ABSTRACT

Title of Dissertation:	30+ YEARS OF LAND COVER AND LAND USE CHANGE IN SOUTH AMERICA
	Viviana Zalles Ballivian, Doctor of Philosophy, 2020
Dissertation directed by:	Professor Matthew C. Hansen, Department of Geographical Sciences

The modification of the Earth's surface constitutes the most impactful way in which humans affect their surrounding environment, with broad and lasting consequences. Changes in land cover accelerate biodiversity loss, contribute to climate change, and affect the provisioning of ecosystem services. Such negative environmental impacts can have important effects on human health and livelihoods. The South American continent, in particular, has undergone significant transformations over the past decade, due in large part to the conversion of natural land to more economically productive land uses, such as crops, pastures, and tree plantations. The agricultural commodities produced in South America are traded and consumed globally, and land will likely continue to be converted if demand for these commodities continues to rise. Despite the environmental and commercial importance of land cover and land use change dynamics in South America, the extent and rates of land change have not yet been thoroughly characterized and quantified. This dissertation aims to advance

scientific knowledge on the extent and rates of change of important land covers and land uses, especially as they relate to the production of agricultural commodities, by leveraging the 34-year Landsat archive of Earth observation data. The general approach employed throughout follows a two-step process of mapping and sampling, in order to provide spatially explicit information on the patterns of land cover/land use change, as well as associated unbiased area estimates. This approach is first employed for the use-case of Brazilian cropland expansion from 2000 to 2014, and results show a near doubling of cropland area, the majority of which (80%) came about through the conversion of existing pastures. The methodology is then repeated at broader thematic, temporal, and geographic scales, resulting in area estimates of changes in cropland, pasture, plantation, natural tree regrowth, semi-natural land, tree cover and degraded tree cover from 1985 to 2018. Altogether, these changes amount to a 60% increase in human impact on natural land over the study period. Finally, an analysis and evaluation of the methodology employed for mapping and sampling when there is a multitude of target classes instead of a single one is provided as an assessment of methodological approaches.

30+ YEARS OF LAND COVER AND LAND USE CHANGE IN SOUTH AMERICA

by

Viviana Zalles Ballivian

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2020

Advisory Committee:
Dr. Matthew C. Hansen, Chair
Dr. Peter V. Potapov
Dr. Stephen V. Stehman
Dr. Christopher O. Justice
Dr. Joseph Sullivan

Acknowledgements

The first time I met Matt, when I was applying to the PhD program, I asked him what the pink stuff in a Landsat pseudo-color image was. Somehow, for some reason, he overlooked my nescience and accepted a freshly minted bachelor of science with zero remote sensing knowledge as his new PhD student. I am eternally grateful to you for giving me the greatest opportunities, for all that you have taught me, for being a friend.

To Peter Potapov, without whom none of this would be possible: thank you for your patience and support. To Steve Stehman, thank you for your always thorough and gentle edits, for helping me with the numbers, for making yourself available. To Chris Justice, thank you for supporting my work and encouraging my path forward and out. To Joe Sullivan, thank you for generously agreeing to serve as the Dean's Representative on my committee.

To my chosen family, Leela and Ale, Andres y Mayayo, mis Pesadillas (todas y cada una), Nico, Antoine, Tina, Sarah: thank you for filling my life with love and support. To those who have helped me and accompanied me closely in this endeavor: Amy (I am forever indebted to you for your endless patience, kindness, generosity; you have taught me so much), Andres (my PhD is split into pre-Andres and post-Andres, gracias un millon de veces por todo), and the entire GLAD team (I am so grateful I get to work with you).

To Goetzi, I know I could have done this without you but I am so glad I did not have to. You lived through it all with me, all six years of it, and I would not have wanted anyone else by my side.

To my mother, a force of nature, a brilliant mind, a strong woman. My father, a guiding light, source of comfort, noblest of souls. To Gabriel and Renata, who teach me – by example – to keep steadily moving forward, always. I am forever your biggest admirer and supporter. This is *my* labor of love dedicated to the four of you, my sine que non, for giving me so much always. Los amo (BB##).

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 **Background of the research**

1.1.1 Importance of land cover and land use change

Land cover/land use change is the single most important way in which humans alter the natural environment (1). Land cover refers to the "biophysical coverage of land", whereas land use refers to "the economic and social purposes for which land is managed", and land cover and land use change refer to changes from one land cover/land use class to another (2). Land cover/land use change has impacts across a host of systems, such as climate, the hydrological cycle, nutrient cycles, biodiversity, and human health and well-being, amongst others. Human appropriation of natural land also brings about important economic benefits. Balancing the costs and benefits of land cover/land use change is essential to ensuring a sustainable future for all.

Changes in land cover have important consequences to climate at regional and global scales by altering fluxes of energy, water, and greenhouse gas emissions (3–5). Land cover changes can modify surface albedo, which affects the reflectance of the Earth's surface and thus impacts surface temperature (6). Land cover changes also alter the hydrological cycle by reducing annual terrestrial evapotranspiration and by increasing the amount of runoff, which affects climate through impacts on precipitation and the onset and length of seasons (7–10). For example, deforestation in the Amazon basin has been shown to increase near-surface air temperature and decrease evapotranspiration, and after a certain threshold, these temperature and

precipitation changes could cause a shift in vegetation to a Cerrado-like ecosystem (11, 12). Perhaps the most important way in which land cover modifications impact climate is through the increase of greenhouse gas emissions. Deforestation, the clearing of peatlands, extensification of rice paddies, ruminant fermentation on expanding pastures, and the production of fertilizers for growing areas of cropland are some examples of land cover/land use change-driven emissions of potent greenhouse gases such as carbon dioxide, methane, and nitrous oxide (13, 14). Emissions from the agriculture, forestry and other land use (AFOLU) sector are estimated to contribute to about 23% of anthropogenic greenhouse gas emissions, making this sector a significant contributor to global climate change (15).

Land cover change also has important impacts on the global Nitrogen (N) cycle due to the large amounts of dinitrogen (N_2) that are fixed either by fertilizer use or by the production of legume crops (such as soybeans) (16). Converting N from unreactive to reactive forms that accumulate in the environment can have cascading effects by increasing plant productivity and accelerating emissions of nitrous oxide from soils, which in turn affects atmospheric systems (16). Land cover/land use change impacts on nutrient cycles extend to the phosphorous (P) cycle as well: land use intensification depletes P reserves through soil erosion, mineralization of soil organic matter, and yield export (17). Nutrient cycles are further disrupted due to leaching and volatilization of agricultural nutrients and pesticides applied to growing areas of crops. For example, leaching of N and P into estuaries, coastal waters, rivers, and streams can cause eutrophication, thus creating hypoxic (low oxygen) or anoxic (no oxygen) conditions that endanger aquatic ecosystems (18).

Land cover/land use change is expected to be the single largest driver of global biodiversity loss by 2100, because of its limiting effects on habitat availability and consequent species extinctions (19). Because the tropical and temperate forests of South America are projected to have the highest rates of land-use change, biodiversity losses are expected to be most acute in South America (19). Species extinction alters ecosystem productivity and decomposition at rates comparable to those of climate warming (20) and has important consequences to the invasibilty, stability and resilience of ecosystems (21).

The climatic, hydrologic, nutrient cycling, and biodiversity impacts of land cover/land use change jeopardize ecosystems' functioning and their ability to provide the ecosystem services on which long-term human well-being is dependent. Food production, access to freshwater, access to clean air, and protection from extreme climatic events such as droughts and floods are only some of the myriad of goods and services provided by well-functioning ecosystems. Land cover/land use change also impacts human well-being by facilitating the emergence and proliferation of diseases. Deforestation, road construction, dam construction, the increase of crop and pasture areas, and extensification of human settlements can provide opportunities for disease transmission by increasing contact between humans and vector populations (22). The current novel coronavirus (COVID-19) outbreak causing mass quarantines across the world was transmitted to humans through zoonotic transmission (23), bringing into stark relief the ways in which animal-transmitted diseases can upend societies and economies across the world. Despite the important tolls on human health that COVID-19 has had to date, they pale in comparison to the impacts of malaria, a

disease that affects hundreds of millions of people worldwide (24). Malaria incidence, a disease that kills roughly half a million people a year, is influenced by land cover/land use changes and by climate change (22, 25, 26). For example, dam construction in some geographies has been shown to be associated with increases in malaria transmission (27, 28).

Despite the evident issues associated with the human appropriation of natural environments through the process of land cover/land use change, the concomitant benefits are indisputable. Land cover/land use changes, along with technological advances that enable the intensification of production, have allowed us to meet growing demands for food, fiber, fuel, and water as global populations skyrocket (29). Growth in the agricultural sector, one of the leading drivers of global land cover/land use change, has permitted increases in food production, reducing the number of undernourished people in the world – though much progress remains to be done (30). Production of agricultural commodities such as soybean, beef, and timber have enabled economic development of countries that produce and trade these goods (31). Further, increasing infrastructure such as roads and dams, which drives land cover/land use change, can yield important social and economic benefits by providing access to markets and services, as well as energy production (32).

Land cover/land use change will continue to occur as long as the economic incentives for it are present and no policy or institutional barriers exist. In reality, the net economic benefits of land cover/land use change are artificially high given that the costs of the loss of ecosystem services associated with these changes are not properly accounted for (29). For example, when a plot of Amazon forest is cleared,

the landowner will benefit financially from the goods produced on the land, but there is no mechanism whereby they are made economically responsible for the loss of biodiversity or the emissions associated with this conversion of land. Evidently, given how ubiquitous land cover/land use change is and how far-reaching its associated negative impacts are, we are all bearing the consequences of it and will continue to do so into the future. However, the costs associated with the loss of ecosystem services are disproportionately borne by the poor, as they are most likely to be directly reliant on those services (for access to water, clean air, and food, for example), and most vulnerable to impacts (disease, extreme weather events, etc.) (29). Without policies and/or market-based mechanisms to address the mismatch between benefits and costs associated with land cover/land use change, rampant anthropogenic appropriation of the natural environment will continue.

1.1.2 South America as a hotspot of land cover and land use change

In South America, rates of land cover/land use change are high and impacts are particularly acute. South America is home to hugely important ecosystems, such as the Amazon, the Cerrado, and the Chiquitania forests. The importance of the Amazon forest cannot be overstated. It is one of the most undisturbed regions remaining on Earth (33), constitutes a massive carbon sink (34), regulates critical hydrologic and climate systems (11), and provides habitat for uncounted numbers of species (35). The Cerrado savanna is also of critical importance as it is a global biodiversity hotspot (36). Other biodiversity hotpsots in South America include the Atlantic forest of Brazil, central Chile, the Tumbes-Choco-Magdalena region of Colombia, and the tropical Andes (36). The Chaco and the Chiquitania ecosystems also make up an

important continuous extent of natural vegetation, one of the largest remaining in South America (37).

In parallel to being an important region in terms of its natural ecosystems, the South American continent is also a critical producer of global commodities, particularly animal feed. Brazil and Argentina are the biggest players in this domain: they are amongst the top exporters of corn (16% and 13% of global exports, respectively), soybean (45% and 5%, respectively), and soybean meal (22% and 40%, respectively) (38). Paraguay, Uruguay, and Bolivia also produce soybeans over extensive areas of land (39–41). South America is also a global leader in the production of meat. Brazil, for example, is among the world's top producers of beef, poultry, and pig meat (42). Brazil also has the largest number of cattle in the world, and Argentina is sixth on the list (42).The South American continent produces 30% of global exports of frozen bovine meat, making it the top-exporting region in the world. Other important agricultural commodity exports from South America include sugarcane, wheat, coffee, wood products, and poultry meat (38). For Brazil, Argentina, Uruguay, Paraguay, and Ecuador, agricultural commodities make up the majority of total exports in terms of absolute value (38).

Land used to produce these agricultural commodities comes from the conversion and appropriation of natural ecosystems. Commodity-driven deforestation is the leading driver of forest loss in Latin America, with important impacts in the Amazon, Cerrado, Chiquitano, and Chaco forests (43). Pressures to increase returns from the agricultural sector are expected to drive further deforestation and conversion of natural lands towards more productive agricultural land uses (44). Brazil and

Argentina are two of the countries with the largest reserves of noncultivated area suitable for cropping (31), further highlighting the potential for continued anthropogenic appropriation of natural land for commodity crop production. Because pasture's suitability requirements are lower than cropland's, the potential for pasture area expansion is even greater (45). Given South America's large tracts of remaining natural vegetation of critical environmental importance, and the expected continued expansion of commodity land uses for economic development, definitive information on the extent and rates of land cover and land use change in South America is urgently needed.

1.1.3 Current state of land cover and land use change characterization in South America

To better understand the dynamics of land cover/land use change, their drivers, and how these changes impact the provisioning of ecosystem services across South America, it is essential that land change be quantified. Remote sensing data from Earth-observing satellites provide us with continuously updated, consistent data across the entire continent (and the world) that enable the mapping and quantification of land cover and land use change. Without satellite imagery, reliable quantification of these dynamics would not be feasible for large geographic and temporal scales. The joint NASA/USGS Landsat program provides us with a continuous archive of globally acquired, free and open 30-m spatial-resolution Earth observation data dating back to 1972. The opening of the Landsat archive in 2008 (46), combined with the increased availability of high-performing computing capabilities, has made it possible to monitor changes to the Earth's surface through time at regional, national,

continental, and even global scales (47). Advances in major remote sensing applications to map and quantify land cover/land use change in South America are reviewed here.

Within the land cover/land use change scientific literature, there is an extensive body of work related to forests, especially within the Amazon. Forests are the primary subject of many remote sensing studies because of their environmental significance and because they are relatively easily mapped compared to other land covers/land uses (47). Information on Brazilian deforestation, in particular, has been readily available since 1997 with the advent of the PRODES (Amazon Deforestation Monitoring Project) operational monitoring system and the DETER (Near Real Time Deforestation Detection) system, both of which are managed by Brazilian National Institute for Space Research (INPE) (48, 49). Large scale monitoring of tree cover extent and change reached its current apogee with the Global Forest Change product. It consists of annually-updated maps of tree cover loss, available globally at 30-m resolution (50). This product, along with the GLAD near-real time tropical forest disturbance alert system (51), constitute perhaps the single most advanced tool for land cover/land use change monitoring available today.

Another important land cover/land use to monitor is cropland. Commodity crop expansion generates important commercial benefits, especially in major production countries such as Brazil and Argentina. However, expansion of cropland area in South America also generates significant externalities associated with the growth of such an intensive land use (52). While important advances have been made to improve agricultural monitoring for the purpose of ensuring food supplies and food

security (53), the objectives of agricultural monitoring for food security differ from those of monitoring cropland in the context of land cover/land use dynamics. For the former, coarser spatial resolution data with high temporal resolution (particularly, Moderate Resolution Imaging Spectroradiometer – MODIS – data) are usually employed, along with ancillary datasets such as weather and rainfall data. These data enable near-real time monitoring of the phenological development and condition of food crops, with coordinated interventions in case of production shortfalls. The Famine Early Warning system (54) is one such application, relying on low latency, regional-scale earth observation data to track crop condition during local growing seasons in developing countries.

In contrast to regional crop condition monitoring, cropland extent and change mapping requires time-series, medium to high spatial resolution data as inputs to map characterizations. Cropland mapping in South America is most advanced in Brazil, due to their institutional commitment to land cover monitoring. Cropland is one of the categories mapped in the TerraClass products developed by INPE for the Amazon and Cerrado regions (55, 56). TerraClass maps are Landsat-based and available for every two-year interval from 2004 to 2014 in the case of TerraClass Amazon and for 2013 alone in the case of TerraClass Cerrado. Other efforts to map and quantify cropland, crop types (for example soybean and sugarcane), and extensification of cropland into forests are numerous, but often limited by their geographic scale (focusing only on a single state or biome), or spatial and temporal resolution (using MODIS data or only focusing on a single or a few years)) (57–66). Outside of Brazil,

cropland extent and change mapping and quantification is similarly limited (40, 67–72).

Mapping and quantification of pasture area—the single most extensive land use in South America—through remote sensing is in its nascency. A few recent studies mapping pasture area focus on relatively small areas, such as the region of Novo Progresso in Para, Brazil (73), the Colombian Amazon (74), and the Paraguayan Chaco (75). Again, studies focusing on Brazil are more advanced, as recent projects have encompassed the entire Cerrado biome (76) and even the entire national territory (77), though only for a single point in time (2013 and 2015, respectively).

Tree plantation mapping and quantification is also limited. A recent study mapped oil palm plantations in South and Central America using MODIS data for 2014 (78), and another focused on *Eucalyptus* plantations in Brazil for 2000 to 2012 using MODIS data as well. By and large, expansion of tree plantation area in South America has not been quantified.

A few studies have sought to characterize land cover at the continental level using remote sensing data. These include a 1-km resolution map for 1995-2000 (79) and a Landsat-derived map for 2010 (80). Two other important studies sought to quantify important land cover/land use change dynamics at the continental level. One by Sy et al. (62) quantified post-deforestation land uses across the continent from 1990 to 2005 by analyzing a systematic sample of approximately 1400 10 X 10 km sample units. The other one is by Graesser et al. (81), wherein cropland and pastureland dynamics were characterized from 2001 to 2013 at multiple scales, the smallest of

which were sub-political hexagon zones of about 1,200 km^2. These continental level products are limited by their low temporal and/or spatial resolution.

The aforementioned studies are limited in one or more of the thematic, spatial, and temporal information domains in characterizing land cover and land use at the continental scale. These studies also lack a crucial component of quantification of land cover/land use change: deriving area estimates from a probability sample of reference data, as has been established by good practice guidance in the scientific literature (82, 83) and widely adopted at the international policy level (84–86). Recommended good practice methods have been developed specifically to address area estimation when remote sensing data are used as the basis of the observations of land cover. The methodology used to estimate land-cover class areas based on a probability sample is statistically rigorous with estimators having the properties of being unbiased with known uncertainty, in contrast to pixel counts, which lack such statistical properties. Maps are always biased and the pixel counting area estimates are not accompanied by quantification of uncertainties. The problems of area estimation based on pixel counts have been highlighted in recent literature, as the following two passages attest:

> "Since the late 1970s the serious risk of bias of direct area measurement or pixel counting on classified or photo-interpreted images is clearly identified and solutions are proposed on the basis of traditional statistical survey techniques." (Gallego et al. 2014, p. 22) (87)

> "Thus, if bias and precision have not been assessed and reported, regardless of the reason, then maps may have little or no utility for scientific inference. The conclusion, which will certainly be disconcerting and unpleasant to some, is that if estimates of population parameters derived from maps cannot be characterized in probabilistic terms the map cannot serve as a basis for scientific inference; from an inferential perspective, the map may be little more than just a pretty picture." (McRoberts 2011, p. 716) (88)

Estimating areas from a probability sample of reference data as opposed to doing so from pixel-counting from a map ensures unbiased area estimates with known uncertainty as quantified by the standard errors associated with the area estimates. Another important advantage of using a sample of reference data to provide area estimates is that additional characteristics of land cover/land use dynamics can be established during the sample interpretation process, thereby adding valuable information that cannot easily be mapped. For example, during sample interpretation we can assign a per-sample-unit date of change, previous land cover type, driver of change, etc., from which we can later derive additional area estimates.

One project that seeks to map and quantify land cover and land use change at biome to national to continental scales addresses the limitations of other studies mentioned thus far. The MapBiomas project (89) is the product of a collaborative Brazilian initiative made up of NGO's, universities, and private companies that have joined together to create an all-encompassing land cover/land use monitoring platform. They have published yearly Landsat-derived maps of land cover and land use for 1985 to 2018, which are openly available through their website (mapbiomas.org). These maps include forests of different types (forest, savanna, mangrove), as well as tree plantations, pasture, cropland, and other natural vegetation types. As part of this project, the LAPIG lab at the Goiás Federal University (UFG) undertook the task of creating a sample-based reference dataset for area estimation and validation of the MapBiomas maps. Because Brazil makes up an important area of the South American continent and is one of the most active regions in terms of

changes in land cover/land use, MapBiomas represents a significant advancement in terms of monitoring capabilities in the region.

However, the project does have a few technical deficiencies. The mapping is done on a per-scene basis, which creates issues in terms of the spatial continuity of the maps. Further, the sample-based analysis does not use consistent land cover/land use definitions across the entire national territory (e.g., what is considered a forest in the Amazon is not the same as what is considered a forest in the Cerrado biome). Also, neither the map nor the sample-based results are framed from a perspective of assessing land cover/land use change directly. Rather, they map and quantify land cover/land use per year and rely on users to assess change by subtracting results, which is not ideal due to the potential of error propagation.

Thus, opportunities to improve on certain aspects of the MapBiomas project remain. Further, no continental-level assessment of land cover/land use change for the temporal scale permitted by the Landsat archive exists, especially as it relates to area estimates derived from a probability sample, as established by good practice recommendations.

1.2 **Research goal and dissertation structure**

Understanding the impact that we have on our surrounding environment is critical in order to ensure a sustainable future, yet the extent of major land cover/land use transformations in one of the most important regions in the world is still unknown. The overarching goal of this dissertation is to advance the scientific knowledge of major land cover/land use dynamics across the South American continent from 1985 to 2018, focusing especially on mapping and quantification of commodity land use

extent and its expansion onto natural land. Specifically, this dissertation seeks to combine established methods for large area land cover/land use characterization using remote sensing data with recommended statistical approaches to determine area estimates of change dynamics and associated uncertainties. Commodity land uses of interest herein are cropland, pastureland, and tree plantations. The general approach used to achieve the stated goal is to start by creating maps of land cover/land use themes, which are then used as stratification for the targeted allocation of a probability sample of reference data. Reference data are then interpreted and area estimates and associated uncertainties are derived. This general approach is applied to the theme of cropland expansion in Brazil in Chapter 2, and to a broader set of land cover/land use themes across South America in Chapters 3 and 4.

The objective of the research presented in Chapter 2 is to employ the joint methodology of mapping and sampling for area estimation to study Brazilian cropland expansion. In this chapter, Brazilian cropland extent in the year 2000 is mapped, as is its subsequent expansion between 2000 and 2014. A stratified random sample of 5,000 units (Landsat pixels) is used as reference data, from which area estimates of cropland expansion by year and by previous land cover type are derived. This chapter serves as a prototype of the general approach applied over a large area for a very relevant land use theme. Thus, its results serve two purposes: it is an example of the applicability of the method, and it advances our understanding of cropland dynamics in Brazil.

In Chapters 3 and 4, the general approach of mapping followed by sampling is applied once again, but at much broader geographic, temporal, and thematic scales.

Thirty-four years of Landsat data are leveraged to map cropland, plantation, natural tree regrowth, and stable land cover. The resulting maps are joined together to stratify the continent and allocate samples based on areas of likely presence or absence of commodity land use. A reference sample of 1,000 units (Landsat pixels) is interpreted, and land cover/land use themes are attributed (beyond those that were specifically targeted through the stratification), thus enabling a continental-scale analysis of land change.

Chapter 3 presents the resulting area estimates of land cover/land use change dynamics across the continent for 34 years. Important classes for which we present estimates of extent and change include tree cover, pasture, cropland, plantation, natural tree regrowth, tree cover degradation, and semi-natural land (a class that is largely missing from the scientific literature). Chapter 3 highlights the growing human impact on natural land cover in South America and how this impact has dramatically increased through the years.

Chapter 4 is a thorough explanation of the methodology employed to leverage the 34-year Landsat archive to map land cover and land uses across South America for the entire study period. Chapter 4 also presents several evaluations of the map and the area estimates, the purpose of which is to assess the performance of the maps, both as a spatially explicit characterization of land cover and land use, and as a tool to facilitate area estimation of continental-scale land cover and land use change. Chapter 4 concludes with a discussion of lessons learned for estimating area for a multitude of land cover/land use themes using a limited set of map strata.

Chapter 5 summarizes this dissertation's contributions to advancing our understanding of land cover/land use change in South America, outlines some of the outstanding issues related to the dissertation research, and presents potential future research directions.

Chapter 2: Near doubling of Brazil's intensive row crop area since 2000[1]

2.1 **Abstract**

Brazil has become a global leader in the production of commodity row crops such as soybean, sugarcane, cotton, and corn. Here, we report an increase in Brazilian cropland extent from 26.0 Mha in 2000 to 46.1 Mha in 2014. The states of Maranhão, Tocantins, Piauí, Bahia (MATOPIBA), Mato Grosso, Mato Grosso do Sul, and Pará all more than doubled in cropland extent. The states of Goiás, Minas Gerais, and São Paulo each experienced >50% increases. The vast majority of expansion, 79%, occurred on repurposed pasture lands, and 20% from the conversion of natural vegetation. Area of converted Cerrado savannas was nearly 2.5 times that of Amazon forests, and accounted for over half of new cropland in MATOPIBA. Spatio-temporal dynamics of cropland expansion reflect market conditions, land use policies, and other factors. Continued extensification of cropland across Brazil is possible and may be likely under current conditions, with attendant benefits for and challenges to development.

2.2 **Introduction**

Growing demands in national and international markets for commodity crops drives increasing production through more intensive management practices,

[1] The presented material has been published in Zalles, V., Hansen, M.C., Potapov, P.V., Stehman, S.V., Tyukavina, A., Pickens, A., Song, X.P., Adusei, B., Okpa, C., Aguilar, R. and John, N., 2019. Near doubling of Brazil's intensive row crop area since 2000. *Proceedings of the National Academy of Sciences*, *116*(2), pp.428-435.

extensification through land conversion, or both. China's soybean imports, for example, increased from just under US$2 billion in 2000 to US$35 billion in 2014 (90). This demand has led to dramatic production increases in countries such as Brazil (91–93), which has become a global leader in the cultivation of soybeans, as well as sugarcane, corn, and cotton (90). Intensification of existing agricultural land uses, such as the conversion of pasture to cropland, and extensification of agro-industrial cropping systems through the conversion of natural vegetation result in numerous externalities, including increased runoff of fertilizers and pesticides, overutilization of freshwater resources, greenhouse gas emissions, and biodiversity loss (52, 94). Knowing where croplands are expanding, their rate of expansion, and the land covers that they are replacing is essential to quantify current and model future environmental impacts. Improved information on cropland extensification also facilitates the study of supply chains and their respective economic and institutional contexts (95).

In Brazil, the topic of cropland expansion is particularly salient. Advances in technology, market liberalization policies, government subsidies, and favorable international prices accelerated the development of the cropland frontier. As production methods matured and soybean proved more profitable than cattle, soybean expansion was accelerated by increasing economies of scale (96–99). Research on land use and land cover change associated with cropland expansion in Brazil is extensive in the literature, but often limited in geographic or thematic scope. The main research focus has been on answering the question of whether crop expansion is a proximate driver of deforestation. Accordingly, there is a strong bias in the research literature towards the Amazon biome and the state of Mato Grosso (60, 91, 100–103),

where the dominant theme is deforestation driven by soybean expansion. The Cerrado biome, a biodiversity hotspot (36, 104), has recently become the focus of attention due to the rapid expansion of cropland in the region of MATOPIBA (acronym for the names of the four states that compose this region: Maranhão, Tocantins, Piauí, and Bahia) (57, 105–107). A number of studies have focused on São Paulo and Goiás, two states in the south-central region of Brazil that have been the site of dramatic expansion of sugarcane for biofuel production (64, 66, 108). However, few studies quantify changes in crop area at the national scale. Furthermore, most of the spatially explicit studies have employed coarse spatial resolution MODIS data (60, 91, 102, 103, 106, 107), limiting accurate cropland area estimation, particularly in the south of the country where relatively smaller field sizes are predominant. A few studies have used census data provided by the Brazilian Institute of Geography and Statistics (IBGE) to characterize changes in cropland area at the national scale, but the last agricultural census was carried out in 2006. Another common data source used (100, 109) is IBGE's SIDRA database, which provides crop areas estimated by experts, and as such are subject to inconsistencies through time. Products such as TerraClass (55, 110), TerraClass Cerrado (56) or Canasat (111) at medium spatial resolution are also limited in temporal and geographic scale. A new project focusing on mapping at biome scale for the entire record of Landsat – MapBiomas (89) – holds promise. However, no published area change studies have followed good practice guidance (82–88) in which a sample of reference data is used to provide an unbiased area estimate of cropland cover expansion accompanied by a standard error that quantifies the uncertainty of the area estimate. As a result, a comprehensive and definitive

national-scale record of yearly land cover changes related to cropland expansion in Brazil is lacking.

Remote sensing data provide a unique resource for measuring such changes consistently over space and through time, facilitating a common understanding between policymakers, civil society, scientists, and private industry. For this study, cropland is defined as the land area under intensively managed, agro-industrial row crops consisting of commodity crops such as soybean, sugar cane, cotton, corn, rice and wheat. We employ 30m spatial resolution Landsat data to estimate cropland extent in the year 2000 and its subsequent expansion through 2014. A probability sample of reference data allows us to report unbiased estimates of national, biome, and state-scale area of crop expansion with associated uncertainties. Through our sample assessment, we disaggregate crop expansion by year and by previous land cover type to produce estimates of temporal trends of area of crop expansion by re-purposing of pastures (defined as lands dominated by herbaceous cover used for grazing livestock) and by conversion of natural vegetation cover. These results represent definitive, precise, and unbiased estimates of national-scale cropland expansion in Brazil.

2.3 **Results**

Cropland extent in the year 2000 in Brazil was 26.0 ± 1.1 Mha (the uncertainty is expressed as ± one standard error of the estimate). In the subsequent 14-year period, cropland expanded by 20.5 ± 1.6 Mha, representing a 79% increase relative to the year 2000 cropland area. We define the states that more than doubled their respective cropland area since 2000 as constituting the cropland frontier: Mato

Grosso, Mato Grosso do Sul, Pará, Bahia, Maranhão, Piauí, and Tocantins (Fig. 2-1a). Cropland loss was limited to 0.7 ± 0.1 Mha for the entire country during the study period. See SI Appendix, Table 1 for accuracy assessment results and SI Appendix, Fig. 2-1 for classification results. See Methods section for detailed information on reference data interpretation.

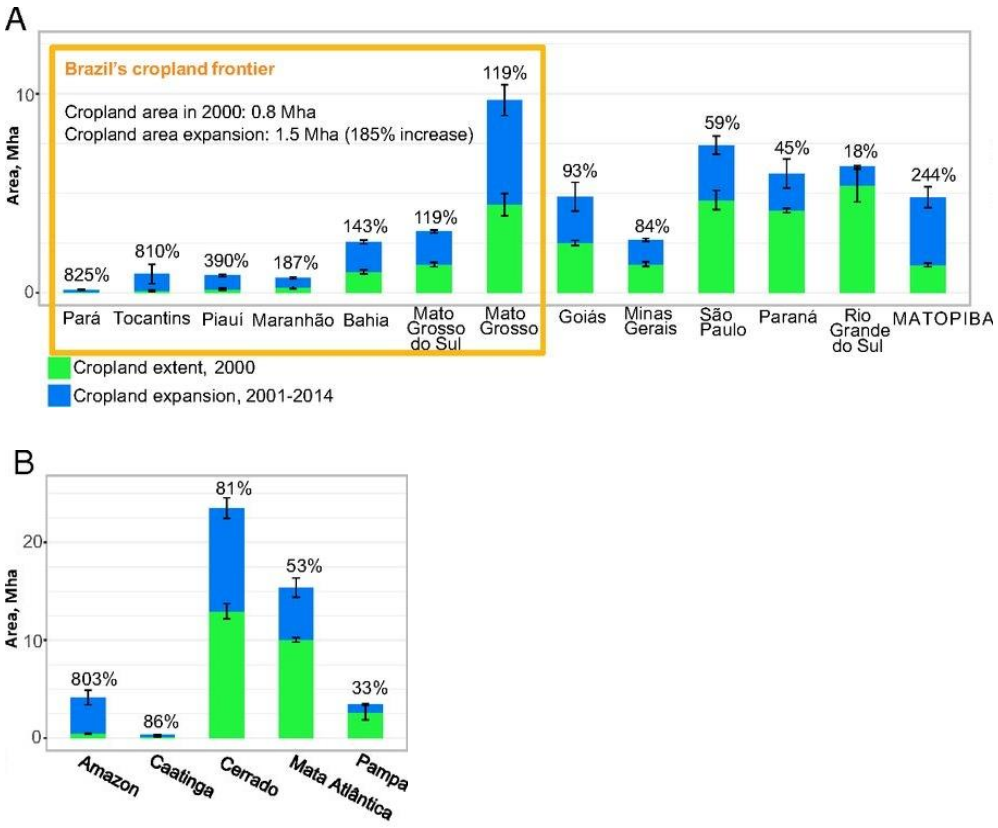


Figure 2-1. Estimated area of cropland extent in 2000 and area of cropland expansion from 2001 to 2014. **(**a) Brazilian states, (b) Brazilian biomes. See SI Appendix, Fig. 2 for the location of states and biomes and for the area included in MATOPIBA. Numbers on top of bars indicate percent increase in cropland area since 2000. Error bars represent ± one standard error. Estimates are presented only for states and biomes that have >10 sample pixels in the "cropland expansion" strata.

The state with the largest area of new cropland was Mato Grosso, with 4.4 ± 0.5 Mha of cropland in 2000 and 5.3 ± 0.8 Mha of cropland gain through 2014.

Cropland expansion in Mato Grosso represents 26% of the total cropland expansion area in the country. The biome with the greatest area of new cropland was the Cerrado, with 10.5 ± 1.0 Mha of additional crop area by the end of the study period (81% increase from 2000). Cerrado cropland expansion represents 52% of the total expansion in the country (Fig. 2-1b).

Brazilian cropland expanded rapidly and peaked during the 2004/2005 growing season, followed immediately by a sudden and pronounced drop in annual expansion area (Fig. 2-2). After a low in 2006/2007, the rate of cropland expansion by 2013/2014 approached that of the 2004/2005 peak. The rapid increase through 2004/2005 and subsequent rapid decline of cropland expansion area was most pronounced in the states of Mato Grosso and MATOPIBA, and the Amazon and Cerrado biomes (SI Appendix, Figs. S2-10 and S2-11). Nearly every state and biome for which we have data available experienced a drop in cropland expansion in 2004 (SI Appendix, Figs. S2-10 and S2-11). Since the decline in the 2004/2005 growing season, the rate of crop expansion has steadily increased in most states, with Mato Grosso do Sul, Minas Gerais, Goiás, and Piauí having the most rapid increase in cropland area after 2005 (SI Appendix, Fig. S2-10). Every state and biome exhibited a peak in cropland expansion between 2011 and 2014 except for Maranhão and the Caatinga biome (SI Appendix, Figs. S2-10 and S2-11).

