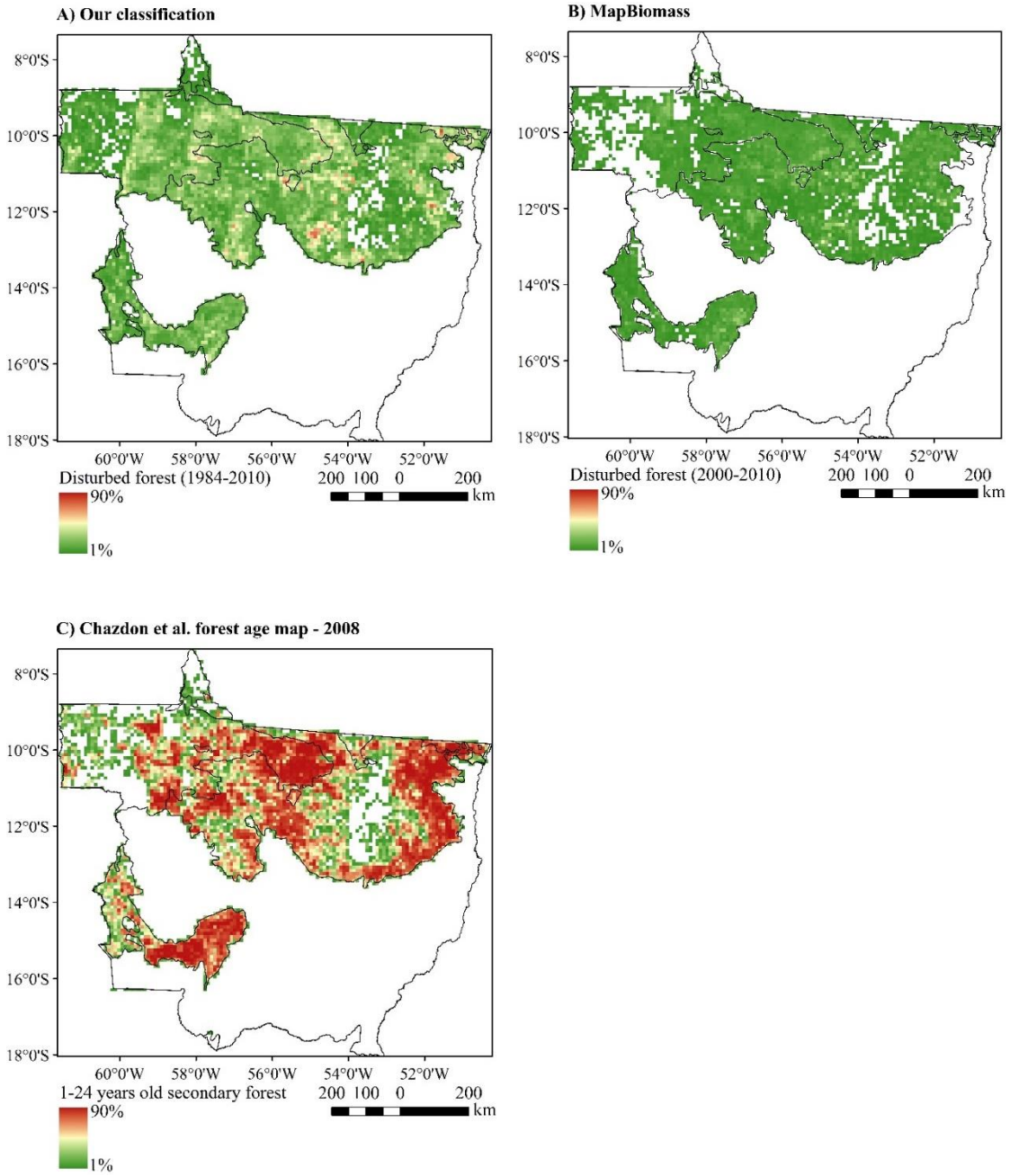


Figure 3.9 Comparison of our classification with MapBiomas land use/cover 2000-2010, and Chadzon et al. 2008 secondary forest age map.

Values represent the percentage of the area of disturbed forests within each grid cell (10*10km). White areas (within study area) represent no disturbed pixels were identified within that grid cell. The disturbed areas are 75285 km^2, 24577 km^2, 246829 km^2 for figure panel A, B, C, respectively.

3.5 Discussion

In this study, we developed a new time-series approach in GEE to map disturbed forests (both forest degradation and post-deforestation regrowth) and intact forests. This approach incorporates random forest machine learning algorithm with multiple Landsat time-series trajectories, which enhances classification power by harnessing differential sensitivities of different time-series. It is flexible with respect to the disturbance patterns it captures. It detects three different disturbances trends (Figure 3.3): 1) single disturbance – time-series have a decrease then increase pattern; 2) multiple disturbances – time-series have multiple increase and decrease signatures pattern; 3) recovery on previous disturbed areas – time-series only have an increase pattern. For example, in this study, it not only maps areas that disturbed and recovering during time-series period (1984-2010), but also captures areas that disturbed before 1984 but following a recovery process after 1984, making our approach more valuable and suitable for distinguishing disturbed forests and intact forests.

Application of our approach in moist/seasonal/dry ecoregions in Mato Gross resulted in high overall classification accuracy, ranging from 81.3% to 86.1% across ecoregions. On one hand, the misclassification of disturbed forests as intact forests may relate to the fast recovery process of secondary regrowth forests whose structural and spectral characteristics could be similar to intact forests after 20-40 years recovery (Aide et al., 2000; Poorter et al., 2016). The degraded old-growth forests recover at even faster rates. For example, it has been shown that about 50% of the canopy opening caused by selective logging becomes closed within one year of regrowth (Asner et al., 2004), making it harder to capture such quick recovery process from remote sensing perspectives. On the other hand, the misclassification of intact forests as disturbed might be because of our sampling of intact forests points which may still include few disturbed old-growth forests, as TerraClass does not map degraded forests. Furthermore, the variation of classification accuracy across ecoregions might be due to the differences of land-use history, land use intensity, severity of disturbance events, soil fertility and texture (Chazdon, 2003) and water

availability (Poorter et al., 2016), which are highly associated with post-disturbance recovery processes and the structure of regrowth forests.

By separating disturbed forests into post-deforestation regrowth forests and degraded forests, we found that approximately two-thirds of disturbed forests were degraded forests, highlighting the importance of effective systems for detecting these. Forest monitoring system should not only focus on clear-cut forest deforestation and recovery, but also degraded forests which may release more than double the amount of carbon than released by deforestation (Baccini et al., 2017). Interestingly, our classification clearly captured straight-line patterns of disturbed forests, which also present a consistent agreement with both PALSAR HV backscatter intensity change and RapidEye very high resolution images (Figure 3.6). Further development of our methodology may provide new opportunities to map selective logging activities at a large regional scale.

The methodology developed in this study dramatically exploits the power of multiple long-term Landsat time-series in the discrimination of disturbed vs. intact forests with support of GEE's massive storage and calculation capability. Unlike previously published single time-series trajectory based approaches (e.g. LandTrendr, VCT, VeRDET) (Cohen et al., 2017), this approach incorporates six different time-series trajectories which generates a much higher classification accuracy than single-trajectory based classification (Figure 3.8A). Also, this approach integrates single year features with scale, location and temporal characteristics derived from time-series trajectories, which significantly enhanced the discriminatory power. Single year features were found to be the least powerful (up to 20% less) for discriminating disturbed pixels compared to the combined use of single year features and other time-series features (Figure 3.8B). Thus, combination of single year and time-series features represents a significant advance on widespread single-year approaches to map previously disturbed forests.

3.6 Conclusion

Our study explored the feasibility of using multiple long time-series Landsat surface reflectance data to map tropical historically disturbed forests as far back as 1984. Using a case study of Mato Grosso moist, seasonal and dry forests, we found that this methodology has high potential in mapping various forested land cover types related to disturbances with an overall accuracy of up to 86.1%. The classification approach developed in this study is capable of capturing not only forest regrowth from forest deforestation (clear-cut), but also forest degradation (partially cut) due to selective logging or other small scale disturbances. Based on TerraClass-2010 forest mask, until 2010, 41% dry forest in Mato Grosso were disturbed, with 28% and 20% of seasonal forest and moist forest disturbed, respectively. By comparing classification from this study with TerraClass-2010 land cover map, we found that up to 18% of area classified as old-growth forest in TerraClass was actually degraded forests, highlighting the importance of including degradation monitoring alongside clear felling monitoring .

Our study clearly demonstrates the potential of extensive time-series of satellite imagery to map historical forest disturbances and recovery processes. More specifically, the discrimination of disturbed forests (both degraded forest and post-deforestation regrowth forest) vs. intact forests was enhanced by simultaneously combining a suite of single date features and time-series characteristics derived from multiple time series of spectral bands and vegetation indices. Our approach is readily applicable to other larger tropical areas, making pan-tropical mapping of forest disturbances and regrowth a highly tangible prospect.

References

Aguiar, A.P.D. et al. 2016. Land use change emission scenarios: anticipating a forest transition process in the Brazilian Amazon. *Global change biology.* **22**(5), pp.1821-1840.

Aide, T.M. et al. 2000. Forest regeneration in a chronosequence of tropical abandoned pastures: implications for restoration ecology. *Restoration ecology.* **8**(4), pp.328-338.

Almeida, C.A.d. et al. 2016. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. *Acta Amazonica.* **46**(3), pp.291-302.

Andersen, H.-E. et al. 2014. Monitoring selective logging in western Amazonia with repeat lidar flights. *Remote Sensing of Environment.* **151**, pp.157-165.

Asner, G.P. et al. 2004. Canopy damage and recovery after selective logging in Amazonia: field and satellite studies. *Ecological Applications.* **14**(sp4), pp.280-298.

Baccini, A. et al. 2012. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change.* **2**(3), pp.182-185.

Baccini, A. et al. 2017. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science.* **358**(6360), pp.230-234.

Barona, E. et al. 2010. The role of pasture and soybean in deforestation of the Brazilian Amazon. *Environmental Research Letters.* **5**(2).

Breiman, L. 2001. Random forests. *Machine Learning.* **45**(1), pp.5-32.

Carreiras, J.M. et al. 2014. Land use and land cover change dynamics across the Brazilian Amazon: insights from extensive time-series analysis of remote sensing data. *PloS one.* **9**(8), pe104144.

Carreiras, J.M. et al. 2017. Mapping major land cover types and retrieving the age of secondary forests in the Brazilian Amazon by combining single-date optical and radar remote sensing data. *Remote Sensing of Environment.* **194**, pp.16-32.

Chazdon, R.L. 2003. Tropical forest recovery: legacies of human impact and natural disturbances. *Perspectives in Plant Ecology, evolution and systematics.* **6**(1), pp.51-71.

Chazdon, R.L. et al. 2016. Carbon sequestration potential of second-growth forest regeneration in the Latin American tropics. *Science Advances.* **2**(5), pe1501639.

Chen, D.Y. et al. 2005. Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near- and short-wave infrared bands. *Remote Sensing of Environment.* **98**(2-3), pp.225-236.

Chen, J. et al. 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. *Isprs Journal of Photogrammetry and Remote Sensing.* **103**, pp.7-27.

Claverie, M. et al. 2015. Evaluation of the Landsat-5 TM and Landsat-7 ETM+ surface reflectance products. *Remote Sensing of Environment.* **169**, pp.390-403.

Cohen, W.B. et al. 2017. How similar are forest disturbance maps derived from different Landsat time series algorithms? *Forests.* **8**(4), p98.

Cohen, W.B. et al. 2018. A LandTrendr multispectral ensemble for forest disturbance detection. *Remote Sensing of Environment.* **205**, pp.131-140.

D'Agostino, R.B. 1970. Transformation to normality of the null distribution of g1. *Biometrika.* pp.679-681.

Foley, J.A. et al. 2005. Global consequences of land use. *Science.* **309**(5734), pp.570-574.

Gorelick, N. et al. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment.* **202**, pp.18-27.

Hansen, M.C. et al. 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science.* **342**(6160), pp.850-853.

Hermosilla, T. et al. 2015. Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. *Remote Sensing of Environment.* **170**, pp.121-132.

Hirschmugl, M. et al. 2017. Methods for Mapping Forest Disturbance and Degradation from Optical Earth Observation Data: a Review. *Current Forestry Reports.* **3**(1), pp.32-45.

Home, O.D.H.N. et al. 2013. LEDAPS calibration, reflectance, atmospheric correction preprocessing code, version 2.

Houghton, R. 2012. Carbon emissions and the drivers of deforestation and forest degradation in the tropics. *Current Opinion in Environmental Sustainability.* **4**(6), pp.597-603.

Huang, C. et al. 2010. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment.* **114**(1), pp.183-198.

Huete, A.R. 1988. A soil-adjusted vegetation index (SAVI). *Remote sensing of environment.* **25**(3), pp.295-309.

Hughes, M.J. et al. 2017. Patch-based forest change detection from Landsat time series. *Forests.* **8**(5), p166.

INPE. 2007. DEGRAD-Mapping of Forest Degradation in the Brazilian Amazon (2007-2016). http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/degrad. *Brazilian National Institute for Space Research.*

Joshi, N.P. et al. 2015. L-Band SAR Backscatter Related to Forest Cover, Height and Aboveground Biomass at Multiple Spatial Scales across Denmark. *Remote Sensing.* **7**(4), pp.4442-4472.

Kayastha, N. et al. 2012. Monitoring wetland change using inter-annual landsat time-series data. *Wetlands.* **32**(6), pp.1149-1162.

Kennedy, R.E. et al. 2007. Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sensing of Environment.* **110**(3), pp.370-386.

Kennedy, R.E. et al. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. *Remote Sensing of Environment.* **114**(12), pp.2897-2910.

Kimes, D.S. et al. 1999. Mapping secondary tropical forest and forest age from SPOT HRV data. *International Journal of Remote Sensing.* **20**(18), pp.3625-3640.

Kohavi, R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Ijcai*: Montreal, Canada, pp.1137-1145.

Kuplich, T.M. 2006. Classifying regenerating forest stages in Amazonia using remotely sensed images and a neural network. *Forest Ecology and Management.* **234**(1-3), pp.1-9.

Langner, A. et al. 2007. Land cover change 2002–2005 in Borneo and the role of fire derived from MODIS imagery. *Global Change Biology.* **13**(11), pp.2329-2340.

Lee, J.S. 1980. Digital Image-Enhancement and Noise Filtering by Use of Local Statistics. *Ieee Transactions on Pattern Analysis and Machine Intelligence.* **2**(2), pp.165-168.

Lee, J.S.H. et al. 2016. Detecting industrial oil palm plantations on Landsat images with Google Earth Engine. *Remote Sensing Applications: Society and Environment.* **4**, pp.219-224.

Lopes, A. et al. 1990. Adaptive Speckle Filters and Scene Heterogeneity. *Ieee Transactions on Geoscience and Remote Sensing.* **28**(6), pp.992-1000.

Lu, D. 2005. Integration of vegetation inventory data and Landsat TM image for vegetation classification in the western Brazilian Amazon. *Forest Ecology and Management.* **213**(1-3), pp.369-383.

MapBiomas. 2015. Project MapBiomas - Collection 2 of Brazilian Land Cover & Use Map Series, accessed on 15/11/2017 through the link: http://mapbiomas.org/pages/database/mapbiomas_collection.

Masek, J.G. et al. 2006. A Landsat surface reflectance dataset for North America, 1990-2000. *IEEE Geoscience and Remote Sensing Letters.* **3**(1), pp.68-72.

Maxwell, S.K. and Sylvester, K.M. 2012. Identification of “ever-cropped” land (1984–2010) using Landsat annual maximum NDVI image composites: Southwestern Kansas case study. *Remote sensing of environment.* **121**, pp.186-195.

Motohka, T. et al. 2014. Using time series PALSAR gamma nought mosaics for automatic detection of tropical deforestation: A test study in Riau, Indonesia. *Remote Sensing of Environment.* **155**, pp.79-88.

Nelson, R.F. et al. 2000. Secondary forest age and tropical forest biomass estimation using thematic mapper imagery. *Bioscience.* **50**(5), pp.419-431.

Olson, D.M. et al. 2001. Terrestrial ecoregions of the worlds: A new map of life on Earth. *Bioscience.* **51**(11), pp.933-938.

Padarian, J. et al. 2015. Using Google's cloud-based platform for digital soil mapping. *Computers & Geosciences.* **83**, pp.80-88.

Pan, Y. et al. 2011. A large and persistent carbon sink in the world's forests. *Science.* **333**(6045), pp.988-93.

Patel, N.N. et al. 2015. Multitemporal settlement and population mapping from Landsat using Google Earth Engine. *International Journal of Applied Earth Observation and Geoinformation.* **35**, pp.199-208.

Pekel, J.-F. et al. 2016. High-resolution mapping of global surface water and its long-term changes. *Nature.* **540**(7633), p418.

Poorter, L. et al. 2016. Biomass resilience of Neotropical secondary forests. *Nature.* **530**(7589), p211.

PRODES, p. 2018. Projecto Prodes: Monitoramento de Floresta Amazonica Brasileira por satelite. http://www.obt.inpe.br/prodes/index.php. *Instituto Nacional de Pesquisas Espaciais*

Schaffer, C. 1993. Selecting a Classification Method by Cross-Validation. *Machine Learning.* **13**(1), pp.135-143.

Shelestov, A. et al. 2017. Exploring Google Earth Engine Platform for Big Data Processing: Classification of Multi-Temporal Satellite Imagery for Crop Mapping. *Frontiers in Earth Science.* **5**, p17.

Skole, D. and Tucker, C. 1993. Tropical deforestation and habitat fragmentation in the Amazon: satellite data from 1978 to 1988. *Science.* **260**(5116), pp.1905-1910.

Souza, C. et al. 2003. Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models. *Remote Sensing of Environment.* **87**(4), pp.494-506.

Team, P. 2017. Planet application program interface: In Space for life on earth. *San Francisco, CA.*

Trisasongko, B.H. 2010. The use of polarimetric SAR data for forest disturbance monitoring. *Sensing and Imaging: An International Journal.* **11**(1), pp.1-13.

USGS. 2018. PRODUCT GUIDE. LANDSAT 4-7 SURFACE REFLECTANCE (LEDAPS) PRODUCT. *United States Geographical Survey, Department of the Interior.*

van der Werf, G.R. et al. 2009. CO2 emissions from forest loss. *Nature Geoscience.* **2**(11), pp.737-738.

Vieira, I.C.G. et al. 2003. Classifying successional forests using Landsat spectral properties and ecological characteristics in eastern Amazonia. *Remote Sensing of Environment.* **87**(4), pp.470-481.

White, J.C. et al. 2017. A nationwide annual characterization of 25 years of forest disturbance and recovery for Canada using Landsat time series. *Remote Sensing of Environment.* **194**, pp.303-321.

Xiong, J. et al. 2017. Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. *ISPRS Journal of Photogrammetry and Remote Sensing.* **126**, pp.225-244.

Chapter 4
Historical Degradation of the Brazilian Amazon (Paper Ⅲ)

Wang, Y., Ziv, G., Adami, M., and Galbraith, D., 2019. Historical Degradation of the Brazilian Amazon. This chapter is being prepared for submission to either *Geophysical Research Letters* or *Environmental Research Letters*.

Abstract

The 5-fold decline of deforestation in the Brazilian Amazon between 2004 and 2012 is a modern-day environmental success story. Since 1988, Brazil's PRODES system has monitored Amazonian deforestation and provided annual estimates of forest loss in the region. In PRODES, patches are considered to be deforested if they have been clear-felled and their total area is at least 6.25 ha. However, these estimates do not consider forest areas that have been degraded, i.e. that have not been totally clear-felled or where forest loss is below the minimum area threshold for deforestation. Indeed, information on degradation in Amazonia is largely restricted to the very recent period (post 2007) or to localised areas. Here, we use multi-decadal Landsat time-series images to provide a comprehensive assessment of historical old growth forest degradation across the Brazilian Amazon over 30 years (1984-2014). Our classification algorithm explicitly considered spatial variation in spectral characteristics across the region. Our classification resulted in very accurate detection of degradation arising from fire, road constructions and small-scale clearings. Our results show that, until 2014, over 246,845 km^2 area of old-growth forests in the Brazilian Amazon (moist forest ecoregion) were degraded, equivalent to an annual degradation area of 8,228 km^2 yr^{-1}. The cumulative area of degradation in the Brazilian Amazon accounted for approximately 10% of total area of old growth forests, 59% of which occurred within a distance of 500 m from the forest edge. Degradation patterns varied by state but were highest in Tocantins, where ~55% of old growth forests

experienced some form of historical degradation. Although these numbers are considerable, they are still likely to represent a conservative estimate, as our approach did not detect low intensity degradation linked to selective logging or defaunation and does not include areas that had been degraded prior to 1984.

4.1 Introduction

Covering an area of 5.5 million km^2, Amazon rainforests are a vitally important component of the Earth System. They are home to one-quarter of global biodiversity (Malhi et al., 2009; Dirzo and Raven, 2003), providing a host of goods and services to society, and have acted as a strong carbon sink over decades (Brienen et al., 2015; Phillips et al., 2008). Brazil contains approximately 64% of remaining Amazonian forests (RAISG, 2012) and is responsible for the majority of Amazon forest loss (Kalamandeen et al., 2018). Deforestation in the Brazilian Amazon has reduced considerably over the last 15 years, falling from 28,000 km^2 yr^{-1} in 2004 to a mean rate of 6,000 km^2 yr^{-1} between 2012 and 2017 (PRODES, 2018), due largely to the success of the PPCDAM project (Maia et al., 2011; Assunção et al., 2013), which introduced new detection and monitoring systems designed to curb forest loss.

National estimates of deforestation based on PRODES consider only patches of forest that have been completely clear-felled and that have attained a minimum size threshold of 6.25 ha. They do not include forests loss below that threshold or forest areas that have been *degraded* but not completely clear-felled. The term 'forest degradation' encompasses degradation arising from multiple drivers, representing a gradient of disturbance intensity. It includes heavy disturbance associated with road construction and mining, as well as fire damage which can span a broad range of intensity from mega-fires to small-scale burnings. Normal forest management such as thinning and harvest (selective logging) were not considered as degradation in this study (de Cambio Climático, 2003). Forest degradation need not necessarily be restricted to anthropogenic drivers but may also include natural disturbance associated with windstorm events, for example (Foley et al., 2007; Espírito-Santo et al., 2014; Espírito-Santo et al., 2010).

Since the early 2000's several studies have sought to quantify forest degradation in Amazonia. However, these have been incomplete as they have focused on studying one driver of degradation in isolation (e.g. logging), have been restricted to short timeframes or have omitted smaller degradation events. For example,

INPE's DETER B system, based on CBERS-4 and AWiFS data has provided information on degradation of primary forests in Amazonia for patches > 3 ha in size since 2015 (Diniz et al., 2015). Prior to this the DEGRAD mapped degradation from 2007 (INPE, 2007-2016) and DETEX detected selective logging from 2009 (INPE, 2009-2015), but only for a short time period. However, despite their undoubted value, these systems provide only a snapshot of degradation in Amazonia as they do not consider forests degraded prior to 2007 or patches below the designated area thresholds. Other studies have used higher-resolution Landsat data but these have also been limited in scope. For example, Asner et al. (2005) used the CLAS algorithm, based on linear spectral mixture models, to map the extent of selective logging over a three-year period (2000-2002), while Souza-Junior et al. (2014) used an analogous approach, also based on Landsat data, to map fire and logging degradation from 2000-2010. More recently, other studies mapping degradation have also been undertaken but these have largely been limited in spatial scale (Hethcoat et al., 2019; Hasan et al., 2019; Bullock et al., 2018). Thus, a fully comprehensive evaluation of the extent of forest degradation in Amazonia, still remains elusive.

Quantifying forest degradation is critically important, given its importance for carbon and biodiversity storage. Forest degradation can result in considerable losses of carbon, the magnitude and permanence of which depends on the underlying nature and intensity of the degradation. For example, low-impact logging without fire has been shown to result in the loss of 4-21% of aboveground carbon, while multiple burning can lead to losses of up to 94% of their aboveground carbon (Longo et al., 2016). Another large field study in eastern Amazonia found that on average, forests that experienced both logging and understory fires stored 40% less aboveground carbon than undisturbed forests (Berenguer et al., 2014). A recent pan-tropical study, based on the comparison of two remote-sensing derived biomass maps, concluded that tropical forests acted as a significant net carbon source between 2003 and 2014, with 69% of the carbon losses due to the degradation (Baccini et al., 2017). However, this was deduced from biomass maps, with no actual determination of degradation occurrence per se. Forest degradation can also markedly affect biodiversity

storage. In a seminal study, Barlow et al. (2016) showed that forest degradation from fire and logging could double the biodiversity loss arising from deforestation at a landscape-level. Thus, an accurate quantification of degradation extent is critical for accurate projections of the impact of anthropogenic change on forest biodiversity.

The emergence of Google Earth Engine as a processing platform that can exploit the power of cloud computing to simultaneously process multiple remote sensing data streams has revolutionised the study of land use change. In Chapter 3, I developed a new GEE-based algorithm whereby a random forest classifier, trained with TERRACLASS data, took multiple remote-sensing time series as inputs to classify forests as either disturbed or undisturbed in the state of Mato Grosso (Wang et al., 2019). Here we extend this approach to the entire Brazilian Amazon (moist forest ecoregion, Dinerstein et al., 2017), using the same multi-decadal 30 m Landsat time-series images (1984-2014) as in Chapter 3 and applying the same algorithm to classify moist primary forests (i.e. those which have not been deforested according to PRODES) as intact vs. degraded in 2014. Our approach captures degradation events which occurred at any point in the 1984-2014 time series in order to address our focal question for this chapter: How much of the extant primary forest in Amazonia has previously been degraded?

4.2 Data and Methods

Our approach to mapping historical degradation in this study involved two main steps: 1) classifying forests as either disturbed or intact using multi-decadal Landsat time-series, following the method developed by Wang et al. (2019) (Chapter 3) and 2) applying the PRODES old growth forest mask to the classification from the first step, thus eliminating previously deforested areas from our analysis. We refine this approach in this chapter to map degradation across the moist forest ecoregion in the Brazilian Amazon domain (Figure 4.1). The moist forest ecoregion follows the RESOLVE global terrestrial ecoregions dataset updated in 2017 (Dinerstein et al., 2017, http://ecoregions2017.appspot.com/), where the dry forests, floodplain, and savanna are excluded in this study.

4.2.1 Landsat time-series trajectories

We combined Landsat 5, 7 and 8 surface reflectance datasets to generate annual time-series trajectories from 1984 to 2014. Imagery from Landsat-5 Thematic Mapper (TM) sensor and Landsat-7 Enhanced Thematic Mapper Plus (ETM+) sensor are compatible in their spectral characteristics for time-series studies (Claverie et al., 2015). However, reflective wavelength differences between Landsat-7 ETM+ and Landsat-8 Operational Land Manager (OLI) mean that transformations are required to ensure temporal continuity between both sensors. All Landsat-8 spectral reflectance and vegetation indices were calibrated using the transformation functions from Roy et al. (2016). Appendix Figure C.1 shows the details of Landsat dataset used in this study.

Using Landsat 5-8 datasets, we generated six different time-series trajectories, including Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index ($NDWI_{2130}$, $NDWI_{1640}$), Soil-Adjusted Vegetation Index (SAVI), and short-wave spectral bands (Band 5 and band 7 for Landsat-5/7, Band 6 and Band 7 for Landsat-8) (Wang et al., 2019). For each of the six time-series trajectories, eleven metrics were computed which describe statistical and temporal features of the time series (minimum, maximum, range, mean, standard deviation, coefficient of variation, kurtosis, skewness, full time-series slope, maximum slope of 5-year intervals, and 2014 values) (Wang et al., 2019). In total, 66 metrics (six time-series trajectories $\times$ 11 metrics) were used as inputs for random forest classification of intact and disturbed forests.

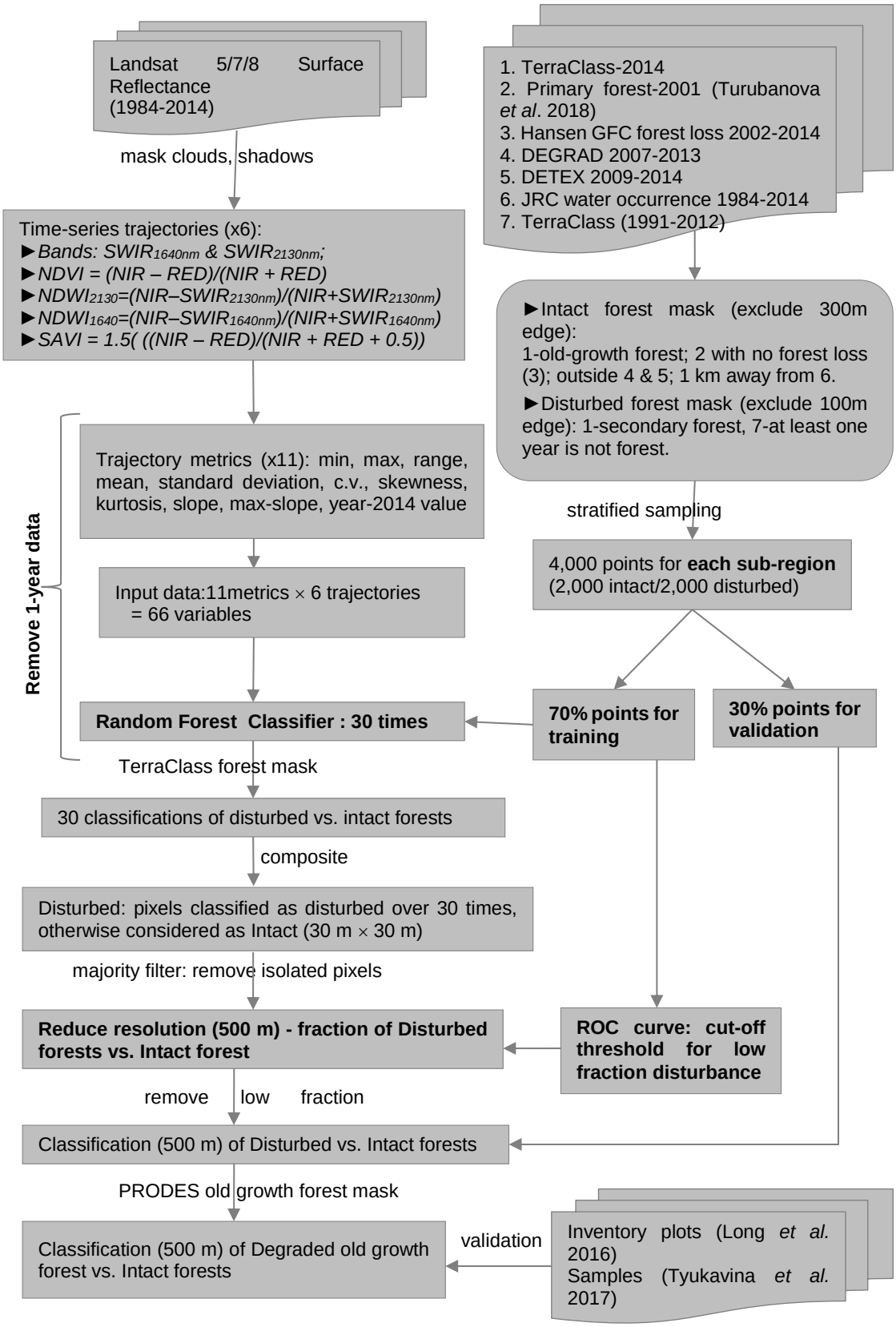


Figure 4.1 Methodological flow chart.
The key elements different from Chapter 3 were highlighted as bold.

4.2.2 Training and validation dataset

We used stratified random sampling algorithm to generate a set of spatially representative points of intact and disturbed forests, 70% of which used to train the random forest classification with 30% remaining points as for validation. For this exercise, intact forest pixels were required to satisfy 5 criteria: (1) defined as old growth forest in TERRACLASS-2014 (Almeida et al., 2016, https://www.terraclass.gov.br); (2) defined in (Turubanova et al., 2018) as primary forest, with no subsequent forest loss until 2014 reported in the University of Maryland Global Forest Cover product (Hansen et al., 2013); (3) be located at least 1 km away from any water body over 1984-2014 (Pekel et al., 2016); (4) not have been designated as degraded since 2007 by the INPE-based DEGRAD (INPE, 2007) and DETEX (INPE, 2009) products; (5) be located at least 300 m away from the edge of any pixels that meet above all criteria (1)-(4), this is to remove the boundary pixels. Disturbed pixels used for training and validation were required to: (1) be classified as secondary forest in TERRACLASS 2014; (2) at least one year be not forest in TERRACLASS 1991, 2000, 2004, 2008, 2010, 2012; (3) be located at least 100 m away from the edge of any pixels that meet above all criteria (1)-(2).

To account for spatial variation in spectral reflectance across the Brazilian Amazon, we divided our study area (moist forest ecoregion) into 34 sub-regions (Figure 4.2) for running the following classification algorithm. Each sub-region was further divided into 9 smaller grids ($1^o \times 1^o$) for stratified sampling to ensure a spatially even distribution of sampled points (Figure 4.2). Within each $1^o \times 1^o$ grid, 500 pixels (30×30 m) were sampled individually for each stratum (i.e. intact forest, disturbed forest). In total, 4,500 intact and 4,500 disturbed pixels were sampled for each individual sub-region, of which 2,000 intact and 2,000 disturbed pixels were subsequently randomly selected as final training and validation dataset. We used 70% of the final training and validation dataset to train the classifier and the remaining 30% served as validation points.

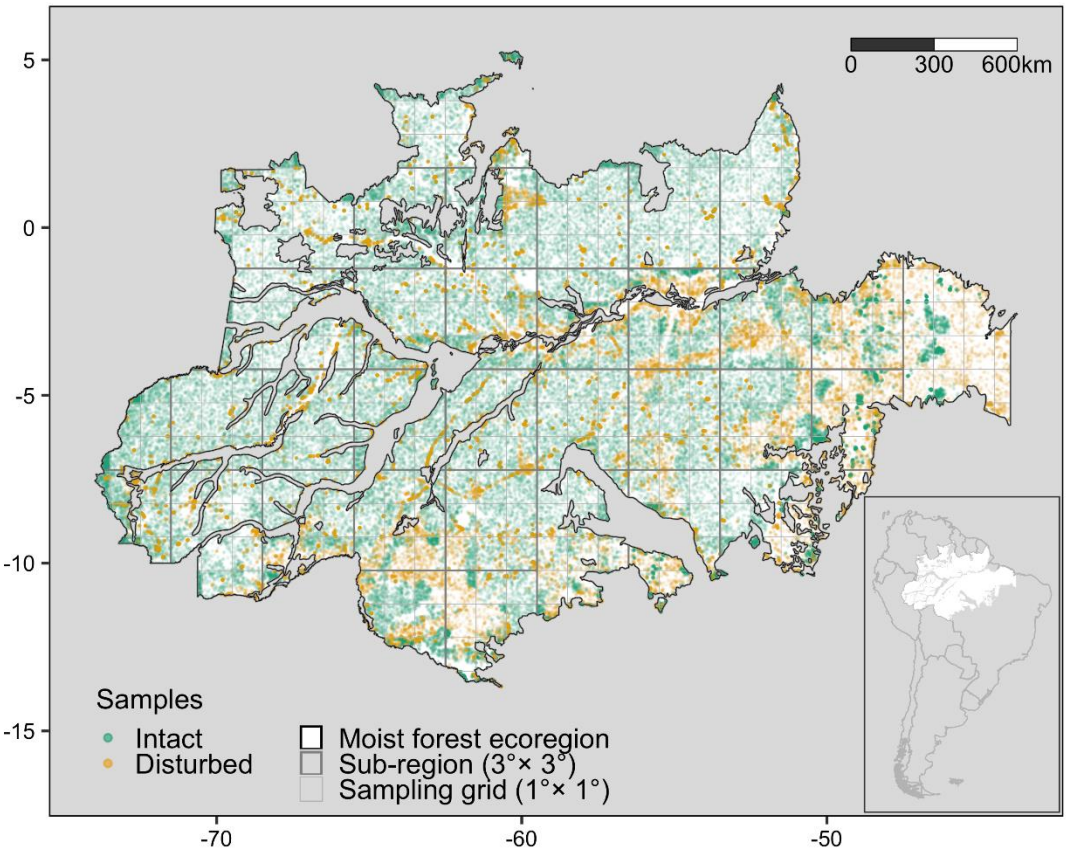


Figure 4.2 Study area.
Sub-regions for classification algorithm, sampling grids, and the distribution of sampled points as training and validation for random forest classifier.