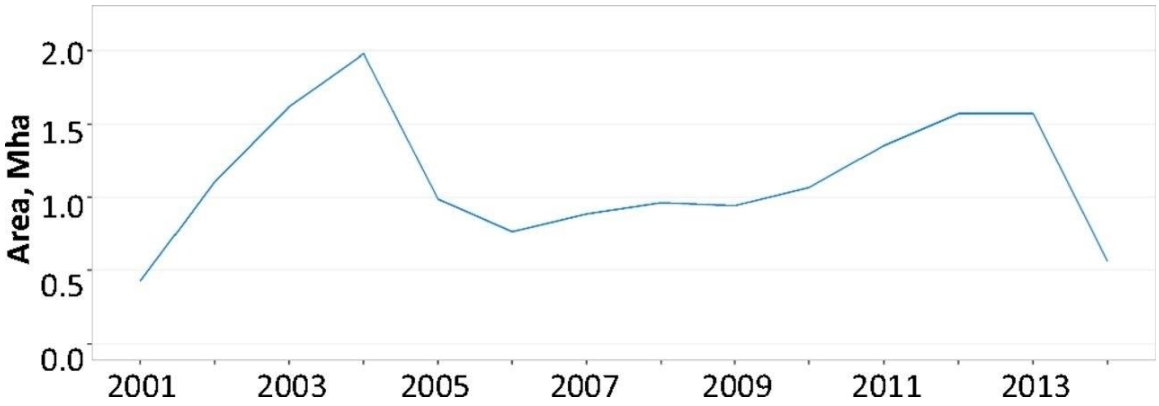


Figure 2-2. Estimated annual cropland expansion area in Brazil from 2001 to 2014. Yearly trends are based on "cropland 2000" and "cropland expansion" strata. Year of expansion corresponds to year of planting (e.g., 2001 corresponds to the 2001/2002 growing season). Sample pixels from the "no cropland" strata add 4.7 ± 1.6 Mha to the total cropland expansion area shown here; this area is not displayed in the figure as it is not representative of yearly trends. See Dataset S1 for tabular data for all strata.

Pasture conversion was the source of nearly 79% of new cropland area in Brazil and 20% the result of conversion of natural vegetation, including Amazon humid tropical forests and Cerrado dry tropical woodlands and savannas. Only 1% of the total expansion area was created through the conversion of tree plantations. The overall proportion of cropland expansion within natural vegetation remained relatively constant at ~20% throughout the study period, though with substantial regional variation. The MATOPIBA region had the largest proportion of natural vegetation conversion to cropland (57% ± 15%), consisting largely of Cerrado conversion (Fig. 2-4 and SI Appendix, Fig. S2-10). In the Amazon biome, 30% ± 2% of new cropland resulted from natural vegetation conversion, primarily of dense humid tropical forests (Fig. 2-4). The southern states of Mato Grosso do Sul, Paráná, Rio Grande do Sul, and São Paulo expanded their cropland area mostly through the conversion of pastures (99%, 99%, 88%, and 93% respectively). Note that the areal

increase of one crop, for example sugar cane (112), at the expense of other row crops would not meet our definition of cropland gain. Summary statistics and time-series graphs of cropland gain for all states and biomes having at least 10 sample pixels in the "cropland expansion" class are shown in Dataset S1 and SI Appendix Table 2, and Figs. 2-4 and 2-5. See SI Appendix, Fig. S2-8 for a list of states and biomes for which we estimate cropland expansion areas.

2.4 **Discussion**

2.4.1 Comparison with existing datasets

Our results differ from existing estimates on cropland area and cropland area expansion in several ways. Table S2-3 of the SI Appendix provides a comparison of the technical characteristics of our results and other available studies and data sources. Our study advances current knowledge on Brazilian cropland extensification due to its spatial extent (we present results at the national level but also disaggregate by states and biomes), its temporal extent (comparable to Mapbiomas), and, most importantly, because it adheres to "good practice" recommendations (82–88) on area estimation and accuracy assessment. Unlike previous research, our study uses a probability sample of reference data for area estimation and provides uncertainty estimates (i.e., standard errors) for the area estimates. Lastly, our results provide information on pasture conversion to cropland, which is largely lacking in the literature.

Official estimates of cropland area available through IBGE's SIDRA database are widely used in the literature to study land use changes in Brazil (59, 92, 100, 113).

These data are not directly comparable to our results because IBGE's area numbers double-count the area of a field if it is double-cropped. Dias et al. (2016) use these numbers in their study and therefore significantly over-estimate cropland land use area in Brazil (SI Appendix,Fig. S2-15). Barona et al. (2010) also cite double-cropping as a possible source of overestimation of cropland area in their analysis.

In order to compare our results to data from the IBGE SIDRA database, we tried to approximate an estimate of cropland cover area based on their "planted area" metric by removing the area of secondary crops as well as areas of crops that do not fit our cropland definition (intensive row crop agriculture). To do this, we started out by adding together the areas of Brazil's most important crops: soy, corn, sugarcane, beans, rice, wheat, manioc, and cotton. These eight crops make up 95% of the total crop planted area in Brazil. We then removed the area of 2nd crop corn as well as 2nd and 3rd crops of beans. Although cotton is also used a secondary crop in crop rotations, data on cotton as a second crop are not available through the IBGE SIDRA database, so we included all of the cotton planted area in our area estimate. We also subtracted wheat area because wheat is a winter crop that is almost exclusively double cropped. Finally, we removed the area of planted manioc because manioc production in Brazil is mostly small-scale and non-intensive, which excludes it from our cropland class definition (it is not produced as an intensive row crop). The result, which we refer to as the IBGE LC estimate (because it is the IBGE cropland Land Cover estimate), corresponds to 35.7 Mha in 2000 and 52.5 Mha (SI Appendix, Fig. S2-15).

These estimates are higher than the ones we present in our study. There are many possible reasons why IBGE LC estimates may differ from ours. Area estimates provided by IBGE are the result of expert surveys and, as such, they are to some degree inherently inconsistent across space and time. Additionally, IBGE does not provide any indication of the accuracy or the uncertainty of their statistics. As a result, IBGE statistics may not always be the most appropriate data source for land cover change studies related to changes in cropland area in Brazil. Indeed, several authors have pointed to the limitations of IBGE statistics and called for the need for higher quality cropland maps for Brazil (59, 92, 100). A thorough description of the comparison between the dataset presented herein and IBGE data is provided in the Appendix of this dissertation).

Another important dataset that holds promise for cropland extent and expansion monitoring is MapBiomas (89). MapBiomas provides Landsat-based maps of land cover disaggregated into five broad categories (and up to fifteen detailed categories) for every year from 1985 to 2017. One of these categories is “farming” which they disaggregate into “pasture”, “agriculture”, and “agriculture or pasture” for areas of confusion between the two. We compared their results from the “agriculture” category to our results and found that their results approach ours. At the national level, we report lower area estimates than they do, in Mato Grosso, their results are similar to ours, and in the Cerrado biome, we report higher area estimates (SI Appendix, Fig. S2-15). Their results for cropland expansion diverge substantially from our results (SI Appendix, Fig. S2-14). The main limitation of the MapBiomas product is that they do not follow current good practice guidance (82–88) which

recommends estimating area from the reference sample observations and assessing accuracy of the mapped land cover change. The latest version of the MapBiomas project (Collection 3.0) does not yet have an accuracy assessment of any type.

Additionally, the TerraClass Amazon and TerraClass Cerrado products provide data on land cover at Landsat resolution for the Amazon and Cerrado biomes respectively. The limitations of these products compared to the results obtained through this study are that 1) maps are available only for certain years; 2) they do not provide accuracy assessment of change classes; 3) they do not employ good practice recommendations (82–88) for area estimation and associated uncertainties. IBGE also provides data on cropland area through their "Systematic Monitoring of Land Use" project (114) , which maps land use and land cover change in Brazil for the years 2000, 2010, 2012, and 2014. This product has the same limitations as the ones listed for the TerraClass products, amongst others (an additional limitation of this product is its minimum mapping unit of 62.5 ha). (SI Appendix, Table S2-3).

Comparisons of cropland expansion, total cropland area, and conversion of natural vegetation to cropland between our study and other studies and datasets (10, 60, 89, 91, 92, 106, 107, 114–116) are provided in the SI Appendix, Figs. S2-14, S2-15, and S2-16.

2.4.2 Trends in cropland expansion

National-scale dynamics of cropland expansion in Brazil from 2000 through 2014 reflect an early peak in the 2004/2005 growing season, followed by a sharp decline and subsequent gradual recovery to near 2004/2005 levels by 2013/2014. Here, we discuss a number of policy, management, and economic factors that may

have played a role in shaping trends of cropland expansion in the region. Establishing cause-and-effect relationships between these factors and the land cover changes discussed requires further research which would be enabled by accurate cropland expansion area estimates such as presented in this study.

The 2005/2006 decline in crop expansion in Brazil coincides with a period of unfavorable market conditions (91, 117–120). A decrease in soybean prices, the appreciation of the Brazilian real relative to the US dollar, and an increase in costs of production linked to high oil prices, caused soy profits to decline dramatically from 2004 to 2005. As a result, farmers in Mato Grosso were faced with negative profit margins for soybean production in 2005 and 2006, which might have disincentivized expansion (Fig. 2-3). Added to these economic factors was a severe drought (120, 121). Our estimates of annual cropland expansion in Mato Grosso closely mirror data on annual soybean profit (Fig. 2-3). The largest residual is related to the period of greatest expansion in 2004, with dramatic declines in both profits and expansion the following year. Peak cropland expansion post-2004 is observed in the 2013/2014 growing season, the year of greatest soybean profit during the study period for Mato Grosso. Morton et al. (2006) and Macedo et al. (2012) have cited soy profitability as a potential influencing factor on trends of forest conversion to cropland (see SI Appendix, Fig. S2-16 for a comparison of their results to results from this study). Our results support this hypothesis.

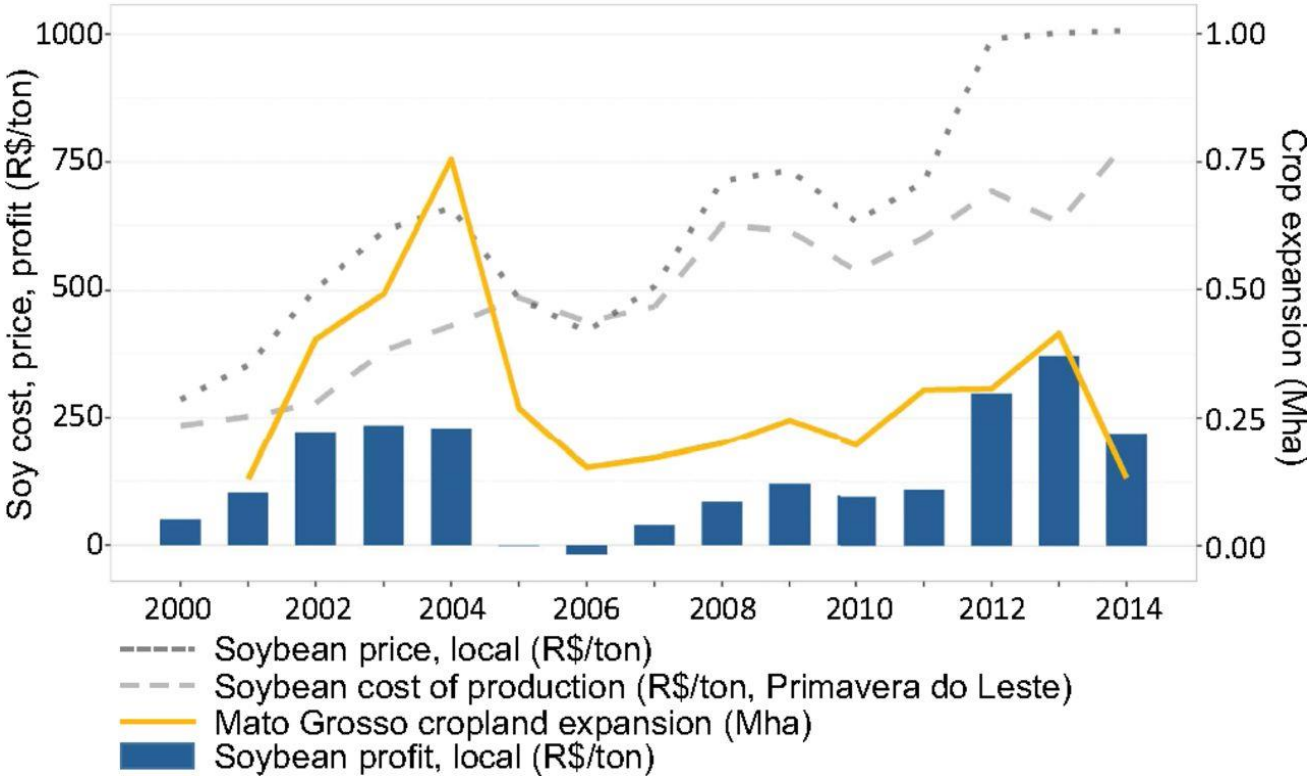


Figure 2-3. Soybean terms of trade in Mato Grosso. Mato Grosso cropland expansion from this study is compared to soybean price and cost of production. Soybean price is the nominal producer price, obtained from FAOSTAT (42). Soybean cost is from CONAB (122). Mato Grosso cropland expansion is derived from the sample-based area estimate for the "cropland 2000" and the "cropland expansion" strata (see Dataset S1 for tabular data). Year of expansion corresponds to year of planting (e.g., 2001 corresponds to the 2001/2002 growing season). FAOSTAT and CONAB data display adapted from Arvor et al. (115).

Humid tropical forests in the Brazilian Amazon have experienced the highest rates of deforestation globally in recent decades (50, 123). Drivers of deforestation include pasture land use for beef production and cropland land use for soybean production. Due to the extraordinary ecological significance of the Amazon biome, international attention and national policies have focused on slowing deforestation, with unprecedented success (59, 107, 109, 117). A number of policy initiatives and supply chain interventions have contributed to the reduction of deforestation in the

Brazilian Amazon. These include an increased capacity for enforcement of the forest code by the government through the implementation of the "Detection of Deforestation in Real Time" (DETER) program in 2004 (124), the implementation of an "Action Plan" (PPCDAm) allowing coordination amongst agencies and ministries at the federal level to combat deforestation in 2004 (59), the rapid expansion of the protected area network starting in 2002 (125), and a successfully implemented multi-stakeholder moratorium on sourcing soybeans from newly deforested lands, starting in 2006 (65, 107, 117).

We find that cropland expansion into forests in the Amazon began to slow in 2004/2005, reflecting a possible response of land owners to policies (and the anticipation of pending policies), market conditions, or both (Fig. 2-4). After 2006, conversion of forests to cropland in the Amazon remained consistently low. This result supports existing findings on the decrease of cropland expansion into deforested areas during this time period (60, 91) and has been linked to the Soy Moratorium (107, 117). At the same time, conversion of pastures to cropland began to rise. The primary target area for the Soy Moratorium, the state of Mato Grosso, experienced decreased clearing of natural vegetation for cropland after 2004 (Fig. 2-4). Cropland expansion within natural vegetation in MATOPIBA, a region that is outside the reach of the Soy Moratorium, did not experience a similar sustained decline, and increased slightly over the study period (Fig. 2-4). The trends in converting pastureland to cropland also differ, with Mato Grosso experiencing a dramatic increase over time following a minimum expansion area within pastureland in 2006/2007.

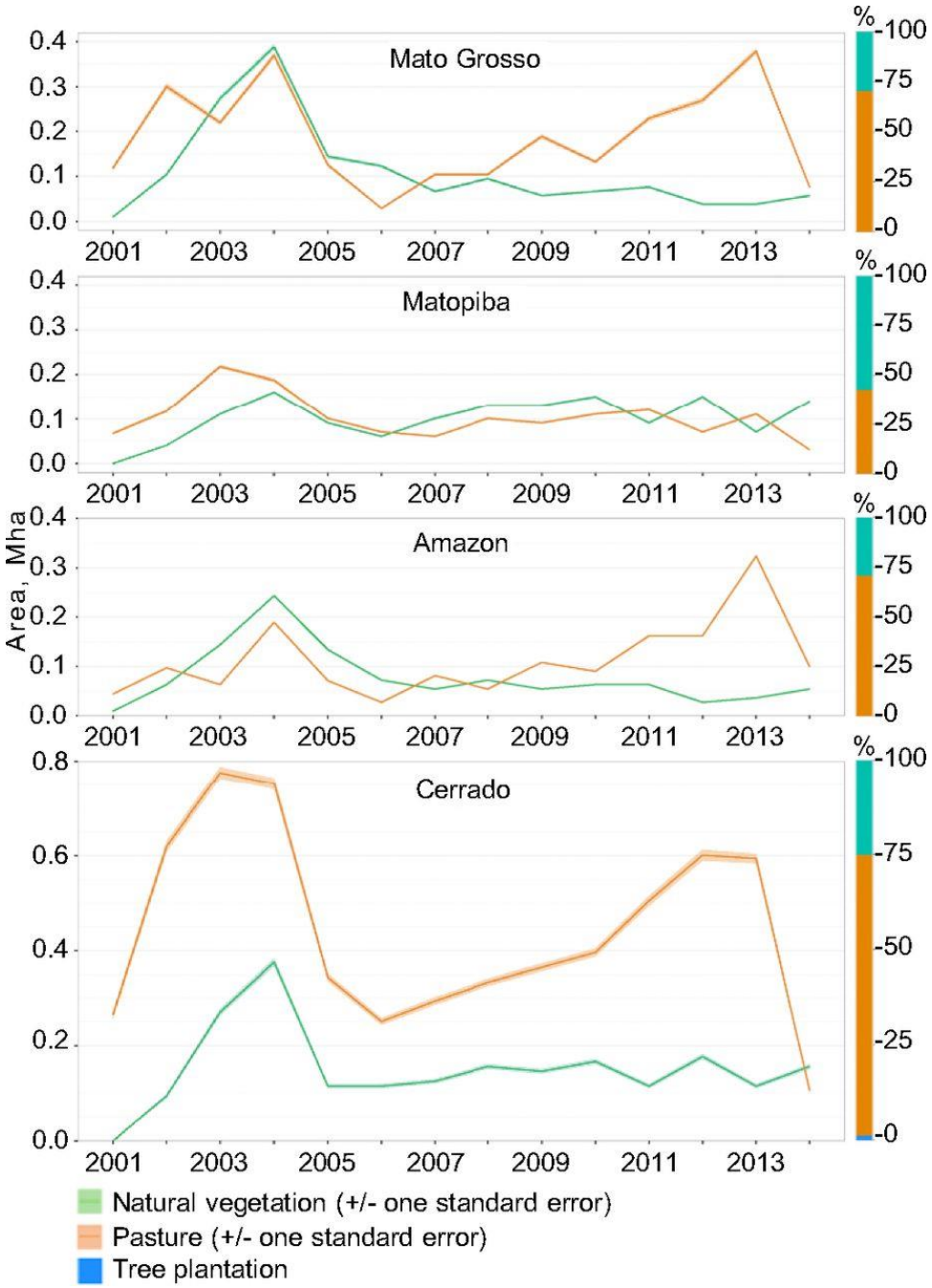


Figure 2-4. Trends in cropland expansion disaggregated by conversion from pasture and natural vegetation for Mato Grosso, Matopiba, the Amazon, and the Cerrado. Bars on the right represent cumulative share of pasture and natural vegetation as source of new cropland for 2001-2014. Trends shown reflect sample-based area estimates of cropland expansion for "cropland 2000" and "cropland expansion" strata. Year of expansion corresponds to year of planting (e.g., 2001 corresponds to the 2001/2002 growing season). Sample pixels from the "no cropland" strata are not displayed. See Dataset S1 (available through the pnas.org website) for tabular data for all strata.

Two possible impacts of the regulatory measures implemented in the Amazon (such as the Soy Moratorium and other public policy initiatives) are observed in Fig. 2-4. First, the ratio of new cropland converted from pastureland versus converted from natural vegetation for Mato Grosso increases from 1.1:1 from 2001-2004 to 4.3:1 from 2011-2014, reflecting the strategy of adding soybean area within already deforested lands. Second, the same ratios for MATOPIBA change from 1.3:1 to 0.7:1, possibly reflecting leakage of cropland expansion pressure to a region that is largely unconstrained by regulatory limits. The potential for leakage of cropland expansion from the Amazon to the Cerrado's MATOPIBA states has been discussed in the literature (106, 107), but there has been limited evidence until now due to the paucity of spatio-temporally consistent cropland datasets for both regions. Determining whether there is a cause and effect relationship between policies aimed at slowing humid tropical deforestation and increased clearing in MATOPIBA requires additional study. It is indeed possible that the conversion of natural vegetation areas in MATOPIBA would have occurred regardless of policies in the Amazon due to favorable market conditions, infrastructure development, or land suitability.

By combining the Global Forest Change maps (50) with the cropland expansion map, we are able to observe regional patterns of forest conversion to cropland over the study period (Fig. 2-5). The resulting map illustrates the decline in the conversion of tree cover (defined as >=5m trees and >=30% tree canopy cover) to intensive cropland within the Amazon after 2005 and a corresponding increase in the conversion of tree cover to cropland within the Cerrado starting in 2006. The spatial pattern and temporal dynamics are confirmed through our probability sample

assessment in estimating natural vegetation cover conversion (Fig.2-6). The conversion of low/no tree cover Cerrado vegetation in both Mato Grosso and MATOPIBA is substantial and not captured in the global forest loss data (SI Appendix, Fig. S2-8). This result highlights the need for spatially explicit maps of natural shrublands and non-woody vegetation cover types in addition to tree cover in assessing the impacts of cropland expansion on natural ecosystems such as the Cerrado.

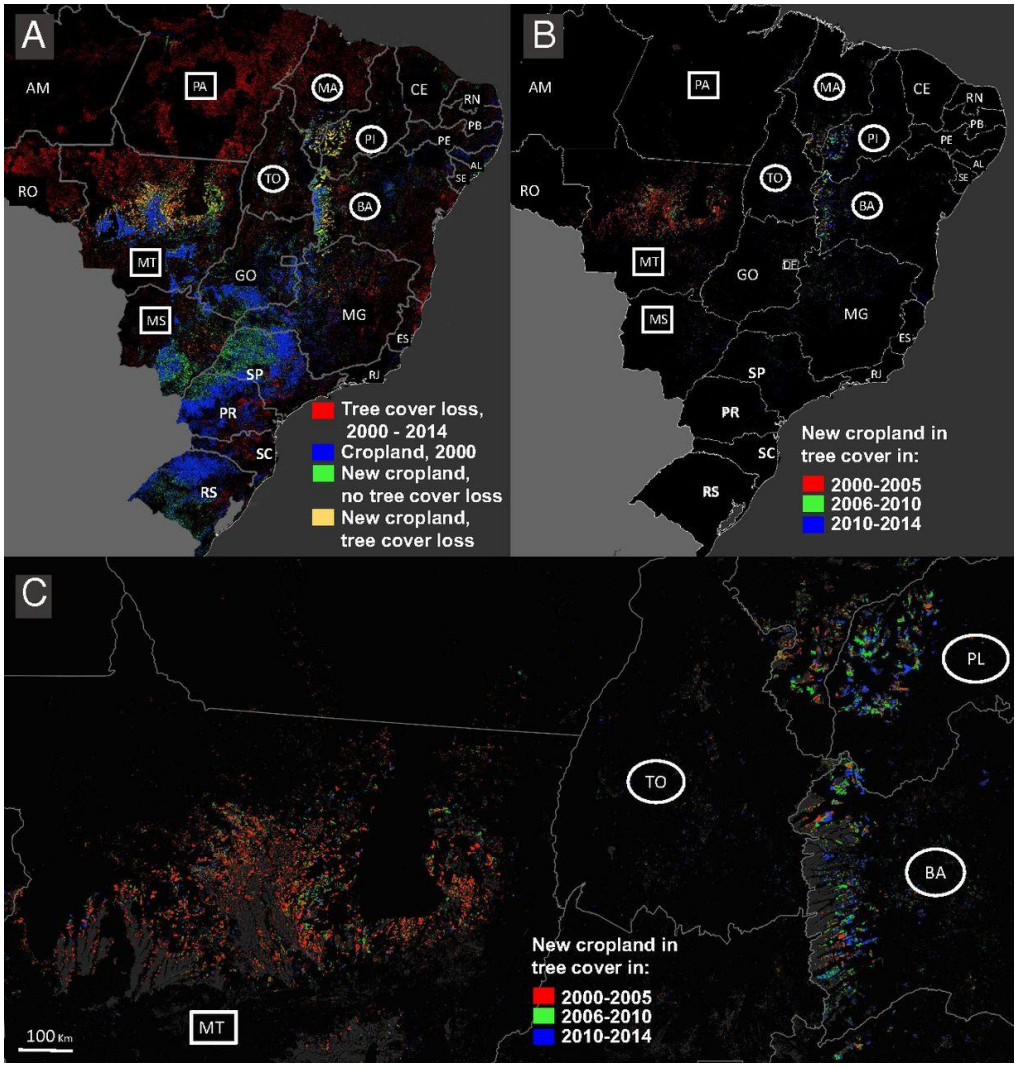


Figure 2-5. Regional patterns of forest conversion to cropland. (a) Cropland extent in 2000 (green), cropland gain outside (blue) and inside of tree cover (magenta) through

2014, and tree cover loss unrelated to cropland expansion (red). (b) Cropland gain inside tree cover disaggregated by epoch. MATOPIBA states are shown in ellipses, and other states having cropland increases of greater than 100% are shown in boxes. (c) Subset of (b) centered on Mato Grosso and MATOPIBA states.

Another factor probably impacting cropland dynamics has been the advent and spread of soybean rust. At the beginning of the study period, Brazilian farmers were 'unaware of the presence' (121) of the fungus which left them unprepared to manage its effects. Year on year increases in lost production reached a peak in 2004 with 4.6M tons of grain lost (121). Formal interventions to limit soy rust included new planting strategies such as the implementation of an annual 90-day soybean-free period starting in 2007 and 2008. Fungicide treatments in combination with double cropping practices and the introduction of new soybean varieties have further reduced soybean rust losses (121). The role of soybean rust in mediating investment in new croplands during the study period must be considered along with other factors.

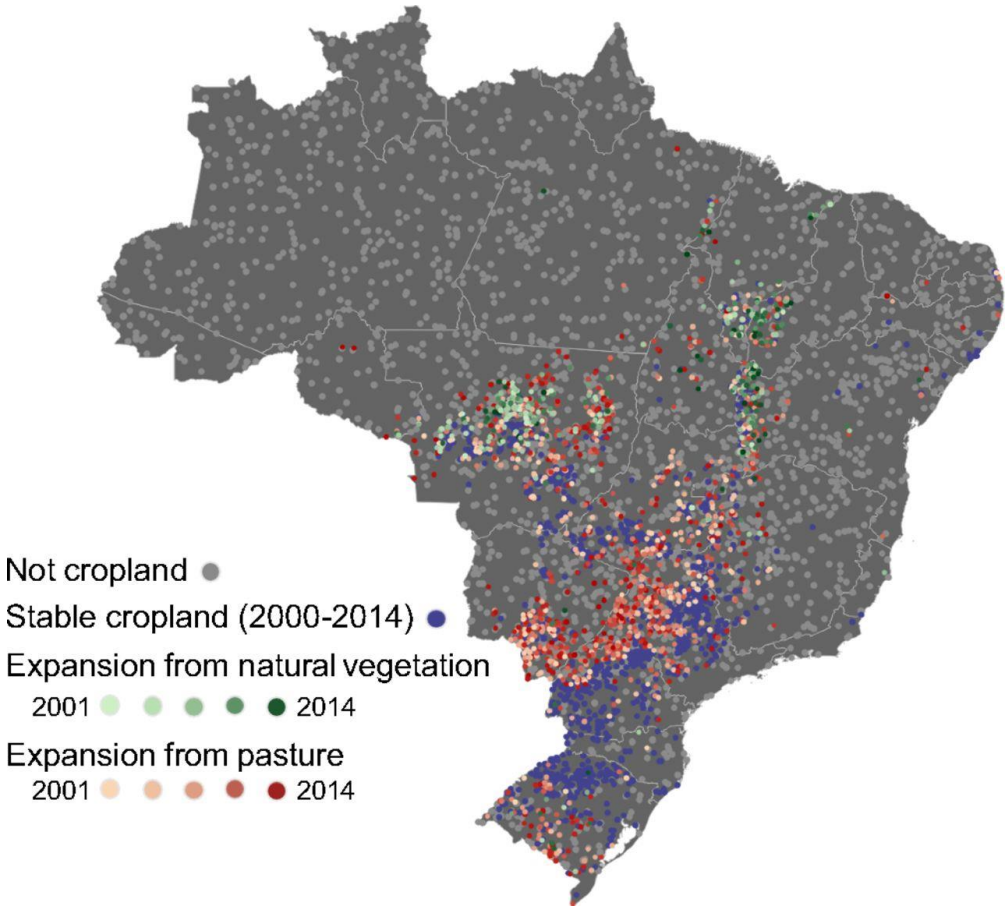


Figure 2-6. Geographic distribution of the 5000 sampled pixels. Classified by reference cropland type (stable/expansion/not cropland), previous land cover type and year of change.

Cropland expansion is not limited to the cropland frontier states where cropland area more than doubled. Even historically established agricultural states experienced substantial increases in crop area. In absolute terms, São Paulo, Goiás, Paraná, and Mato Grosso do Sul each followed Mato Grosso and MATOPIBA in area of new cropland. The Mata Atlântica biome with 5.4 ± 1 Mha of new cropland was second to the Cerrado in total area of cropland area increase, reflecting a dramatic repurposing of pasture land uses. Just over 1% of Mata Atlântica cropland expansion consisted of conversion of natural vegetation. However, cropland expansion in

Brazil's southern states has been linked to deforestation in the Amazon through the displacement of cattle ranching activities (126, 127), which would indirectly increase the environmental costs of this type of land cover change. Results for the Mata Atlântica and Pampas reveal that despite substantial intensification in recent years (92, 93), cropland extensification remained a potential pathway for increased crop production across Brazil during 2000-2014. States experiencing nascent agricultural investment, such as Roraima and Amapá (128), represent the next potential frontier of Brazilian cropland expansion (we do not have cropland area estimates for these regions because they did not have substantial enough cropland areas during our study period).

As Brazil's agricultural sector grows in response to internal and global market demands, accurate and transparent geospatial data depicting this dynamic are needed. In this study, we have presented unbiased and precise estimates of Brazilian cropland expansion area nationally and at the scale of major production states and biomes. These methodologically consistent estimates along with our corresponding spatio-temporal data (i.e., maps of 2000 cropland and 2000-20014 cropland expansion) contribute to enhanced understanding of the economic, policy, social and environmental drivers and outcomes of the rapid and large-scale expansion of agro-industrial land uses. Our results for the dynamic time period of 2000-2014 reflect the dramatic growth of commodity crop land use in Brazil driven primarily by repurposing pasture land and converting natural vegetation. Extending these analyses to the beginning of the Landsat record (circa 1984) and forward in time will provide estimates and data that can be used to gain further insight regarding the response of

cropland expansion to market conditions, disease, and other factors, as well as the impact of land use policies in redistributing expansion pressures.

2.5 **Methods**

Landsat time-series data were used to map Brazil into the following categories: 2000 cropland, 2000-2014 cropland gain, and no cropland. The mapped classes were used as an input to a stratified random sample of reference data consisting of MODIS, Landsat and GoogleEarth imagery, in order to estimate the area of year 2000 cropland and 2000-2014 cropland expansion.

2.5.1 Landsat data

Two sets of Landsat data were used to create the maps: all available Landsat 7 ETM+ data for 1999 to 2001 and all available Landsat 7 ETM+ and Landsat 8 OLI for 2011 to 2014. All the images were downloaded from the USGS Earth Resources Observation and Science Center (EROS) in the L1T terrain-corrected format. Inputs for the land cover classification were derived from spectral bands that are not as sensitive to atmospheric contamination and scattering(129): red (ETM+ 0.630–0.690 μm and OLI 0.630–0.680 μm), near infrared (ETM+ 0.775–0.900 μm and OLI 0.845–0.885 μm) and two shortwave infrared – SWIR1 (ETM+ 1.550–1.750 μm and OLI 1.560–1.660 μm) and SWIR2 (ETM+ 2.090–2.350 μm and OLI 2.100–2.300 μm). Blue (ETM+ 0.45–0.52 μm and OLI 0.45–0.51 μm) and green (ETM+ 0.525–0.605 μm and OLI 0.525–0.600 μm) bands were only used for quality assessment (QA) of viable observations. The thermal band (ETM+ 10.40–12.50 μm and Landsat 8's

Thermal Infrared Sensor 10.60–11.19 μm) was used for QA and for creating rank-based multi-temporal metrics.

2.5.2 Topography data

Ninety meter resolution Shuttle Radar Topography Mission (SRTM)(130) digital elevation model (DEM) data were also used as an input for classification. The elevation layer was re-projected via cubic spline to 0.00025° to match the Landsat resolution. Slope and aspect calculated from this elevation layer were used as additional inputs.

2.5.3 Auxiliary data for image interpretation

Time series of 16-day MODIS NDVI (131) composites and Google Earth high resolution imagery were used only for interpretation of training set and reference samples. The high temporal frequency of MODIS reflecting crop phenology helped to distinguish between crop and pasture pixels.

2.5.4 Landsat data processing

Landsat data processing was undertaken independently for both data sets (1999-2001 and 2011-2014) following methods developed for global data processing (132) : First, we converted raw digital numbers to top-of-atmosphere (TOA) reflectance and brightness temperature using established methods (133). Second, we used a set of existing quality assessment models (existing sets of bagged decision trees) to get a per-pixel QA flag for cloud, shadow, haze, and water detection. Third, we applied a radiometric normalization using a cloud-free surface reflectance MODIS

composite as a normalization target. The mean bias per band between the MODIS and Landsat TOA reflectance was calculated and successively applied to adjust Landsat reflectance. Lastly, we corrected for cross-track reflectance anisotropy by regressing the bias between Landsat TOA and MODIS surface reflectance against the Landsat scan angle. The slope and intercept of this regression were used to correct reflectance values per band, per image. These steps are part of an established Landsat processing system that has been successfully applied in a number of studies (50, 132, 134).