4.2.3 Classification of intact vs. disturbed forests

We used random forest classification to distinguish between disturbed forests and intact forests based on the forest mask (primary and secondary forests) from TERRACLASS-2014. The random forest classifier was run individually for each sub-region, using 200 trees and a bagging fraction of 0.632. Due to the high frequency of cloud occurrence in the Amazon, clouds and cloud shadow effects still exist for certain years even after cloud masking and compositing processes. To reduce the influence of clouds on our classification, we ran the random forest classifier 30 times for each sub-region, removing one year of data at a time from 1984 to 2013 in each iteration. If the pixel was classified as ‘disturbed’ across all 30 classification iterations, the pixel was considered ‘disturbed’. A post-

classification noise filter (3×3 pixels, majority rule) was subsequently applied to remove single isolated pixels.

4.2.4 Classification aggregation

Considering the inherent edge effects of forest disturbances (Harper et al., 2005; dos Santos et al., 2015) and for comparability with other products, we aggregated the original classification (30×30 m) into a coarser classification map (500×500 m) and reclassified the aggregated pixel as being either disturbed or intact. To do this, we computed the fraction of the larger grid cell occupied by disturbed 30×30 m pixels as classified by our algorithm, and related these to a sub-region specific disturbance cut-off threshold, below which the aggregated pixel was classified as intact and above which it was classified as disturbed. The cut-off threshold was computed using ROC (receiver operating characteristic) curves and was found to vary considerably across individual sub-regions, ranging from 0.02 to 0.16 (Appendix Figure C.3). ROC curves were constructed based on the training points used to train the original classifier (~1400 intact points and 1400 disturbed points in each sub-region), by plotting the true positive rate (TPR, also known as sensitivity) against the false positive rate (FPR, fall-out) at various threshold settings. The true positive rate (TPR) is the probability of intact points classified as ‘intact’. The false positive rate (FP) is the probability of disturbed points classified as ‘intact’. The best cut-off is % disturbance value where the true positive rates are highest and false positive rates are lowest. The final classification map (500×500 m) of intact forests and disturbed forests was validated using the remaining 30% of previously sampled pixels which were not used for training the machine learning classification algorithm.

4.2.5 Mapping degradation and validation

Our algorithm detects both secondary forests re-growing on previously clear-felled land and degraded old-growth forests, not classed as deforested by PRODES. To quantify degradation in old-growth forests, including both natural and anthropogenic disturbances except for selective logging (see Introduction on

page 102 for the definition in detail), we applied the 2014 primary (old growth) mask from PRODES, which excludes all deforested areas identified up to that year.

To validate actual degraded old-growth forests, we used forest status (intact and degraded) derived from 153 inventory plots across the Brazilian Amazon published by Longo et al. (2016). Of 153 inventory plots, 85 were intact forests and 68 had been degraded by burning that occurred from 1987 to 2013.

Additionally, to test the ability of our algorithm to capture different types of degradation, we compared our results to the results of 283 visually interpreted samples indicating various forms of degradation, from Tyukavina et al. (2017). These samples were visually interpreted by two experts using annual Landsat composite images during 1990-2013 and, when available, high-resolution imagery from Google Earth (Tyukavina et al., 2017). The samples cover ten different types of degradation events including both anthropogenic and natural disturbances. Of 283 samples, the majority experienced small-scale clearing (117), fire (61) or logging (72), and the remaining 33 samples were related to road construction, wind blowdowns, river flooding and other disturbances. The spatial distribution of the inventory plots and the samples are shown in Appendix Figure C.2.

4.3 Results

4.3.1 Overall disturbance classification accuracy

Our localised classification algorithm yielded very high overall accuracies, in relation to the validation datasets outlined in section 4.2.2, across all 34 sub-regions. The overall accuracies were all over 0.94, with the kappa statistics ranging from 0.88 to 0.99 (Figure 4.3). The producer's accuracies of both intact forests and disturbed forests were high, with ranges of 0.95-1.00 and 0.92-1.0, respectively, same as the user's accuracies.

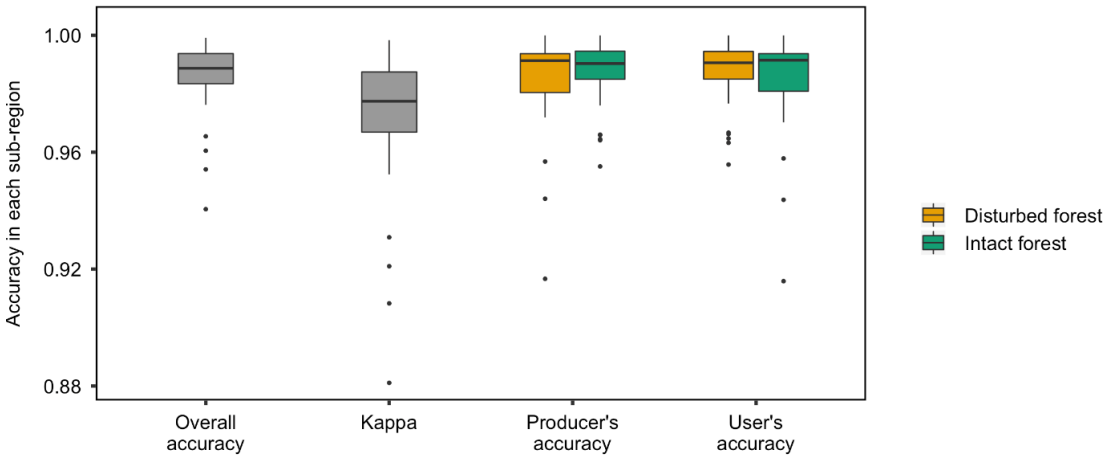


Figure 4.3 The accuracy of the classification of intact forests vs. disturbed forests. Across all sub-regions, the overall accuracy: 0.9857; kappa: 0.9713; producer's accuracy: 0.9845 (disturbed forest), 0.9868 (intact forest); user's accuracy: 0.9867 (disturbed forest), 0.9846 (intact forest). Validation data for the disturbed forest stratum were based on TERRACLASS secondary forests (Section 4.2.2: Training and validation dataset).

4.3.2 Accuracy of degraded old growth forests vs. intact forests

As our overall disturbance accuracy assessment presented above is restricted to secondary forests and does not specifically encompass degraded old-growth forests, we further validated our classification of degraded old growth forests and intact forests using the available inventory plots (Longo et al., 2016) and visual interpretation samples (Tyukavina et al., 2017). We find that all intact forest inventory plots according to Longo et al. (2016) were classified as intact forests, thus our producer's accuracy for intact forests was 100% (Table 4.1). Furthermore, our classification algorithm effectively detected degradation resulting from fire (over 73% accuracy), and from combined fire and conventional logging (74%) (Table 4.1). When compared to degradation data from Tyukavina et al. (2017), we also find that our algorithm captured more intense degradation arising from anthropogenic activities such as road construction (70%), mining (100%), dam constructions (100%), and small-scale clearings (80%) (Table 4.1). However, our algorithm performed less well at detecting natural disturbances such as wind blowdowns (33%) or river flooding (50%) and could not adequately detect logging, especially reduced impact logging. Thus, our classification

algorithm captured the more important and more intense degradation types very well but not lower intensity degradation.

Table 4.1 Validation of intact forests and degraded old growth forests, and the type of degradation.

Inventory Plots (*Longo et al.* 2016)	Intact forest (map)	Degraded (map)	Total plots	Producer's accuracy
Intact forest	85	0	85	1.00
Burning	11	30	41	0.73
Conventional logging & Burning	7	20	27	0.74
Samples (*Tyukavina et al.* 2017)	Intact forest (map)	Degraded (map)	Total samples	Producer's accuracy
Burning	19	42	61	0.69
Logging	54	18	72	0.25
Small-scale clearing	23	94	117	0.80
Mining	0	1	1	1.00
Road	3	7	10	0.70
Dam	0	2	2	1.00
Wind	8	4	12	0.33
River	0	3	6	0.50
Other nature disturbances	0	2	2	1.00

4.3.3 Distribution of old growth forest degradation.

Although degradation of old-growth forests is widespread across the entire Brazilian Amazon (Figure 4.4), the majority of degraded areas occur along rivers and near deforested areas. In the eastern Brazilian Amazon, large areas of old growth forest are heavily degraded with little intact forest remaining. Small focal areas of degradation can be seen in the north and northwest of the Amazon, where deforestation rates are much lower than in the southern and eastern sections of the Brazilian Amazon. These may potentially result from large wind blow-down and high frequency convective storms that are more marked in this region than in others (Araujo et al., 2017; Negron-Juarez et al., 2018; Espírito-Santo et al., 2014).

Our results (Table 4.2) show that within the moist forest ecoregion of the Brazilian Amazon, about 246,845 km^2 of old growth forests were degraded between 1984 and 2014, with the annual degradation rate of 8,228 km^2 yr^{-1}. The state of Pará contributed up to 39% of the total area of degradation, following by the state of Amazonas (22%). However, when compared with the total area of old growth forests, degraded forests in Pará only accounted for 13% of total old growth forest area, and only 4.56% for the state of Amazonas. Tocantins experienced the most severe relative degradation, and was the only state where the area of degraded old-growth forest exceeded the area of intact forests. Other states with large relative areas of degraded primary forest include Maranhão (40%), Rondônia (21%) and Mato Grosso (18%).

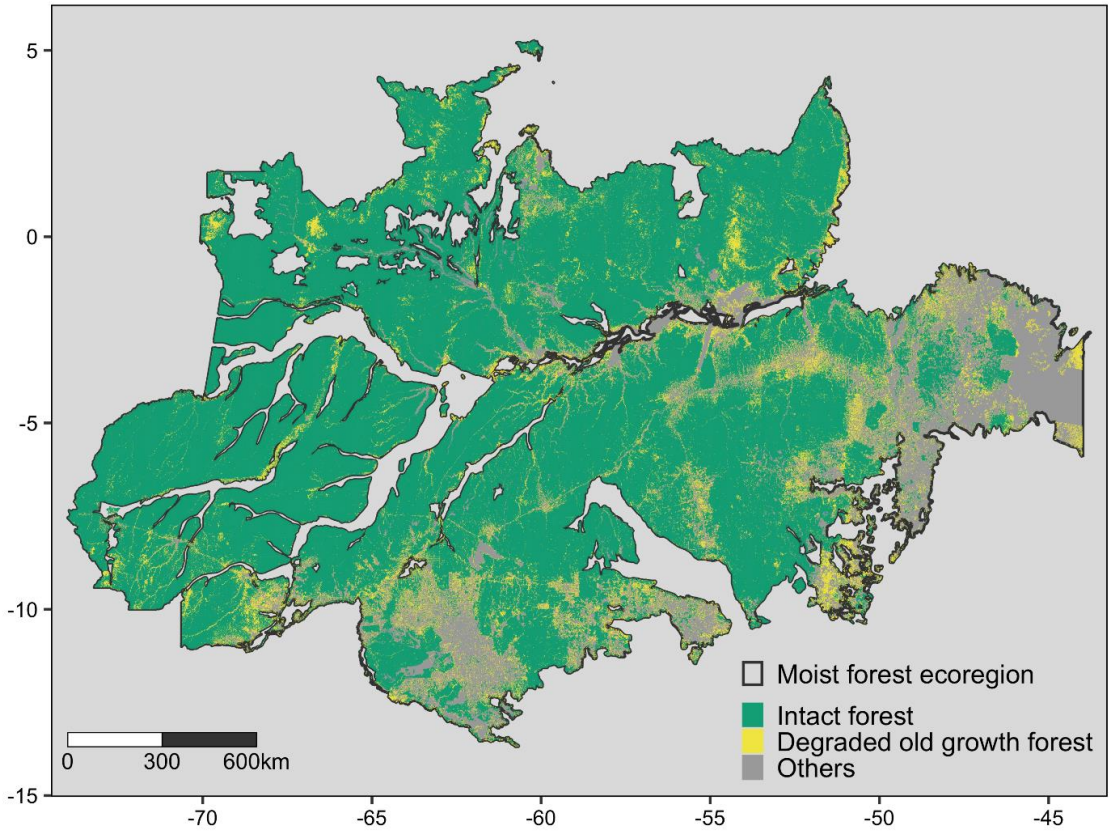


Figure 4.4 Classification (500 m × 500 m) of degraded old growth forests vs. intact forest for the year 2014 in the Brazilian moist forest ecoregion.
Degraded old growth forests represent cumulative area of degradation over the period of 1984-2014. Others include secondary forests and non-forest areas.

Table 4.2 The area (km^2) of intact forest and degraded old growth forests in Brazilian Amazon (moist forest ecoregion) in 2014.
The degraded old growth forests account for the cumulative degradation occurred from 1984 to 2014.

States	Intact forest (A)	Degraded forest (B)	B/(A+B)	B/sum(B)	Degradation (yr^{-1})	Degradation within 500m to the edge
Tocantins	578.04	715.24	55.30%	0.29%	23.84	93%
Maranhão	13,812.24	9,113.06	39.75%	3.69%	303.77	84%
Rondônia	89,189.30	24,326.25	21.43%	9.85%	810.88	82%
Mato Grosso	93,638.13	21,228.90	18.48%	8.60%	707.63	73%
Acre	103,259.43	18,095.75	14.91%	7.33%	603.19	45%
Pará	664,901.90	96,020.37	12.62%	38.90%	3,200.38	63%
Roraima	116,233.28	14,625.91	11.18%	5.93%	487.53	33%
Amapa	94,223.47	7,548.03	7.42%	3.06%	251.60	43%
Amazonas	1,154,159.01	55,171.15	4.56%	22.35%	1,839.04	47%
Brazilian Amazon	2,329,994.80	246,844.67	9.58%		8,228.16	59%

4.3.4 Linkage between degradation and deforestation.

We find that deforestation and degradation are generally closely linked. As seen in Figure 4.4, there is a clear spatial association between the intensity of degradation and deforestation. Relative degradation is also highest in the states that have lost more of their old-growth forest cover (Figure 4.5). Further insights into the anthropogenic contribution to degradation can be gleaned by considering the distance of a degradation event to a forest edge, which in most cases is a deforested patch, but can also be a natural edge such as a river. We find that across the Brazilian Amazon, over 80% of our mapped degradation pixels were within 2.8 km distance to the forest edge (Figure 4.6) and that 59% of degradation took place within 500 m of the forest edge. Moreover, the proportion of the degradation occurring within 500 m of the forest edge increased in line with the state-level old-growth forest degradation status. For example, in the heavily degraded states of Tocantins and Maranhão, 93% and 84% respectively of degradation occurred within 500 m of the forest edge, while in states with low relative degradation such as Amapá and Amazonas, <50% of total degradation

occurred within 50% of the forest edge, implying a potentially greater contribution of natural disturbances in driving degradation in these states.

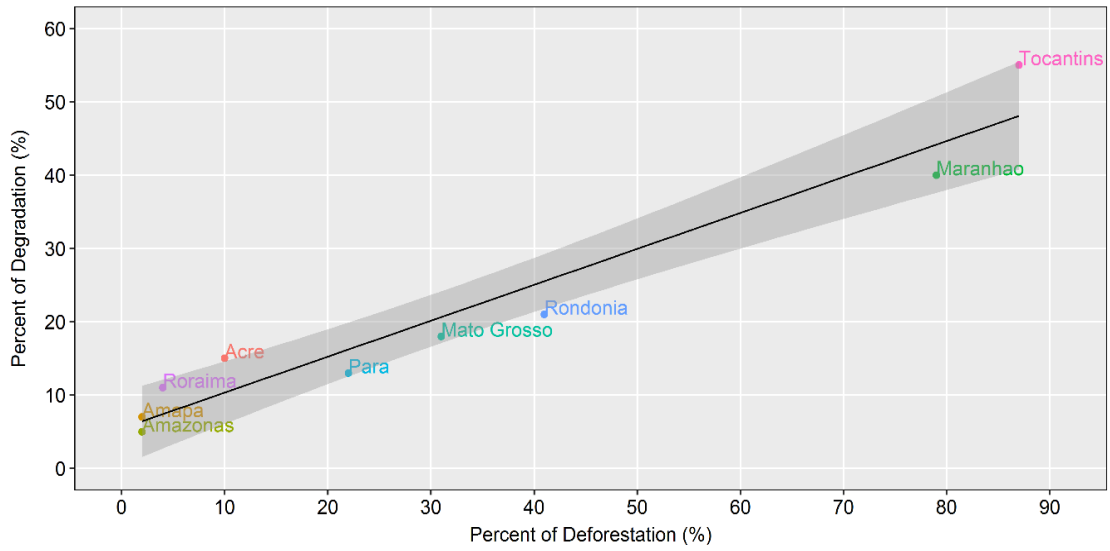


Figure 4.5 The relationship between our mapped degradation and deforestation by states. The degradation (%) is the proportion of total degradation relative to the area of old growth forests (Table 4.2, column 'B/(A+B)'). The deforestation (%) represents the proportion of total deforestation until 2014, computed as '(total deforested area) / (total deforested area + area of old growth forest)'. The shaded grey indicates OLS (ordinary least squares) linear regression with the line as best fit (adjusted R2 = 0.9264).