2.5.5 Metrics creation

Multi-temporal metrics allow us to capture phenological changes in vegetation within a consistent and standardized spatio-temporal feature space (50, 135). They facilitate regional-scale mapping using Landsat data despite variability in observation counts. Landsat processing steps are performed at the image level, while metric creation is a per-pixel process. Two sets of multi-temporal metrics were created using the data from each time period (1999-2001 and 2011-2014). To create one of these sets, we started by pooling together all cloud-free observations and ranking them based on: (1) band reflectance value, (2) Normalized Difference Vegetation Index (NDVI), (3) Normalized Difference Water Index (NDWI), and (4) brightness temperature. We created two types of metrics: rank-based metrics and average-based metrics. Rank-based metrics represent the minimum, maximum, and 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles of surface reflectance for the red, near infrared, and both shortwave bands and for the NDVI and NDWI indices for each rank method. Average-based metrics represent the averages for the following percentile intervals for each rank method: min-10^{th}, 10-25^{th}, 25-50^{th}, 50-75^{th}, 75-90^{th}, 90-max, min-max,

10-90th, and 25-75th. Additional metrics were derived by applying a moving average filter to all existing metrics using a 3x3 kernel. Once we had obtained a multi-temporal metric set for each time period, a third metric set was created by taking the difference of the corresponding average-based metrics from both time periods. These metric sets, along with the DEM and slope layers, were used as inputs for the classifications. In total, about 650 metrics were used for the cropland 2000 classification and about 1350 for the cropland expansion classification.

2.5.6 Classification

For this study, we created two separate map products: a map of cropland extent in Brazil for the year 2000, and a map of cropland expansion in Brazil from 2000 to 2014. We define the "cropland" land cover as areas of intensive row crop agriculture. To create the cropland expansion map, we targeted expansion between 2000 and 2014 as a class, as opposed to deriving cropland change from post-classification of annual maps of cropland from 2001-2014. Post-classification comparisons can lead to significant inaccuracies because of the confusion between real land cover change and apparent change due to misclassification errors. Both maps were created using supervised bagged classification tree models (136). Training data were manually labeled using Landsat cloud-free mosaics. Google Earth data and MODIS NDVI time-series data were used as additional inputs to aid interpretation. Classification trees work by recursively splitting the training dataset into increasingly homogenous groups until a certain purity criterion is met. Seven bagged classification trees were used per model, each derived from a random sample of 20% of the total training data set in order to avoid over-fitting. The cropland extent map for the year

2000 was created using the 1999-2001 metric set as independent variables for the classification. To create the cropland expansion map, we used all three metric sets. Both classifications were done iteratively, by checking the classification results and adding more training in areas where the results were poor. Obtaining models that produced sufficiently accurate results needed several iterations because of the large spectral differences between different crop types, agricultural practices, and crop calendars, as well as because of confusion with other land cover types such as pasture and shrubland. See the SI Appendix for further information on which metrics were most important for classification.

2.5.7 Accuracy assessment and sample-based area estimation

All maps contain errors, which is why land cover area estimates should be based on a probability sample of reference data (82–88). Aside from producing unbiased area estimates and associated uncertainties, sampling allows us to add value to land cover change studies by including attributes regarding date and type of change (123, 137, 138). Information in this study related to previous land cover type and year of change was attributed through sample interpretation and not by using auxiliary land cover maps. Other studies (10, 60, 106, 107) have used maps of deforestation or land cover to determine previous land cover type and derive areas of conversion from forest to cropland, but the results from such studies are prone to bias of area estimates caused by map classification error (82–88).

For our study, a stratified random sample of 5,000 30x30m pixels was selected. Crop area in 2000, total crop expansion from 2000 to 2014, and crop expansion per year from 2000 to 2014 were estimated from this sample. The three

strata used in the sampling design were cropland 2000, cropland expansion, and no cropland (all pixels that were not included in the previous two categories), where the stratum to which a pixel was assigned was determined from the 2000 cropland and the 2000-2014 cropland expansion maps. The cropland expansion stratum was allocated 2,000 sample pixels to reduce the standard errors of the area estimates of expansion by year and by previous land cover type. The remaining 3,000 sample pixels were allocated evenly between the cropland 2000 and no cropland strata. Map accuracy and sample-based area estimates were calculated from the confusion matrix (82, 83).

The reference class label for each sampled pixel was determined based on expert interpretation of annual cloud-free Landsat image composites for 2000 to 2014, MODIS NDVI time-series, and Google Earth high resolution imagery time-series, as available. A web interface was built to aggregate the different sources of data for each sample pixel (SI Appendix, Fig. S2-12 and S2-13). Each sample pixel was labeled as one of four classes: stable cropland (the pixel belonged to the cropland class every year from 2000 to 2014), cropland expansion (the pixel was not cropland in the year 2000 but it became cropland in any of the following years), cropland loss (the pixel was cropland in the year 2000 but it changed to a different land cover in any of the following years), or not cropland. We consider “cropland 2000” and “stable cropland” to be equivalent classes, because the amount of cropland loss over the 14-year time period was found to be negligible. If the sample pixel exhibited cropland expansion, we also recorded the year of expansion and the previous land cover type (pasture, natural vegetation, or tree plantation).

Spectral, temporal and spatial/contextual information domains of the reference remote sensing data facilitated interpretation. For example, pastures have a higher albedo than natural savanna vegetation due to the effects of grazing pressure at the per pixel scale. However, distinguishing pasture from herbaceous Cerrado natural vegetation (such as Campo Limpo grasslands) can be challenging when using only per-pixel spectral data. To facilitate discrimination, we examined landscape context, such as the presence of paddock or pasture boundaries, roads and watering holes (high spatial resolution data provide more definitive evidence for more detailed features such as watering holes). Landscape context was also the primary information source for discriminating forestry land use from natural forests. For pixels that exhibited a land cover transition from forest to pasture to cropland, we assigned forest as the previous land cover type if three or fewer years passed between the pasture to cropland transition. Otherwise, those pixels were labeled as conversion from pasture. All area estimates reported throughout this paper are sample-based, and have known uncertainties (i.e., standard errors) following good practice recommendations for estimating area (82–88). See the SI Appendix for detailed results describing accuracy of the map used to create the sampling strata, along with an assessment of our sample interpretations against a dataset of field-verified samples.

2.5.8 Global Forest Change map

To better understand the spatio-temporal patterns of cropland expansion into previously forested areas, we combined our cropland expansion map with the Global Forest Change (GFC) map from Hansen et al. (50) The GFC map shows forest loss

(defined as a stand-replacement disturbance) at 30 m resolution, and is disaggregated by year of loss event from 2001 to 2014. As previously mentioned, area estimates related to year of change and previous land cover type were derived from sample interpretation alone and not from the combination of our cropland maps with the GFC map. The combination of our cropland maps with the GFC map does provide a spatial representation of where cropland expansion was most likely to have occurred. This spatial display augments the sample-based area estimates which quantify the cropland expansion area, but do not indicate where this expansion is occurring.

2.6 **Supplemental information appendix**

2.6.1 Classification results and accuracy assessment

Classification results can be observed in Supplementary Fig. S2-7. Our sample-based accuracy assessment estimates overall map accuracy at 97.4% (SE=0.2%) (Supplementary Table 1). However, our map overestimates year 2000 cropland area, with user's accuracy of 89.9% (SE=3.5%), and producer's accuracy of 71.5% (SE=1.2%). User's and producer's accuracies of the cropland expansion class are more balanced – at 72.0% (SE=1.0%) and 71.2% (SE=5.3%) respectively. Map area of cropland expansion is 20.3 Mha, and sample-based area is 20.5 ± 1.6 Mha, demonstrating that the map does not under- or overestimate cropland expansion area.

Only 32 out of the 5000 sampled pixels were labeled as "cropland loss." All of these pixels had been mapped as cropland in the year 2000 and were counted as correctly classified pixels for the purpose of the confusion matrix.

Despite the very large number of metrics used for each classification, the

classification tree models decrease the majority of the training data's deviance based on a few select metrics. For the cropland 2000 classification, 74% of the overall deviance in the training data is decreased with only twenty of the >600 metrics used. The top metric, which corresponds to the average of the 75-90th percentile of the red band, an indicator of bare ground, provided 40% of the deviance decrease for this classification. For the cropland expansion classification, 78% of the deviance decrease is accounted by the top twenty metrics used by the classification tree, and 54% of the deviance decrease is explained by the top metric alone (the difference between the 2000 and the 2014 averages of the top 10% of SWIR2). In both cases, the classification tree captured cropland by targeting the spectral signatures of both high vegetation and bare ground, which is indicative of the agricultural cycle. Mapping of agricultural areas typically makes use of MODIS data because of its high temporal resolution, which provides NDVI time series needed for cropland characterization based on phenological responses. We get around the limitation of Landsat's lower temporal resolution by targeting the variation between the vegetated state and the unvegetated state typical of the agricultural cycle. Variations of our method using time-series metrics for cropland monitoring can be found for MODIS in (139–141), and for Landsat in (142–144).

2.6.2 Assessment of sample interpretations

We were able to assess our sample interpretations against a dataset of field-verified samples. A stratified, two-stage cluster sampling design (consisting of 45 20x20 km blocks, each containing 20 randomly selected sample polygons corresponding to a Landsat pixel footprint) targeting soybean area was used to collect

field data throughout Brazil in January 2017. We then used MODIS data to compare the spectral signatures of the field samples during the 2016/2017 growing season with those of the sample pixels we interpreted as cropland, pasture, and natural vegetation during the 2000/2001 growing season (SI Appendix, Fig. S2-17). Note that in addition to being from a different year, these data are not co-located. The plot illustrates how a competent interpreter can assign land cover and land use categories with spectral signatures consistent with those labeled in situ. The largest difference in distribution is in the cropland class. This is likely due to the fact that the 2016/2017 field-based random sample was stratified targeting soybean and not cropland, meaning most cropland field samples are likely to be soybean. However, interpreted samples from 2000/2001 are stratified on our cropland class which also includes sugarcane, cotton, rice and other crops that have spectral responses, particularly in the near-infrared, that are significantly lower than that of soybean.

2.6.3 Supplementary figures and tables

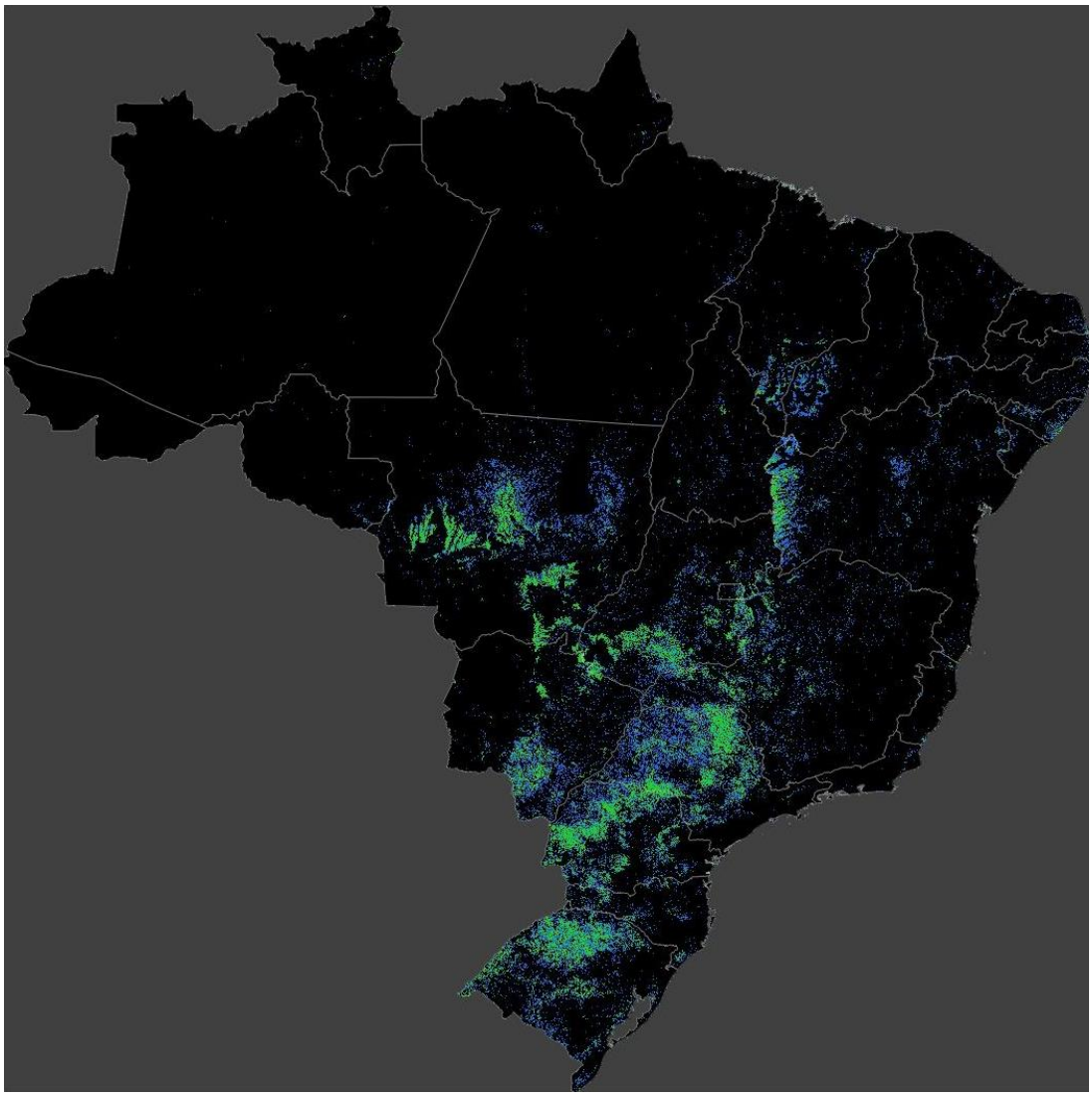


Figure S2-7 Classification result. In green, cropland extent in the year 2000. In blue, cropland expansion through 2014.

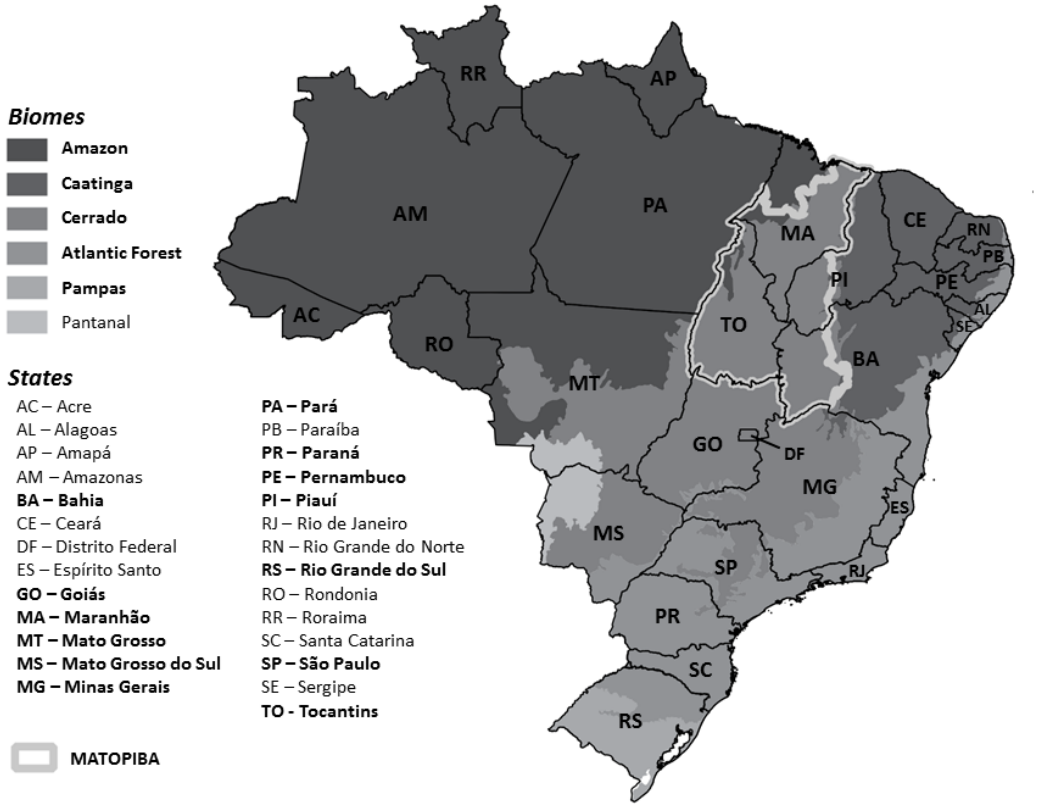


Figure S2-8 Brazilian states, biomes, and MATOPIBA. States and biomes for which we are able to report sample-based area estimates are in bold

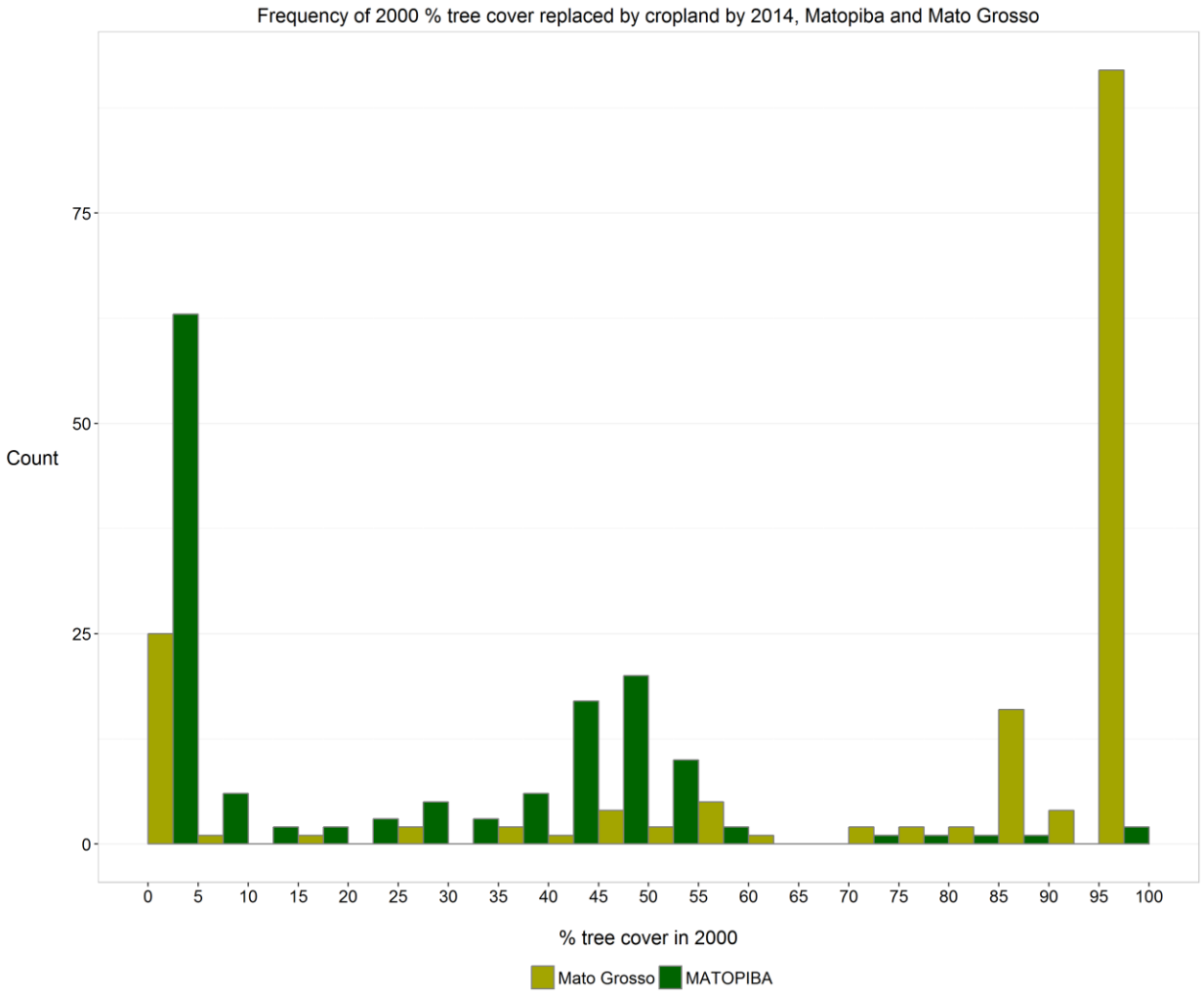


Figure S2-9 Antecedent tree cover for samples of naturally vegetated lands converted to cropland for Mato Grosso and Matopiba. 2000 tree cover percent was obtained from the Global Forest Change maps(50)

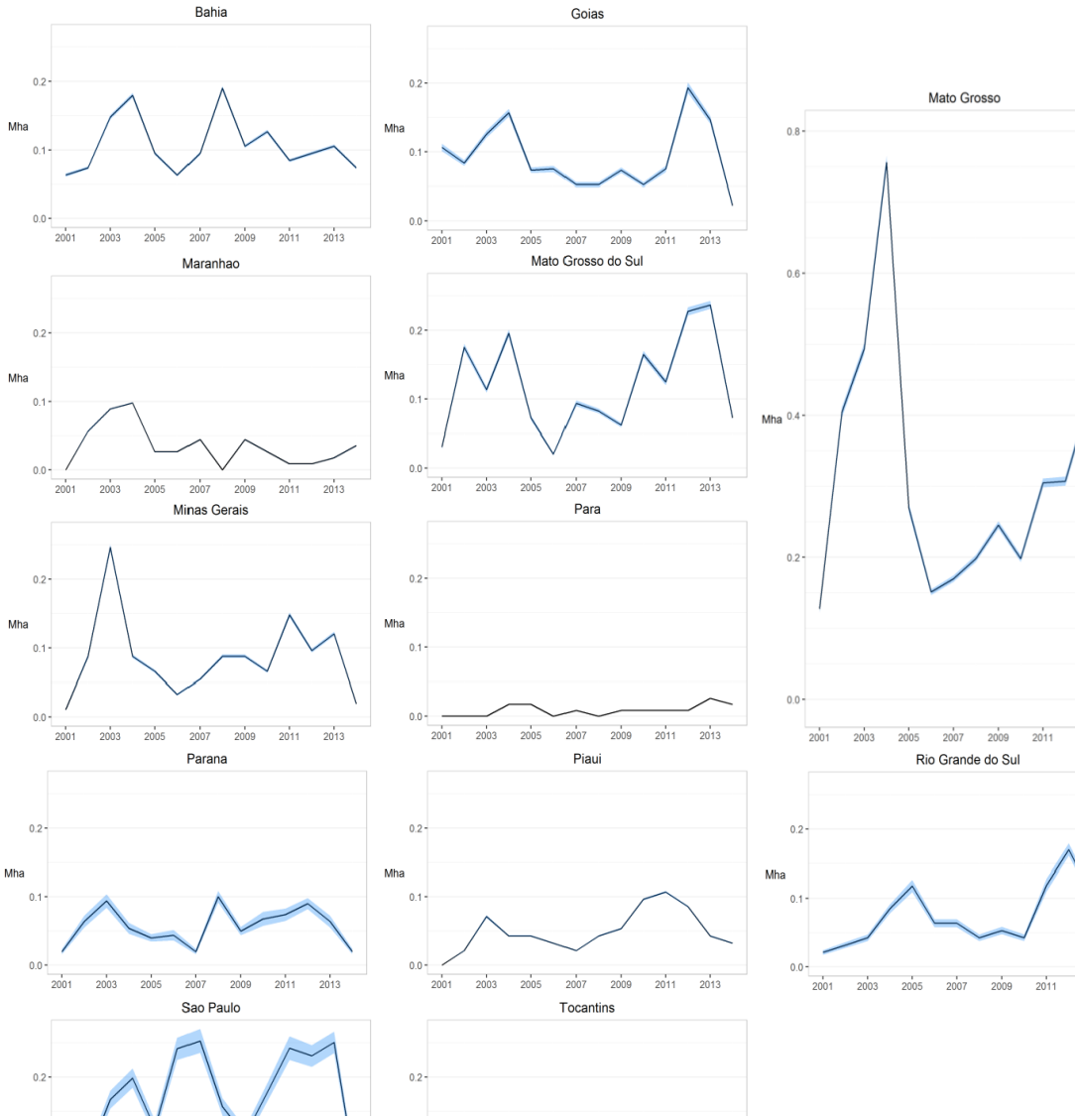


Figure S2-10 Trends of cropland expansion per state. Sample-based cropland expansion area estimates for "cropland 2000" and "cropland expansion" strata (+/- one standard error) per state with >10 samples in "cropland expansion" strata. Year of expansion corresponds to year of planting (e.g. 2001 corresponds to the 2001/2002 growing season). Samples from the "no cropland" strata are not displayed. See Supplementary Table 3 for tabular data for all strata.

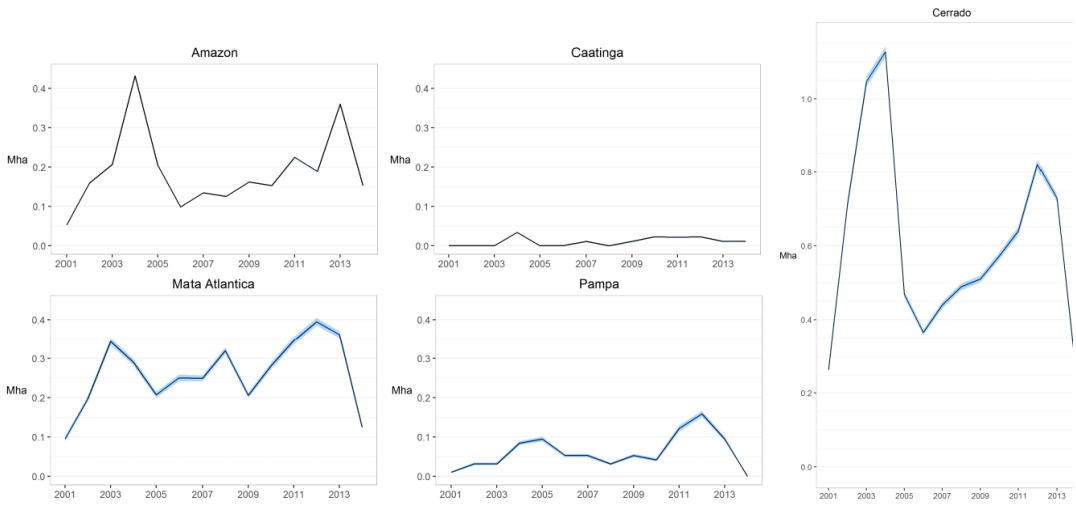


Figure S2-11 Trends of cropland expansion per biomes. Sample-based cropland expansion area estimates for "cropland 2000" and "cropland expansion" strata (+/- one standard error) per biome with >10 samples in "cropland expansion" strata. Year of expansion corresponds to year of planting (e.g. 2001 corresponds to the 2001/2002 growing season). Samples from the "no cropland" strata are not displayed. See Supplementary Table 3 for tabular data for all strata.

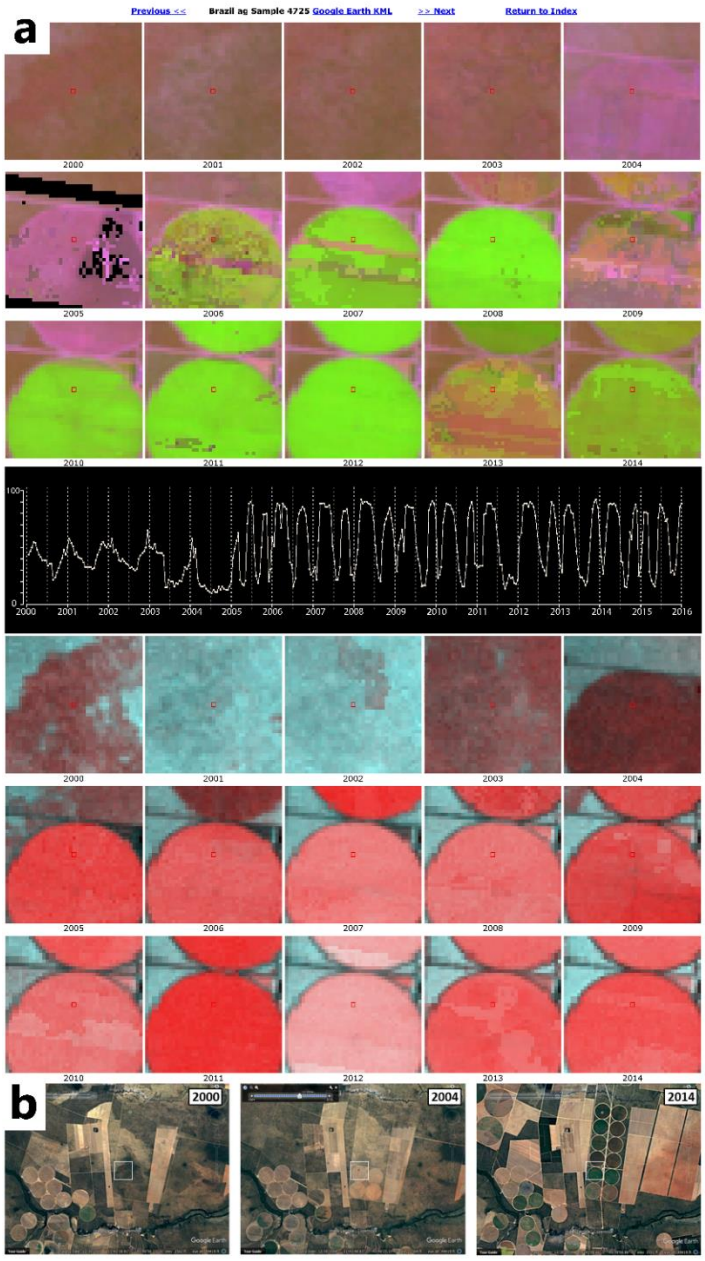


Figure S2-12 Sources of remote sensing data used for sample interpretation shown through the example of a sample pixel. (a) shows web interface for an example sample pixel. Composites on top are annual cloud-free Landsat composites with SWIR 1- NIR - red loaded in RGB from 2000 to 2014. Bottom composites are annual cloud-free Landsat composites where yearly maximum NDVI is loaded in the red band, and yearly minimum NDVI is loaded in the green and blue bands from 2000 to 2014. Time series graph shows MODIS 16-day NDVI time series. The link on the top right downloads a .kml file which allows the interpreter to visualize the sampled pixel on Google Earth. This example shows conversion from natural vegetation to cropland. (b) shows Google Earth imagery corresponding to this sample pixel for the

beginning of the time period (2000), year of change from natural vegetation to cropland (2004), and end of the time period (2014). White boxes on (b) are 1.1 x 1.1 km and correspond to the size of a Landsat subset on the web interface.

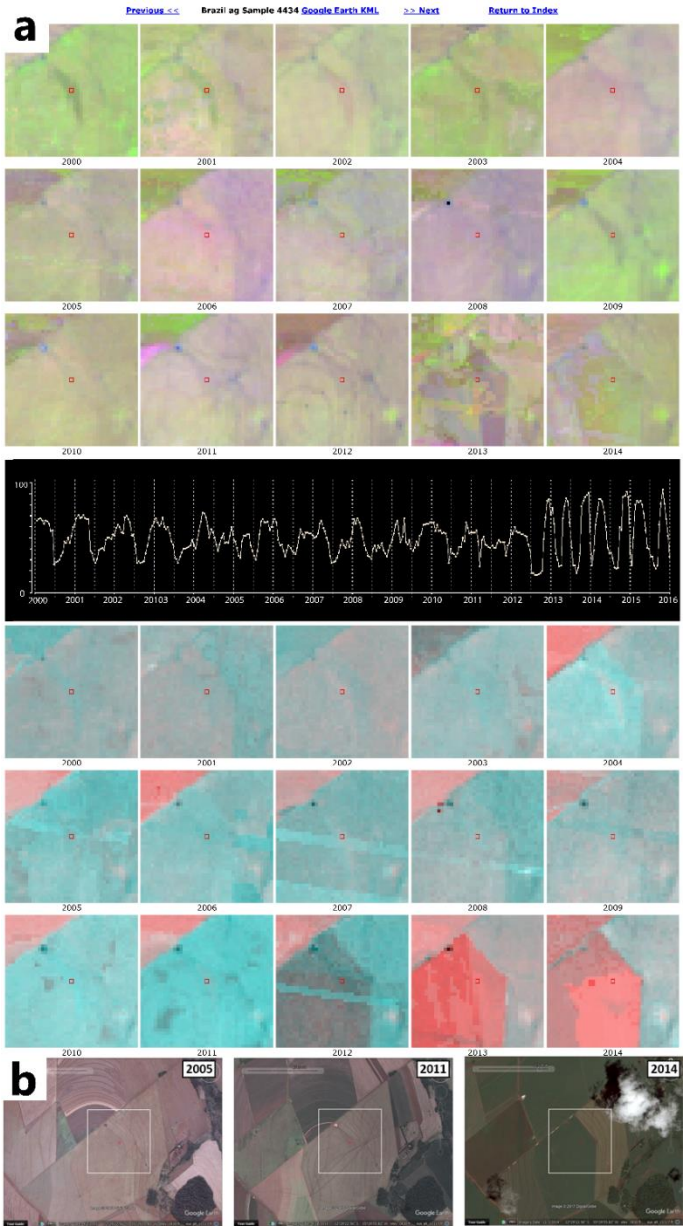


Figure S2-13 Sources of remote sensing data used for sample interpretation shown through the example of a sample pixel. (a) shows web interface for an example sample pixel. Composites on top are annual cloud-free Landsat composites with SWIR 1- NIR - red loaded in RGB from 2000 to 2014. Bottom composites are annual cloud-free Landsat composites where yearly maximum NDVI is loaded in the red band, and yearly minimum NDVI is loaded in the green and blue bands from 2000 to 2014. Time series graph shows MODIS 16-day NDVI time series. The link on the top

right downloads a .kml file which allows the interpreter to visualize the sampled pixel on Google Earth. This example shows conversion from pasture to cropland. (b) shows Google Earth imagery corresponding to the sample pixel for the earliest available time (2005), year before change (2011), and end of the time period (2014). White boxes on (b) are 1.1 x 1.1 km and correspond to the size of a Landsat subset on the web interface.