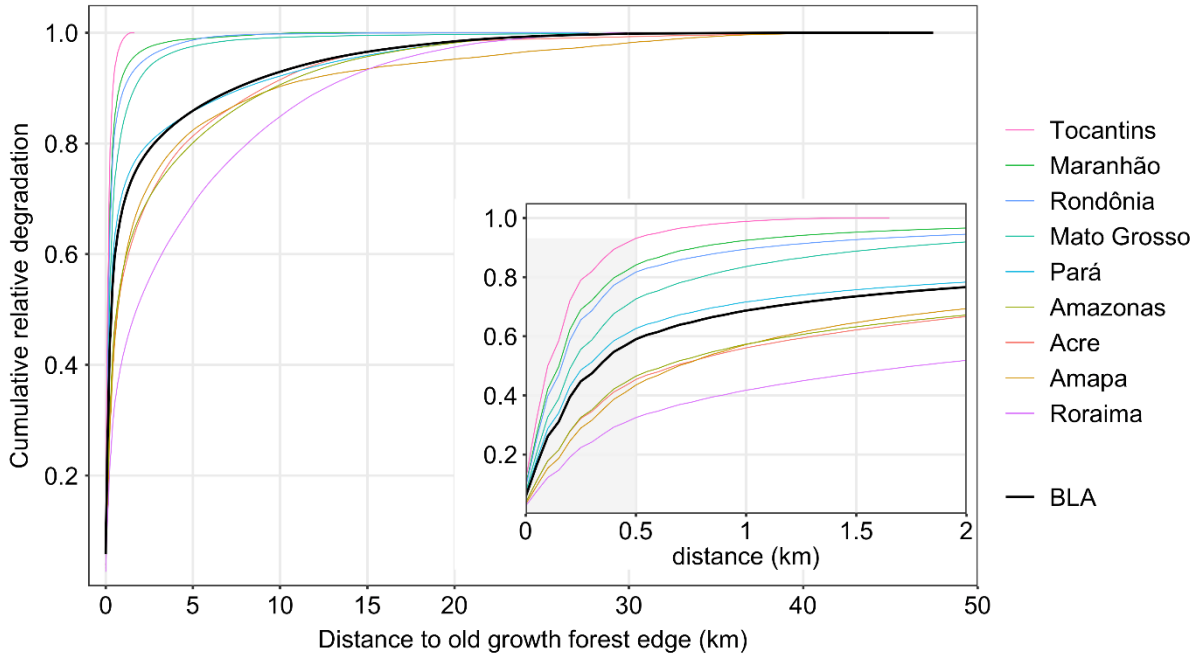


Figure 4.6 The distance of degraded old growth forest to old growth forest edge. In the Brazilian Amazon moist forest ecoregion (BLA), 59% of the degraded areas were within 500 m distance to the forest edge.

4.4 Conclusion and Discussion

This study is the first to consider cumulative historical degradation over a multi-decadal timeframe (i.e. 31 years, 1984-2014) in the Brazilian Amazon. Over this timeframe, we find that 246,845 km^2 (9.58%) of old-growth moist forest in the Brazilian Amazon experienced some form of degradation. This amounts to an annual degradation rate of 8,228 $km^2\ yr^{-1}$ over the timeframe of our study. This estimate is comparable with the sum area of degradation from DEGRAD (2007-2014, 5835 $km^2\ yr^{-1}$) and selective logging from DETEX (2009-2014, 4153 $km^2\ yr^{-1}$), for the same domain as our study area. This estimate is also within the range of degradation rates reported in other studies. For example, Souza Jr et al. (2013) report degradation rates of 5,081 $km^2\ yr^{-1}$ between 2001-2010, Nepstad et al. (1999) estimate degradation rates of 10,000-15,000 $km^2\ yr^{-1}$ in the 1990's. In the state of Rondônia, we estimated an average degradation rate of 811 km^2

yr^{-1} which is also comparable to estimates (< 1000 km^2 yr^{-1}) from a recent study (Bullock et al., 2018).

We also spatially compared our classification of degraded old growth and intact forests (Figure 4.4) with an aboveground carbon density loss map from Baccini et al. (2017) (Figure 4.6). For comparison, we applied the same old growth forest mask to the carbon density loss map. In their study, Baccini et al. (2017) considered any forest that lost biomass to be degraded. Assuming this definition of degradation, over 479,000 km^2 area of old growth forests were degraded during 2003-2014. This is equivalent to an annual degradation rate of 43,553 km^2 yr^{-1}, approximately 4.5 times larger than our estimates. Moreover, there are considerable spatial differences between our map and the map derived from Baccini et al. (2017). Most notably, the Baccini et al. (2017) map reports large degradation in areas of the Northern Amazon with little history of deforestation (e.g. northwest Amazonas), which our classification suggests is not the case.

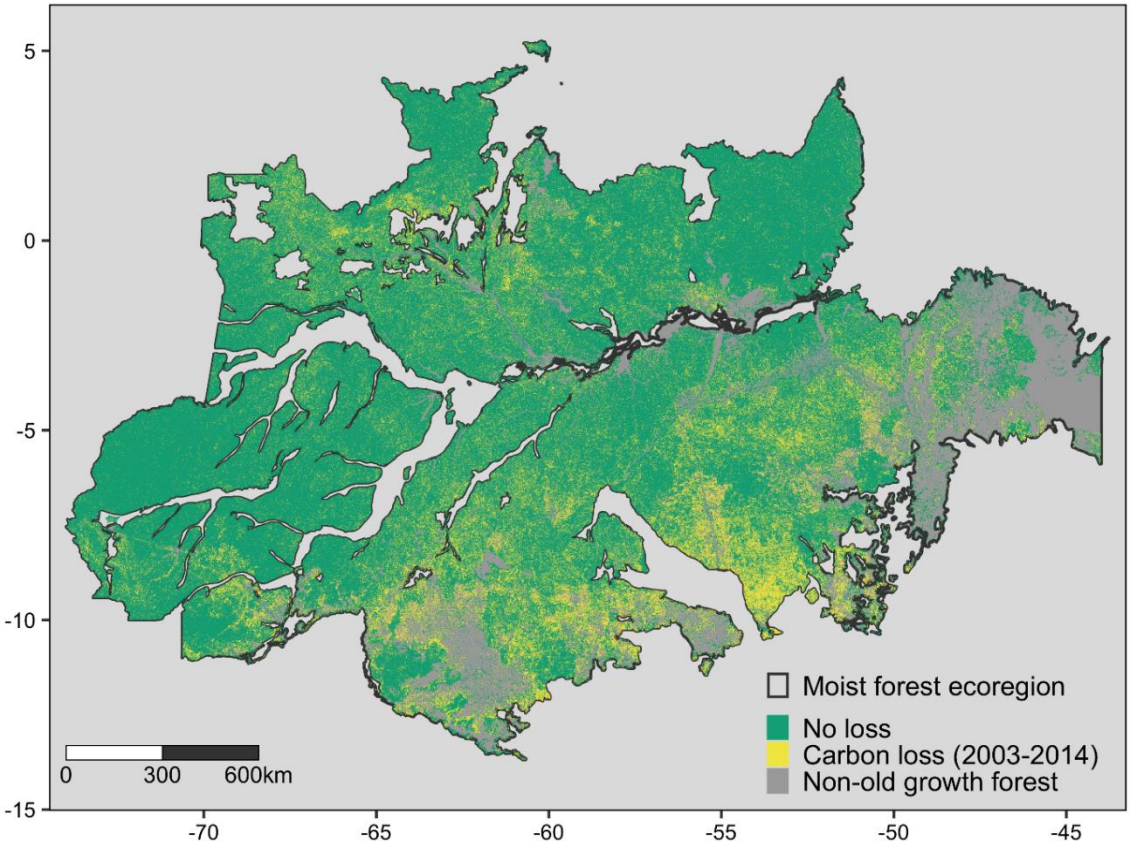


Figure 4.7 Carbon density loss (Baccini et al., 2017) of old growth forests in the Brazilian Amazon.
Between 2003 and 2014, approximately 479,081 km^2 area of old growth forests had carbon loss, which was equivalent to the annual loss area of 43,553 km^2 yr^{-1}.

Despite being in broad agreement with other reported degradation rates in the Brazilian Amazon, there are several reasons to suggest that our figures are conservative and underestimate the full extent of historical degradation in the Brazilian Amazon. First, our algorithm does not consider old-growth forest degradation prior to 1984. Second, although our classification algorithm effectively detected degradation resulting from more intensive anthropogenic disturbances (e.g. fire, road construction, small-scale clearing) with high accuracies, it did not capture low-intensity degradation such as that associated with selective logging. This form of degradation is believed to be extensive across the Amazon. For example, Asner et al. (2005) reported logged areas in the Brazilian Amazon of between 12,000-19,000 km^2 yr^{-1}. Third, our algorithm was such that the random forest classifier was run 30 times, removing one year

of data from time-series trajectories in each iteration. The final ‘disturbed’ pixels were pixels that were identified as ‘disturbed’ by all the 30 random forest classifiers. Although this approach minimise background noise, it potentially neglects disturbances which are followed by very fast recovery, i.e. within two years.

Our study provides a unique historical perspective on degradation in Amazonia. However, it does not provide information in trends in degradation over time, as our algorithm was not designed to produce annual estimates. A future next step therefore is to study the temporal evolution of degradation in the Amazon and the extent to which this relates to trends in deforestation. Over the timeframe of our study, approximately 445,000 km^2 of forest in the Brazilian Amazon was clear-felled (including drier forests excluded from our analysis). Our study reveals a very close relationship between deforestation and degradation. However, it is not clear how this coupling has changed over time. While deforestation rates in the Amazon have fallen over time, the temporal pattern of degradation is unknown. However, to fully characterize degradation in the Brazilian Amazon, we will have to develop our algorithm to better detect selective logging. This may be possible through linear spectral mixture models, as used in the current version of DETER-B (Diniz et al., 2015), which does have the capacity to capture logging activities.

References

Almeida, C.A.d. et al. 2016. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. *Acta Amazonica.* **46**(3), pp.291-302.

Araujo, R.F. et al. 2017. Regional distribution of large blowdown patches across Amazonia in 2005 caused by a single convective squall line. *Geophysical Research Letters.* **44**(15), pp.7793-7798.

Asner, G.P. et al. 2005. Selective logging in the Brazilian Amazon. *science.* **310**(5747), pp.480-482.

Assunção, J. et al. 2013. DETERring deforestation in the Brazilian Amazon: environmental monitoring and law enforcement. *Climate Policy Initiative.* pp.1-36.

Baccini, A. et al. 2017. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science.* **358**(6360), pp.230-234.

Barlow, J. et al. 2016. Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. *Nature.* **535**(7610), p144.

Berenguer, E. et al. 2014. A large-scale field assessment of carbon stocks in human-modified tropical forests. *Global change biology.* **20**(12), pp.3713-3726.

Brienen, R.J. et al. 2015. Long-term decline of the Amazon carbon sink. *Nature.* **519**(7543), p344.

Bullock, E.L. et al. 2018. Monitoring tropical forest degradation using spectral unmixing and Landsat time series analysis. *Remote Sensing of Environment.* p110968.

Claverie, M. et al. 2015. Evaluation of the Landsat-5 TM and Landsat-7 ETM+ surface reflectance products. *Remote Sensing of Environment.* **169**, pp.390-403.

de Cambio Climático, P.I. 2003. Definitions and Methodological Options to Inventory Emissions from Direct Human-Induced Degradation of Forests and Devegetation of Other Vegetation Types [Penman, J. y colaboradores (directores de la publicación)]. *The Institutefor Global Environmental Strategies (IGES), Japón.*

Dinerstein, E. et al. 2017. An ecoregion-based approach to protecting half the terrestrial realm. *BioScience.* **67**(6), pp.534-545.

Diniz, C.G. et al. 2015. DETER-B: The new Amazon near real-time deforestation detection system. *Ieee journal of selected topics in applied earth observations and remote sensing.* **8**(7), pp.3619-3628.

Dirzo, R. and Raven, P.H. 2003. Global state of biodiversity and loss. *Annual review of Environment and Resources.* **28**.

dos Santos, M.N. et al. 2015. Lidar-based assessment of forest edge effects across a degraded landscape in the Brazilian Amazon. In: *Embrapa Territorial-Artigo em anais de congresso (ALICE)*: In: CONFERENCE ON LIDAR APPLICATIONS FOR ASSESSING AND MANAGING FOREST ….

Espírito-Santo, F.D. et al. 2014. Size and frequency of natural forest disturbances and the Amazon forest carbon balance. *Nature communications.* **5**, p3434.

Espírito-Santo, F.D. et al. 2010. Storm intensity and old-growth forest disturbances in the Amazon region. *Geophysical Research Letters.* **37**(11).

Foley, J.A. et al. 2007. Amazonia revealed: forest degradation and loss of ecosystem goods and services in the Amazon Basin. *Frontiers in Ecology and the Environment.* **5**(1), pp.25-32.

Hansen, M.C. et al. 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science.* **342**(6160), pp.850-853.

Harper, K.A. et al. 2005. Edge influence on forest structure and composition in fragmented landscapes. *Conservation biology.* **19**(3), pp.768-782.

Hasan, A.F. et al. 2019. Cumulative disturbances to assess forest degradation using spectral unmixing in the north-eastern Amazon. *Applied Vegetation Science.*

Hethcoat, M.G. et al. 2019. A machine learning approach to map tropical selective logging. *Remote sensing of environment.* **221**, pp.569-582.

INPE. 2007. DEGRAD-Mapping of Forest Degradation in the Brazilian Amazon (2007-2016). http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/degrad. *Brazilian National Institute for Space Research.*

INPE. 2009. DETEX project for mapping selective logging (2009-2015) by INPE (Brazilian National Institute for Space Research) http://www.dpi.inpe.br/arquivos_pime/DETEX_BR163_Claudio.pdf.

Kalamandeen, M. et al. 2018. Pervasive rise of small-scale deforestation in Amazonia. *Scientific reports.* **8**(1), p1600.

Longo, M. et al. 2016. Aboveground biomass variability across intact and degraded forests in the Brazilian Amazon. *Global Biogeochemical Cycles.* **30**(11), pp.1639-1660.

Maia, H. et al. 2011. Avaliação do Plano de Ação para Prevenção e Controle do Desmatamento na Amazônia Legal: PPCDAm: 2007-2010.

Malhi, Y. et al. 2009. Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest. *Proceedings of the National Academy of Sciences.* **106**(49), pp.20610-20615.

Negron-Juarez, R.I. et al. 2018. Vulnerability of Amazon forests to storm-driven tree mortality. *Environmental Research Letters.* **13**(5), p054021.

Nepstad, D.C. et al. 1999. Large-scale impoverishment of Amazonian forests by logging and fire. *Nature.* **398**(6727), p505.

Pekel, J.-F. et al. 2016. High-resolution mapping of global surface water and its long-term changes. *Nature.* **540**(7633), p418.

Phillips, O.L. et al. 2008. The changing Amazon forest. *Philosophical Transactions of the Royal Society B: Biological Sciences.* **363**(1498), pp.1819-1827.

PRODES, p. 2018. Projecto Prodes: Monitoramento de Floresta Amazonica Brasileira por satelite. http://www.obt.inpe.br/prodes/index.php. *Instituto Nacional de Pesquisas Espaciais*

RAISG. 2012. Amazonia under Pressure. p10.

Roy, D.P. et al. 2016. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote sensing of Environment.* **185**, pp.57-70.

Souza Jr, C. et al. 2013. Ten-year Landsat classification of deforestation and forest degradation in the Brazilian Amazon. *Remote Sensing.* **5**(11), pp.5493-5513.

Turubanova, S. et al. 2018. Ongoing primary forest loss in Brazil, Democratic Republic of the Congo, and Indonesia. *Environmental Research Letters.* **13**(7), p074028.

Tyukavina, A. et al. 2017. Types and rates of forest disturbance in Brazilian Legal Amazon, 2000–2013. *Science advances.* **3**(4), pe1601047.

Wang, Y. et al. 2019. Mapping tropical disturbed forests using multi-decadal 30 m optical satellite imagery. *Remote sensing of environment.* **221**, pp.474-488.

Chapter 5
Conclusion and Discussion

5.1 Discussion

This thesis has conducted a comprehensive analysis of the fate of previously deforested lands in the Brazilian Amazon and historical forest degradation in the region. This work is the first to provide spatiotemporal information on the dynamics of secondary forests across the entire Brazilian Amazon and to map cumulative degradation of old-growth forests over a multi-decadal timeframe. Both of these findings contribute directly towards a better understanding of forest dynamics in Amazonia.

5.1.1 Summary of thesis findings

In Chapter 2, I used the 30-meter resolution TERRACLASS time-series dataset (i.e. 2000, 2004, 2008, 2010, 2012, 2014) to track the fate of secondary forests in the Brazilian Amazon over 14 years, providing the first estimates of secondary forest loss for the region. Secondary forest loss was quantified using a sampling-based approach, combined with visual interpretation of 4,665 points randomly sampled from TERRACLASS classification maps. This analysis revealed two distinct phases of secondary forest loss in Amazonia: 1) a marked decline in secondary forest loss between 2000-2008 in line with decline in primary forest loss and 2) a rapid increase of secondary forest loss between 2008-2014, despite stabilization of primary forest loss. Overall, the proportion of total forest loss accounted for by secondary forests rose from 37±3% in 2000 to 72±5% in 2014. This phenomenon occurred across the entire Brazilian Amazon and was not driven simply by increasing secondary forest area but instead reflects a conscious preferential shift from cutting primary forests to the clearance of secondary forests. However, the total loss of secondary forests from 2000 to 2014 has cost a carbon sequestration opportunity of 2.59-2.66 Pg C, which is equivalent approximately 18 years of Brazil's fossil fuel emissions (Bank, 2014)

In Chapter 3, with the support of Google Earth Engine (GEE) and the availability of multi-decadal Landsat imagery, I developed a novel methodology, based on integration of multiple time series of vegetation indices and reflectance data from specific Landsat bands, which is able to map forest disturbance using a machine learning classification algorithm. This approach not only correctly classified areas of secondary forest, as confirmed by TERRACLASS, but also was capable of detecting forest degradation in old-growth forests. This approach was tested in three different ecoregions (moist forest, seasonal forest and dry forest ecoregions) of Mato Grosso state with high overall accuracies of up to 86%.