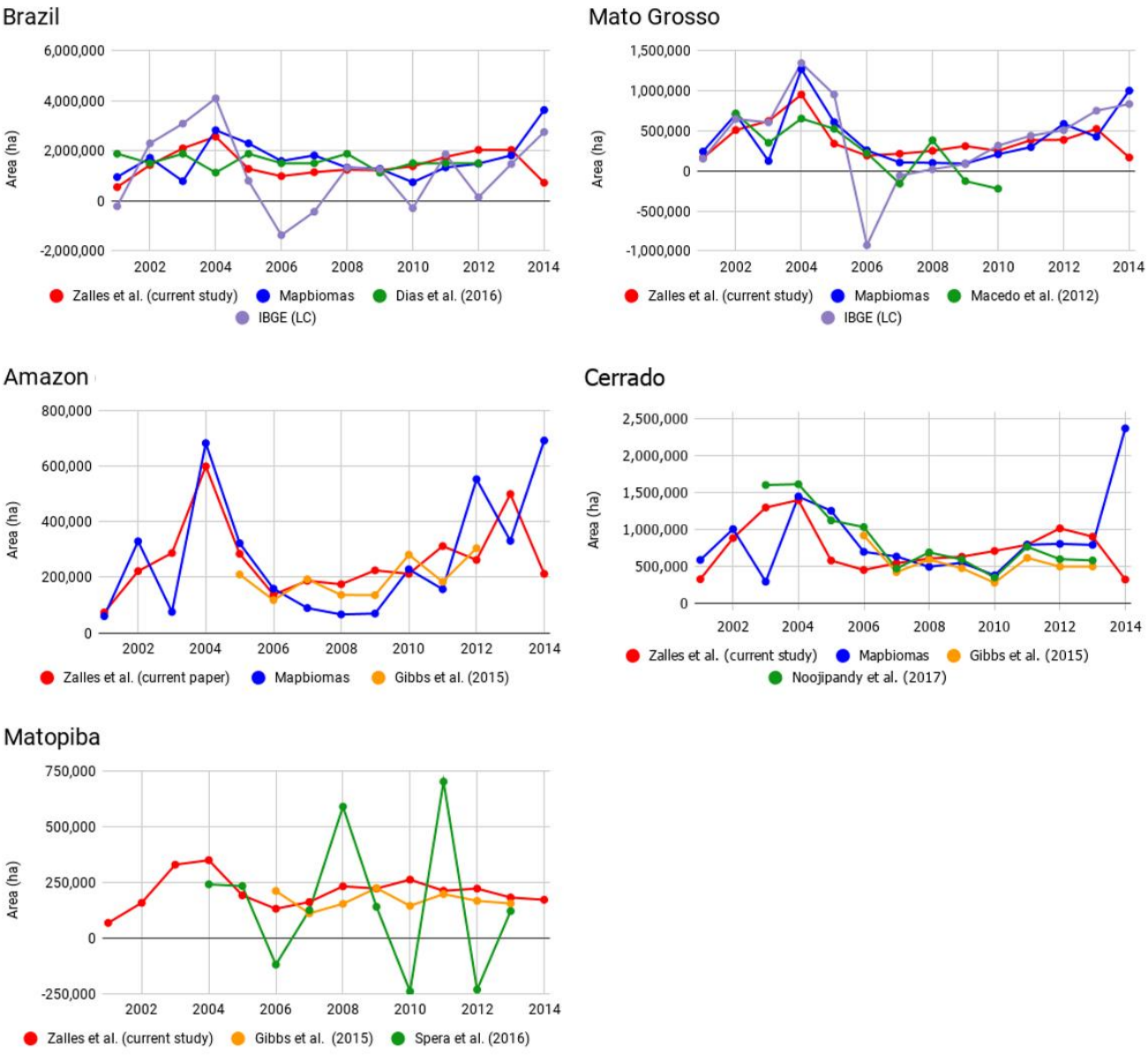


Figure S2-14 Comparison of cropland expansion area between current study and others (10, 89, 106, 107, 114, 116). "IBGE LC" (IBGE cropland Land Cover) area corresponds to 1st season corn, 1st season bean, soy, cotton, sugarcane, and rice areas from the IBGE SIDRA database. For comparison purposes, area of the cropland

expansion coming from the "no cropland" stratum was distributed across all years proportionally to the area of expansion from the "cropland expansion" and "cropland 2000" strata combined.

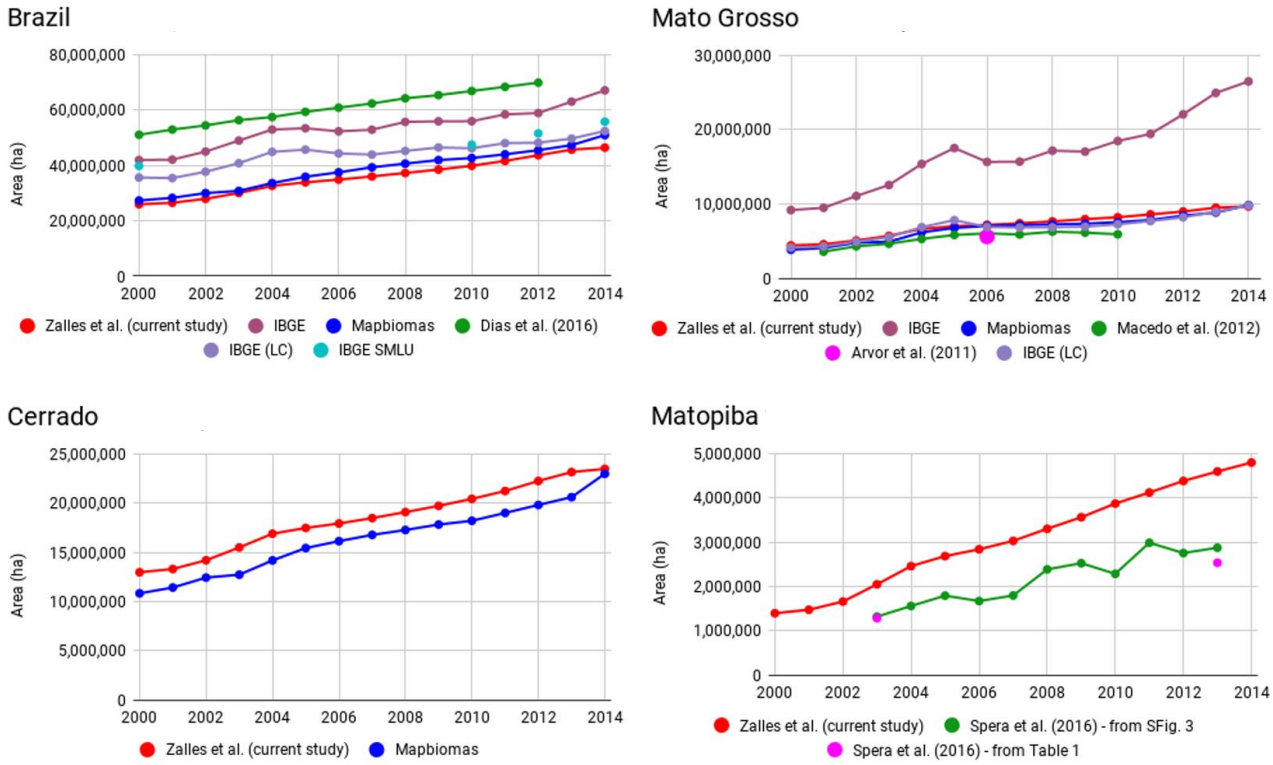


Figure S2-15 Comparison of total cropland area between current study and others (10, 89, 91, 92, 115, 116). “IBGE” data corresponds to area of soy, corn, sugarcane, cotton, rice, manioc, beans, and rice from the IBGE SIDRA database. “IBGE LC” (IBGE cropland Land Cover) area corresponds to 1st season corn, 1st season bean, soy, cotton, sugarcane, and rice areas from the IBGE SIDRA database. IBGE SMLU corresponds to IBGE’s Systematic Monitoring of Land Use project. Spera et al. (2016) report different areas in Table 1 and in Supplementary Figure 3 of their study. Both are included. For comparison purposes, area of the cropland expansion coming from the "no cropland" stratum was distributed across all years proportionally to the area of expansion from the "cropland expansion" and "cropland 2000" strata combined.

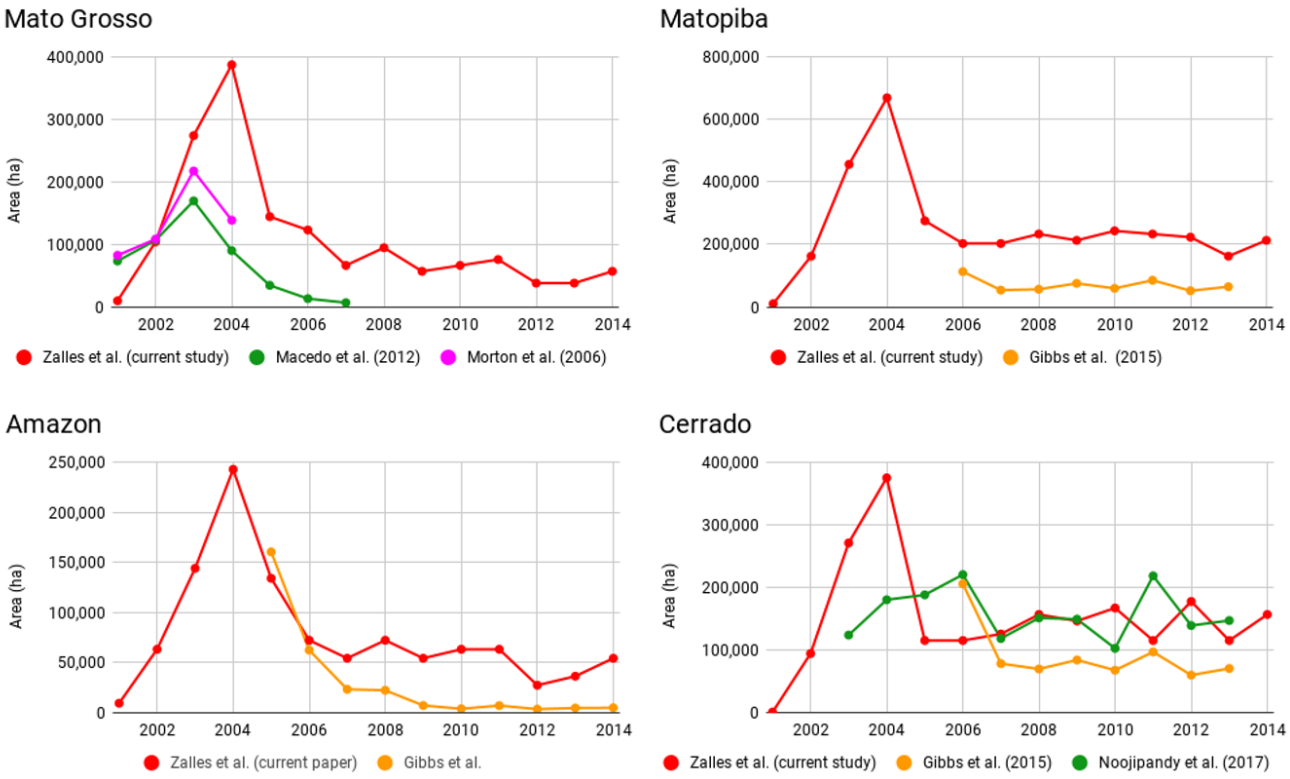


Figure S2-16 Comparison of natural vegetation conversion to cropland area between current study and others (91, 106, 107, 145). For comparison purposes, area of the cropland expansion coming from the "no cropland" stratum was distributed across all years proportionally to the area of expansion from the "cropland expansion" and "cropland 2000" strata combined.

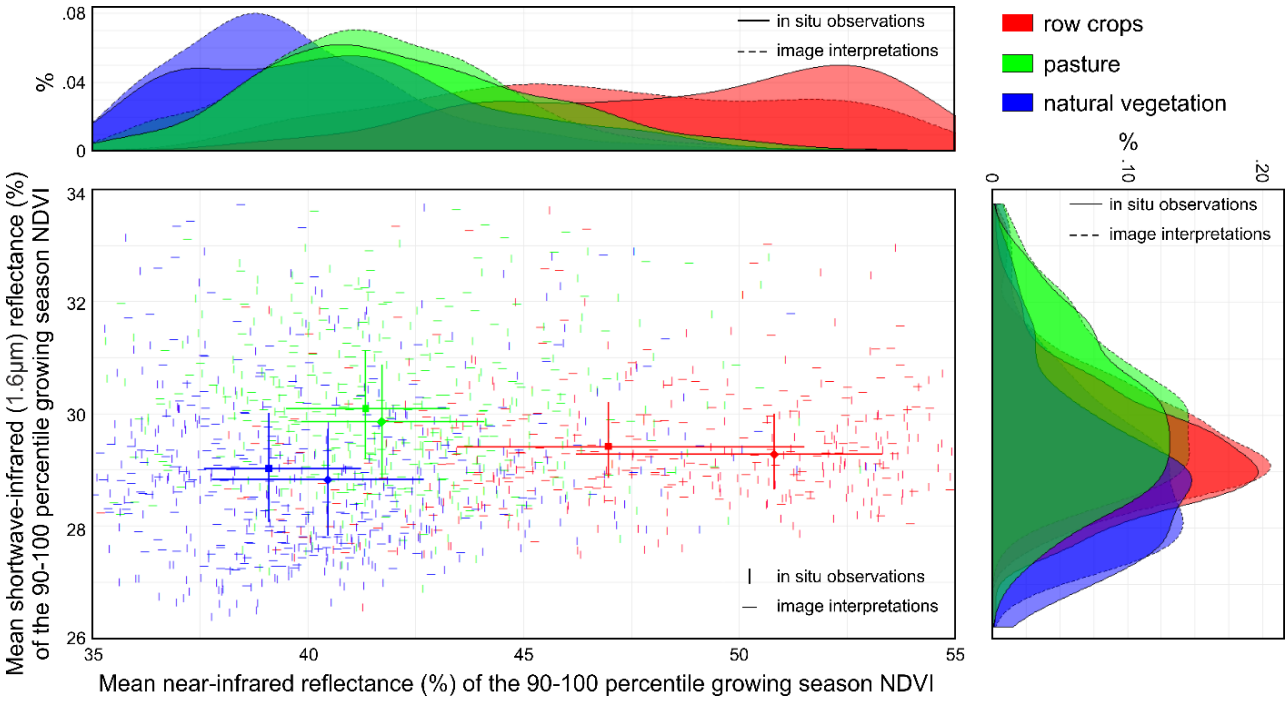


Figure S2-17 Comparison of separate probability-based samples of in situ observations and interpretations using multi-source earth observation data using a common reference of growing season MODIS data. For in situ samples, MODIS data are from the 2016/2017 growing season (November-April) as part of a field campaign estimating soybean cultivated area. For multi-source interpreted samples (this study), MODIS data are from the 2000/2001 growing season and consist of a subset of samples from the 2000 cropland and 2000-2014 cropland gain strata. The x-axis corresponds to the mean MODIS near-infrared reflectance (%) of the 90-100 percentile growing season NDVI. The y-axis corresponds to the mean MODIS shortwave-infrared (1.6µm) reflectance (%) of the 90-100 percentile growing season NDVI. These time-series metrics represent the near-infrared and shortwave-infrared reflectance of each sample at the time of peak vegetation greenness. Bars on the scatterplot correspond to the 25th and 75th percentiles of the respective distributions.

		Reference				
Map		Cropland 2000	Cropland Expansion	Not Cropland	Total	User's accuracy (SE)
	Cropland 2000	2.81	0.15	0.98	3.93	71.5 (1.2)
	Cropland expansion	0.07	1.75	0.62	2.43	72.0 (1.0)
	Not cropland	0.25	0.56	92.83	93.64	99.1 (0.2)
	Total	3.12	2.46	94.42	Overall accuracy: 97.4 (0.2)	
	Producer's accuracy (SE)	89.9 (3.5)	71.2 (5.3)	98.3 (0.1)		

Table S2-1 Confusion matrix for cropland 2000 and cropland expansion validation. Values shown are % of the study area.

	Cropland extent, 2000		Crop expansion (2001-2014)	
	Area	**SE**	**Area**	**SE**
Amazon	459,489	67,719	3,691,086	735,555
Bahia	1,049,878	103,800	1,500,385	89,041
Brazil	26,007,115	1,109,094	20,482,178	1,577,001
Caatinga	172,698	57,448	148,056	41,306
Cerrado	12,978,620	776,110	10,513,213	1,044,223
Goias	2,504,490	126,930	2,315,471	725,427
Maranhao	259,189	52,952	484,523	49,791
Mata Atlantica	10,027,647	217,241	5,356,662	978,253
Mato Grosso	4,428,442	556,760	5,255,843	770,853
Mato Grosso Do Sul	1,407,947	106,198	1,672,412	75,507
Matopiba	1,395,333	95,010	3,401,240	524,577
Minas Gerais	1,435,500	123,200	1,210,611	81,690
Pampa	2,601,462	729,403	862,150	57,678
Para	14,792	14,792	121,981	22,586
Parana	4,127,574	113,881	1,854,300	734,782
Piaui	177,227	51,524	690,899	59,235
Rio Grande Do Sul	5,381,663	820,042	959,638	61,673
Sao Paulo	4,658,543	479,890	2,753,393	466,900
Tocantins	103,483	39,113	838,209	478,874
Pernambuco	53,371	35,302	33,439	13,218

Table S2-2 Area and standard error estimates. Estimates of area and SE of cropland extent in 2000 and cropland expansion through 2012 for all regions with >10 samples in the expansion strata. All area and standard error estimates are in hectares.

Study	Map data source	Study region	Study period	Provides some form of accuracy assessment	Follows "good practice" recommendations[+]			
					Probability sample	SE/CI of accuracy metrics reported	Uses reference sample data for area estimation	Estimates SE/CI for area estimates
Morton et al. (2006)	MODIS	Deforested areas in Mato Grosso	2001 - 2004	✓	×	×	×	×
Macedo et al. (2012)	MODIS	Mato Grosso	2001 - 2010	✓	×	×	×	×
Arvor et al. (2011)	MODIS	Mato Grosso	2006/2007	✓	×	×	×	×
Gibbs et al. (2015)	MODIS	Amazon, Matopiba, Cerrado	2001 - 2014	✓	×	×	×	×
Spera et al. (2016)	MODIS	Matopiba	2003 - 2013	✓	×	×	×	×
Noojipady et al. (2017)	MODIS	Cerrado	2003 - 2013	✓	×	×	×	×
Graesser and Ramankutty (2017)	Landsat	South America	2000/2001	✓	×*	×	×	×
IBGE SMLU	MODIS	Brazil	2000, 2010, 2012, 2014	×	×	×	×	×
TerraClass Amazon	Landsat	Brazilian Legal Amazon	2004, 2008, 2012, 2014	✓	×	×	×	×
TerraClass Cerrado	Landsat	Cerrado	2013	✓	✓	×	×	×
Mapbiomas (3.0)	Landsat	National level and every biome, state and municipality	1985-2017	×	×	×	×	×

Zalles et al. (current study)	Landsat	National level, 5 biomes, 13 states, and Matopiba	2000 - 2014	✓	✓	✓	✓	✓

Table S2-3 Comparison of studies on cropland area in Brazil (10, 55, 56, 60, 63, 89, 91, 106, 114, 115) to the current study. "SE/CI" refers to standard error/confidence interval. *Graesser and Ramankutty provide two accuracy assessments. The information on this chart reflects the accuracy assessment of their thematic map. [+] Per (82, 84–86, 146)

Chapter 3: Rapid expansion of human impact on natural lands of South America[2]

3.1 **Abstract**

Human-induced land cover and land use change impacts climate, biodiversity, nutrient cycling, and other environmental systems. In South America, recent expansion of commodity land uses has underpinned significant regional economic development at the expense of natural land covers and associated ecosystem services. Here we show an estimated a 60% increase in human impact on natural land cover through conversion or modification since 1985. Over the 34-year study period, the area of natural tree cover decreased by 16%, and pasture, cropland and plantation land uses increased by 23%, 160%, and 288%, respectively. Natural tree cover loss was the dominant change dynamic, with 287Mha either cleared or degraded, and 68Mha converted to pasture. A substantial area of disturbed natural land cover totaling 55Mha had no discernable land use, representing land that is degraded in terms of ecosystem function but not economically productive. Results here illustrate the extent of ongoing human appropriation of natural ecosystems in South America, which intensify threats to ecosystem scale functions. Informed and enforced policies are needed to better balance economic development with the maintenance of critical ecosystem services.

[2] The material in this chapter was co-authored. Co-authors include Matthew C. Hansen, Peter Potapov, Stephen V. Stehman, Diana Parker, and Amy Pickens.

3.2 Introduction

Our improved ability to monitor changes on the Earth's surface via time-series satellite-based earth observations is timely given increasing rates of human-induced environmental change(29, 147, 148). Growing global populations, increased levels of development, and resulting greater interconnectedness, have led to heightened demand for goods such as food, timber, minerals, and fuels, the production of which requires the transformation and appropriation of natural ecosystems(149). Since the costs associated with the loss and degradation of natural ecosystems and the services they provide are largely not taken into account in national balance sheets, a full accounting of the consequences of this trade-off is missing(150, 151). Critically, up-to-date and accurate data on the extent of human appropriation of the natural environment is not readily available. Baseline data on rates of land cover and land use change are required to formulate policy responses meant to balance economic development with the maintenance of critical ecosystem services(29, 149).

Land use change is the dominant factor altering the global land surface(148) and the most significant of anthropogenic impacts on the environment(1), with implications for ecosystem functioning, biogeochemical cycles, and biophysical processes, amongst others. The degradation of natural land covers also has important consequences, albeit less well-quantified(152). Nowhere have these changes been as extensive in recent decades as on the South American continent, home to the largest tracts of pristine forests as well as to leading exporters of agricultural commodities such as beef, soy, and sugarcane(147, 148). Tensions between expanding the area used for the production of these goods and preserving important natural ecosystems

have long existed and led to efforts to map and quantify changes in land cover and land use in order to better understand the dynamics and drivers of change(149).

In the presented study, we employ over 30 years of time-series satellite data to estimate the area of human impact and land cover and land use change across South America. We first created maps of cropland, natural land cover, tree regrowth, and tree plantations using Landsat earth observation data. We then used these maps to select a stratified random sample of 1,000 pixels and labeled the land cover/land use class of each sample pixel for every year from 1985 to 2018 to create reference data. Area estimates of major land change dynamics across South America along with associated uncertainty measures were derived from the reference sample data, according to current international reporting guidelines(83, 86, 153).

3.3 **Results**

Results are divided into three sections. We first present results in terms of the overall human impact on natural land (Fig. 3-1). We then report area estimates of change between 1985 and 2018 for specific land cover and land use classes (Fig. 3-2 and Fig. 3-3). Finally, we provide area estimates of the most important from-to land cover/land use dynamics (Fig. 3-4). Each section is subdivided into results at the continental scale, administrative scale, and ecozone scale (154).

3.3.1 Human impact on natural land

Continental scale results: We define the human impact on natural land cover as the combined area of land use, semi-natural land cover, degraded natural tree cover, and secondary forest, as can be observed at the Landsat spatial resolution (see

hierarchy of land cover/land use, Fig. S3-8). Human impact on natural land cover in South America increased by 60% since 1985, reaching 713.7 ± 32.3 Mha in 2018 (the uncertainty is expressed as ±1 SE of the estimate) (Fig.3-1). This corresponds to a 20% decrease in unaltered natural land cover in a 34-year time period. The area of land under intensive land uses (defined as crop, pasture, plantations, built-up area) increased by 56%, totaling 465.7 ± 25.9 Mha in 2018. The amount of land under low intensity, small-scale, abandoned, or ephemeral land uses (collectively referred to as semi-natural land) increased by 66%, totaling 114.5 ± 16.3 Mha in 2018. By 2018, the area of natural tree cover that had been degraded by fire at some point over the study period was 68.9 ± 18.9 Mha, with an additional 64.6 ± 17.8 Mha degraded by selective logging or other types of biomass removal. Altogether, 17% of 2018 tree cover extent in South America has been affected by a degradation event at some point since 1985.

Administrative scale results: In Brazil, human impact increased by 64%, totaling 398.8 ± 19.7 Mha by 2018. The majority of this change consisted of gains in intensive land use, which amounted to 102.4 Mha. In the broad region we refer to as the Pan-Amazon excluding Brazil (Colombia, Venezuela, Peru, Bolivia, Ecuador, Suriname, French Guiana, and Guyana combined), human impact increased by 174% reaching 131.6 ± 17.8 Mha in 2018. Argentina, on the other hand, had much lower rates of new human appropriation of natural land cover: the human impact on natural land grew by 23% to reach 130.0 ± 14.7 Mha in 2018.

Ecozone scale results: South America's tropical rainforest ecozone, which is primarily made up of the Amazon rainforest, experienced an increase in human

impact on natural land of 134%, increasing to 155.5 ± 19.2 Mha in 2018. In this ecozone, 8% of natural tree cover was degraded by 2018, with a majority of degradation (77%) due to factors other than fire, such as logging. Human impact increased in the tropical moist deciduous forest ecozone by 61%, much of it due to increases in intensive land use. Unlike the tropical rainforest ecozone, the vast majority of tree cover degradation in tropical moist deciduous forests (which consists mainly of the Cerrado and humid Chaco biomes) was due to fire, which accounted for 87% of degraded tree cover. Human impact in the tropical dry forest ecozone (Caatinga and dry Chaco biomes) increased by 88%, totaling 89.5 ± 11.8 Mha in 2018. By 2018, both the tropical moist deciduous forest ecozone and the tropical dry forest ecozone had more land impacted by human activity than intact natural land, a reversal from 1985.

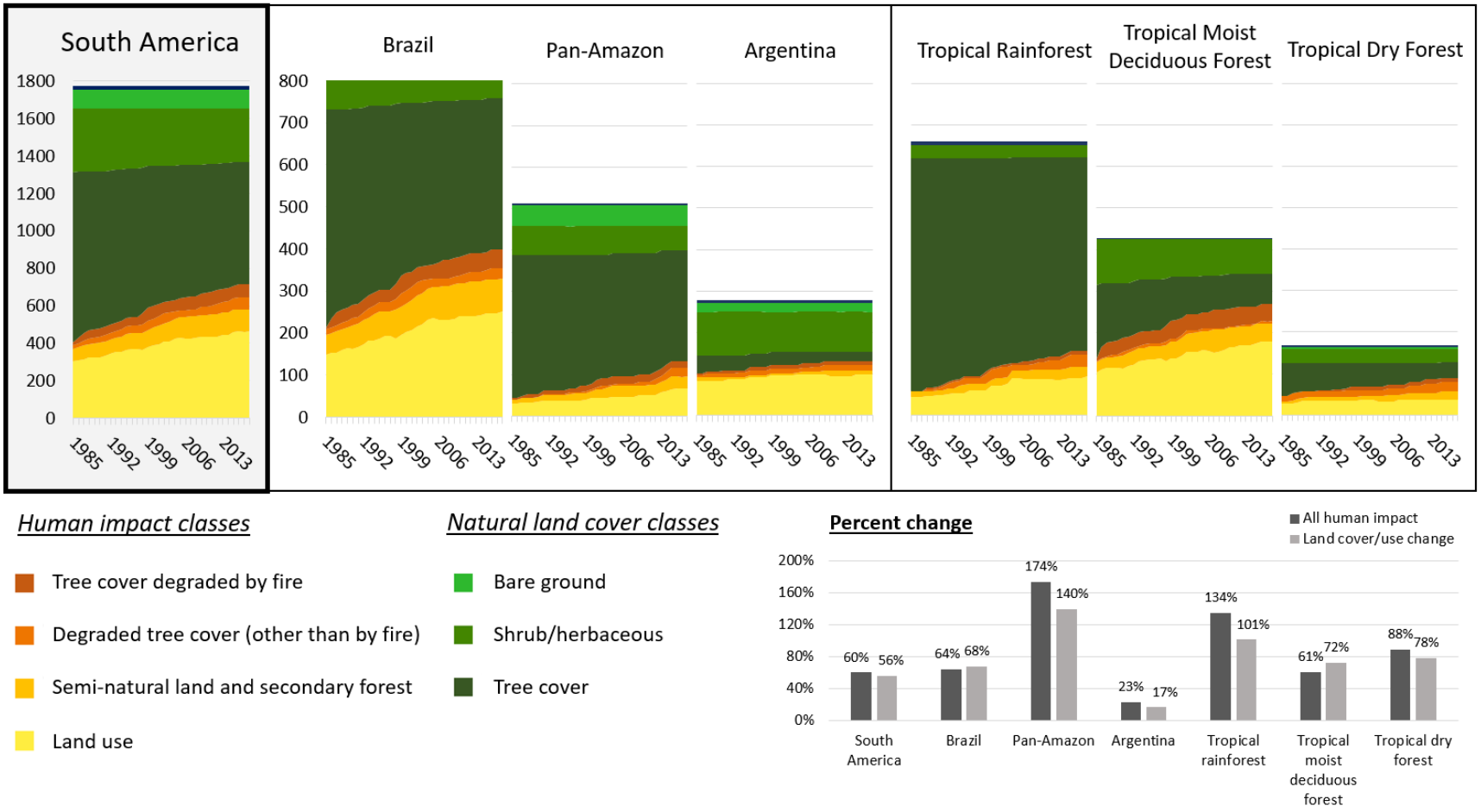


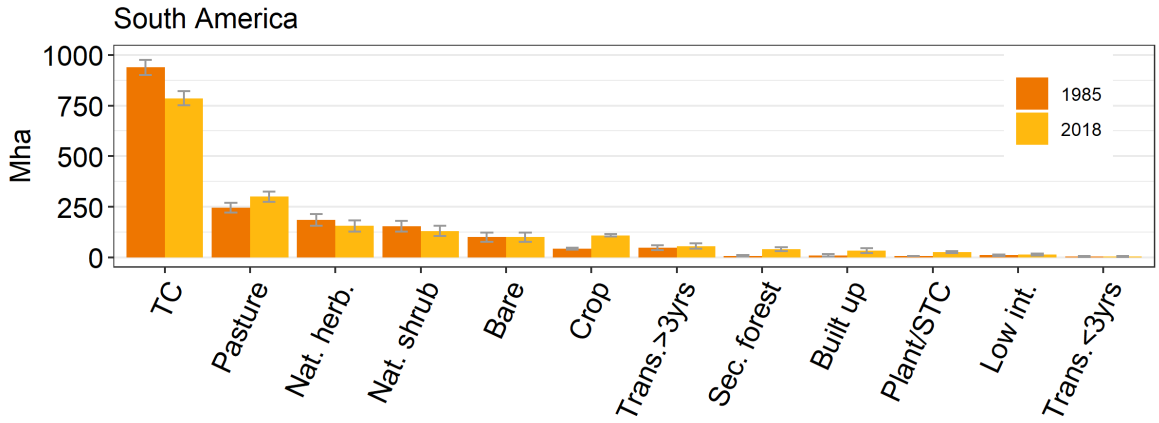
Figure 3-1 Human impact across South American regions categorized by degree of modification or conversion and the intensity of land use and natural land cover type. Inset on lower right shows percent of change over time period for all human impact classes and for land cover/use change classes only ("semi-natural land and secondary forest" and "land use" categories).

3.3.2 Per class land cover/land use change, 1985 to 2018

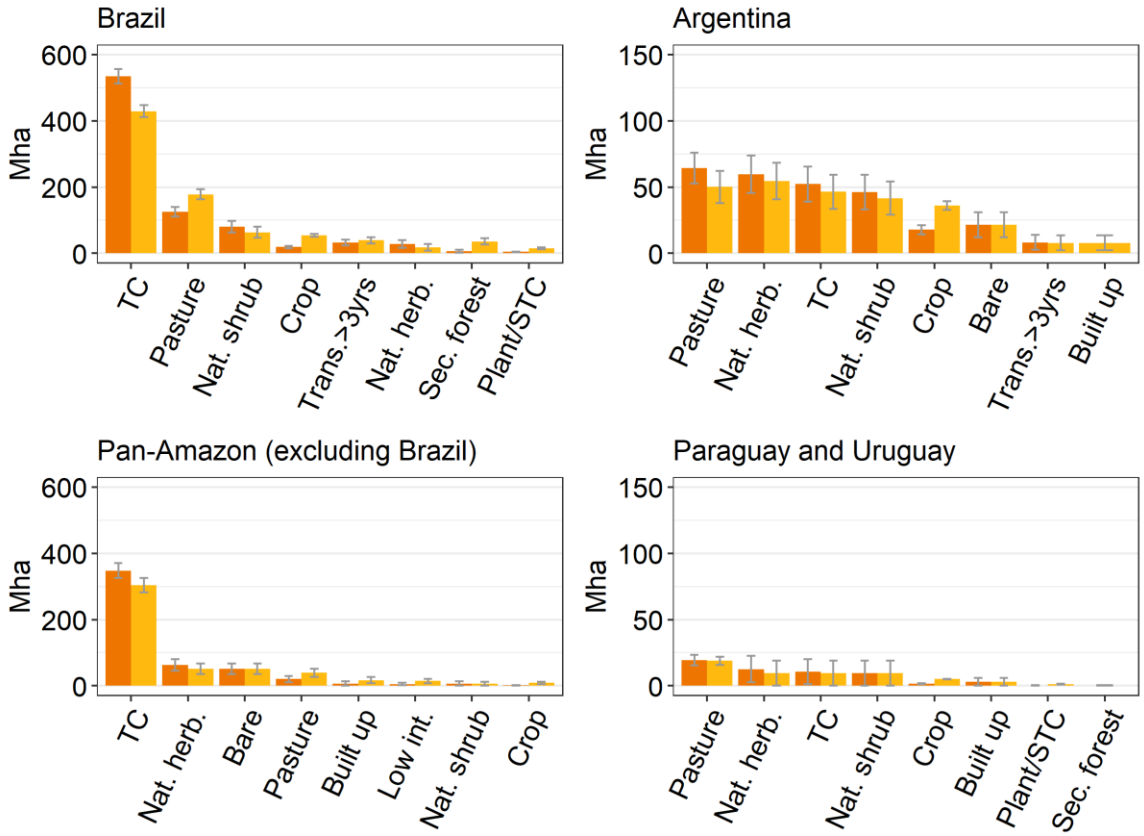
Continent: Nearly one fifth of the South American continent underwent some type of land cover/land use change over our study period (regardless of whether the change caused an increase in human impact). The single most significant land change dynamic was the loss of natural tree cover, which totaled 153.6 ± 16.8 Mha, representing half of all continental-level change (Fig.3-2). Tree cover decreased steadily at a rate of 6.4 Mha/year from 1985 until 2004, at which point the rate of loss slowed considerably to an average of 2.3 Mha/year (Fig. 3-3). Forest loss in Brazil experienced a dramatic decline following 2004, which largely persisted through 2018, driving both national and continental trends. The reduction was due to the enforcement of land use regulations, and greater scrutiny and action on improving supply chains and granting of credit, all of which drove a decrease in new deforestation (117). By 2018, 16% of South America's 1985 tree cover area was lost. The second largest category of change consisted of pasture gain, which totaled 129.4 ± 16.7 Mha. Pasture loss was also a major contributor to change, totaling 74.2 ± 7.9 Mha. From 1985 to 2004, pasture area increased at a rate of 3.5 Mha/year, after which the area of pasture stabilized at around 300 Mha (Fig. 3-3). Cropland area grew consistently across the continent at an average rate of 2.0 Mha/year, resulting in an increase in area by a factor of 2.6 over the study period (cropland gain was 70.0 ± 5.0 Mha while cropland loss was 3.4 ± 1.0 Mha). The combined area of tree plantations and short tree crops nearly quadrupled since 1985, reaching 25.6 ± 4.6 Mha by 2018. Natural tree regrowth observed during the study period totaled 37.8 ± 9.6 Mha. Transitional land, corresponding to land where a change in cover due to human

intervention is observable, but with no signs of an established land use following (see SI), accounted for 55.4 ± 12.7 Mha by 2018. From 1985-2018, the estimated loss of transitional land was 28.6 ± 8.4 Mha compared to an estimated gain of 36.8 ± 9.7 Mha.

a.


South America
Mha
1000
750
500
250
0
1985
2018
TC
Pasture
Nat. herb.
Nat. shrub
Bare
Crop
Trans.>3yrs
Sec. forest
Built up
Plant/STC
Low int.
Trans.<3yrs

b.


Brazil
Mha
600
400
200
0
TC
Pasture
Nat. shrub
Crop
Trans.>3yrs
Nat. herb.
Sec. forest
Plant/STC
Argentina
Mha
150
100
50
0
Pasture
Nat. herb.
TC
Nat. shrub
Crop
Bare
Trans.>3yrs
Built up
Pan-Amazon (excluding Brazil)
Mha
600
400
200
0
TC
Nat. herb.
Bare
Pasture
Built up
Low int.
Nat. shrub
Crop
Paraguay and Uruguay
Mha
150
100
50
0
Pasture
Nat. herb.
TC
Nat. shrub
Crop
Built up
Plant/STC
Sec. forest

c.

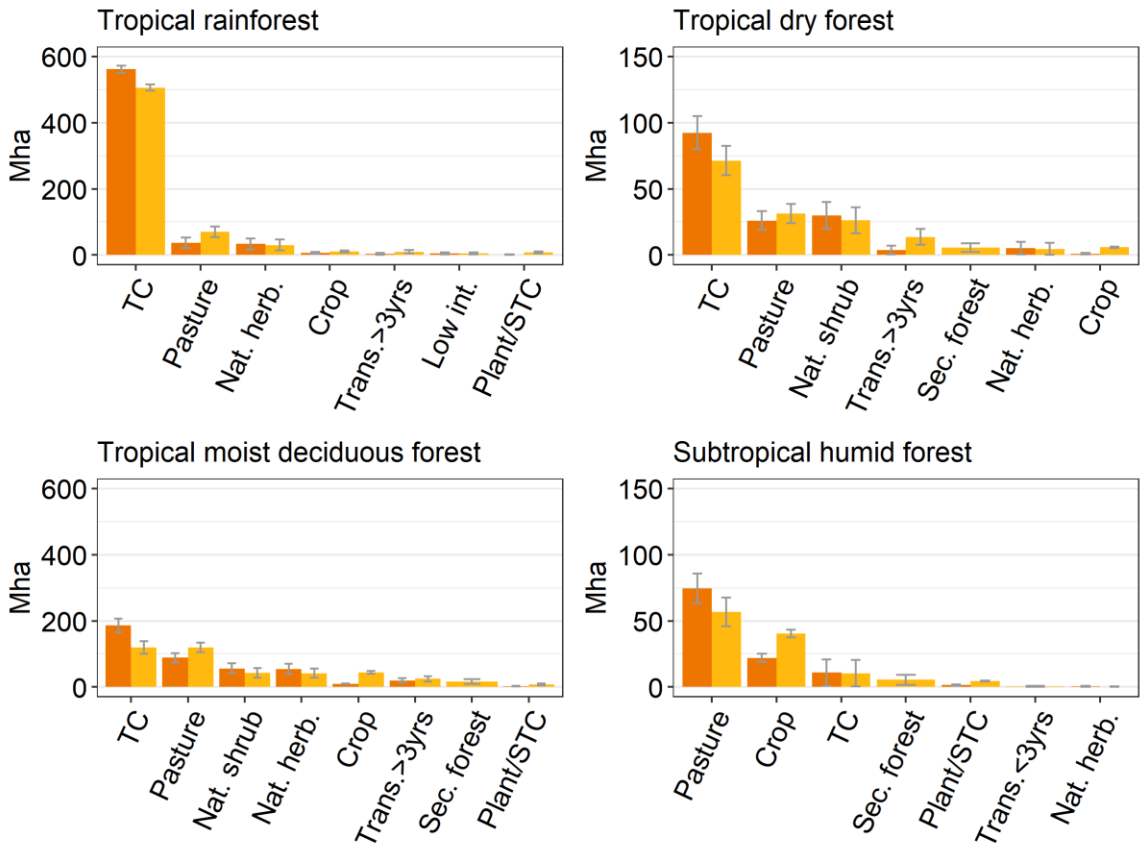


Figure 3-2 South America's land cover and land use: 1985 vs. 2018. Results shown for **(a)** the continent, **(b)** administrative regions, and **(c)** ecozones. Land cover and land use classes of tree cover ("TC"), pasture, natural herbaceous ("Nat. herb"), natural shrub ("Nat. shrub"), bare ground ("Bare"), cropland ("Crop"), transitional for >3 years ("Trans.>3yrs."), secondary forest ("Sec. forest"), built-up, plantation/short tree crops ("Plant/STC"), low intensity ("Low int."), and transitional ≤3 years ("Trans. <3yrs") are ranked by 2018 area within each region. Note the upper bound of the charts on the left is 600 Mha, and of those on the right is 150 Mha.

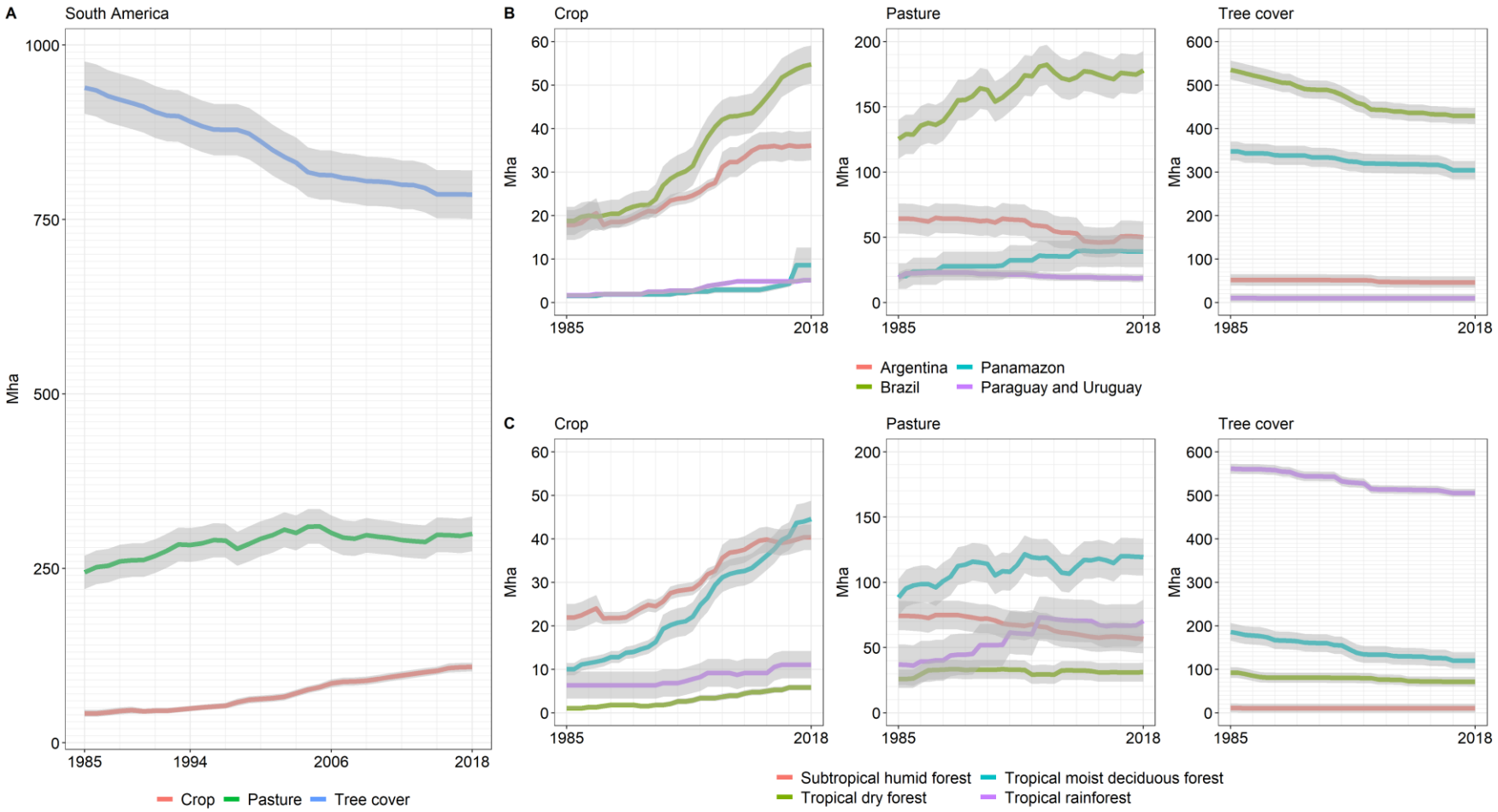


Figure 3-3 Land cover/land use trends of cropland, pasture, and tree cover area for **(A)** the continent, **(B)** administrative regions, and **(C)** ecozones. The shaded area corresponds to 1 SE of the estimate. Note the upper bound of the charts for crop, pasture, and tree cover vary.

Administrative: Brazil, the largest country in South America, exhibited the most extensive land use change. The area of natural tree cover in 2018 was 429.13 ± 18.6 Mha, a 20% decrease in extent since 1985. Tree cover loss in Brazil drove the continental-level trend, with an inflection point in 2004 after which loss slowed substantially. The rate of loss before 2004 averaged 4.0 Mha/year and declined to an average of 0.2 Mha/year after 2004. There were also substantial gains and losses of pasture (96.2 ± 12.6 Mha and 43.6 ± 6.6 Mha, respectively), netting a 42% increase in pasture area over the study period. Total pasture area increased at an average rate of 2.7 Mha/year until 2004, after which it slowed to 0.3 Mha/year. An important factor in the leveling of pasture expansion is intensification, including cultivated pastures, semi-containment and other practices(93, 155). Cropland area nearly tripled since 1985, gaining 38.0 ± 3.4 Mha and losing 2.0 ± 0.8 Mha, while the combined area of plantations and short tree crops increased by a factor of 3.4 resulting in 14.9 ± 3.1 Mha by 2018. The net increase in area of transitional land in Brazil from 32.0 ± 8.5 Mha in 1985 to 39.3 ± 9.4 Mha in 2018 masked an underlying gross change dynamic as gross loss (23.3 ± 7.3 Mha) and gross gain (30.6 ± 8.5 Mha) largely compensated each other. Argentina had lower levels of change than Brazil. The most significant change was the doubling of cropland area, which increased steadily until 2010, after which it stabilized at around 36 Mha, and a 22% net decrease in pasture area. Pasture area in Argentina was stable at around 64 Mha until 2003, when it started to decrease. The stabilization of continental-level pasture area is not due to a lack of land use change, but rather to the fact that regional losses and gains partially compensate each

other: Brazil and the Pan-Amazon region's gains are compensated by losses in Argentina, and the combined region of Paraguay and Uruguay.

Ecozones: Natural tree cover loss constituted the largest change in both the tropical rainforest and the tropical moist deciduous forest ecozones. The tropical rainforest lost 55.4 ± 8.5 Mha, representing a 10% net decrease from 1985 tree cover extent, and the tropical most deciduous forest lost 66.4 ± 11.3 Mha, representing 36% of 1985 extent. The rate of tree cover loss slowed markedly after 2004 in both these ecozones. Pasture area increase was also substantial in both of these ecozones: in the tropical rainforest ecozone, pasture area increased steadily until 2004, after which it stabilized at around 70 Mha. Total pasture gain was 43.6 ± 8.4 Mha, representing a net increase of 91%. There were considerable losses and gains of pasture in the tropical moist deciduous forest ecozone (60.4 ± 10.0 Mha of gain, 29.5 ± 5.2 Mha of loss**)**, resulting in a net increase of 34%. Loss of pasture area in the subtropical humid forest ecozone compensated for gains in other ecozones, resulting in stable pasture area at the continent level after 2004. The tropical moist deciduous forest ecozone was the epicenter of cropland expansion in South America, as its crop area nearly quintupled during the study period, increasing from 10.0 ± 1.4 Mha to 44.6 ± 4.2 Mha. Cropland area also increased significantly in the subtropical humid forest ecozone (19.8 ± 1.4 Mha of cropland gain), netting an increase of 84% over its 1985 extent. The rate of increase in this ecozone was nearly half that of the tropical moist deciduous forest: 0.6 Mha/year compared to 1.0 Mha/year.

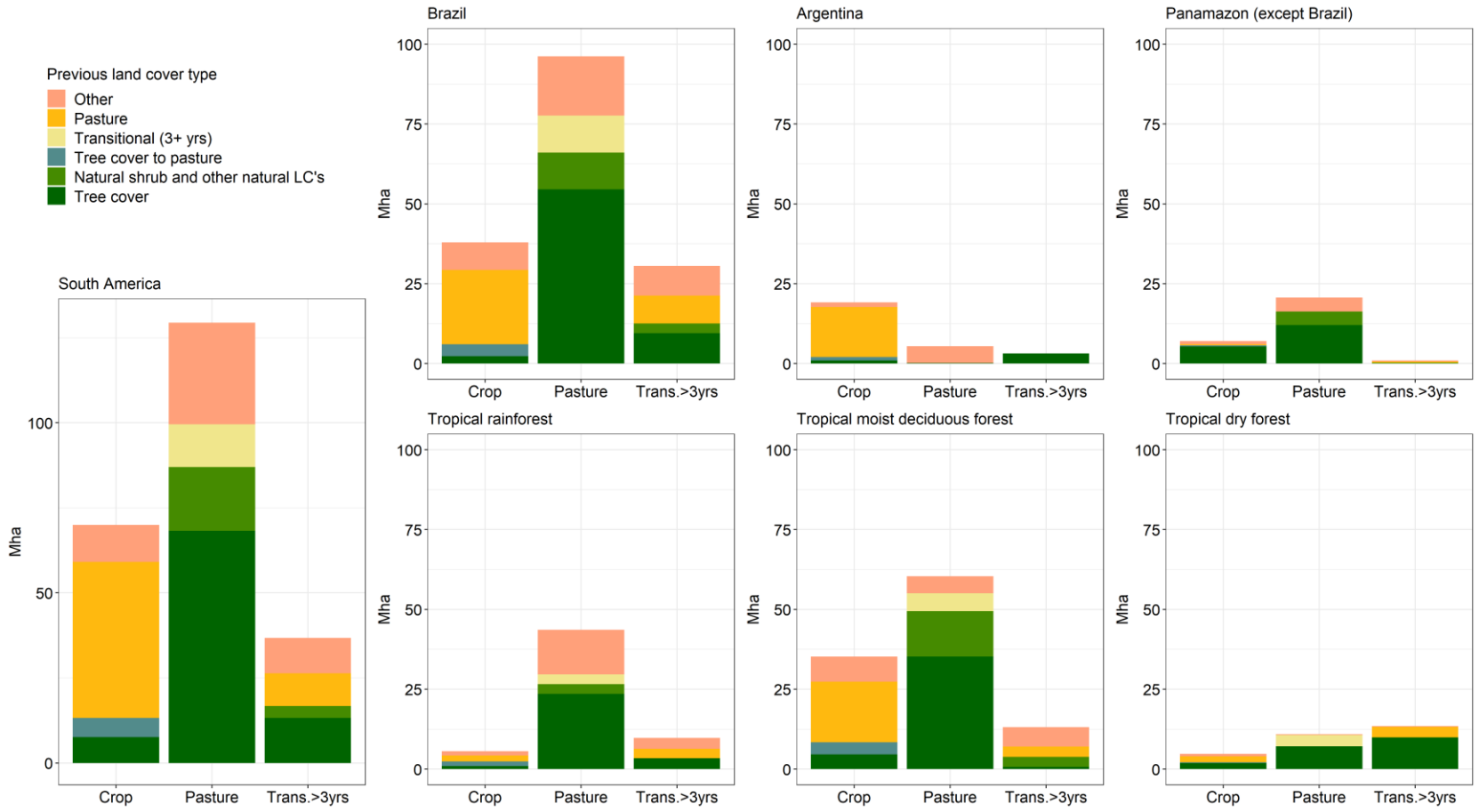


Figure 3-4. Land cover/land use transitions: Area of new cropland, pasture, and transitional land (>3 years) per region. Colors represent previous land cover type. The areas shown here correspond only to areas having experienced a single transition over the time period, with the exception of the "tree cover to pasture to crop" category. Land that underwent more than one land cover/land use change event over the time period are included in the "Other" category.

3.3.3 From-to land cover/land use transitions, 1985 to 2018

All land cover/land use transitions reported here correspond to areas having experienced a single transition over the time period, with the exception of the "tree cover to pasture to crop" category. For example, the area of tree cover to pasture conversion includes only land that was tree cover in 1985, pasture in 2018, and underwent a single land cover change event throughout the time period. Land that underwent more than one land cover/land use change event over the time period are included in the "Other" category in Figure 3-3.

Continent: The most prevalent land cover transition in South America was the conversion of natural tree cover in 1985 to pasture in 2018, which totaled 68.2 ± 12.8 Mha (Fig. 3-4). Following pasture, the most common 2018 fate of 1985 tree cover was transitional land, which equaled 13.2 ± 6.1 Mha. Tree cover was also converted to cropland, albeit at lower levels (7.6 ± 3.3 Mha). The estimated area of land that was converted from tree cover to pasture to crop during the study period was 5.6 ± 1.2 Mha. The majority of new crop area was created through the conversion of pastures (45.9 ± 4.1 Mha). Pasture abandonment led to an increase in transitional land (9.6 ± 5.2 Mha).

Administrative: Within administrative boundaries, significant transitions included: tree cover conversion to pasture both in Brazil (54.5 ± 10.7 Mha) and in the Pan-Amazon region (12.0 ± 6.5 Mha); tree cover to transitional land in Brazil (9.5 ± 4.9 Mha); and pasture to crop both in Brazil (23.2 ± 3.5 Mha) and in Argentina (15.7 ± 1.4 Mha).

Ecozones: Tree cover to pasture conversion was most prevalent in the tropical moist deciduous forest ecozone (35.26 ± 8.6 Mha), followed by the tropical rainforest ecozone (23.6 ± 7.2 Mha), and the tropical dry forest ecozone (7.1 ± 4.5 Mha). Additional pasture area in the tropical moist deciduous forest ecozone came through the conversion of natural shrubs and natural herbaceous cover (14.2 ± 6.0 Mha). This ecozone also had the greatest increase in crop area, largely due to pasture conversion (19.0 ± 3.3 Mha), but also in part through the direct conversion of tree cover (4.6 ± 2.9 Mha) and the conversion of pasture that had previously been tree cover (3.8 ± 1.0 Mha). The subtropical humid forest ecozone was characterized by the conversion of pasture to crop (18.3 ± 1.4 Mha).

3.4 **Conclusion**

Humans have dramatically altered environments through the conversion of natural vegetation to pasture, crop, and other land uses, and the degradation of natural vegetation through logging, fire, and other disturbance types. In South America, the increase in human impact on natural land cover over the past 34 years has averaged 9.5 Mha per year. Given that extensive lands remain suitable for further expansion of commodity land uses (156) and that recent expansion has occurred even in low-suitability areas (71, 102, 157, 158), we may expect further loss of natural land cover, threatening the maintenance of ecosystem services for major biomes such as the Amazon, Cerrado, Chaco, and Chiquitania (159–163). Given the role of the Amazon in maintaining regional climate (12), results here are a clarion call to improved land use policy formulation, implementation, and enforcement.

In South America, human impact on land is most directly due to the conversion of natural vegetation into pastureland. Beef production in South America has historically been characterized by its extensive rather than intensive nature (155, 164). It has been estimated that the productivity of pastures in Brazil is less than half of their carrying capacity (165), despite significant gains more recently (155, 164). Croplands replacing pasturelands is a widespread phenomenon and reflects the overall intensification of land use across South America (see SI). Given the tripling of cropland area over the study period and the likely continued expansion of the commodity crop footprint (166, 167), an understanding of the interplay of crop and pasturelands is needed (81). While policies such as the soy moratorium in Brazil have proven to be nominally effective (117, 168, 169), leakage effects and the displacement of other land uses such as pastures into deforestation frontiers illustrate the need for more comprehensive monitoring (170).

Beyond intensive land uses, there are 55.4 ± 12.7 Mha of land that have been converted from a natural state, but are not used for any discernable economic purpose. This long-lasting transitional land category should be of particular concern to policymakers and other stakeholders, as it represents lands that have been compromised in term of the provisioning of ecosystem services, but are also relatively unproductive in terms of economic output. Monitoring natural land cover from initial disturbance to possible conversion is necessary to fully understand land use pathways, including establishment of tenure, land banking, and eventual production of high value commodities.

3.5 **Methods**

Efforts to quantify the human impact on the Earth's surface by mapping land cover and land use are often limited for a variety of reasons: 1) they employ low spatial resolution data for mapping (79, 171, 172); 2) they report results for a single or a few years (80, 173); 3) they focus on a single land cover theme (50, 174); or 4) they look at only at a single region (55, 56). Further, most studies do not follow good-practice recommendations (83, 86, 153) suggesting that area estimates be derived from a probability sample.

In the presented study, we overcome these limitations by providing area estimates of the total human impact on natural land as well as rates of land cover/land use change in South America from 1985 to 2018. To achieve this, we first created maps of cropland, natural land cover, tree regrowth and tree plantations using Landsat data. We then used these maps to draw a stratified random sample of 1,000 pixels and employed reference data to label the land cover/land use class of each sample pixel for every year from 1985 to 2018. Area estimates of major land change dynamics across South America and associated uncertainty measures were derived from the reference sample data.

3.5.1 Landsat data processing

All Landsat TM, ETM+, and OLI/TIRS observations from January 1st, 1985 until December 31st, 2018 were downloaded and processed using the methodology described in Potapov et al. (2019) and Potapov et al. (2020)(175, 176). Processing steps include conversion to at-sensor radiance, per-pixel quality assessment,

reflectance normalization, and aggregation into 16-day composites. The 16-day composites were then used as inputs to time-series metrics for annual and multi-year mapping tasks. Metrics are statistical derivatives of time-series imagery that represent a generalized feature space appropriate for large area mapping(132, 135, 177).

3.5.2 Multi-epochal metrics

An additional metric set was created using the annual phenological metrics described above as inputs. Instead of using annual change detection metrics, we created this new set of metrics in order to capture land cover change across the entire 34-year time period, regardless of timing. To create this multi-epochal metric set, we selected a set of metrics from the annual metric set: the second highest, second lowest, and mean of the 25th to 75th percentile of observations of the red, NIR, SWIR 1 (1.6 μm), SWIR 2 (2.2 μm), and NDVI. We then aggregated each of these into five-year epochs (six five-year epochs and one four-year epoch corresponding to 2015-2018). Epochs were employed in order to capture land changes over relatively short time frames and to normalize the 34-year input data, which vary dramatically in data richness over the study period. For each band and each epoch, we retained the maximum and the minimum value. At this stage, for each of the selected annual metrics, we had seven values corresponding to the five-year maxima of each epoch and seven values corresponding to the five-year minima of each epoch. From these seven values, we extracted eight new statistics: 1) maximum value; 2) minimum value; 3) mean value; 4) amplitude; 5) sum of the amplitude of change between each consecutive epoch; 6) maximum decrease between two consecutive epochs; 7) maximum increase between two consecutive epochs; and 8) sum of maximum

increase and maximum decrease between two consecutive epochs. Refer to Fig. 4-1 for a schematic representation of the inter-epochal metric set.

Yearly metrics are not appropriate for mapping natural tree regrowth and tree plantations because mapping a gradual process such as the growth of trees requires longer time-series. The multi-epochal metric-set captures changes in surface reflectance over the entire time-period, enabling us to effectively detect changes that may be too slow or too subtle to measure over a shorter time span. A single map of natural regrowth and a single map of plantation were created for the entire study period, and these include any tree growth event (whether natural or within a man-made plantation, respectively) regardless of permanence. We also used the same multi-epochal metric-set to create a single map of stable natural land cover for the entire continent for our study period. The stable natural land cover class contains enormous spectral variation because it includes many land cover types, including forests (of varying cover and height), shrublands, wetlands, grasslands, and even bare ground. This class is predominantly characterized by its spectral stability throughout the 30+ year time-period. The multi-epochal metric-set enabled us to target this stability in order to discriminate between stable natural land cover and all other land cover/land use classes. As with the cropland maps, we used a bagged classification tree algorithm and manually labeled Landsat training data to create each of these three maps.

3.5.3 Stratification and sample allocation

We joined the aforementioned maps as well as an additional map of surface water(178) to create eight strata for sampling. The crop maps were aggregated into

two strata: one maximum extent of all crop maps from 1985 to 1994 and another for all crop maps from 2016 to 2018. This approach was undertaken in order to target baseline crop areas as well as areas of crop gain through time. Additionally, we created a separate stratum for natural land cover within the Amazon biome by overlapping our natural land cover map with the tropical rainforest ecozone(154). Once all layers were joined, the remaining area was assigned to a final stratum of "leftover" land, which corresponded largely to all land uses excluding cropland and plantations (predominantly pasture and settlements) (Fig. S3-5).

A stratified random sample of 1,000 units was drawn (Fig. S3-6). Each unit corresponds to a single Landsat pixel. The distribution of sample pixels can be found in Table S1. Reference data for interpretation consisted of annual Landsat mosaics, bimonthly Landsat mosaics, time-series of NDVI, SWIR I, and red reflectance, and high-resolution Google Earth imagery wherever available. A user interface was created to aggregate all reference data and to collect interpretation data (see Fig. S3-9). For each sampled pixel, we recorded the land use/land cover class for every year in the study period and every visible degradation event, as well as the confidence level of our interpretation. The interface is openly available so that anyone may perform the interpretation independently to evaluate the results presented here: https://indus.umd.edu/SAM_1985_2018_samples. For detailed information regarding the land cover/land use hierarchical legend used for interpretation and the class definitions, refer to Fig. S3-8 and Table S3-2. Area estimates were derived from the sample using the R survey package (179) to implement the stratified sampling formulas following good practice recommendations (82, 83, 86, 153, 180) . We

compared our results to those derived from a large sample for Brazil obtained from the MapBiomas/LAPIG (89) land cover and land use change project and obtained comparable results (See Chapter 4, Fig. 4-4 and Fig. 4-10).

3.5.4 Degradation

From our sample, we estimated the cumulative amount of land that was impacted by a degradation event since 1985. Depending on the ecosystem, degraded natural vegetation may fully recover and revert to its natural state after some time. Time-series of Landsat imagery allow us to identify certain degradation events (logging, fire, edge effects), but assessing fine-scale impacts on structure or species composition is not feasible. Thus, we cannot rely on these data to assess whether a sampled pixel recovered from a degradation event. Although we refer to this land as degraded land, it should be interpreted as land that has been affected by an observable degradation event at some point during the study period.

An important limit of our method is that it likely omits a portion of degradation events, leading to the underestimation of the amount of land that was affected by a degradation event since 1985. Omissions can happen if a degradation event occurred at a time that did not coincide with a satellite overpass, or if a degradation event had impacts that are not detectable by the Landsat instrument. Both the fine spatial scale and ephemeral nature of degradation events limit their comprehensive quantification using Landsat data (181, 182). We also cannot reliably detect degradation events that occurred before 1985. See SI for comparison of our results to those of a recent study (181).

3.5.5 Human impact on natural land cover

In order to estimate the human impact on natural land cover, we aggregated our land cover/land use classes into four categories based on the degree of modification or conversion and the intensity of land use. At the top of the human impact scale is the complete and permanent conversion of land from its natural state to intensive economic land uses (e.g., cropland, pasture, plantation, infrastructure, etc.). In between intensive land use and intact natural land cover is a second category combining semi-natural lands and secondary forests. This second category includes land that has been subjected to human modification, but is not discernably economically productive. Semi-natural lands consist of 1) transitional lands, for example poorly maintained or abandoned land uses such as pastures that feature natural vegetation recovery, such as woody encroachment, without the complete establishment of a forest canopy and, 2) low intensity, small-scale land uses that may be cyclical such as shifting cultivation within the Amazon forest. The remaining two categories consist of natural tree cover that has experienced degradation, or canopy and biomass loss due to human-induced disturbance. These last two categories are distinguished from each other by whether the initial degradation was caused by fire or caused by logging or other types of biomass removal. Degraded tree cover can typically only be detected by detecting the degradation event itself, so the amount of degraded tree cover at the beginning of the study period is underestimated. To compensate for this, we use the 1988 degradation areas for 1985 to 1987 for the purpose of estimating the total change in human impact over the time period.

Human impact as we define it here differs from the human footprint defined by Sanderson et al. (2002)(183) and others(184) in that they consider a continuum of human influence on the natural environment by including proxies such as population density (to incorporate the degree of impacts on land) and accessibility (to incorporate potentially impacted areas). It also differs from the "human appropriation of net primary production" (HANPP) indicator, which models how energy in the form of biomass is appropriated by human actions(147, 185). Our results, on the other hand, focus only on measurable direct human impacts on the land at a 30m spatial resolution, and do not include distance or buffer measures, nor vegetation modeling.

3.6 **Supplementary information**

3.6.1 Comparison of our degradation area estimates with Bullock et al.

While there are differing definitions of forest degradation (138), we define it as the modification of tree cover that results in a partial loss of canopy cover and biomass. Data on degraded tree cover are still not extensive in the literature because of the difficulties associated with monitoring degradation using remotely-sensed satellite data. From the perspective of earth observation data, degradation is sub-pixel in spatial extent and ephemeral in terms of observational discrimination, posing significant challenges for its quantification. A recent degradation study (186) estimated the area of forest degradation and natural disturbance for the Amazon Ecoregion from 1995 to 2017 using Landsat data and good practices (83) for area estimation. The study estimated 36.63 ± 2.05 Mha of forests in the Amazon were affected by degradation and natural disturbance during the study period. Our study of

degradation starts in 1985 and by 2017, we estimate there to have been 39.2 ± 12.5 Mha of degraded tree cover in the Tropical rainforest ecozone (which largely overlaps with the Amazon Ecoregion), 30.0 ± 11.0 Mha of which occurred in the period corresponding to the study of Bullock et al.

3.6.2 Agricultural intensification in South America

Results from our study show a dramatic proportional increase in row crops and a stabilization of pastureland area, both indicating an increased intensification of land use. One way to assess the intensification of agricultural activities in South America is by looking at the change in economic output that is produced per unit of land. To do this, we used the value added from the agricultural sector for each country (187) and compared it to the increase in the area used for agriculture. In our study, agricultural area corresponds to the combined area of cropland, pastureland, short tree crops, and tree plantations (Figure S3-10). For all reportable areas, the increase in value added has outpaced the increase in agricultural area, confirming the intensification of agricultural production. Peru and Chile achieve much higher rates of return per hectare than Brazil or Argentina. Both Peru and Chile are producers and exporters of highly profitable cash crops, such as grapes, wine and apples in the case of Chile (188), and asparagus, grapes, and avocados in the case of Peru (189). Despite the fact that Peru and Chile lead the way in terms of maximizing profit per hectare, Argentina and especially Brazil are far ahead in terms of absolute profits derived from agricultural production. In these graphs, we see reflected two distinct agricultural strategies playing out: the strategy of producing high-margin crops in reduced areas versus that of producing vast quantities of less profitable crops, largely

feedstocks, over very large areas. The latter strategy is a function of the availability of large areas of land for agricultural extensification. The ability to dramatically expand the area under production results in much higher overall returns. Production is also a function of improving yields, and like Peru and Chile, intensification of agriculture in Brazil and Argentina has increased over time, with Brazil demonstrating greater relative gains. The value added used for this comparison includes forestry, hunting, and fishing, as well as cultivation of crops and livestock production. The inclusion of fishing may skew these results, especially for Chile because of the importance of fish exports relative to other agricultural exports.