In Chapter 4, I extended the above approach to a larger scale, using the same multi-decadal 30 m Landsat time-series images (1984-2014) as in Chapter 3 and applying the same algorithm to classify moist old growth forests (i.e. those which have not been deforested according to PRODES) as intact vs. degraded in 2014. The classification resulted in very accurate detection of degradation arising from fire, road construction, small-scale clearings and natural disturbances such as river flooding and wind-throws. The results show that, until 2014, over 246,845 km^2 area of old-growth forests in the Brazilian Amazon (moist forest ecoregion) were degraded, accounting for 9.58% of total area of old growth forests. Across states, the intensity of degradation was found to be very closely linked to the deforestation in the region since the relative degradation was also higher in the states that have lost more of their old-growth forest cover. Furthermore, almost 60% of degradation occurred within 500 m distance to the forest edge. This further confirmed the considerable contributions of anthropogenic drivers to the area of degradation in the Brazilian Amazon. However, further research is needed to better quantify the individual contributions of anthropogenic and natural disturbances to degradation in Amazonia.

5.1.2 Findings in a broader context of forest monitoring

My results from Chapter 2 suggest that the accelerated loss of secondary forests in the Brazilian Amazon reflects a preferential behaviour shift from the deforestation of primary forests to the clearance of secondary forests in the region. This result suggests that despite demand for new pastureland (the main fate of deforested primary and secondary forests in the Amazon), farmers chose to intensify cutting of secondary forests rather than increase the deforestation of primary forests. While this points to the effectiveness of policies designed to curtail deforestation, it also raises awareness of the need for policies to allow regeneration of forests on previously deforested lands.

However, some key gaps in our knowledge remain. First and foremost, there is currently no accurate spatial distribution of secondary forests outside of the Brazilian Amazon, as provided by TERRACLASS. However, it may be possible to use the available secondary forest information from the Brazilian Amazon to train some remote sensing based machine learning algorithms and apply them to other tropical regions where no secondary forest information available. Additionally, there would be great value in expanding the PRODES and TERRACLASS methodologies to other South American countries, so that deforestation and re-growth statistics can be produced in a consistent manner across the entire Amazon.

In Chapter 2, I developed a method that is able to use the time-series Landsat spectral information of secondary forests to train a machine learning (i.e. random forest) classification algorithm to detect the old-growth forest degradation. This method could potentially be extended to a larger context towards a pan-tropical or global level of degradation mapping, given the availability of adequate training data.

The recent published land use and land cover dataset from MAPBIOMAS (*https://mapbiomas.org/en*) contains the annual land use and land cover classification map since 1985, but the feasibility of using this dataset to quantify secondary forest loss and historical degradation is unclear. First, the land use

information for secondary forest and degradation is limited since MAPBIOMAS provides five collections and only collection 2.0, which is available since 2000, includes the land use categories of secondary forest and degradation. Second, the classification of MAPBIOMAS for each year is independent and does not consider the time-series spectral characteristics. This might not affect the classification between forest and non-forest land uses (e.g. agriculture), but would potentially underestimate secondary forest and degradation because of their rapid recovery processes. It also results in temporal inconsistencies. Thus, I believe that MAPBIOMASS dataset would definitely contribute the understanding of land use and land cover changes in Amazonia, but, in its current configuration, probably not for the dynamics of secondary forest and historical degradation.

Currently, DETER-B (INPE-CRA, regional centre Amazon) provides real-time deforestation and degradation alerts for the Brazilian Environment Agency (Ibama) to work upon. It detects both clear-cut deforestation and degradation arising from fire, selective logging, mining and other disturbances, and has acted as an effective forest monitoring system for the region (Assunção et al., 2013). However, DETER-B is based on the visual interpretation of 56 m resolution satellite images, thus is limited for identifying the small-scale disturbances (<1 ha), and it's less cost-effective comparing with the machine learning classification algorithms. However, the spectral mixture modelling which DETER-B is based upon is capable of detecting selective logging. Future lines of work for further developing the algorithm developed for this thesis might be to 1) include image fractions (e.g. soil and vegetation) in our classification algorithm to better map lower-intensity disturbances and 2) revise the algorithm to enable annual detection of degradation and thus track trends in degradation over time.

For secondary forests monitoring, TERRACLASS is currently the only project in Brazil that provides the information on the fate of previously deforested lands, which includes secondary forests. So far, TERRACLASS has produced the land use/land cover classification of post-primary forest deforestation areas over 24 years (i.e. 1990, 2000, 2004, 2008-2014 bi-annually). However, TERRACLASS

involves a huge effort based largely on the visual interpretation of satellite imagery, and there is still no near-real time monitoring system for secondary forests as DETER-B only detect the deforestation and degradation occurred in primary forests. Given existing TERRACLASS products, it may be possible to train a classifier in Google Earth Engine that can produce annual estimates of secondary forest loss for a fraction of the total effort that goes into producing existing TERRACLASS products. My thesis also contributes directly to TERRACLASS efforts by providing a robust classification error analysis of TERRACLASS, which the product had not been subjected to thus far.

5.1.3 The implication of findings

The accelerated loss of secondary forests and historical degradation of old growth forests reported in this thesis have significant implications.

First, our results have direct implications for policy commitments that the Brazilian government has agreed to. Brazil has committed to restore 120,000 km^2 of forest land by 2030 as part of its Nationally Determined Contribution (NDC) for the Paris Agreement (UNFCCC, 2015). A cost-effective way to achieve this would be to allow part of its existing Amazonian secondary forest area (235,718 ± 7,773 km^2) to recover naturally. However, over 180,329 $\pm$ 11,760 km^2 of secondary forests were cut over a 14-year period (2000-2014). If the accelerated loss of secondary forests continues, meeting the NDC goal would be a challenge. My findings suggest that an appropriate monitoring and management system for secondary forests in the region is necessary.

Moreover, primary forest from PRODES has been heavily degraded. This thesis estimated that, until 2014, approximately 10% of the old growth forests (i.e. primary forest in PRODES) in the Brazilian Amazon were actually degraded, with an average of 8,228 km^2 affected annually. Although still conservative, these estimates highlight the further extent to which forests in the Brazilian Amazon are affected by human activity beyond deforestation. Many forests reported as degraded actually go on to be deforested. Thus, much of the degradation can in

fact be considered as ‘committed’ deforestation, after clearance coalesce into the minimum patch size threshold considered by PRODES.

5.1.4 Advances of future remote sensing prospects

Remote-sensing based machine learning is very useful to map forest disturbances, but the lack of available ground-truth data has hampered its further application over larger areas. However, visual interpretation could potentially be used as a substitute for ground-true data, especially with the development of interpretation tools such as TimeSync, Collect Earth (CE) and Google Earth. The recent released web-based visual validation tool - Collect Earth Online (CEO) is easier and more user-friendly than Collect Earth desktop version, which also has linked to all the remote sensing data from Google Earth Engine (GEE) and the Very High Resolution (VHR) from Digital Globe and Bing Aerial. VHR images and the enhanced web-based technology have opened up new possibilities for the role of visual interpretation in forest observation (Schepaschenko et al., 2019).

Landsat imagery has become freely accessible since 2008 and considerably improved the science and operational applications (Zhu et al., 2019). With the new data becoming free available from Senntinel-2 (10-60 m resolution), the combination of Landsat-8, Sentinel-2A, and Sentinel-2B provides a global median average revisit interval of 2.9 days (Li and Roy, 2017). This opens up the new possibility of mapping highly dynamic forest disturbances especially low intensity selective logging and very small-scale clearings. The coming launch of Landsat-9 and Landsat-10 (Wulder et al., 2019) will further expand the advantages of using remote sensing technology for earth observations.

Unlike optical satellite images, remote sensing using laser Lidar or Radar has the advantage of being unaffected by the presence of clouds, but has been limited by the small areas of data coverage. However, the new released and free available GEDI Lidar data (Global Ecosystem Dynamics Investigation project) which provides global precise measurements of the 3D structure of forest canopy will greatly advance our ability to characterize forest disturbances, carbon cycling and biodiversity dynamics. And from 2022, a radar satellite (BIOMASS, satellite

planned for launch by the European Space Agency) will also begin providing 3D data on forest structure and forest biomass at the global level.

The cloud-based geospatial analysis platform - Google Earth Engine (GEE) has made the planetary-scale remote sensing analysis much easier and faster (Gorelick et al., 2017) (). GEE consists of a multi-petabyte analysis-ready data catalog housing a large repository of publicly available geospatial datasets which includes observations from a variety of satellites and environmental and climate variables. All the analysis conducted in this thesis were through GEE. Using GEE allows me to process all the 30-m original TERRACLASS classification data across entire Brazilian Amazon, and map the forest disturbances through machine learning and 31-years Landsat images in the region. Now with the new added higher resolution (10 m) Sentinel-2 images and radar data in GEE, monitoring land use and land cover change through GEE would be a highly promising prospect.

5.2 Conclusion

Besides deforestation of primary forests in the Amazon, the clearance of secondary forests and old growth forest degradation have posed additional challenges to the region, which have been much less researched. The accelerated increase of secondary forest loss in the Brazilian Amazon has considerably exceeded deforestation of primary forest since 2008, accounting for approximately (72 $\pm$ 5) % of the total forest loss between 2012-2014. From 2000 to 2014, about 180,329 $\pm$ 11,760 km^2 area of secondary forests was cleared, causing a lost carbon sequestration opportunity of up to 2.66 Pg C which is equivalent to 18 years of Brazil's fossil fuel emissions. Degradation of old growth forests has contributed another 8,228 km^2 yr^{-1} area of damage to the region, resulting in up to 10% of the old growth forests being degraded until 2014. Our work provides baseline numbers for the formal inclusion of these processes in estimates of the carbon balance of the Brazilian Amazon and for conservation and land use planning.

Despite the accelerated secondary forest loss and considerable forest degradation, until 2014, the Brazilian Amazon still has in excess of 235,718 ± 7,773 km^2 of secondary forests and over 2.3 million km^2 of intact forests. Manging these resources sustainably so as to maximise and maintain their conservation value is crucial to the local, regional and global ecosystems. This will entail ensuring the continued functionality of efforts to curb deforestation of primary forests (PRODES, DETER-B) but also a coordinated system for monitoring and managing secondary forests.

References

Assunção, J. et al. 2013. DETERring deforestation in the Brazilian Amazon: environmental monitoring and law enforcement. *Climate Policy Initiative.* pp.1-36.

Bank, W. 2014. CO^2 emissions (metric tons per capita).

Gorelick, N. et al. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment.* **202**, pp.18-27.

Li, J. and Roy, D. 2017. A global analysis of Sentinel-2A, Sentinel-2B and Landsat-8 data revisit intervals and implications for terrestrial monitoring. *Remote Sensing.* **9**(9), p902.

Schepaschenko, D. et al. 2019. Recent advances in forest observation with visual interpretation of very high-resolution imagery. *Surveys in Geophysics.* pp.1-24.

UNFCCC. 2015. Nationally Determined Contributions (NDCs).

Wulder, M.A. et al. 2019. Current status of Landsat program, science, and applications. *Remote sensing of environment.* **225**, pp.127-147.

Zhu, Z. et al. 2019. Benefits of the free and open Landsat data policy. *Remote Sensing of Environment.* **224**, pp.382-385.

Appendix A

Supplementary information for Chapter 2 (Paper I)

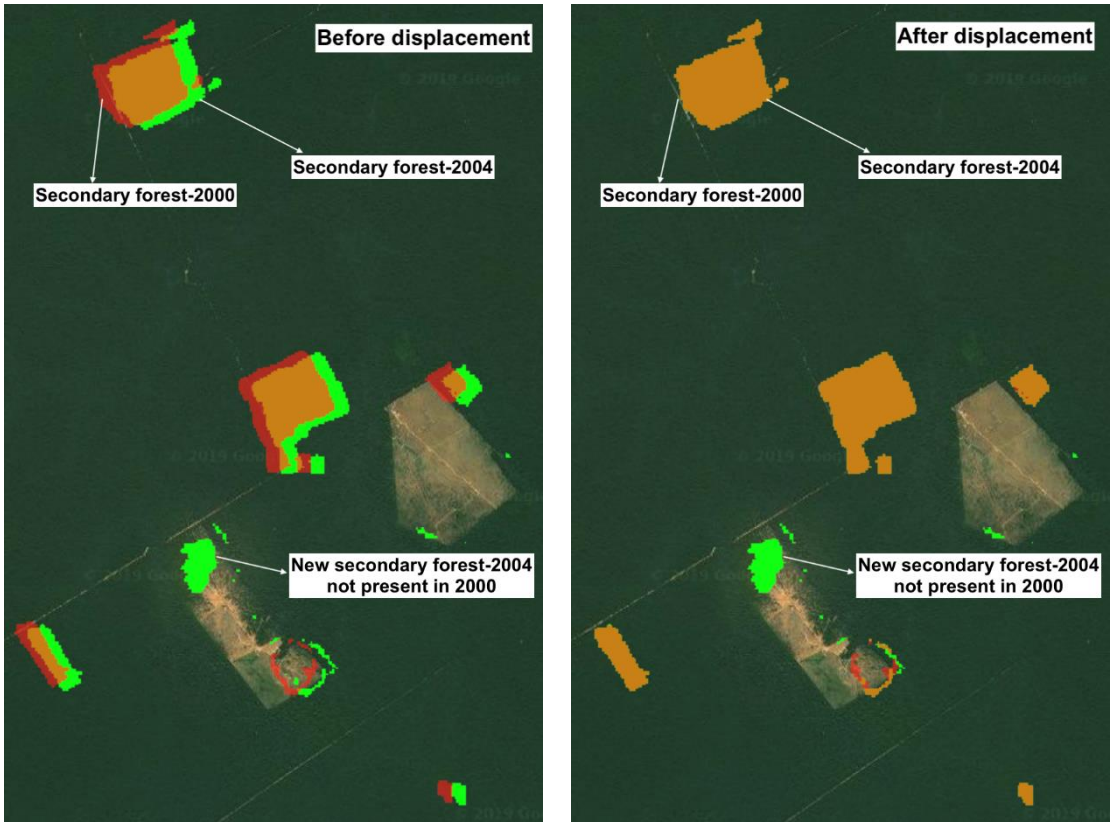


Figure A.1 Example of misalignment between TERRACLASS-2000 and 2004 and comparison of before and after application of displacement algorithm.

The left panel demonstrates the misalignment between TERRACLASS-2000 and subsequent TERRACLASS years (2004-2014) and right panel demonstrates the correction of the misalignment following application of the 'displacement' algorithm in Google Earth Engine to register TERRACLASS-2000 to TERRACLASS-2004.

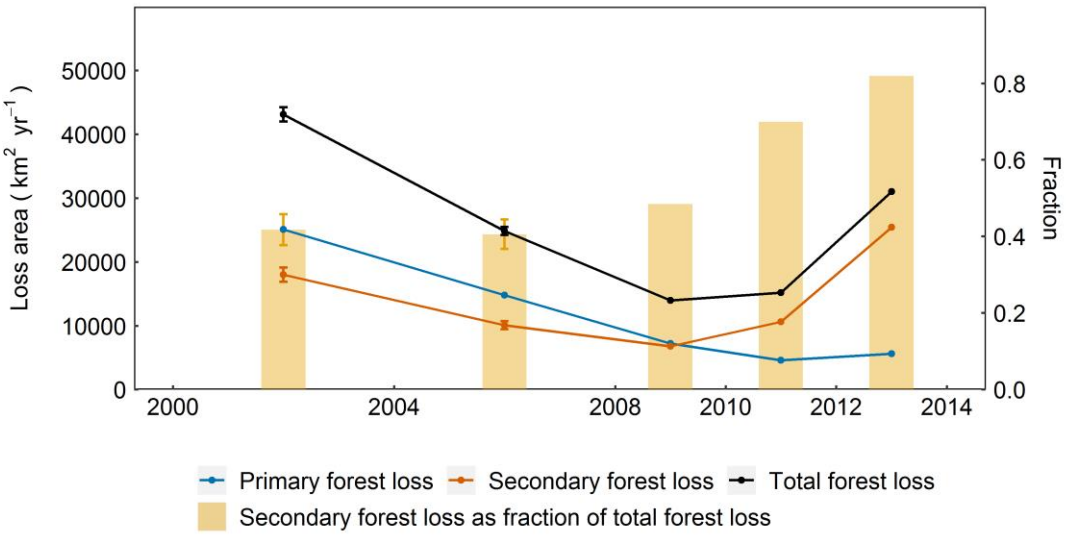


Figure A.2 Map-based estimates of annual primary and secondary forest loss in the Brazilian Amazon from 2000-2014.

Total forest loss is the sum of primary and secondary forest loss. Time-interval corrections were applied to account for missed secondary forest loss in 4-year intervals (i.e. 2000-2004, 2004-2008). Error bars for the first two intervals denote the error associated with the interval correction.

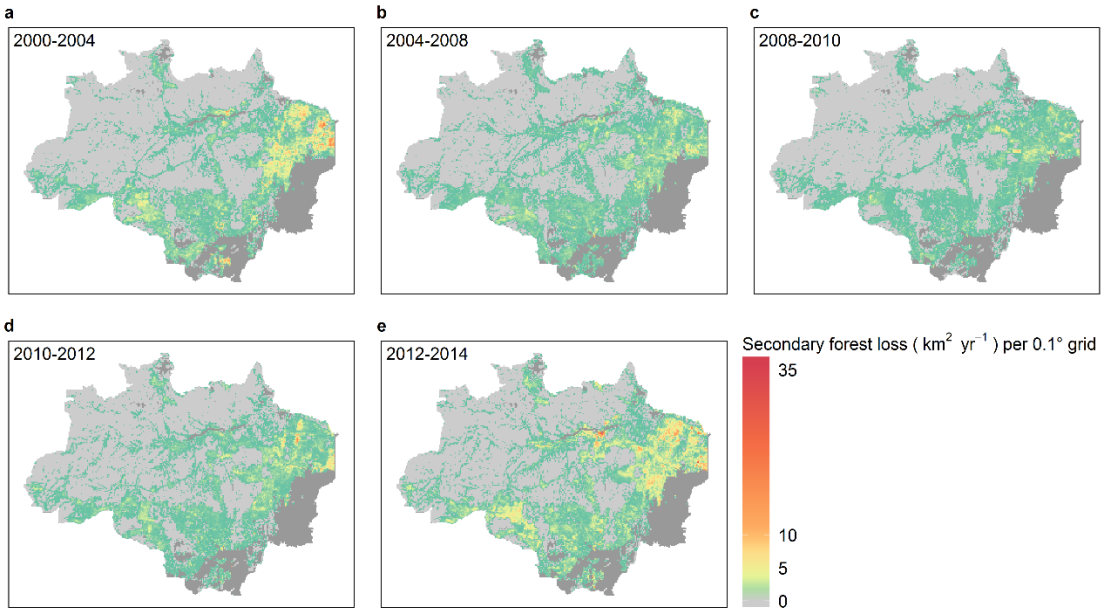


Figure A.3 Spatial distribution of the area of secondary forest loss in the Brazilian Amazon from 2000 to 2014.