3.6.3 Supplementary figures and tables

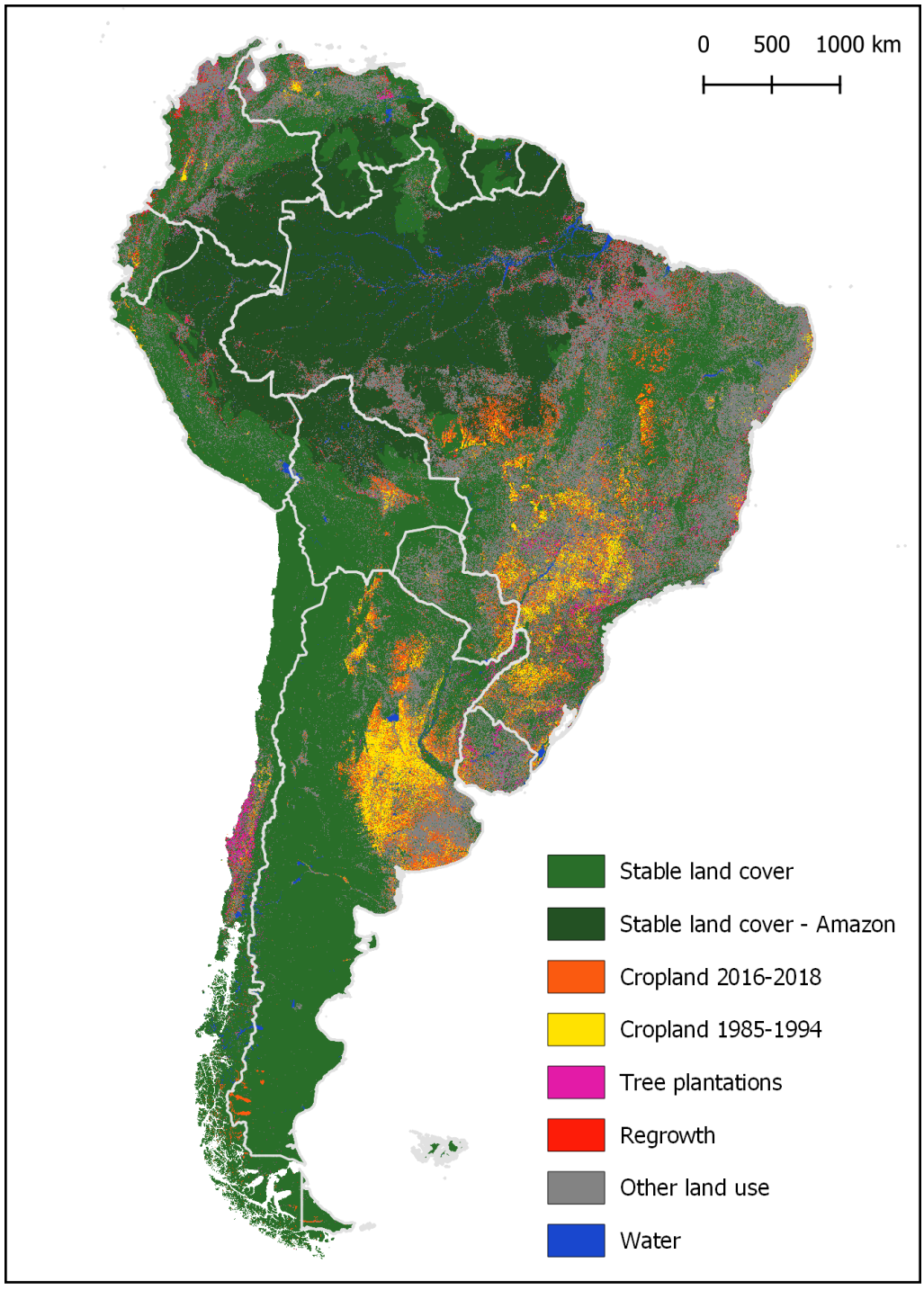


Figure S3-5 Study area and sampling strata

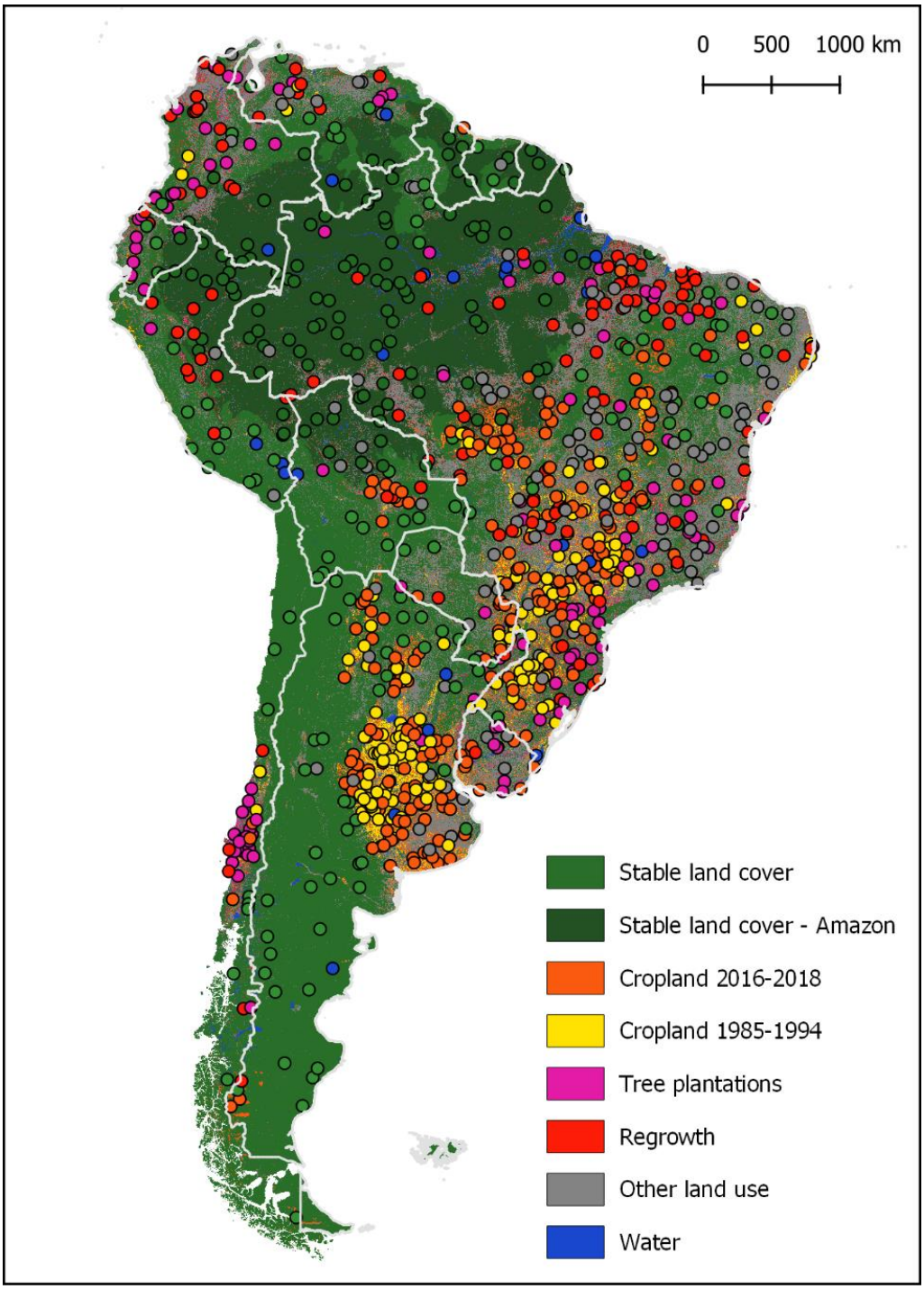


Figure S3-6 Sampled pixels by mapped strata

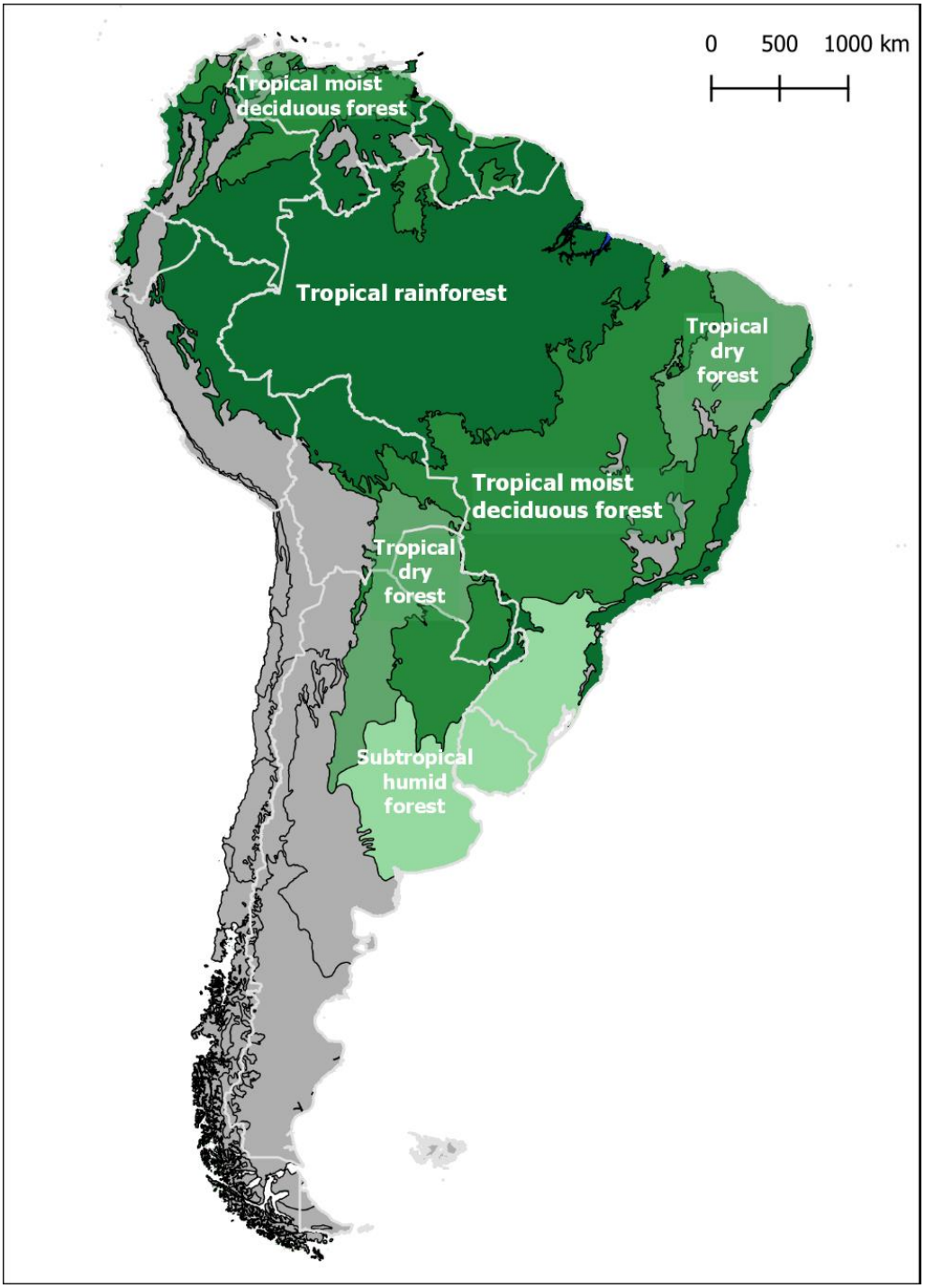


Figure S3-7 Country and FAO ecological zone boundaries. We report area estimates for ecozones colored in shades of green: Tropical rainforest, Tropical moist deciduous forest, Tropical dry forest, and Subtropical humid forest.

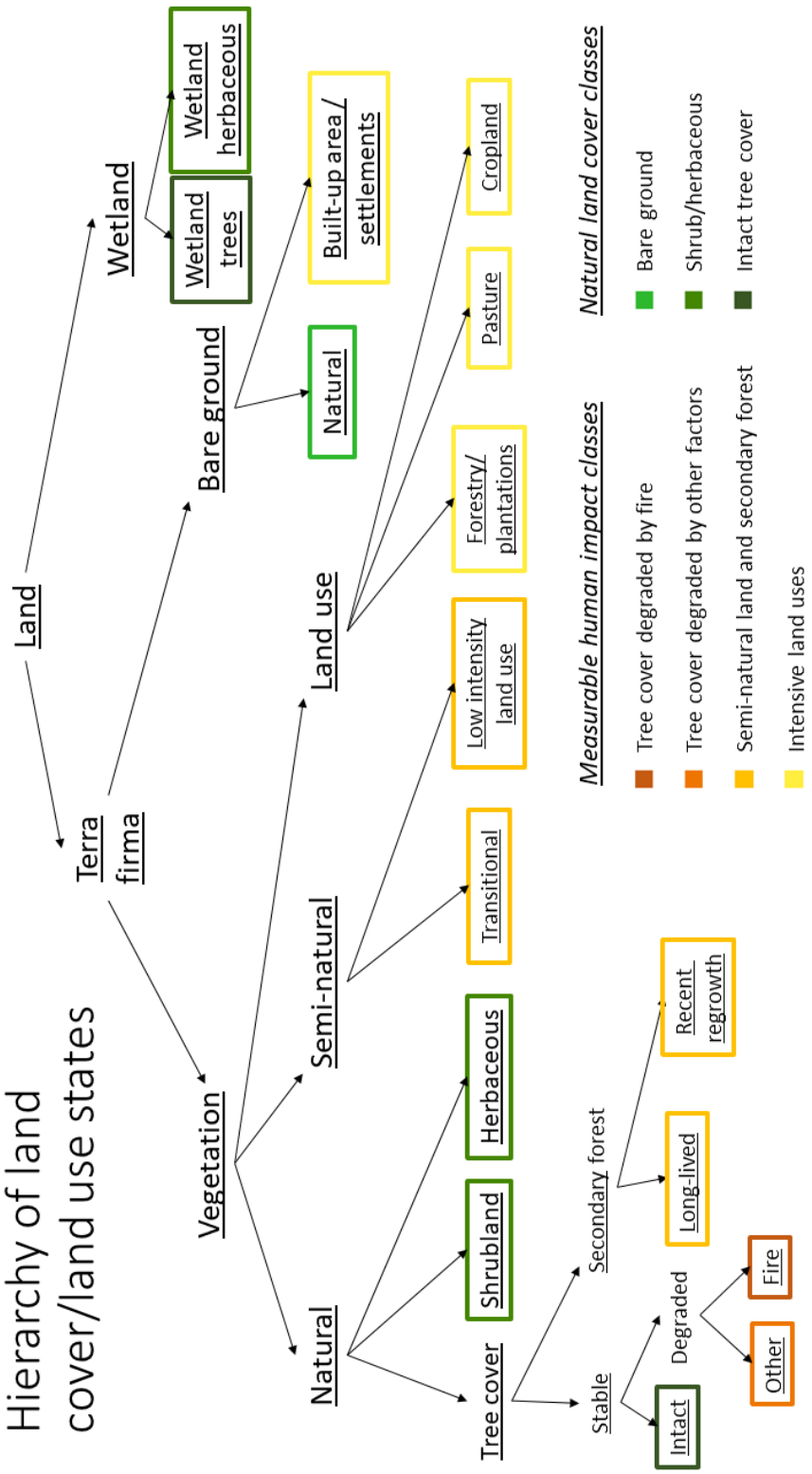


Figure S3-8 Hierarchical legend used for sample interpretation and for reporting.

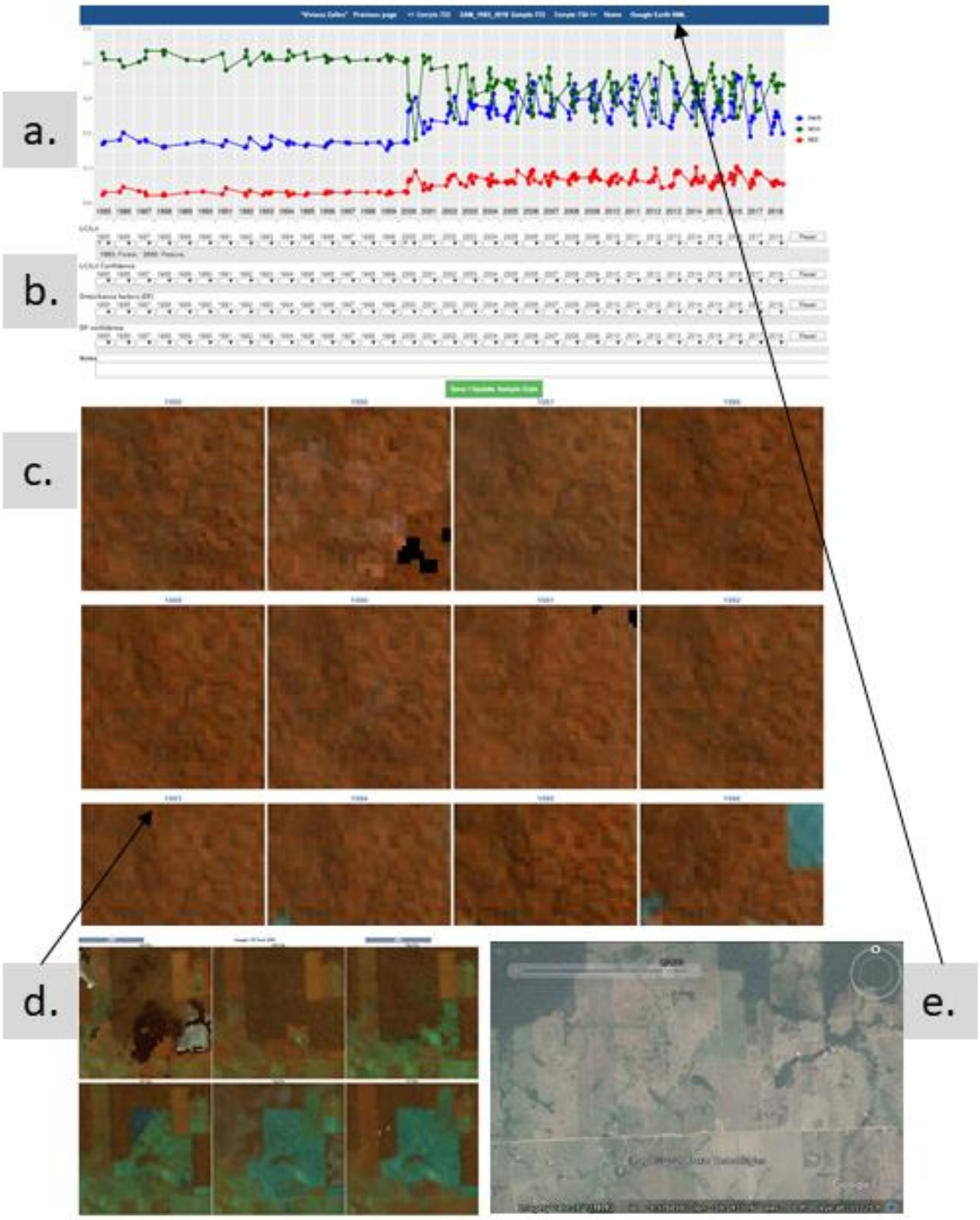


Figure S3-9 Sample interpretation interface. For each sampled pixel, the following data were available to facilitate interpretation: **a.** Time-series of 16-day NDVI, SWIR I, and red reflectance; **b.** drop-down menus for selecting land cover/land use type, degradation event, and confidence level of the interpretation; **c.** annual Landsat composite in NIR/SWIR1/SWIR2 in RGB; **d.** Bi-monthly composites in NIR/SWIR1/SWIR2 in RGB; and **e.** .kml of sampled pixel for Google Earth visualization. Interface access: https://indus.umd.edu/SAM_1985_2018_samples

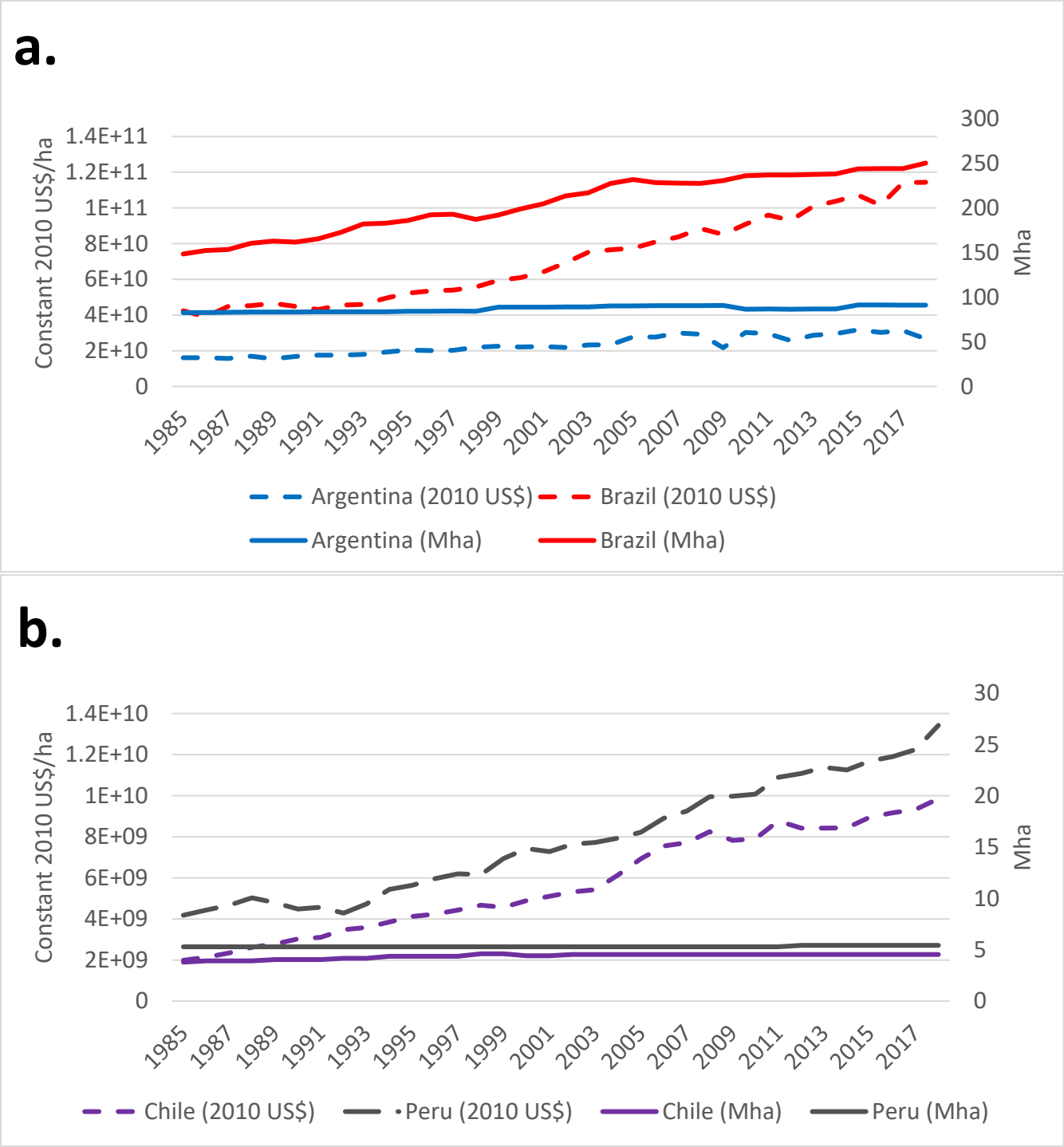


Figure S3-10 Value added from agriculture (includes forestry, hunting, fishing, crops, and livestock production) in constant 2010 US$ and agricultural area (includes pasture, cropland, short tree crops, and tree plantations) in Argentina and Brazil **(a)** and Chile and Peru **(b)**.

Stratum	Stratum size (pixels)	Sample size (pixels)
Other land use	5,094,171,779	120
Stable land cover - outside Amazon	10,394,998,299	124
Stable land cover - Amazon ecoregion	6,666,978,196	115
Regrowth	425,036,362	127
Plantation	267,397,540	115
Recent cropland	956,503,823	232
Early cropland	647,344,973	143
Water	230,196,241	24

Table S3-1 Sample distribution

Table S3-2 Definitions, distinguishing characteristics, and example sample pixels for each land cover/land use class used in this study. Example sample pixels can be visualized at indus.umd.edu/SAM_1985_2018_samples

Land cover/land use class	Definition	Distinguishing characteristics	Example sample pixel number
Tree cover	Canopy cover >10% and height >5m and no evidence of current or past land use or disturbance.	Red reflectance ≤ 5% for parklands and 3% for forests, high NDVI values	117, 794
Wetland trees	Tree cover (>10% canopy cover and height >5m) on lands transitional between terrestrial and aquatic systems where the water table is usually at or near the surface or the land is covered by shallow water.	Landscape spatial pattern revealing hydrological features, and generally low red and SWIR reflectances. Tree cover where there is intermittent flooding, hydrological topology features present, or proximity to rivers	502
Long-lasting secondary forest	Naturally regenerated tree cover (>30% canopy cover and height >5m) in 1985 displaying evidence of past modification.	Tree cover with high and uniform near-infrared reflectance compared to intact forests.	83 (from 1985 to 2007)
New secondary forest	Naturally regenerated tree cover (>30% canopy cover and height >5m) that experienced a land cover modification during the study period.	Very high near-infrared reflectance characteristic of uniform canopy recovery and often mono-specific stands of colonizing tree	632

		species. Tree cover following a different land cover	
Natural herbaceous	Lands covered by herbaceous vegetation and no evidence of current or past land use or disturbance.	High SWIR values, red reflectance ≥4000, low NDVI	672
Natural shrub	Lands covered by woody vegetation (>1m and <=5) and no evidence of current or past land use or disturbance.	Red reflectance higher than for tree cover, and NDVI lower than for tree cover, dependent on % cover	114, 537
Wetland herbaceous	Herbaceous vegetation on lands transitional between terrestrial and aquatic systems where the water table is usually at or near the surface or the land is covered by shallow water.	Landscape spatial pattern revealing hydrological features, and generally low SWIR reflectances. Herbaceous vegetation where there is intermittent flooding, hydrological topographic features present, or proximity to rivers	548
Bare ground	Lands of exposed soil, sand, rocks, or snow.	High SWIR values, low NDVI	251

Cropland	Intensive row crop agriculture.	Large within-year variation in SWIR, NDVI, and red reflectance, repeating from year to year. Homogeneous cover and presence of rows.	155
Pasture	Herbaceous land cover used for grazing.	Generally lower NDVI and brighter red and SWIR reflectances than croplands. Herbaceous cover, planted or otherwise, presence of watering holes, cow trails visible in high resolution imagery.	759
Short tree crops	Perennial agriculture produced on planted trees and shrubs	Uniform spectral response, medium red and SWIR reflectances. Stable tree cover or shrub cover planted in rows.	491
Plantation forests	Intensively managed planted forest that at maturity is composed of one or two species, has one age class, and has regular tree spacing	Uniform spectral response, low red and SWIR reflectances. Tree cover planted in rows with cycles of clear-cutting.	233
Built-up area / Settlements	Urban areas, settlements, roads, mining.	Bright reflectance values and low NDVI. Geometric patterns, stable but heterogenous reflectance	974

		because of the variety of land covers included.	
Low intensity land use	Very low intensity, small-scale land uses. Includes rotational agriculture.	Small-scale, can have shrub, herbaceous, or bare ground reflectance. In the presence of rotational agriculture, periods of bare ground and herbaceous cover interspersed with shrubs and tree cover.	171
Transitional land	Vegetation not planted by humans nor in its original, intact state: land that is in the process of being converted to a land use or land that has been abandoned and is in the process of returning to its natural state.	Shrub or herbaceous cover, delimited area suggesting property limits, preceding or following a distinguishable land use.	955 (from 2005 to 2018)

Chapter 4: South America land cover/land use change mapping and area estimation – Evaluation of the methods and results[3]

4.1 Abstract

Maps of land cover and land use extent and change provide crucial information that enables a better understanding of associated impacts to natural ecosystems. They are also a key input for land use planning and management decisions. South America is home to extensive natural ecosystems, including large tracts of pristine rainforest, as well as growing pressures associated with increased demand for agricultural commodities. However, detailed land cover and land use extent and change maps for South America are not currently available. Likewise, unbiased area estimates associated with these changes are lacking, but urgently needed to provide definitive information on trends. Here, we present a continental map of land cover/land use for South America produced using Landsat data from 1985 to 2018. Mapped themes include cropland, plantation, natural tree regrowth, other land uses, and stable land cover (inside and outside the Amazon). The map is used to define strata to implement a stratified random sample of 1,000 Landsat pixels from which area estimates of important land cover/land use changes classes are produced. The continental map is evaluated in several ways, first via accuracy assessment. We find that cropland and stable land cover were mapped accurately, while plantations and regrowth less so. Second, our area estimates for Brazil were comparable to those resulting from an analogous study that employed a substantially

[3] The material in this chapter was co-authored. Co-authors include Matthew C. Hansen, Peter Potapov, Stephen V. Stehman, and Leandro Leal Parente (LAPIG lab, Universidade Federal de Goiás).

larger sample size. Third, using the maps to define strata for the sample allocation yielded substantial reductions in standard errors relative to simple random sampling for estimates of cropland and plantation areas, while other categories suffered modest increases in standard errors. We present lessons learned from using a single map to stratify the allocation and interpretation of sample-based area estimates for a multitude of target classes.

4.2 Introduction

Although humans have altered the global land surface for millennia, we have only recently been able to measure these changes directly through the analysis of satellite imagery. The longest continuous record of earth observation data from satellites is the Landsat program, with the first satellite launched in 1972. The Landsat program was first announced by the United Stated Secretary of the Interior Stewart Udall in 1966 as the "beginning of a great decade in land and resource analysis" 31. This was an understatement, as the now nearly 50 year record of Landsat acquisitions consists of global systematic imaging and free and open data distribution, policies that allow researchers to study and measure long term land dynamics from local to global scales. Recent advances in computing capabilities and developments in classification algorithms enable the analysis of larger areas and for the entire multi-decadal time-series. Large-scale and even global datasets characterizing long-term surface change at 30-m spatial resolution are now emerging due to the combination of all these factors (50, 89, 174, 175, 178).

Efforts to map land cover and land use change in South America using remotely sensed data abound, and yet significant gaps remain. Tree cover and tree

cover loss are the most studied land cover themes in the continent with the Brazilian Amazon the most studied region in large part due to governmental efforts to monitor deforestation in the Legal Amazon(48). Since 2013, tree cover loss can be monitored beyond the Amazon thanks to the continuously updated Global Forest Change product(50). Other regions and other themes are far less well-characterized. Significantly, information on important commodity land uses such as crops, tree plantations, and pasture are sorely lacking. Again, Brazil is the exception: a recent project called MapBiomas provides yearly maps of land cover and land use for the entire national territory for 1985 to 2018 at Landsat-resolution. Although highly valuable, the project only covers Brazil, thus missing significant land cover/land use dynamics in other important regions of the continent such as the Chaco and Chiquitania ecosystems. Because the causes and consequences of land cover/land use change can cut across national borders, continental-scale land cover/land use maps are critical to providing a better understanding of the spatial patters of change.

Chapter 3 of this dissertation provided area estimates for important land cover and land use change dynamics across the entire South American continent for 1985 to 2018, with a specific emphasis on commodity land uses. Those results were derived from a probability sample that was stratified based on a land cover/land use map of South America. In this chapter, the data and methodology used to create this map are thoroughly described. Further, this chapter provides a multipronged evaluation of the map and its utility as a spatially explicit characterization of land cover/land use and as an effective mechanism for facilitating the area estimation of continental scale land cover/land use extent and change. We provide an analysis of the classification tree

models employed to map the strata. We then estimate the accuracy of the mapped classes using the stratified random sample as our reference data to determine the quality of the map. The area estimates derived from the sample stratified by the mapped classes are then compared to those of the MapBiomas project which used an analogous stratified sample with a substantially larger sample size. Finally, we determine the sampling efficiency provided by the use of the map as a stratifier by determining the sample size that would have been required to achieve the reported standard errors had a simple random sample been employed. We conclude by discussing the lessons learned from the method in the context of a limited set of map strata used in the estimation of a multitude of land cover and land use themes.

4.3 **Data processing and metrics creation**

4.3.1 Landsat data pre-processing

All available Landsat Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+), and Operational Land Imager/Thermal Infrared Sensor (OLI/TIRS) scenes from January 1st 1985 to December 31st 2018 for the South American continent were downloaded from the USGS EROS data center. The total number of Landsat scenes downloaded and processed for this project totaled nearly half a million. The data were processed using an established methodology that has been developed and successfully applied for a number of projects (50, 137, 142, 175, 176).

The downloaded Landsat data were radiometrically calibrated and orthorectified (Level 1 Tier 1 collection). The first step in the pre-processing chain was the conversion or raw digital numbers (DN) to top of atmosphere (TOA)

reflectance. This conversion was performed using the equations and coefficients described by Chander et al. (133). The TOA reflectance values were then used as inputs to a per-pixel quality assessment to determine which observations were suitable for analysis. Our quality assessment models consisted of a set of decision tree classifiers that determined the per-pixel likelihood of contamination by cloud, haze, snow, shadow, and whether the pixel corresponded to a land or water observation.

Clear-sky land observations were then used to perform image normalization in order to correct for atmospheric contamination and BRDF effects. Radiometric normalization was necessary to ensure consistency of reflectance values through space and time, so that classification models could be successfully applied over large areas and over different years. The normalization was performed using a reflectance target created from a growing season average reflectance of all 16-day MODIS composites from the MODIS 44C product(190) from 2000 to 2011. Pseudo-invariant objects were selected by excluding pixels where the difference in reflectance values for the red and shortwave infrared bands for the Landsat image and MODIS normalization target was more than 0.1. The per-band median bias between MODIS and Landsat pseudo-invariant objects was calculated and regressed against the distance from the Landsat ground track. This per-band linear regression model was then applied to all pixels within the Landsat image to adjust for surface anisotropy as a function of cross track pixel location. This method was a relative normalization of Landsat to a top of canopy MODIS reference reflectance and did not constitute an absolute correction.

Finally, the data were aggregated into 16-day composites on a per-pixel basis using the best available observation for each pixel (this was determined by the QA flags attached all pixels within defined compositing periods). Each step in this Landsat data processing chain has been thoroughly described by Potapov et al.(132, 175, 176). In addition to Landsat data, we also used digital elevation data from the Shuttle Radar Topography Mission (SRTM). Slope and elevation metrics were derived and resampled to the Landsat pixel grid.

4.3.2 Multi-temporal metrics

Multi-temporal metrics are values derived from time-series of observations and have been used extensively for land cover classification using AVHRR, MODIS, and Landsat data (50, 135, 175, 191, 192). Multi-temporal metrics represent a generic feature space that allowed us to apply classification algorithms over large areas despite the variability in observation frequency through space and time. Different metric sets can be created depending on the mapping objective: if mapping static land cover classes, we used annual phenological metrics; if mapping land cover change, we used multi-year change detection metrics (175, 176). In this study, our first level metrics were annual and derived from gap-filled time series of 16-day composites.

To create the annual phenological metric set, we used all Landsat spectral reflectance bands as well as the brightness temperature band. We also computed normalized ratios for the following band pairs: shortwave infrared (SWIR 1.6 μm)/near-infrared (NIR), blue/green, blue/NIR, blue/red, green/red, green/NIR, and SWIR (1.6 μm)/SWIR (2.2 μm). Time-series for each band were ranked in two different ways: first, they were ranked on the spectral band reflectance or index

values. From these ranked observations, we extracted the highest/lowest, second to highest/lowest, and median reflectance values for each band and index, as well as the mean of all observations in each quartile (0-25%, 25-50%, 50-75%. We also extracted averages for all observations between 1) the second lowest value and the 50th percentile of observations, 2) the 50th percentile of observations and the second highest value, 3) the second lowest value and the 25th percentile of observations, 4) the 75th percentile of observations and the second highest value, 5) the average between the 25th and 75th percentiles of observations, 6) the average of all values, and 6) the average of all values except the minimum and maximum values. Second highest value metrics were meant to mitigate against the presence of poor quality observations missed by the quality assessment models that can deleteriously impact maximum and minimum-based metrics. We also ranked all observations based on corresponding values of NDVI and then again based on corresponding values of brightness temperature. For each of these, we extracted all the same metrics mentioned above. Once all metrics were computed, we were able to create amplitude metrics for each, including the differences between: 1) the maximum value and the minimum value, 2) the second maximum and the second minimum, 3) the means of the first two quartiles and the second two quartiles, and 4) the means of the first quartile and the fourth quartile (132, 175, 176).

4.3.3 Multi-epochal metrics for tree cover dynamics and natural land cover extent

An additional metric set was created using the annual phenological metrics described above as inputs. Instead of using annual change detection metrics, we created this new set of metrics to capture land cover change across the entire 34-year

time period, regardless of timing. To create this multi-epochal metric set, we selected a set of metrics from the annual metric set: the second highest, second lowest, and mean of the 25^{th} to 75^{th} percentile of observations of the red, NIR, SWIR 1 (1.6 μm), SWIR 2 (2.2 μm), and NDVI. We then aggregated each of these into five-year epochs (six five-year epochs and one four-year epoch corresponding to 2015-2018). Epochs were employed to capture land changes over relatively short time frames and to normalize the 34-year input data, which varied dramatically in data richness over the study period. For each band and each epoch, we retained the maximum and the minimum value. At this stage, for each of the selected annual metrics, we had seven values corresponding to the five-year maxima of each epoch and seven values corresponding to the five-year minima of each epoch. From these seven values, we extracted eight new statistics: 1) maximum value; 2) minimum value; 3) mean value; 4) amplitude; 5) sum of the amplitude of change between each consecutive epoch; 6) maximum decrease between two consecutive epochs; 7) maximum increase between two consecutive epochs; and 8) sum of maximum increase and maximum decrease between two consecutive epochs. Refer to Fig. 4-1 for a schematic representation of the inter-epochal metric set.

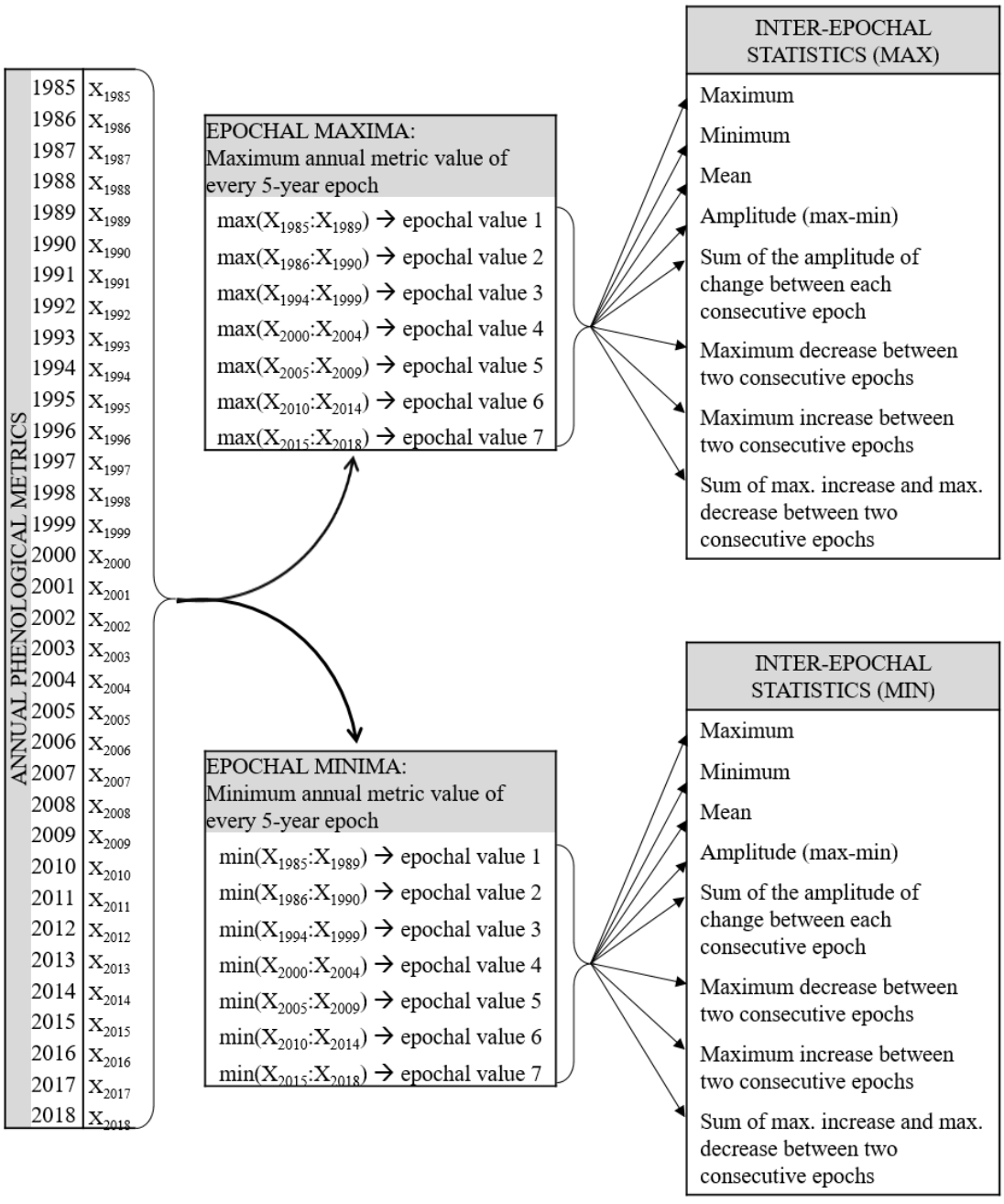


Figure 4-1. Schematic representation of the multi-epochal metric set. Annual phenological metrics used include the penultimate highest value, penultimate lowest value, and interquartile mean of annual red, NIR, SWIR 1 (1.6 μm), SWIR 2 (2.2 μm), and NDVI. X_i represents the value of a given annual phenological metric for year *i*. The inter-epochal metric set is composed of all "max" and "min" inter-epochal statistics for each of the 15 selected annual phenological metrics, totaling 240 new metrics encompassing 34-years of spectral reflectance.

4.4 **Methods**

4.4.1 Creating map strata for sample allocation

We created four independent land cover products to stratify the South American continent with a focus on commodity land uses for which a feasible prospect for mapping existed. Our first mapped theme was stable land cover, or lands that did not exhibit spectral change over the time-series. While some land uses are stable spectrally, the goal was to create a stratum for which commodity land uses were absent, thereby improving the efficiency of the sample-based estimates (i.e., reducing standard errors). The remaining strata contained our target commodity land use themes. Chapter 2 demonstrated the capability for mapping cropland at national scale over a long-time series, and we extended that general approach to the continental scale and 34-year TM/ETM+/OLI imaging period. We also added plantation land use as a second commodity theme, given its growing importance in terms of land use change as well as the rather unambiguous spectral dynamics of forestry and other tree crops over time. We also added regrowth as an individual map theme in order to separate natural forest recovery dynamics from commercial ones. Arguably, the most significant commodity land use in terms of natural vegetation loss is pasture. However, pasture is a challenging land use to map, and instead of mapping, we opted to create an 'other land use' stratum within which we expected pasture to be present. Cropland was mapped using the annual phenological metric set, and individual cropland maps were created for each year from 1985 to 2018. Stable land cover, tree plantations, and tree regrowth were mapped using the

aggregate multi-epochal metrics, which were designed to highlight change over the entire 34-year time-period regardless of when it occurred.

To create each map, training data were collected using a manual pixel-labeling approach based on visual interpretation of Landsat imagery and aided by the use of high spatial resolution Google Earth imagery. The training dataset for each map consisted of pixels labeled as belonging to the target class and pixels labeled as not belonging to the target class (e.g. crop and non-crop, plantation and non-plantation, etc.). We then used a bagged decision tree algorithm to classify the entire continent into a per-pixel probability of belonging to the target class. Decision tree algorithms are a type of machine learning algorithm that recursively partitions the training dataset into increasingly homogeneous subsets using a deviance (entropy) measure (136). The algorithm recursively splits the dataset maximizing the decrease in deviance at every split, and continues until the splitting procedure reduces the deviance by less than a user-defined threshold (in this case, 1% of the initial training data set deviance). Decision trees are distribution-free, nonlinear, and able to handle missing data (193). Because decision trees tend to overfit, we used bootstrap aggregation (bagging) (194) to ensure a more stable and accurate classification model. Our final map classifications were each derived using 21 bagged tree models with each model created from a 20% simple random sample with replacement of the training dataset. The output of the classification was a per-pixel probability layer, which we thresholded at 50% to create a binary map for each land cover/land use theme. We used an iterative approach for each classification, meaning that after training data collection, we ran the decision tree algorithm, assessed the results, and

continued adding training data in order to correct classification errors. We repeated this cycle until the resulting classification was considered to conform to known geographic land cover and land use distributions.

A map was created for each of the stable land cover, tree plantations, and tree regrowth classes using the multi-epochal metric set and a bagged decision tree model. The stable land cover map represents land on which no significant land cover/land use changes occurred from 1985 to 2018. Most of this land is natural vegetation or natural land cover (such as bare ground), but there are some areas of very low intensity land use that are also included within this map, such as the natural grasslands used as pasture of the Pampas biome. The stable land cover map was disaggregated into stable land cover within the Amazon region (mostly tropical rainforest) and stable land cover outside of it. The tree plantation and tree regrowth maps represent tree plantations and natural tree regrowth that may have been present or expanded at any given time from 1985 to 2018, regardless of whether harvest or clearing occurred by the end of the study period.

Cropland, on the other hand, was mapped on a yearly basis using the annual phenological metric set. A stable model of cropland classification was created by employing three years of training data (2016, 2017, and 2018) and their respective phenological metric sets. The resulting model was then applied to each year's metric set from 1985 to 2018, which resulted in 34 annual cropland maps. Data gaps during the early years of the study period caused classification problems, leading to poor results. However, when aggregated together, the maps are representative of areas where crops were present at some point during our study period over the entire

continent. Similar to the epochal maps used for stable land cover, plantations, and regrowth, we aggregated the cropland maps into epochs to overcome data limitations in the early Landsat TM-only data record, which lacks the systematic acquisition strategies of later ETM+ and OLI/TIRS sensors. Specifically, we created 1985-1994 and 2016-2018 maximum cropland extent maps as inputs for continental-scale stratification in combination with the stable land cover, plantation and regrowth maps.

The five maps derived from the Landsat metrics, with stable land cover divided into Amazon and non-Amazon, were aggregated to create a single continental-level map of South American land cover and land use. When combining the layers, the following priorities were given: 1. Maximum cropland 2016-2018; 2. Maximum cropland 1985-1994; 3. Tree plantations; 4. Tree regrowth; 5. Stable land cover within the Amazon; 6. Stable land cover outside the Amazon. We added a surface water map (178) which took priority over all other classes. Land that remained unclassified corresponds to all other land use classes, and is our proxy for pasture, the single largest land use in South America. Refer to Fig. 4-2 for a schematic representation of the workflow used to create the different maps, including the leftover theme of 'other land use'. The final map can be seen in Fig. S3-5 (Chapter 3).

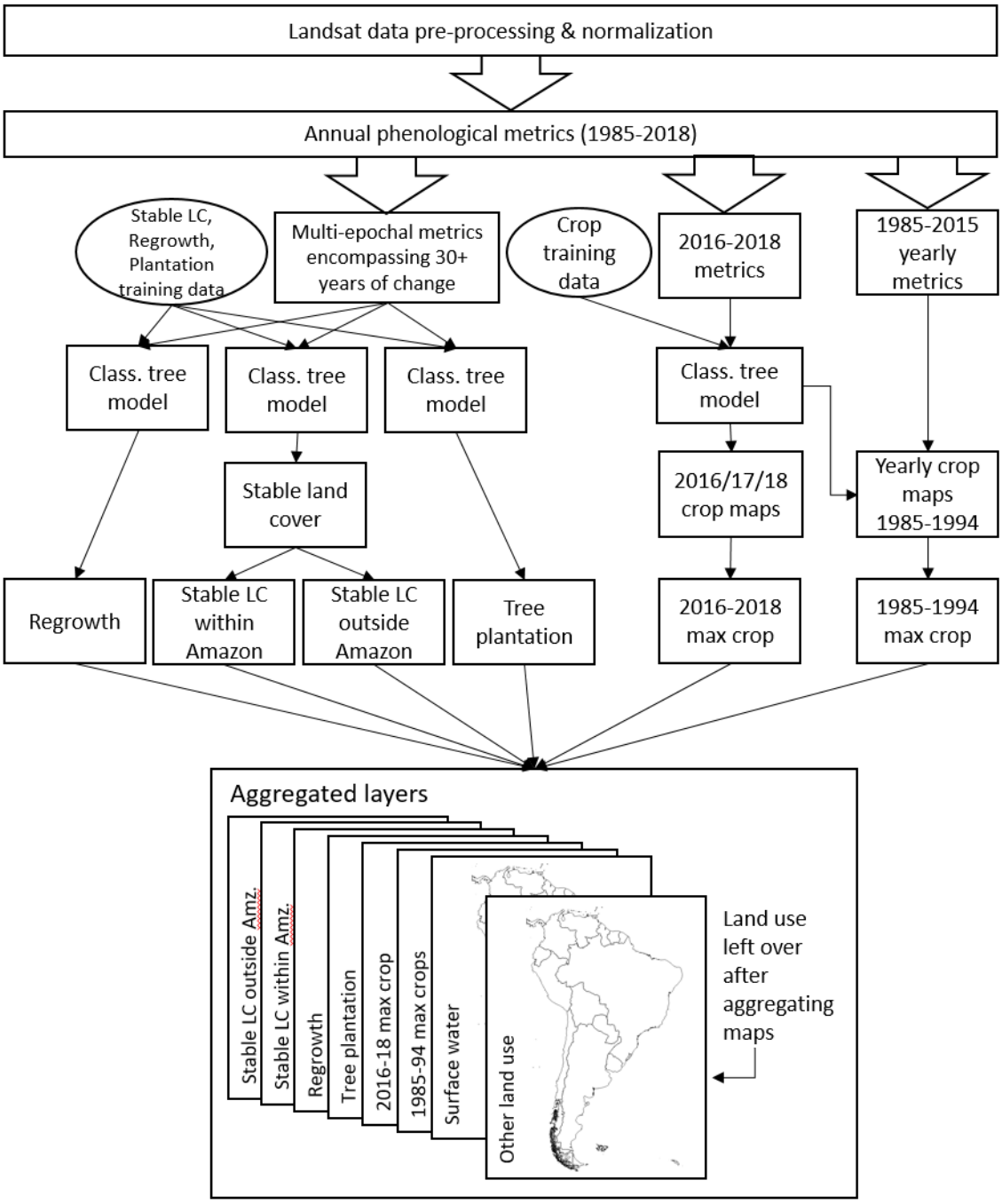


Figure 4-2 Schematic representation of the workflow used to create cropland, tree plantation, tree regrowth, and stable land cover maps.

4.4.2 Evaluations

The main purpose of this study was to assess the performance of the maps, both as a spatially explicit characterization of land cover and land use, and as a tool to improve precision of area estimates of continental scale land cover and land use change. The final map served as a stratification for allocating the reference sample to better target our land cover and land use themes of interest. Reference data were collected from a stratified random sample of 1000 sample pixels and used for estimating land cover and land use extent and change areas for South America. These data and associated results are described in Chapter 3 with the sample distribution shown in Table S3-1 and Fig. S3-6.

Evaluation of the maps as inputs to reference data allocation and area estimation consisted of the following four analyses: 1) assessment of the decision tree model outputs, 2) accuracy assessment of overall, user's and producer's accuracies per mapped land cover and land use theme, 3) comparison of area estimates with MapBiomas data for Brazil, and 4) quantification of sampling efficiencies per mapped land cover and land use theme.

The assessment of the classification tree model outputs was done by analyzing the overall reduction of deviance per input time series metric per target land cover and land use theme. The accuracy of the map was assessed using reference data from a stratified random sample. Reference labels were compared to mapped strata themes and overall, user's (commission error) and producer's (omission error) accuracies calculated.

The third evaluation consisted of a comparison between our area estimates with area estimates for Brazil produced by the MapBiomas project. MapBiomas is a collaborative Brazilian initiative made up of NGO's, universities and private companies whose goal is to create an all-encompassing land cover/land use monitoring platform. They have published yearly Landsat-derived maps of land cover and land use for 1985 to 2018, which are openly available through their website (mapbiomas.org). As part of this project, the Image Processing and GIS Laboratory (LAPIG) at the Federal University of Goiás (UFG) undertook the task of creating a sample-based reference dataset for area estimation and validation of the MapBiomas maps. Because the MapBiomas project covers a significant part of our study area (the entire national territory of Brazil), covers the same study period (1985-2018), and uses a good-practice (83) approach to area estimation (area estimates are derived from a probability sample, as opposed to pixel-counts), it is an ideal dataset with which to compare our results.

In order to estimate land cover/land use areas, LAPIG used a two-level stratified random sample approach. The strata were determined based on a rectangular grid dividing the country into 127 strata (the first level of stratification) which were overlapped with a map of slope classes that served as a second level of stratification (six slope classes: plain, smooth wavy, wavy, wavy strong, mountainous, and craggy). Their stratified random sample consisted of 85,152 sample pixels (sampling unit is a Landsat pixel) and a team of 19 analysts worked for eight months to interpret all sample pixels.

Comparing area estimates from our sample to the MapBiomas/LAPIG sample is difficult because each study employed different response designs. Our land cover/land use classes and definitions are consistent across the entire study region and are based on the hierarchical legend in Fig. S3-8 (Chapter 3). Their land cover/land use categories and definitions are different from the ones we used and can vary depending on the biome (e.g., what is defined as "forest" is not the same in the Amazon and the Cerrado biomes) (195). To compare results, we merged MapBiomas classes to derive new classes that would be consistent across the two studies (Table 1). From these aggregate classes we compared the MapBiomas estimates with our results using mean absolute relative difference and mean absolute difference over all years between annual MapBiomas area estimates and those from this study.

New merged classes	**Land cover/land uses included**	**MapBiomas/ LAPIG classes**	**Classes in this study**
Natural tree cover, shrub, wetland	Tree cover, natural woody vegetation, wetlands (trees and herbaceous)	Forest formation, savanna formation, other non-forest natural vegetation	Tree cover, long-lasting secondary forest, new secondary forest, natural shrub, wetland trees and wetland herbaceous
Herbaceous (natural and pasture) and non-natural shrub	Terra firma herbaceous cover, transitional lands, smallholders, pasture	Grassland, pasture	Pasture, natural herbaceous, transitional land, low intensity land use.
Crops	Cropland (intensive row crop agriculture)	Annual crops, semi-perennial crops	Crops
Plantation	Tree plantation	Tree plantation	Plantation
Perennial crops	Perennial agriculture (crops growing on trees and shrubs)	Perennial crops	Short tree crops

Table 4-1 New merged classes for comparison with MapBiomas/LAPIG area estimates

The final evaluation consisted of assessing the efficacy of the map strata in targeting our land cover and land use change themes of interest in terms of the reduction in standard errors of the area estimates. Estimating area using a stratified random sampling design has been shown to provide gains in efficiency in terms of the number of sample units needed to achieve target uncertainty estimates, as compared to using a simple random or a systematic sampling design (196, 197). Using the population mean of the area estimates in Chapter 3, we can calculate the number of sample units that would have been required to match the standard error of our estimates had we used a simple random sample. Given that a primary objective of the study was quantifying the growth of commodity land uses, we expected the area estimates of croplands and plantations to benefit from the stratified sampling approach. The other important land use, pasture, was not targeted directly, but it was expected that the 'other land use' stratum might convey some benefit in facilitating precise area estimation. Efficiency gains for estimating pasture area were difficult to predict in this scenario. Area estimates of the stable land cover and regrowth classes were similarly expected to benefit from the mapped strata. As for land cover and land use themes not directly related to the strata, gains in efficiencies would largely reflect their correlation with the themes of the mapped strata.

4.5 **Results**

4.5.1 Decision tree model outputs

Four different decision bagged tree models were used to derive the classes from which the final map was composed. Each of these models can be analyzed in terms of the percent per-metric contribution to overall decrease in deviance. Table 4.2 shows the metrics that were most useful in reducing deviance, as well as the total reduction in deviance by the end of the splitting procedure for each model. Ideally, the top metrics provide a significant reduction in deviance, indicating that the classes are relatively easily separable and can therefore be accurately mapped.

Table 4-2 Deviance reduction provided by the three most important metrics of each decision tree model and overall deviance reduction per model.

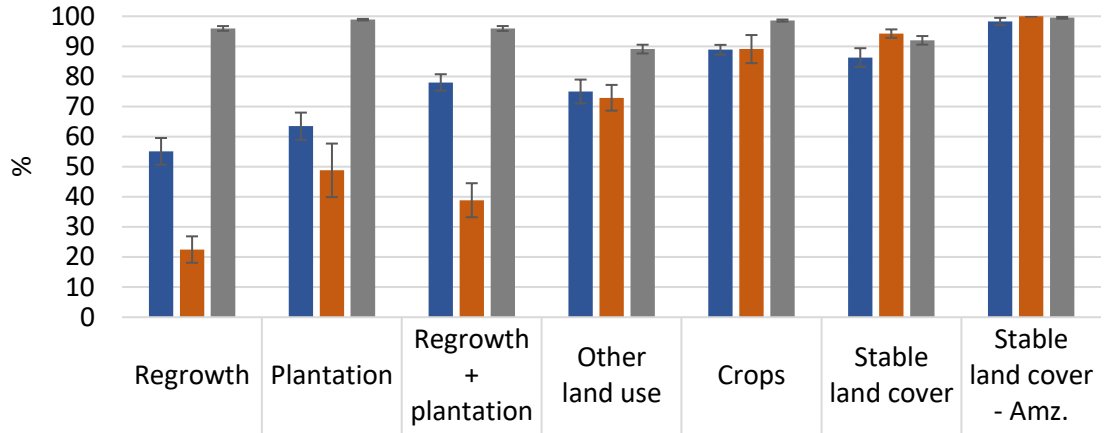
Stable land cover	**Deviance decrease (%)**
Annual metric-interquartile near-infrared reflectance, mean of epochal minima	39.0
Annual metric-interquartile mean NDVI, amplitude of epochal maxima	5.4
Annual metric-penultimate lowest NDVI, amplitude of epochal maxima	3.8
Top 3 metrics	48.3
All metrics	98.5
Plantation	**Deviance decrease (%)**
Annual metric-interquartile mean of shortwave infrared 1.6 μm, amplitude of epochal maxima	19.3
Annual metric-penultimate lowest NDVI, maximum of epochal minima	9.4

Annual metric-penultimate lowest shortwave infrared 1.6 μm, sum of absolute values of maximum inter-epochal increase and decrease	2.5
Top 3 metrics	*31.3*
All metrics	*99.7*
Regrowth	**Deviance decrease (%)**
Annual metric-interquartile mean of shortwave infrared 2.2μm, amplitude of epochal maxima	22.0
Annual metric-penultimate lowest NDVI, maximum of epochal minima	6.9
Annual metric-interquartile mean of shortwave infrared 1.6 μm, maximum of absolute values of inter-epochal decrease	4.5
Top 3 metrics	*33.4*
All metrics	*66.6*
Cropland	**Deviance decrease (%)**
Annual penultimate maximum shortwave infrared 2.2μm	44.6
Annual maximum near-infrared / green normalized index	16.2
Annual near-infrared of maximum NDVI	6.0
Top 3 metrics	*66.8*
All metrics	*99.5*

4.5.2 Accuracy

We computed user's (indicating errors of commission), producer's (indicating errors of omission), and overall accuracy for the following classes within the aggregated continental map: regrowth, plantation, combined regrowth and plantation, other land use, crops (early and recent crops aggregated into a single class), stable land cover, and stable land cover in the Amazon. The resulting accuracies are shown in Figure 4-

3, and reflect a wide range of accuracies, largely correlated with stratum size, except for 'other land use', which was not directly mapped using the decision tree algorithm.



	Regrowth	Plantation	Regrowth + plantation	Other land use	Crops	Stable land cover	Stable land cover - Amz.
User's accuracy	55.1	63.5	78.0	75.0	88.9	86.3	98.3
Producer's accuracy	22.5	48.8	38.9	72.9	89.1	94.2	100.0
Overall accuracy	96.0	98.9	95.9	89.1	98.6	92.0	99.5

Figure 4-3 Accuracy results for regrowth, plantation, combined regrowth and plantation classes, other land use, combined cropland classes, stable land cover, and stable land cover within the Amazon. Error bars represent 1 SE.

4.5.3 Comparison with MapBiomas area estimates

Figure 4-4 shows the comparison between the MapBiomas/LAPIG area estimates and those from this study in terms of the mean absolute percentage difference and the mean absolute difference between the two estimates. We find that the mean absolute relative difference between our estimates and MapBiomas (using the difference divided by the MapBiomas estimate to produce the relative difference) is below 10% for all except for short tree crops (perennial crops) class. The mean absolute relative

difference for that class is 35%, but due to its small size, this difference represents roughly only one million hectares.

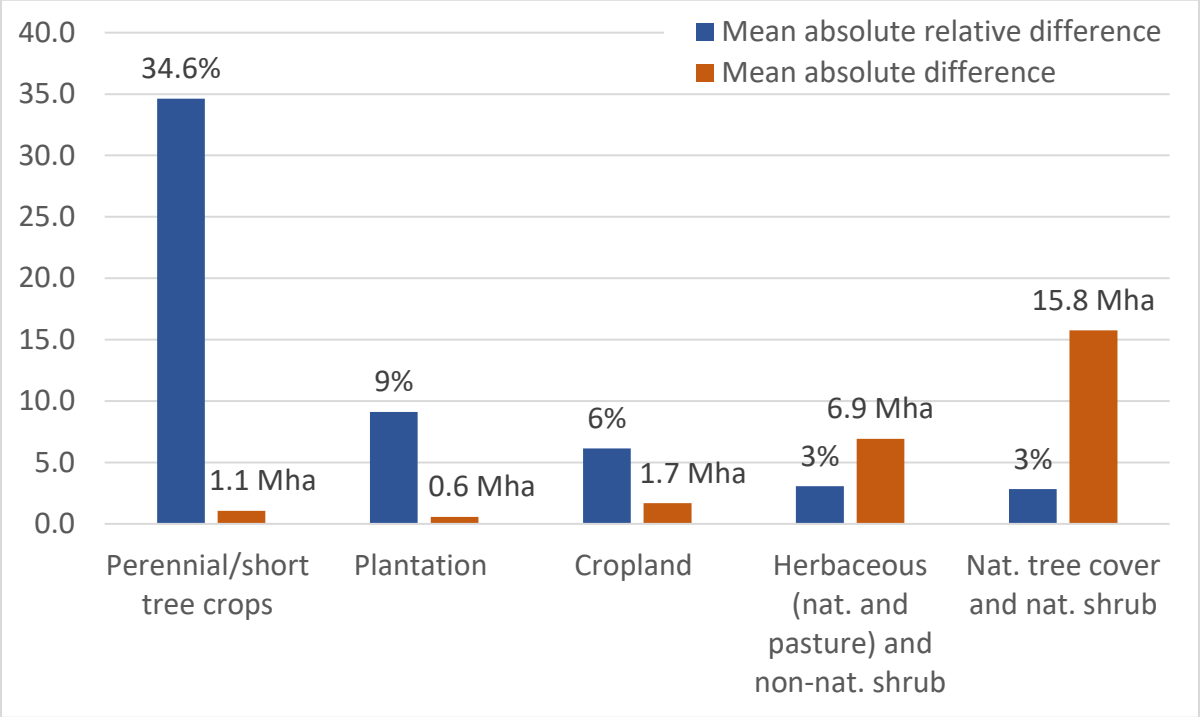


Figure 4-4 Mean absolute relative difference and mean absolute difference between MapBiomas area estimates and those from this study.

4.5.4 Efficiency

Figure 4-5 illustrates the change in standard errors of area estimates for land cover/land use extent and change classes that largely correspond to the mapped classes. For the targeted commodity strata, specifically cropland and plantations, substantial gains in efficiency were achieved, while nearly all other estimates exhibited modest losses in efficiency.

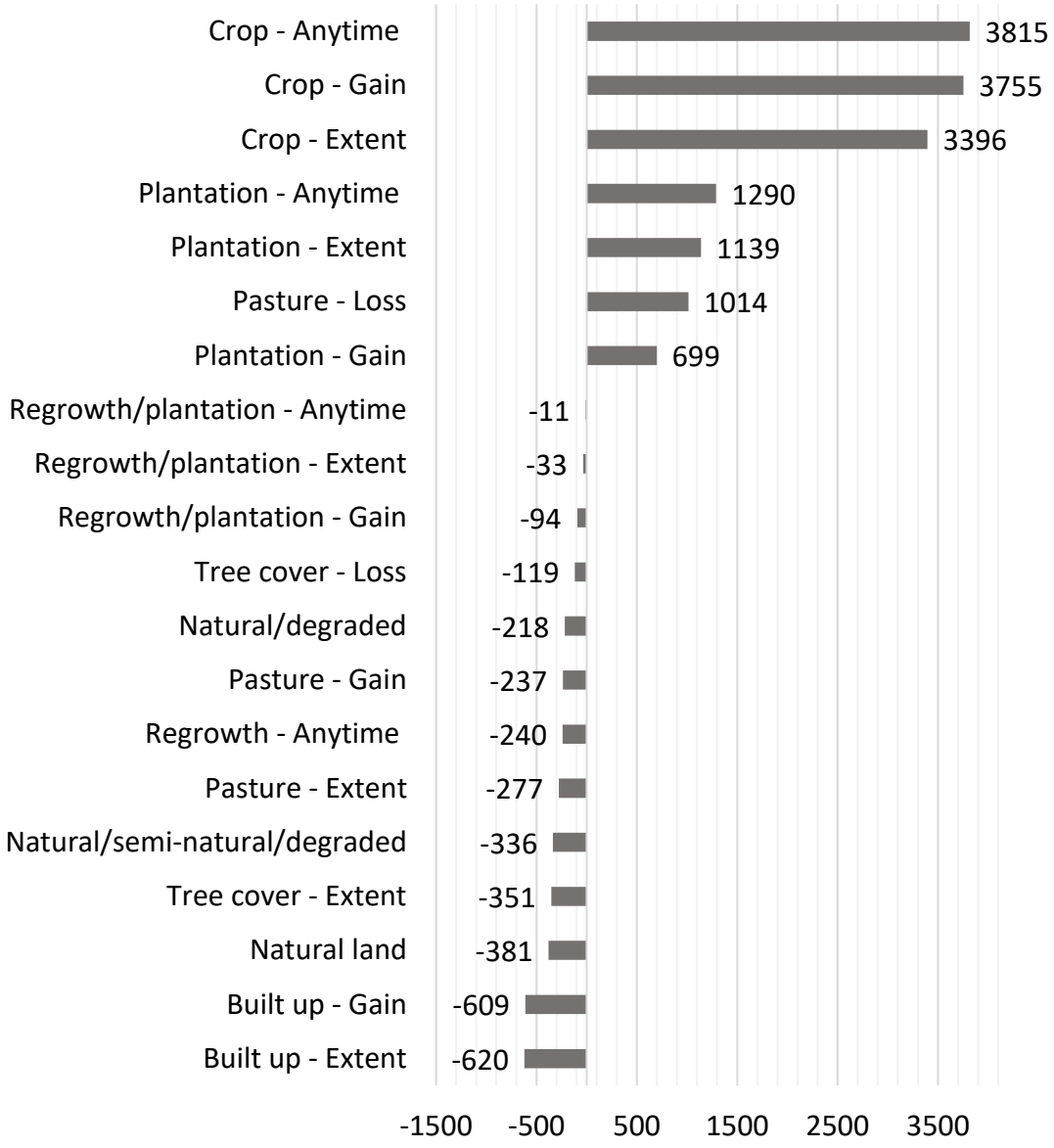


Figure 4-5 Difference in the number of sample pixels needed to obtain the standard errors presented for each area estimate had a simple random sample been used instead of a stratified random sample. Values shown indicate the required additional or fewer samples compared to the stratified n=1000 sample pixels. "Extent" refers to total area in 2018, the end of the study period. "Anytime" refers to whether the class was present at any time between 1985 and 2018. "Natural land" is composed of tree cover, natural shrub, natural herbaceous, and bare ground.

4.6 **Discussion**

4.6.1 Top metrics for classification

One of the advantages of classification trees is that they allow us to easily see which metrics were most useful for separating the map classes. Interpreting the biophysical meaning of these metrics allowed us to evaluate and confirm that the classification was indeed targeting the class of interest. This exercise was also useful for providing insight into the spectral signatures that best characterized each land cover type so that metric sets that target those signatures can be designed for future applications. The analysis of classification tree models also provides valuable information regarding the separability of the classes, which is a good indicator of the difficulty of mapping the class and of the likely accuracy that can be achieved.

4.6.1.1 Stable land cover

Stable land cover, and every other land cover discussed hereafter, was mapped using the multi-epochal metric set which encompasses the entire 1985-2018 time-period (as opposed to cropland, which was mapped using the annual phenological metrics). The most important metric for classification of stable land cover was the mean of the epochal minima of the interquartile mean (25-75% average of annual observations) reflectance in the near-infrared (NIR). This metric alone provided a 39% decrease in root deviance in the classification tree model (Table 2). For vegetated areas, the annual interquartile mean should be elevated. Areas lacking dense vegetation, and by definition not associated with commodity land uses such as cattle production, row crops, or plantations, have low NIR values for the interquartile annual mean. Such areas are largely stable over time and can be split off into a

homogeneous cluster using this metric (Fig. 4-6). The second and third most important metrics both capture variation in NDVI through the years. They are, respectively, the amplitude of the epochal maxima of the interquartile mean NDVI for each year and the amplitude of the epochal maxima of the penultimate lowest NDVI for each year. These metrics are important in the model as stable vegetated land is characterized by having relatively low interannual change in NDVI throughout the time series. For the false-color metric composite of Figure 4-6, stable land cover appears largely blue, with low red and green values for the NDVI amplitude metrics.