Lighter grey represents no secondary forest loss. Darker grey represents non-forest areas (e.g. savannas). Time interval corrections were applied in the first two intervals (i.e. 2000-2004, 2004-2008).

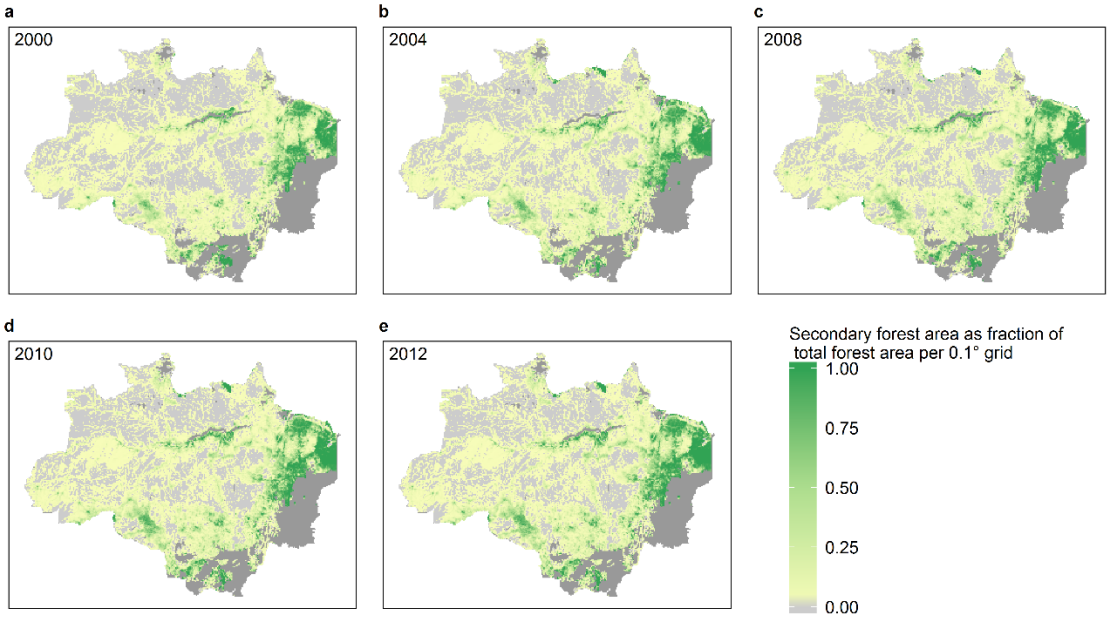


Figure A.4 Distribution of secondary forest area as fraction of total forest area. Lighter grey represents areas that have no secondary forests. Darker grey represents non-forest areas (e.g. savannas).

Table A.1 Strata names, strata weights (w_i) and sample sizes (n_i) used for sample-based estimation of forest loss rates.

In total, 933 pixels (30 m) were sampled for each time interval. For the two-year intervals (i.e. 2008-2010, 2010-2012, 2012-2014), 50 and 75 points were allocated to the smaller strata(Olofsson et al., 2014; Arévalo et al., 2019), and the remaining sample points were allocated proportionally to the strata of stable primary forest and stable others. For the four-year intervals (i.e. 2000-2004, 2004-2008), because of the increase in the proportions of each strata, 75 and 100 points were allocated to the smaller strata, and the remaining sample points were allocated to the stable primary forest stratum.

Strata	2000-2004		2004-2008		2008-2010		2010-2012		2012-2014	
	w_i	n_i	w_i	n_i	w_i	n_i	w_i	n_i	w_i	n_i
Stable primary forest (outside buffer)	0.7004	433	0.7056	433	0.7624	537	0.7721	538	0.7448	538
Primary forest loss	0.0256	75	0.0151	75	0.0037	50	0.0023	50	0.0029	50
Stable secondary forest	0.0269	75	0.0341	75	0.0500	75	0.0510	75	0.0461	75
Secondary forest loss	0.0143	75	0.0080	75	0.0034	50	0.0054	50	0.0130	50
Secondary forest gain	0.0097	75	0.0156	75	0.0061	50	0.0076	50	0.0112	50
Stable Others	0.1046	100	0.1234	100	0.1367	96	0.1358	95	0.1319	95
Stable primary forest (within buffer)	0.1184	100	0.0982	100	0.0377	75	0.0257	75	0.0502	75

Table A.2 Error matrix of sample-based estimates for period 2000-2004.
18 sampled points were excluded due to being mosaic pixels for which it was not possible to determine a specific land cover or for which no clear satellite imagery available.

n_{ik}	Reference								Producer's accuracy
Mapped	Stable PF	PF loss	Stable SF	SF loss	SF gain	Stable others	n_i	w_i	
Stable PF	433	0	0	0	0	0	433	0.7004	1.00
PF loss	5	70	0	0	0	0	75	0.0256	0.93
Stable SF	0	0	71	3	0	0	74	0.0269	0.96
SF loss	0	0	11	49	0	5	65	0.0143	0.75
SF gain	0	0	29	2	38	4	73	0.0097	0.52
Stable others	0	0	3	0	0	93	96	0.1046	0.97
Buffer-(stable PF)	96	2	0	0	0	1	99	0.1184	0.97
User's accuracy	0.81	0.97	0.62	0.91	1.0	0.90			Overall accuracy: 0.8240

Error Matrix populated by estimated proportions of area using (eq. 2.3) in methods, standard error (eq. 2.4) and area estimates (eq. 2.4 – 2.5). Our total study area A = 3924375.63 km^2.

$\hat{p}_{ik}$	Reference					
Mapped	Stable PF	PF loss	Stable SF	SF loss	SF gain	Stable others
Stable PF	0.7004	0	0	0	0	0
PF loss	0.0017	0.0239	0	0	0	0
Stable SF	0	0	0.0258	0.0011	0	0
SF loss	0	0	0.0024	0.0108	0	0.0011
SF gain	0	0	0.0039	0.0003	0.0050	0.0005
Stable others	0	0	0.0033	0	0	0.1013
Buffer-(stable PF)	0.1149	0.0024	0	0	0	0.0012
Sum ($\hat{p}_{.k}$)	0.8170	0.0263	0.0354	0.0121	0.0050	0.1042
$S(\hat{p}_{.k})$	0.0022	0.0018	0.0022	0.0010	0.0006	0.0023
Estimated areas (km^2) ($\hat{A}_k \pm S(\hat{A}_k)$)	3206236.88 ±8559.74	103153.68 ±7219.74	138729.18 ±8448.97	47677.65 ±3952.95	19809.99 ±2240.58	408768.26 ±8960.23
Annual loss/gain (km^2 yr^{-1})		25788.42 ±1804.94		11919.41 ±988.24	4952.50 ±560.14	

Table A.3 Error matrix of sample-based estimates for period 2004-2008.
16 sampled points were excluded due to being mosaic pixels for which it was not possible to determine a specific land cover or for which no clear satellite imagery available.

n_{ik}	Reference								Producer's accuracy
Mapped	Stable PF	PF loss	Stable SF	SF loss	SF gain	Stable others	n_i	w_i	
Stable PF	433	0	0	0	0	0	433	0.7056	1.00
PF loss	2	70	0	0	1	2	75	0.0151	0.93
Stable SF	0	0	74	0	1	0	75	0.0341	0.99
SF loss	0	0	6	51	4	11	72	0.0080	0.71
SF gain	0	0	25	3	31	3	62	0.0156	0.50
Stable others	0	0	2	3	1	94	100	0.1234	0.94
Buffer-(stable PF)	100	0	0	0	0	0	100	0.0982	1.00
User's accuracy	0.81	1.00	0.69	0.89	0.82	0.85			Overall accuracy: 0.8211

Error Matrix populated by estimated proportions of area using (eq. 2.3) in methods, standard error (eq. 2.4) and area estimates (eq. 2.4 – 2.5). Our total study area A = 3924375.63 km^2.

$\hat{p}_{ik}$	Reference					
Mapped	Stable PF	PF loss	Stable SF	SF loss	SF gain	Stable others
Stable PF	0.7056	0	0	0	0	0
PF loss	0.0004	0.0141	0	0	0.0002	0.0004
Stable SF	0	0	0.0337	0	0.0005	0
SF loss	0	0	0.0007	0.0057	0.0004	0.0012
SF gain	0	0	0.0063	0.0008	0.0078	0.0008
Stable others	0	0	0.0025	0.0037	0.0012	0.1160
Buffer-(stable PF)	0.0982	0	0	0	0	0
Sum ($\hat{p}_{.k}$)	0.8042	0.0141	0.0431	0.0101	0.0101	0.1183
$S(\hat{p}_{.k})$	0.0003	0.0004	0.0021	0.0022	0.0017	0.0030
Estimated areas (km^2) ($\hat{A}_k \pm S(\hat{A}_k)$)	3156115.72 ±1106.56	55145.78 ±1713.30	169126.88 ±8089.54	39770.72 ±8637.66	39767.36 ±6583.01	464449.17 ±11806.66
Annual loss/gain (km^2 yr^{-1})		13786.45 ±428.32		9942.68 ±2159.41	9941.84 ±1645.75	

Table A.4 Error matrix of sample-based estimates for period 2008-2010.
10 sampled points were excluded due to being mosaic pixels for which it was not possible to determine a specific land cover or for which no clear satellite imagery available.

n_{ik}	Reference								Producer's accuracy
Mapped	Stable PF	PF loss	Stable SF	SF loss	SF gain	Stable others	n_i	w_i	
Stable PF	537	0	0	0	0	0	537	0.7624	1.00
PF loss	10	32	0	0	0	7	49	0.0037	0.65
Stable SF	0	0	72	1	0	0	73	0.0500	0.99
SF loss	0	0	10	34	2	3	49	0.0034	0.69
SF gain	0	0	5	0	33	6	44	0.0061	0.75
Stable others	0	0	0	0	0	96	96	0.1367	1.00
Buffer-(stable PF)	75	0	0	0	0	0	75	0.0377	1.00
User's accuracy	0.86	1.00	0.83	0.97	0.94	0.86			Overall accuracy: 0.8711

Error Matrix populated by estimated proportions of area using (eq. 2.3) in methods, standard error (eq. 2.4) and area estimates (eq. 2.4 – 2.5). Our total study area A = 3924375.63 km^2.

$\hat{p}_{ik}$	Reference					
Mapped	Stable PF	PF loss	Stable SF	SF loss	SF gain	Stable others
Stable PF	0.7624	0	0	0	0	0
PF loss	0.0007	0.0024	0	0	0	0.0005
Stable SF	0	0	0.0493	0.0007	0	0
SF loss	0	0	0.0007	0.0024	0.0001	0.0002
SF gain	0	0	0.0007	0	0.0046	0.0008
Stable others	0	0	0	0	0	0.1367
Buffer-(stable PF)	0.0377	0	0	0	0	0
Sum ($\hat{p}_{.k}$)	0.8009	0.0024	0.0507	0.0031	0.0047	0.1382
$S(\hat{p}_{.k})$	0.0002	0.0003	0.0008	0.0007	0.0004	0.0004
Estimated areas (km^2) ($\hat{A}_k \pm S(\hat{A}_k)$)	3143093.86 ±836.59	9391.80 ±988.05	198967.89 ±3028.68	12080.85 ±2834.33	18425.74 ±1620.44	542415.49 ±1517.38
Annual loss/gain (km^2 yr^{-1})		4695.90 ±494.02		6040.43 ±1417.17	9212.87 ±810.22	

Table A.5 Error matrix of sample-based estimates for period 2010-2012.
11 sampled points were excluded due to being mosaic pixels for which it was not possible to determine a specific land cover or for which no clear satellite imagery available.

n_{ik}	Reference								Producer's accuracy
Mapped	Stable PF	PF loss	Stable SF	SF loss	SF gain	Stable others	n_i	w_i	
Stable PF	538	0	0	0	0	0	538	0.7721	1.00
PF loss	2	45	0	0	0	2	49	0.0023	0.92
Stable SF	0	0	74	1	0	0	75	0.0510	0.99
SF loss	0	0	5	41	0	2	48	0.0054	0.85
SF gain	0	0	10	1	24	7	42	0.0076	0.57
Stable others	0	0	1	0	0	94	95	0.1358	0.99
Buffer-(stable PF)	75	0	0	0	0	0	75	0.0257	1.00
User's accuracy	0.87	1.00	0.82	0.95	1.00	0.90			Overall accuracy: 0.8850

Error Matrix populated by estimated proportions of area using (eq. 2.3) in methods, standard error (eq. 2.4) and area estimates (eq. 2.4 – 2.5). Our total study area A = 3924375.63 km^2.

$\hat{p}_{ik}$	Reference					
Mapped	Stable PF	PF loss	Stable SF	SF loss	SF gain	Stable others
Stable PF	0.7721	0	0	0	0	0
PF loss	0.0001	0.0021	0	0	0	0.0001
Stable SF	0	0	0.0503	0.0007	0	0
SF loss	0	0	0.0006	0.0046	0	0.0002
SF gain	0	0	0.0018	0.0002	0.0044	0.0013
Stable others	0	0	0.0014	0	0	0.1343
Buffer-(stable PF)	0.0257	0	0	0	0	0
Sum ($\hat{p}_{.k}$)	0.7979	0.0021	0.0541	0.0055	0.0044	0.1359
$S(\hat{p}_{.k})$	0.0001	0.0001	0.0017	0.0008	0.0006	0.0015
Estimated areas (km^2) ($\hat{A}_k \pm S(\hat{A}_k)$)	3131395.29 ±260.93	8390.71 ±361.08	212502.79 ±6591.65	21514.88 ±2971.53	17125.08 ±2316.17	533446.88 ±5911.69
Annual loss/gain (km^2 yr^{-1})		4195.36 ±180.54		10757.44 ±1485.76	8562.54 ±1158.09	

Table A.6 Error matrix of sample-based estimates for period 2012-2014.
10 sampled points were excluded due to being mosaic pixels for which it was not possible to determine a specific land cover or for which no clear satellite imagery available.

n_{ik}	Reference								Producer's accuracy
Mapped	Stable PF	PF loss	Stable SF	SF loss	SF gain	Stable others	n_i	w_i	
Stable PF	538	0	0	0	0	0	538	0.7448	1.00
PF loss	3	47	0	0	0	0	50	0.0029	0.94
Stable SF	0	0	73	0	0	1	74	0.0461	0.99
SF loss	0	0	7	30	3	4	44	0.0130	0.68
SF gain	0	0	7	0	36	4	47	0.0112	0.77
Stable others	0	0	1	0	0	94	95	0.1319	0.99
Buffer-(stable PF)	74	1	0	0	0	0	75	0.0502	0.99
User's accuracy	0.87	0.98	0.83	1.00	0.92	0.91			Overall accuracy: 0.8862

Error Matrix populated by estimated proportions of area using (eq. 2.3) in methods, standard error (eq. 2.4) and area estimates (eq. 2.4 – 2.5). Our total study area A = 3924375.63 km^2.

$\hat{p}_{ik}$	Reference					
Mapped	Stable PF	PF loss	Stable SF	SF loss	SF gain	Stable others
Stable PF	0.7448	0	0	0	0	0
PF loss	0.0002	0.0027	0	0	0	0
Stable SF	0	0	0.0455	0	0	0.0006
SF loss	0	0	0.0021	0.0088	0.0009	0.0012
SF gain	0	0	0.0017	0	0.0086	0.0010
Stable others	0	0	0.0014	0	0	0.1305
Buffer-(stable PF)	0.0495	0.0007	0	0	0	0
Sum ($\hat{p}_{.k}$)	0.7945	0.0034	0.0506	0.0088	0.0095	0.1333
$S(\hat{p}_{.k})$	0.0007	0.0007	0.0018	0.0009	0.0009	0.0017
Estimated areas (km^2) ($\hat{A}_k \pm S(\hat{A}_k)$)	3117832.68 ±2652.18	13189.70 ±2652.18	198522.48 ±7003.04	34655.61 ±3610.30	37195.03 ±3372.57	522980.14 ±6626.40
Annual loss/gain (km^2 yr^{-1})		6594.85 ±1326.09		17327.81 ±1805.15	18597.51 ±1686.28	

Table A.7 Comparison of sample-based estimates with map-based estimates of annual forest loss of primary forest vs. secondary forest over 2000-2014 across the Brazilian Amazon.

Errors on map-based estimates are only from time interval corrections. Errors on sample-based estimates include both standard errors and interval corrections. Interval corrections were only applied to the four-year intervals (i.e. 2000-2004, 2004-2008) for secondary forest loss and gain.

Time interval		Secondary forest loss (km^2 yr^{-1})	Primary forest loss (km^2 yr^{-1})	Total forest loss (km^2 yr^{-1})	Secondary forest loss as fraction of total forest loss
2000-2004	sample	15276.44±1366.77	25788.42±1804.94	41064.86 ± 2264.04	0.37 ± 0.03
	map	18002.70±1112.67	25115.20	43117.90 ± 1112.67	0.42 ± 0.04
2004-2008	sample	12742.98±2298.56	13786.45 ± 428.32	26529.43 ± 2338.13	0.48 ± 0.05
	map	10080.21±623.02	14771.19	24851.40 ± 623.02	0.42 ± 0.04
2008-2010	sample	6040.43 ± 1417.17	4695.90 ± 494.02	10736.33 ± 1500.80	0.56 ± 0.06
	map	6768.77	7190.60	13959.37	0.48
2010-2012	sample	10757.44±1485.76	4195.36 ± 180.54	14952.80 ± 1496.69	0.72 ± 0.03
	map	10613.72	4568.28	15182.00	0.70
2012-2014	sample	17327.81±1805.15	6594.85 ± 1326.09	23922.66 ± 2239.89	0.72 ± 0.05
	map	25414.12	5619.72	31033.84	0.82

Table A.8 Secondary forest (SF) standing area, area loss rates and proportional loss rates by age group.

The numbers in brackets denote time interval corrections for 2004-2008. Secondary forest loss for interval 2000-2004 was not included as it is not possible to discriminate ages for the first census interval (2000-2004).

Time interval	TERRACLASS (TC) based SF age definition	SF Age at start of interval (years)	SF area at start of interval (km^2)	SF loss ($km^2\ yr^{-1}$)	SF loss rates (% yr^{-1})
2004-2008	Non-SF in TC-2000, but SF in TC-2004.	0-4	59833.11	4837.83 ± 299.01	8.09 ± 0.50
	SF in TC-2000-2004.	> 4	105553.86	5242.38 ± 324.01	4.97 ± 0.31
		Sum	165386.97	10080.21 ± 623.02	6.09 ± 0.38
2008-2010	Non-SF in TC-2004, but SF in TC-2008.	0-4	75795.34	2972.55	3.92
	Non-SF in TC-2000, but SF in TC-2004-2008.	4-8	44734.28	1560.80	3.49
	SF in TC-2000-2004-2008.	> 8	89192.43	2235.42	2.51
		Sum	209722.05	6768.77	3.23
2010-2012	Non-SF in TC-2008, but SF in TC-2010.	0-2	25260.17	1915.38	7.58
	Non-SF in TC-2004, but SF in TC-2008-2010.	2-6	69850.24	4474.60	6.41
	Non-SF in TC-2000, but SF in TC-2004-2008-2010.	6-10	41612.69	2389.07	5.74
	SF in TC-2000-2004-2008-2010.	>10	84821.59	1834.68	2.17
		Sum	221444.68	10613.72	4.79
2012-2014	Non-SF in TC-2010, but SF in TC-2012.	0-2	31484.28	7299.95	23.19
	Non-SF in TC-2008, but SF in TC-2010-2012.	2-4	21429.42	3329.72	15.54
	Non-SF in TC-2004, but SF in TC-2008-2010-2012.	4-8	60901.03	7954.88	13.06
	Non-SF in TC-2000, but SF in TC-2004-2008-2010-2012.	8-12	36834.55	3542.58	9.62
	SF in TC-2000-2004-2008-2010-2012.	>12	81052.23	3286.98	4.06
		Sum	231701.52	25414.12	10.97

Table A.9 Map-based estimates of the fate of annual secondary forest loss over 2000-2014 across the Brazilian Amazon.

Time interval corrections were applied for 2000-2004 and 2004-2008 intervals.

Secondary forest loss by post-land uses (km^2 yr^{-1})

	2000-2004	2004-2008	2008-2010	2010-2012	2012-2014
Pasture	17189.74±1062.43	9589.58 ± 592.69	6190.98	9676.41	23655.65
Agriculture	382.62 ± 23.65	334.50 ± 20.67	265.86	286.25	548.18
Mining	15.42 ± 0.95	21.79 ± 1.35	14.28	50.27	92.09
Urban	43.28 ± 2.68	52.89 ± 3.27	24.49	53.05	63.15
Reforestation	66.38 ± 4.10	55.41 ± 3.42	167.64	56.91	129.05
Others	305.25 ± 18.87	26.04 ± 1.61	105.52	490.82	925.99
sumLoss	18002.70±1112.67	10080.21±623.02	6768.77	10613.72	25414.12

Secondary forest gain by previous land uses (km^2 yr^{-1})

	2000-2004	2004-2008	2008-2010	2010-2012	2012-2014
Pasture	11140.78 ± 965.10	16678.67±1444.83	11560.42	13956.75	20606.90
Agriculture	36.58 ± 3.17	48.63 ± 4.21	17.64	77.59	64.41
Mining	89.96 ± 7.79	37.90 ± 3.28	0.97	24.28	57.77
Urban	38.71 ± 3.35	7.18 ± 0.62	0.38	0.21	0.60
Others	117.81 ± 10.21	386.71 ± 33.50	45.48	123.47	527.48
Deforestation	165.53 ± 14.34	1501.18 ± 130.04	270.72	693.98	578.30
Primary forest	6646.12 ± 575.74	4451.56 ± 385.63	714.63	757.70	1053.18
Reforestation	24.90 ± 2.16	20.04 ± 1.74	19.85	108.16	182.39
sumGain	18260.39±1581.85	23131.88±2003.86	12630.09	15742.14	23071.02

Net change = sumGain – sumLoss

	2000-2004	2004-2008	2008-2010	2010-2012	2012-2014
Net change (mean)	257.69	13051.67	5861.32	5128.42	-2343.09

Table A.10 Map-based estimates of the fate of annual secondary forest loss by age categories across the Brazilian Amazon.

	2008-2010			2012-2014		
Area (km^2 yr^{-1})	0-4 years	4-8 years	>8 years	0-4 years	4-8 years	>8 years
Pasture	2668.91	1437.74	2084.32	10184.10	7344.87	6126.68
Agriculture	195.35	33.20	37.31	151.15	214.16	182.87
Mining	2.74	3.62	7.92	23.33	28.87	39.89
Urban	4.58	6.70	13.20	16.99	8.00	38.15
Reforestation	70.56	52.49	44.59	57.35	35.99	35.72
Others	30.41	27.04	48.07	196.74	322.99	406.26
sumLoss	2972.55	1560.80	2235.42	10629.67	7954.88	6829.57

Table A.11 Annual secondary forest loss predicted by null model analysis.
The null model assumes biased sampling without replacement given the available areas of secondary/primary forests, the total forest loss and the preferential bias towards cutting secondary forest instead of primary forests. The available areas of secondary/primary forests at the beginning of each interval were computed by the sum of stable secondary forest and the secondary forest loss within each interval. The preferential bias was derived from the first interval (2000-2004). The lower bounds of secondary forest loss for the null model were based on the lower total forest loss and lower available forest areas as derived from the sampling-based estimates. The upper bounds of secondary forest loss for the null model were based on the higher total forest loss and higher available forest areas as derived from the sampling-based estimates.

Time interval	Primary forest area at start of interval (km^2)	Secondary forest area at start of interval (km^2)	Total forest loss (km^2/interval)	Secondary forest loss (km^2 yr^{-1}) (null model)	Secondary forest loss as fraction of total forest loss (null model)
2000-2004	3309390.56 ±11197.94	199834.95 ±10063.51	164259.45 ±9056.15	15276.46 (14098.65-16475.60)	0.372 (0.363-0.380)
2004-2008	3211261.50 ±2039.57	220098.79 ±12246.41	106117.68 ±9352.50	11435.21 (10173.76-12729.97)	0.431 (0.421-0.441)
2008-2010	3152485.67 ±1294.65	211048.74 ±4148.05	21472.65 ±3001.61	5013.76 (4281.95-5755.09)	0.467 (0.464-0.470)
2010-2012	3139786.01 ±445.49	234017.67 ±7230.48	29905.59 ±2993.39	7328.26 (6507.11-10287.68)	0.490 (0.484-0.496)
2012-2014	3131022.37 ±3750.75	233178.09 ±7878.88	47845.31 ±4479.77	11503.93 (10287.68-12743.17)	0.481 (0.474-0.487)

References

Arévalo, P. et al. 2019. Continuous monitoring of land change activities and post-disturbance dynamics from Landsat time series: A test methodology for REDD+ reporting. *Remote Sensing of Environment.* p111051.

Olofsson, P. et al. 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment.* **148**, pp.42-57.

Appendix B

Supplementary information for Chapter 3 (Paper II)

Table B.1 Annual number of Landsat surface reflectance (SR) images used in this study. In total, there are 11483 images. For 2001 and 2002, images were from Landsat-7, otherwise from Landsat-5.

Year	Image No.	Coverage of study area (%)	Year	Image No.	Coverage of study area (%)	Year	Image No.	Coverage of study area (%)
1984	261	99.74	1993	468	99.87	2002	715	99.93
1985	257	99.61	1994	298	99.81	2003	303	99.79
1986	381	99.87	1995	325	99.92	2004	583	99.91
1987	407	99.90	1996	483	99.89	2005	536	99.96
1988	384	99.89	1997	295	99.73	2006	529	99.83
1989	331	99.81	1998	303	99.87	2007	498	99.88
1990	386	99.79	1999	399	99.90	2008	543	99.86
1991	368	99.97	2000	467	99.84	2009	551	99.86
1992	308	99.86	2001	641	99.85	2010	463	99.90

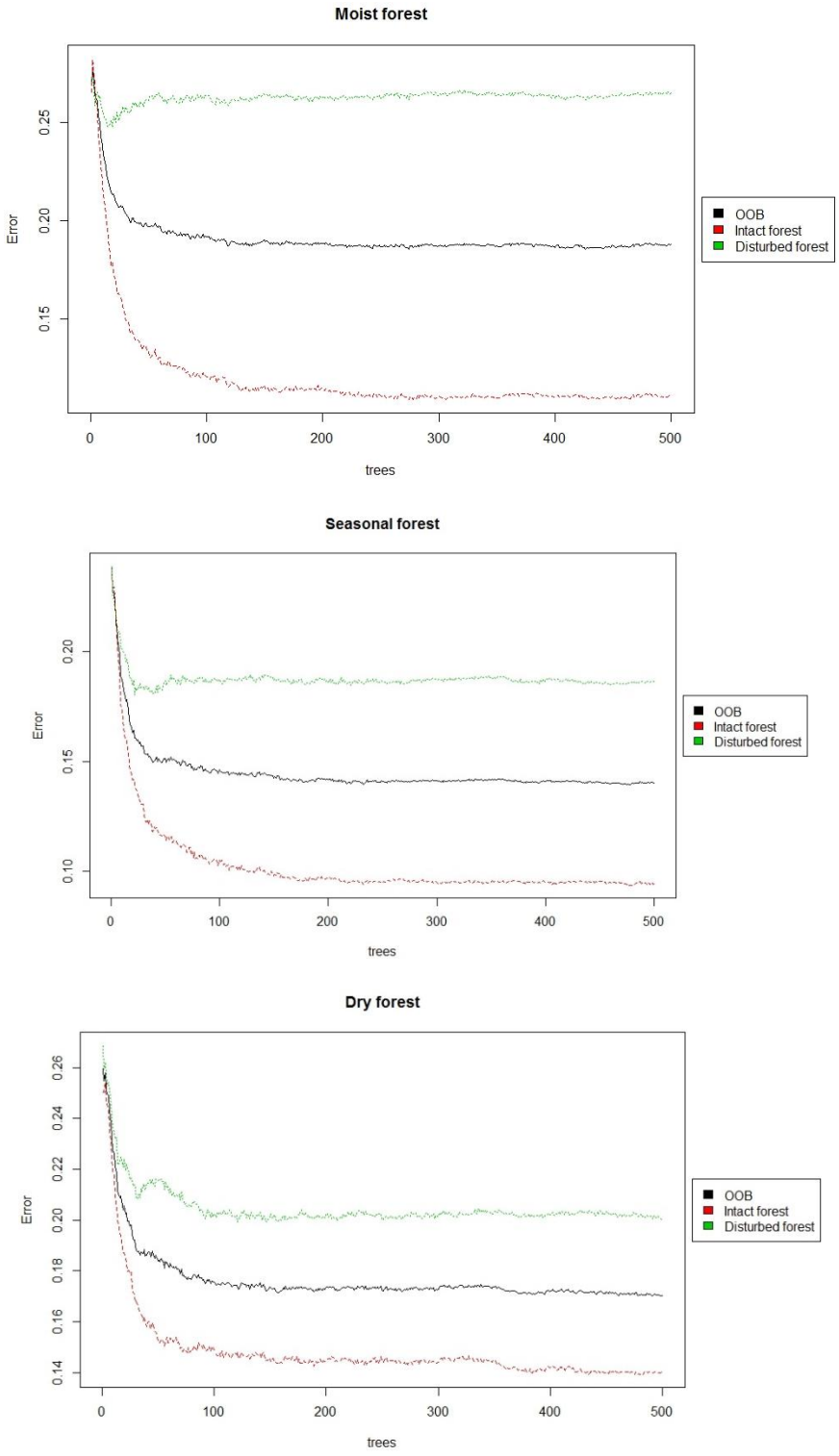


Figure B.1 The change of random forest classification errors with the number of trees. Black: overall classification out of bag (OOB) error; Red: classification error for intact forests; Green: classification error for disturbed forests. For classification in this study, we used 500 trees.

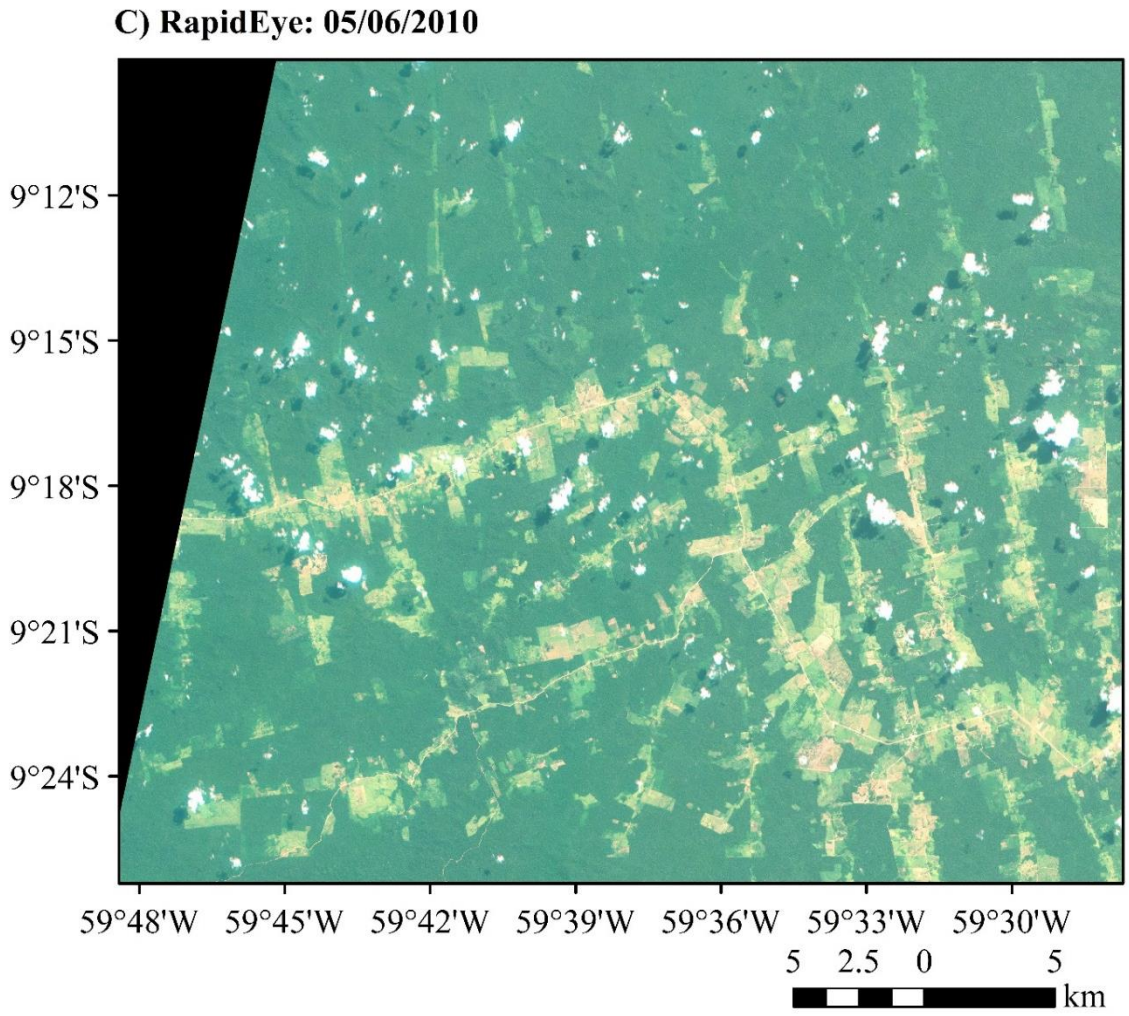


Figure B.2 RapidEye true-colour composite image (Team, 2016) for high resolution image interpretation validation of our classification map corresponding to area 1 (Moist forest) in Figure 3.5.

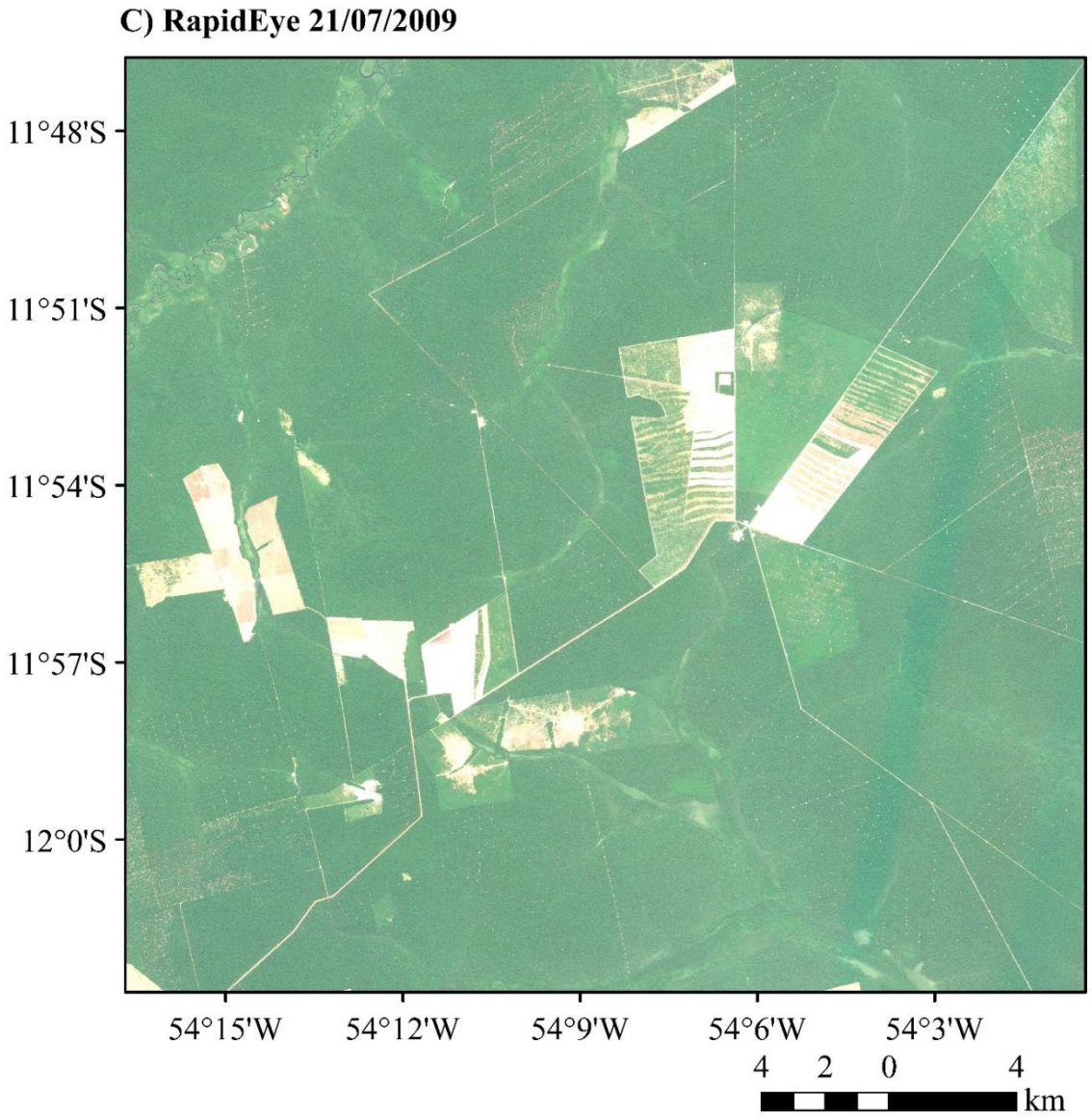


Figure B.3 RapidEye true-colour composite image (Team, 2016) for high resolution image interpretation validation of our classification map corresponding to area 2 (Seasonal forest) in Figure 3.6.

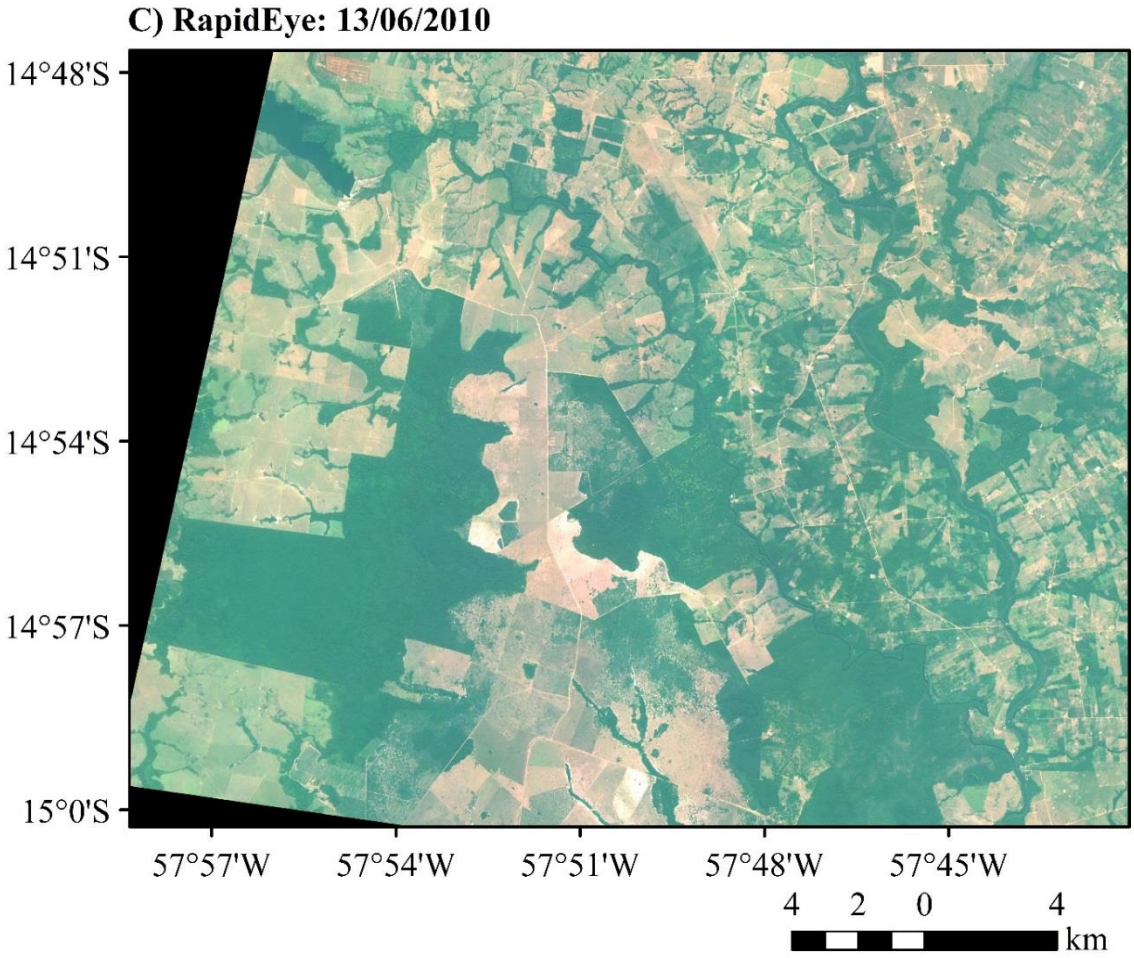


Figure B.4 RapidEye true-colour composite image (Team, 2016) for high resolution image interpretation validation of classification map corresponding to area 3 (Dry forest) in Figure 3.7.

Besides ten-fold cross validation, we also validated our classification with 31 ground-truth intact forest plots and 91 high-resolution imagery validated disturbed forest points, a subset of which were subsequently validated in the field. Intact forest points are old-growth forest plots linked to the PPBIO (Brazilian Program for Biodiversity Research) (Pezzini et al., 2012) and PELD “Cerrado-Amazon Forest Transition: ecological and socio-environmental bases for Conservation” forest plot network. For disturbed forest points, we used the SPOT-validated points of secondary vegetation and regeneration with pasture from TerraClass, which were generated from SPOT-5 High Geometric Resolution (2.5m spatial resolution) data in “panchromatic” mode (Almeida et al., 2016). Both intact forest points and SPOT-validated points are not included in developing TerraClass itself. The distribution of these additional validation data for each ecoregion and valiation results are shown in Appendix Figure B.5 and AppendixTable B.2.

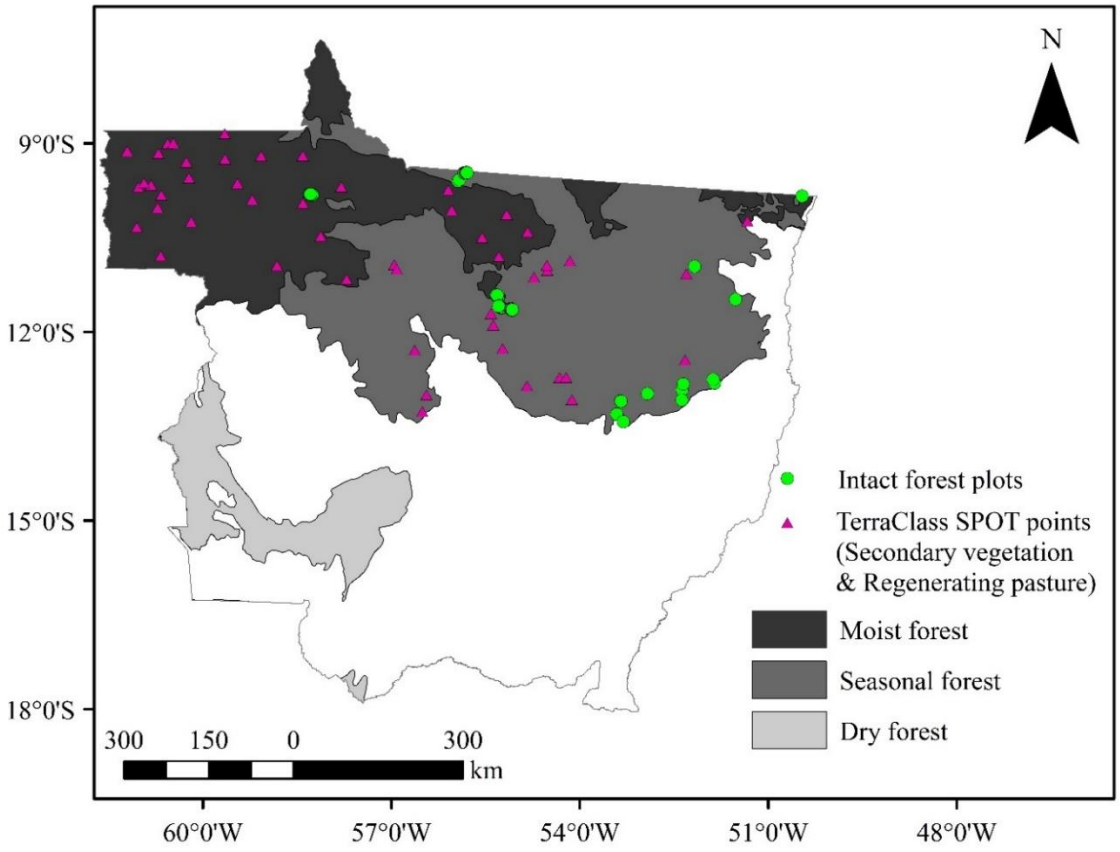


Figure B.5 Study area and additional validation points.
Intact forest plots (18 for seasonal forest and 13 for moist forest) are from Biodiversity Research(PPBio) Information System (Pezzini et al., 2012) and other field sites. TerraClass SPOT points (59 for seasonal forest and 32 for moist forest) are validated points of secondary vegetation and regeneration with pasture that were used for TerraClass products validation. Some points overlap due to the coarse visualization scale.

Table B.2 Further point-scale validation accuracy based on the additional intact forest points and SPOT validated disturbed forest points.
Refer to points in Figure B.5, which were not used to produce the classification map.

Regions	Overall accuracy	Producer's accuracy		User's accuracy		Kappa statistic
		Intact forests	Disturbed forests	Intact forests	Disturbed forests	
Moist forest	0.956	1.0	0.938	0.867	1.0	0.897
Seasonal forest	0.909	0.881	1.0	1.0	0.72	0.776
Dry forest	---	---	---	---	---	---

References

Almeida, C.A.d. et al. 2016. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. *Acta Amazonica.* **46**(3), pp.291-302.

Pezzini, F. et al. 2012. The Brazilian Program for Biodiversity Research (PPBio) Information System. *Embrapa Roraima-Artigo em periódico indexado (ALICE).*

Team, P. 2016. *Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA*.

Appendix C

Supplementary information for Chapter 4 (Paper III)

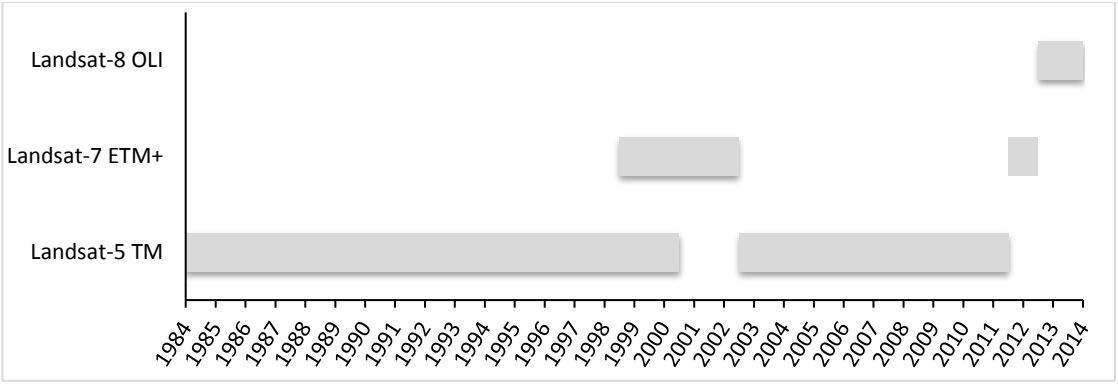


Figure C.1 Landsat surface reflectance dataset.
Landsat-5 images in 2001 contain striping artifacts and have low coverage of the study area in 2002. Landsat-7 contain data gaps across imagery scenes since its Scan Line Corrector (SLC) failed in 2003. However, Landsat-7 is the only dataset available in 2012. Sensor calibration was applied between Landdsat-8 and Landsat-5/7 (Roy et al., 2016).

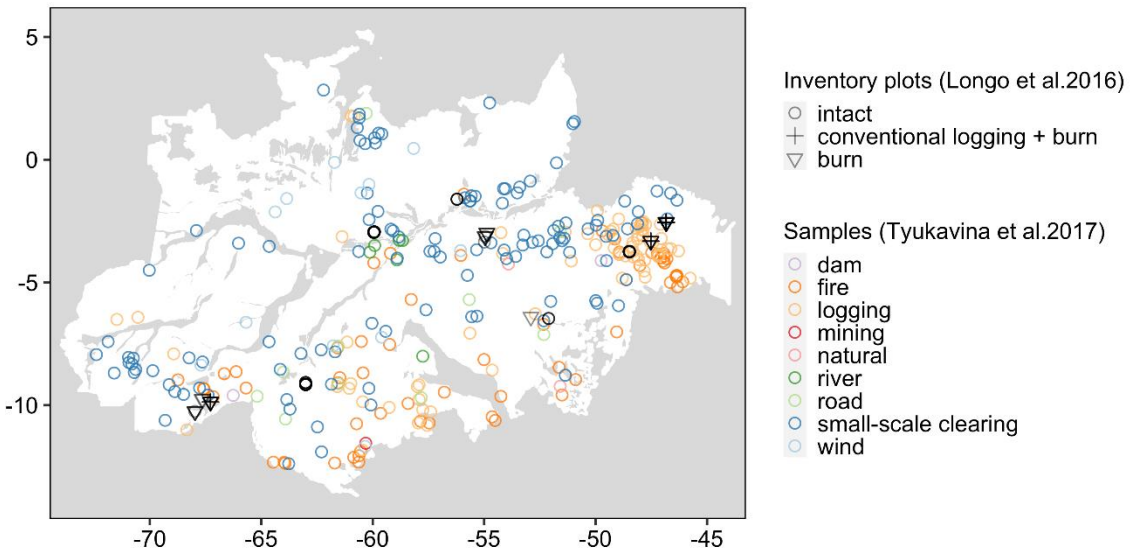


Figure C.2 The distribution of primary forest disturbance samples from Tyukavina et al. (2017) and the inventory plots from Longo et al. (2016).

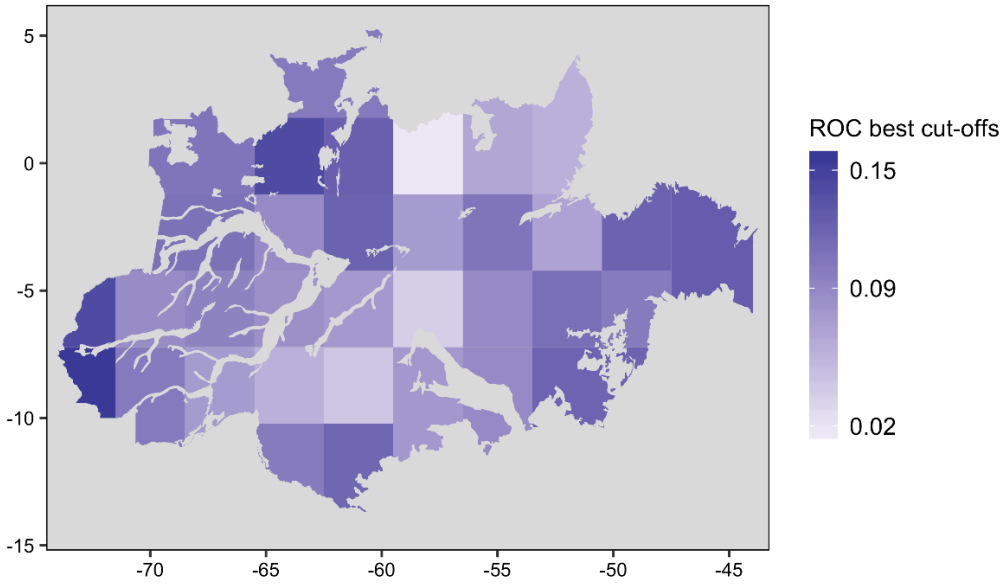


Figure C.3 ROC curve (receiver operating characteristic curve) best cut-off thresholds for classification aggregation.

The thresholds were estimated based on the classification training points (about 1,400 intact points and 1,400 disturbed points) from stratified random sampling.

References

Longo, M. et al. 2016. Aboveground biomass variability across intact and degraded forests in the Brazilian Amazon. *Global Biogeochemical Cycles.* **30**(11), pp.1639-1660.

Roy, D.P. et al. 2016. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote sensing of Environment.* **185**, pp.57-70.

Tyukavina, A. et al. 2017. Types and rates of forest disturbance in Brazilian Legal Amazon, 2000–2013. *Science advances.* **3**(4), pe1601047.