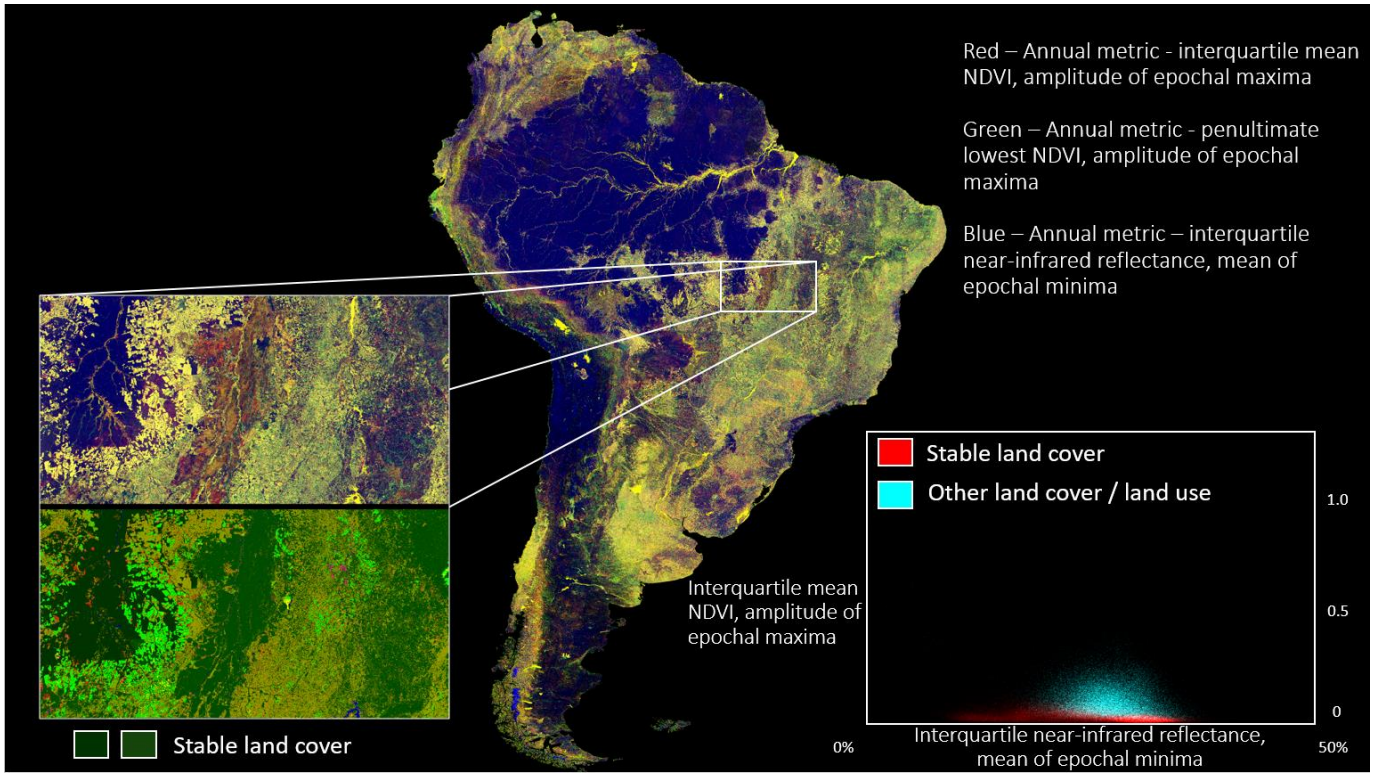


Figure 4-6 Top metrics used for stable land cover classification mapped in R-G-B space. At lower right is a density plot of all mapped pixels as a function of the top two metrics: interquartile mean NDVI, amplitude of epochal maxima (y-axis) and interquartile near-infrared reflectance, mean of epochal minima (x-axis).

4.6.1.2 Plantations

Land covered by tree plantations, defined here as intensively managed planted trees, was most separable from other land cover/use classes thanks to the amplitude of the epochal maxima of the interquartile mean reflectance of the SWIR (1.6 μm) band. This band provided a 19% deviance decrease in the classification tree model. The overall deviance decrease for the entire tree added up to 99.7% though the top three metrics accounted for only 31.3% (Table 2). Amplitude metrics represent variations in reflectance, and high changes in SWIR (1.6 μm) reflectance are indicative of forest disturbance, in this case harvesting operations. Lands for which this metric is high have undergone significant variation from bare to vegetated cover between epochs. Plantations are harvested in cycles that can last from just under a decade to three or more decades, which is why this metric is particularly useful for discriminating this land use. In comparison to the cropland classification, the initial metric's low percent deviance decrease and the relatively flat contribution of subsequent metrics indicates a lack of features with high separability for the plantation class and a likely lower accuracy.

The second most important metric in the classification was the maximum of the epochal minima of the penultimate lowest NDVI. Lands with consistently low vegetation cover exhibit low reflectance values for this metric, which separates non-treed land covers. The third most important metric, the sum of the absolute values of the maximum inter-epochal increase and maximum inter-epochal decrease of the penultimate minimum annual SWIR (1.6 μm), is high for plantation land uses, as it reflects both the short-term gain and loss in canopy cover in managed forests. For the

false-color metric composite of Figure 4-7, plantations appears white, with high values for all three best metrics.

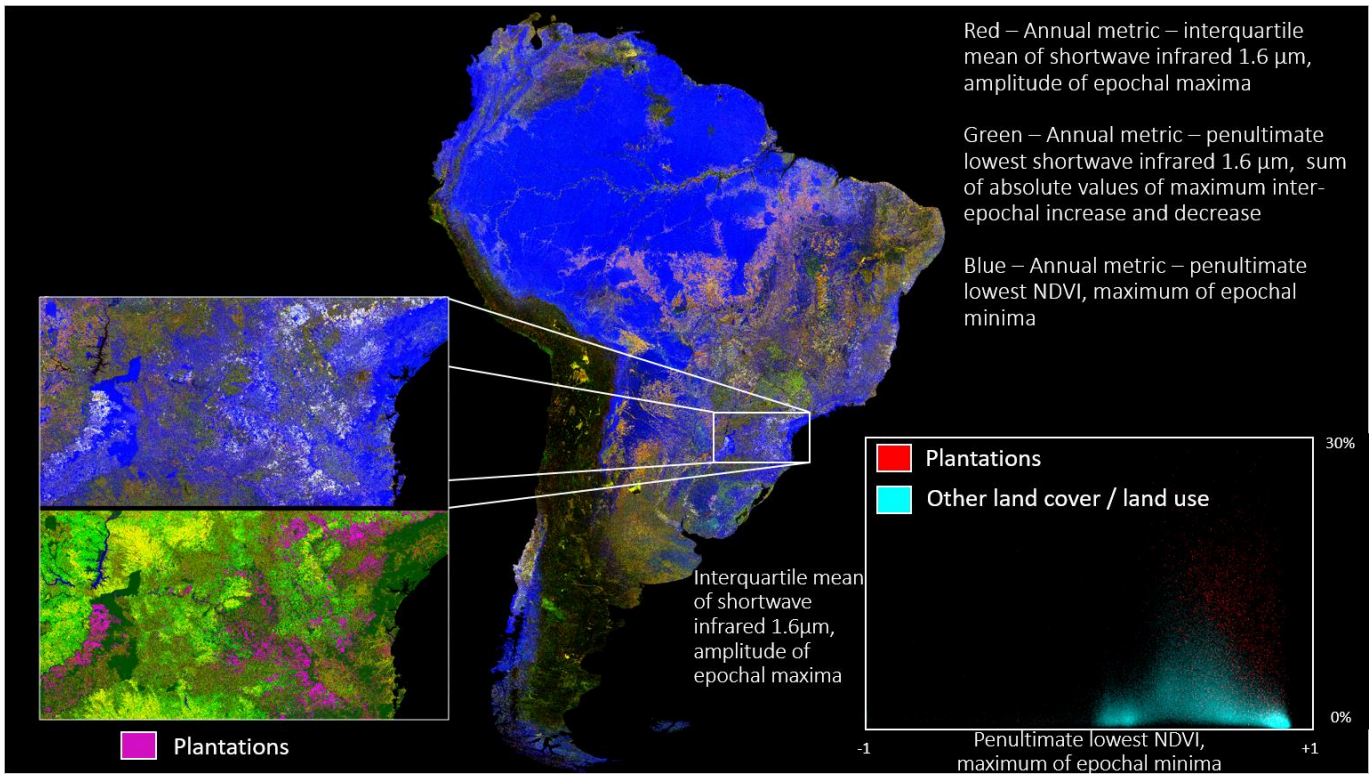


Figure 4-7 Top metrics used for plantation classification mapped in R-G-B space. At lower right is a density plot of all mapped pixels as a function of the top two metrics: Interquartile mean of shortwave infrared (1.6µm), amplitude of epochal maxima (y-axis) and penultimate lowest NDVI, maximum of epochal minima (x-axis).

4.6.1.3 Regrowth

The tree regrowth classification model relied on similar metrics as the classification for plantations, but did not perform as well. Figure 4-8 shows that the two most important metrics in the model were not effective in separating regrowth and non-regrowth pixels. The top metric corresponds to the amplitude of the epochal maxima of the interquartile mean of SWIR (2.2 µm) reflectance. This metric is

analogous to the top metric for the plantation classification. A high value for this metric signifies significant change from bare ground or low vegetation to tall canopy vegetative cover between epochs, which is characteristic of natural tree regrowth. The second best metric was the same as for the classification of tree plantations: the maximum of the epochal minima of the penultimate lowest NDVI. The third metric is an inter-epochal maximum decrease in SWIR (1.6 μm), a measure of canopy cover gain. These three metrics combined provided 33.4% of the overall deviance decrease for the tree. The total deviance decrease added up to only 66.6% (Table 2). This very low number indicates that the classes, as presented in the training data, are not separable.

The density plots of the top performing metrics for the regrowth classification and the plantation classification in Fig. 4-7 and Fig. 4-8 show how similar the mapped regrowth and plantation classes are. The low accuracy of the regrowth class, and to a lesser degree the plantation class, is a consequence of the low separability of these classes. Most likely, the training data provided to the model were not suitable. One metric that has been proven to be effective in monitoring tree cover and height is red reflectance (198, 199), so a regrowth classification model that better leverages decreases in red reflectance as an indicator of increasing tree height would likely yield better results. The lack of this metric in characterizing regrowth may be a function of training data, or due to the fact that both regrowth and plantations are spectrally similar, and other metrics are required to differentiate the two. For the false-color metric composite of Figure 4-8, the metrics mimic to a degree plantations,

but are not readily observable due to their overall inseparability and the finer scale nature of the regrowth class.

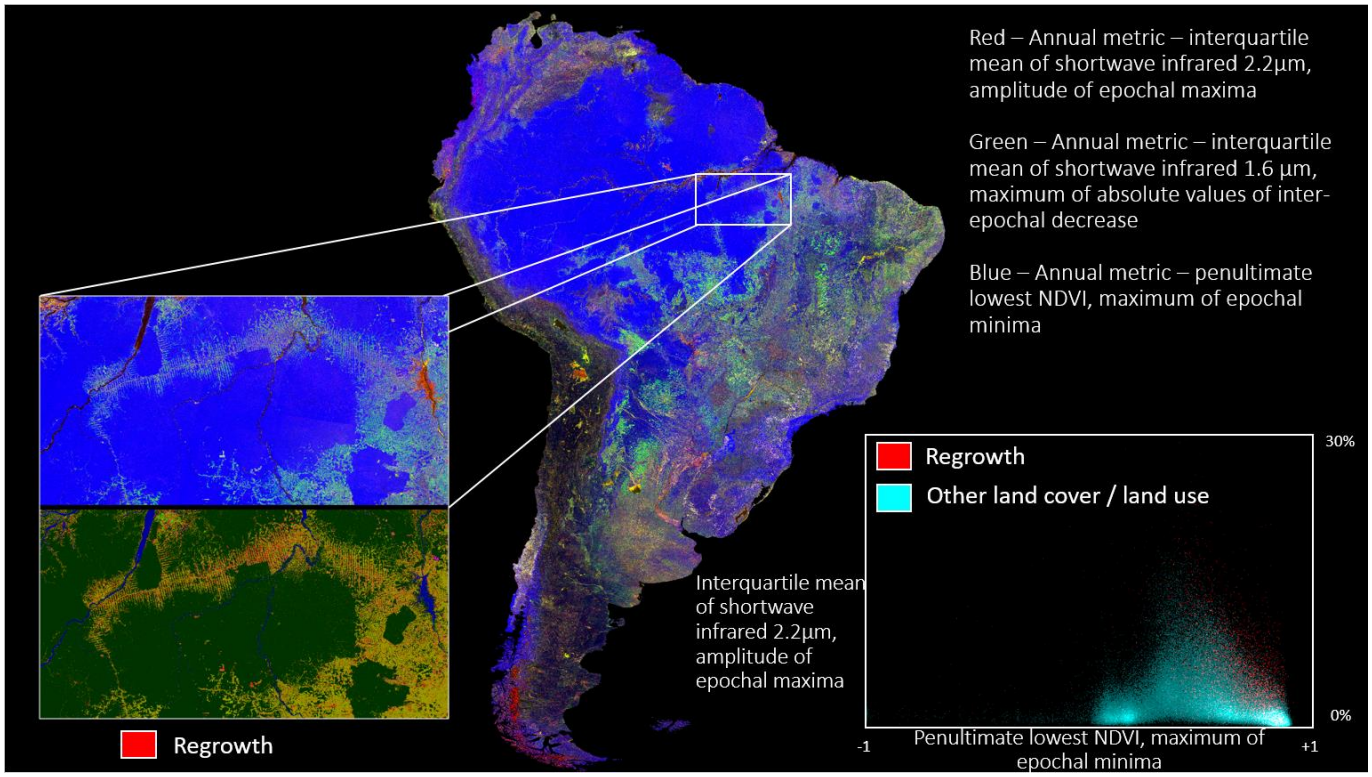


Figure 4-8 Top metrics used for regrowth classification mapped in R-G-B space. At lower right is a density plot of all mapped pixels as a function of the top two metrics: Interquartile mean of shortwave infrared (2.2µm), amplitude of epochal maxima (y-axis) and penultimate lowest NDVI, maximum of epochal minima (x-axis).

4.6.1.4 Cropland

Cropland is a land use that is characterized by its changing cover, which varies from bare soil (after harvest and during the fallow season) to vegetated one or many times a year (depending on the crop and cropping frequency). The decision tree model exploited this particular characteristic to separate cropland from other land covers/uses. The top metric for the cropland classification was the annual penultimate maximum SWIR (2.2 µm) reflectance, which is high for bare soil. This metric alone

provided 44.6% of the overall deviance decrease for the tree (Table 4-2). The second most valuable metric was the annual maximum of the normalized ratio of NIR and green, which is high for vegetated areas (much like NDVI). The NIR reflectance for the observation corresponding to the maximum annual NDVI was the third most important metric in the classification model and characterized by high reflectance values over densely vegetated areas such as intensive row crops like soybean. These three metrics together accounted for 66.8% of the total deviance decrease, indicating that the cropland class was relatively easily separable. Land having high reflectance values for these three metrics, which are distinguishable in white in Fig. 4-9, correspond to land covered by both bare soil and dense vegetation, which corresponds well with cropland.

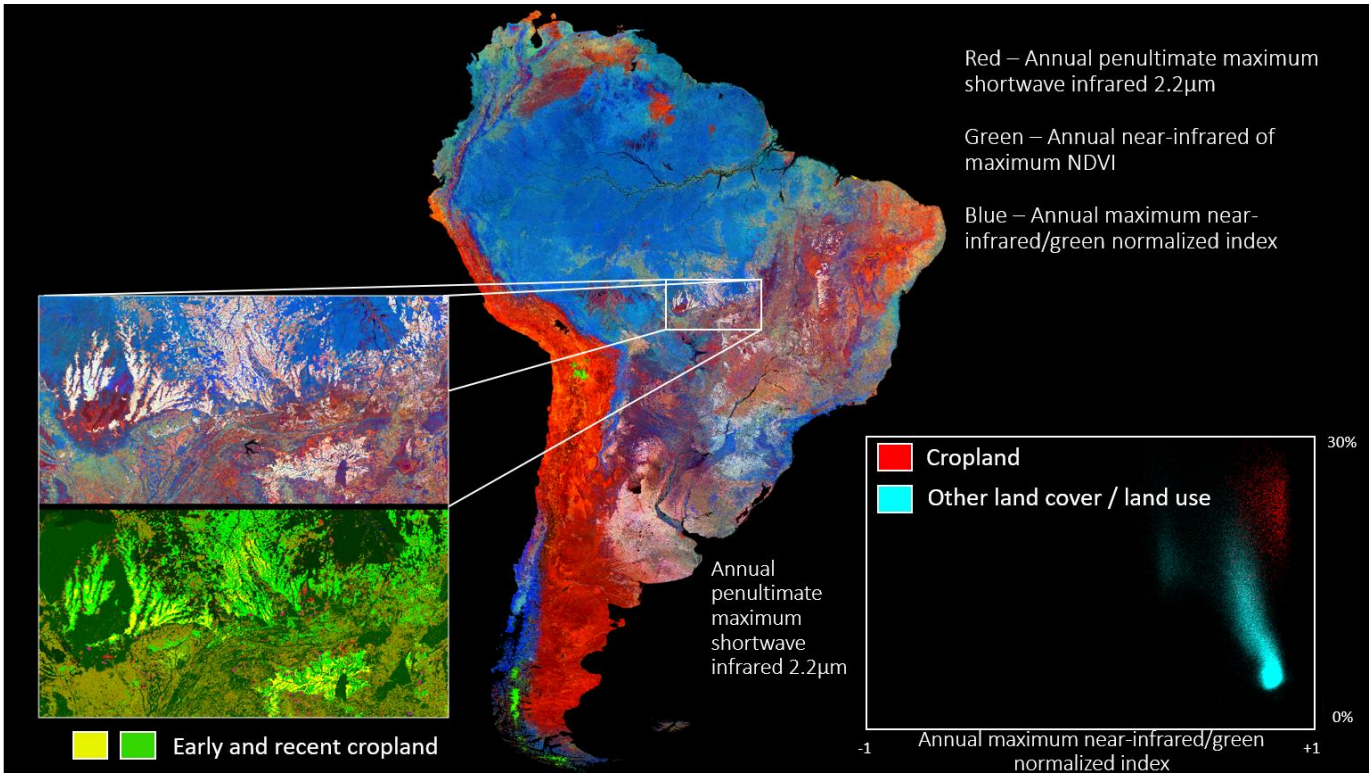


Figure 4-9 Top metrics used for cropland classification mapped in R-G-B space. At lower right is a density plot of all mapped pixels as a function of the top two metrics: Annual penultimate maximum shortwave infrared (2.2µm) (y-axis) and annual maximum near-infrared / green normalized index (x-axis).

4.6.2 Evaluation of the map and the derived stratified sample

4.6.2.1 Accuracy:

Crops, stable land cover, and stable land cover within the Amazon had high overall, user's and producer's accuracies. While crops and stable land outside of the Amazon did not have imbalanced commission and omission errors, stable land did (86.3 ± 3.1% vs. 94.2 ±1.4%), meaning that the map overestimated this category. The "Other land use" class had moderate, but balanced user's and producer's accuracies (75.0 ± 4.0% and 72.9 ± 4.3% respectively). Regrowth, plantation, and combined regrowth and plantation categories had the lowest user's and producer's

accuracies. Our results indicate that the regrowth map missed nearly 80% of natural tree regrowth (22.5 ± 4.4% producer's accuracy) and that nearly half of the mapped regrowth was misclassified (55.1 ± 4.4% user's accuracy). The plantation class had slightly better accuracies (63.5 ± 4.5 % user's and 48.8 ± 8.9% producer's accuracies). There was considerable confusion between the plantation and regrowth classes. When these two were combined, the user's accuracy increased relative to each individual class to 78.0 ± 2.7%, but the producer's accuracy decreased in relation to the plantation class's individual producer's accuracy. This indicates that there were not enough reference sample pixels correctly classified as regrowth or plantation to compensate for the increased proportion of reference sample pixels that were omitted from both classes.

4.6.2.2 Comparison with MapBiomas area estimates

While we do not have access to uncertainty measures for all of the MapBiomas estimates, we can assume they are more precise than ours due to their much larger sample size. For the purposes of this comparison, we assumed MapBiomas area estimates to have no uncertainty. MapBiomas area estimates were within one standard error of our estimates for most years, and within two standard errors for all years, illustrating an overall strong correspondence over the study period (Figure 4-10). The mean absolute percentage difference and mean absolute difference results illustrated in Figure 4-4 also revealed a strong correspondence between the two studies. The comparison showed that despite the relatively small sample size employed in our study, we were able to achieve similar results to those derived from a sample 85 times larger than ours. The concurrence of results from two very different

sampling designs supports the general, shared conclusions of the two studies, specifically the dramatic absolute and relative gain in cropland, the dramatic relative gain in plantations and short tree crops, the stabilization in terms of area of pasture, and the continued loss of natural tree cover.

4.6.2.3 Efficiency

Our results show that if a simple random sampling design had been used, nearly five times as many sample pixels would have been needed in order to match the precision of the area estimates for crop gain and total crop extent at the end of 2018 that we report in Chapter 3. Pasture loss was highly associated with cropland gain – pasture to cropland conversion was a significant land cover transition in South America – so the "recent cropland" stratum that targeted gain in cropland also targeted pasture loss. This dynamic explains the 100% efficiency improvement of our sampling design in relation to a simple random sample for pasture loss.

We also see considerable gains in efficiency using the stratified approach for the classes related to plantations (plantation gain, plantation extent in 2018, and area of plantations at any time during the 34-year time period). Interestingly, the user's and producer's accuracies of the plantation class were comparatively low, but the sampling efficiency was high. This result demonstrated the benefit of even modestly accurate maps as targeting mechanisms for rare classes. As opposed to plantations, we observed a loss in efficiency for the regrowth class using a stratified random sample compared to a simple random sample. The very low user's and producer's accuracies reflected the poor quality of the map, a performance insufficient to provide any sampling efficiency, in contrast to plantations. The sampling inefficiency for

estimating natural tree regrowth may be especially driven by the large errors of omission in the class, since errors of omission tend to have an outsize impact on area estimates if the strata in which they occur are highly weighted (181).

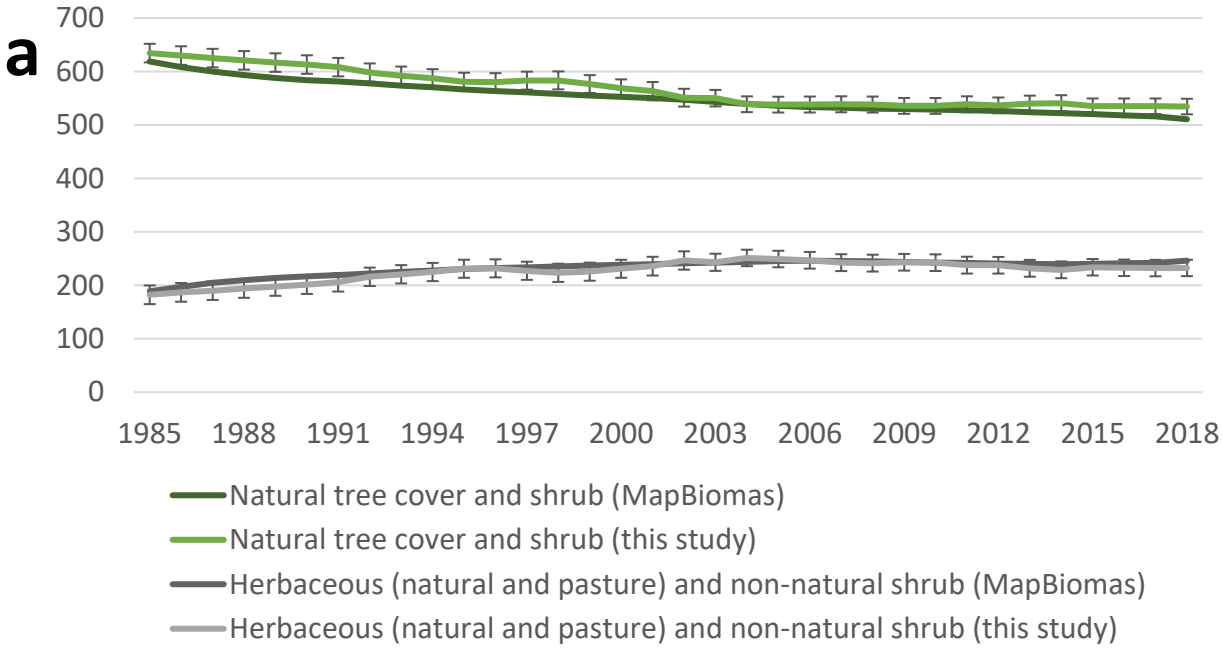
The mapped stable land cover strata targets land cover that has not undergone any land cover conversions over the entire study period. This stratum was largely composed of natural land (vegetated and non-vegetated), including degraded tree cover, but also included some areas of very low intensity land use that were spectrally indistinguishable from natural land. If we compare the precision of the area estimate for natural and degraded tree cover obtained using the stratified sample to that derived from a simple random sample, we observed a 22% decrease in efficiency. This means that had we used a simple random sample, we would have needed 782 sample pixels instead of 1000 to get area estimates with the same precision. This result contrasts with that for plantations. The high overall, user's and producer's accuracies for stable land cover did not result in a sampling efficiency due to the large areal extent of the stable land cover classes in comparison to plantations. The larger proportion of the study area for a given class, the more competitive simple random sampling will be even when using a highly accurate map product. If we examine only natural land, (excluding degraded tree cover) or if we look at the combined area of natural land, degraded tree cover, and semi-natural land (as defined in the land cover/land use hierarchy in Fig. S3-8 and in Table S3-2), the efficiency loss is greater. We would need less than 700 sample pixels in each case to attain similar uncertainties. The natural land, degraded tree cover and semi-natural land categories represent lands not associated with intensive land use. Our greatest efficiency for

such lands was found in the natural land plus degraded tree cover aggregate class. The stable land cover stratum was designed to exclude land use and not degradation. The fact that stable land cover plus degradation performed best of the semi-natural to natural combined categories supported this definition as applied.

Area estimates for classes that did not correspond to any of the mapped strata had lower precision than those than would be expected if we had used a simple random sample. Most notably, our stratified random sample presented efficiency losses for large classes like tree cover loss, tree cover extent, pasture gain and pasture extent (-12%, -35%, -24%, and -28% smaller sample size, respectively, would have yielded equal area estimates). Classes that were small and not targeted by the strata, such as built-up area, were most heavily penalized by the stratified approach (62% loss in efficiency).

Our stratified random sampling approach yielded dramatic gains in efficiency relative to a simple random sample for some classes and comparatively modest losses for others. A stratified approach is likely to yield best results when there is one or very few target classes that are rare, of high importance, and well-mapped. In this study, we were interested in estimating the area of all major land covers and land uses (many target classes) over an extended period of time. Some of the target classes were well-targeted by the strata (e.g. cropland and cropland gain), while others were poorly mapped (e.g. natural tree regrowth) or not targeted at all (e.g. pasture gain). Further, some of the classes were very large (e.g. tree cover extent), which makes them better suited for area estimation via a simple random sample as compared to small (rare) classes. A simple random sample may also be preferable for area

estimation if the targeted change dynamic is difficult to map, which was the case with regrowth and may have been the case of pasture gain in South America. Our results illustrate some of the trade-offs that merit consideration when choosing the most appropriate sampling strategy. Our study's focus was to elucidate the role of commodity land use expansion as a driver of land use change in South America and our strata designed to target these commodity land uses, especially for cropland and plantations. The important efficiency gains that resulted from the stratified approach in these classes did well in documenting land use expansion, largely compensating for the less-significant losses for non-targeted large classes.



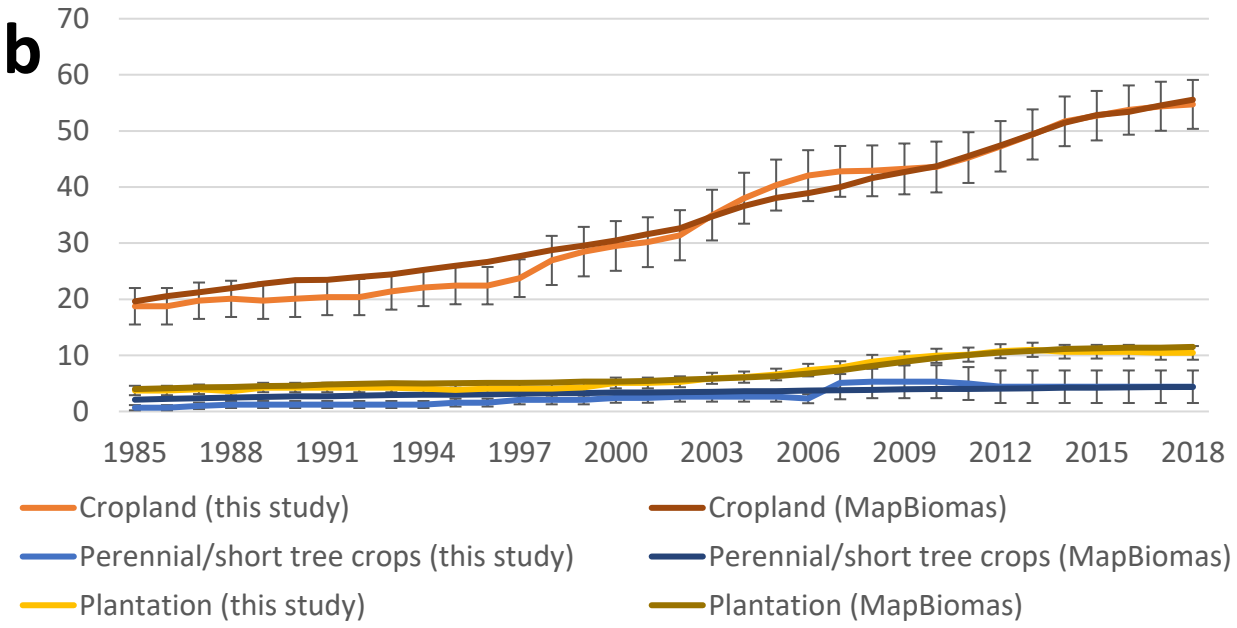


Figure 4-10 Comparison of MapBiomas/LAPIG area estimates to those from this study. Error bars represent one SE. Error bars for MapBiomas/LAPIG are not visible because the error is too small (this is a function of the very large sample size). **a.** Comparison of "natural tree cover and shrub" and "herbaceous and non-natural shrub" classes. **b.** Comparison of cropland, plantation, and perennial/short tree crop classes.

4.7 **Conclusion**

Maps of land cover/land use extent and change provide information on the spatial distribution of the themes at hand, which is crucial to understanding associated drivers and impacts and to facilitate better land use planning. This is especially true in relation to commodity land uses across the South American continent, given that agricultural commodity production is the leading cause of land cover/land use change in the region. Additionally, maps can serve as the basis of stratification to better target the specific land cover/land use classes of interest, therefore improving the precision of the area estimates.

The land cover/land use map of South America created here was highly accurate in depicting areas of cropland extent over the 34-year study period, an important class due to the growing significance of commodity crops such as soy, corn, and sugarcane in the continent and global marketplace. The mapped cropland class also provided the largest gains in sampling efficiency: nearly 5,000 sample pixels would have had to be interpreted to achieve the uncertainty of the provided area estimates had a simple random sample been used. Additionally, our estimated cropland area in 2018 of 54.7 ± 4.4 Mha matched the MapBiomas estimate of 54.6 1.4 Mha. We were able to calculate the standard error of the MapBiomas estimate for cropland in 2018 using sample-based data they agreed to share with us. The efficiency provided by the mapped cropland stratum is further evidenced by comparing the sample size needed to provide an area estimate with a standard error of 1.4 Mha. To achieve the uncertainty of the MapBiomas cropland estimate, for which they interpreted over 85,000 pixels, we would have to increase our sample size to only 1,421 pixels – a 60-fold reduction in sample size.

Another important commodity land use is plantations. The mapped plantation class had significantly lower accuracy, yet it provided significant efficiency gains in estimating area, with resulting area estimates comparable to those of MapBiomas. Tree regrowth, on the other hand, had a very low accuracy and stratifying by this class resulted in a loss of efficiency in area estimation. Both stable land cover classes (inside and outside the Amazon), which were created and included in the stratification to isolate areas of no land cover change and thus better target land cover/land use change, had high accuracies, but did not result in a gain of efficiency for estimating

natural, undisturbed land. Other classes that were not correlated with any of the mapped classes did not benefit from a sampling efficiency. For example, we expected the "other land use" class to be a good proxy for pasture and therefore provide a sampling efficiency. However, equally precise area estimates for pasture extent and gain would have been produced if we had used a simple random sample and a smaller sample size.

The two-tiered approach of mapping and sampling used in this study revealed some strengths and weaknesses associated with the strategy of stratifying to produce area estimates. A key takeaway is that it may not be as useful to stratify if the target class is very large. Further, the relationship between the map's accuracy and the sampling efficiency is not very clear, as exemplified by the tree plantation and the tree regrowth classes, but it is possible that errors of omission are the most important drivers of sampling efficiency losses associated with low accuracy strata. Lastly, because stratification entails tradeoffs between classes (gains in efficiency for one class may lead to losses for another), having several classes of interest simultaneously warrants careful consideration and prioritization of the study's objectives. A preferred sampling approach is likely a stratified random sampling design that includes more and accurate strata directly related to all relevant land cover / land use extent and change themes. An initial equal allocation of the sample per stratum could be used to calculate initial estimates of variance per stratum, and then the initial sample could be augmented to improve the sample allocation based on the initial variance estimates per stratum. Using such an approach, an optimal allocation objective could be employed to reach absolute or relative uncertainty goals per land theme.

Studies that seek to produce area estimates for many classes at once are currently rare in the literature, but as the practice of sampling to derive area estimates becomes more widespread, it is likely that these types of studies will become more common. Future studies should focus on advancing methodologies and providing guidance on which sampling strategies are best suited for difference land cover and land use change monitoring scenarios.

Chapter 5: Conclusion

5.1 **Advances in understanding of land cover/land use change in South America contributed by this dissertation research**

This dissertation sought to combine established methods for mapping and sampling using remote sensing data to characterize land cover and land use change in South America. The goal was to advance our understanding of the major land cover and land use dynamics taking place across the continent, especially as they relate to the expansion of land uses associated with the production of agricultural commodities such as soy, corn, sugarcane, beef, and wood products. To do this, a two-stage approach of first creating a map of target land cover/land use themes to then use as strata for selecting a random sample of pixels from which unbiased area estimates can be derived was employed. In Chapter 2, we employed the approach for the use-case of Brazilian cropland. In Chapters 3 and 4, we employed it for the entire South American continent and for a wide range of land cover/land use themes. Chapter 4 also provides a thorough evaluation of using this approach for meeting the aforementioned objectives. Several novel findings resulted from this dissertation research.

Chapter 2 resulted in the finding that cropland area nearly doubled in Brazil between 2000 and 2014, mainly because of the repurposing of pastures (80% of new cropland) rather than conversion of natural vegetation (20%). Area of converted Cerrado savannas was nearly 2.5 times that of Amazon forests, reflecting the potential leakage of cropland expansion pressure to a region that is largely unconstrained by regulatory limits. There was a slowdown in conversion of Amazon forests to cropland

beginning in the 2004/2005 crop season, which may have been due to a possible response of land owners to policies (and the anticipation of pending policies), market conditions, or both. This study constitutes the first source of Brazilian national-level cropland maps and area derived from satellite data in the scientific literature.

Building on the successful application of the methodology employed in Chapter 2, Chapter 3 aimed at producing analogous results for a much broader geography, time period, and set of land cover/land use themes. The results in Chapter 3 outline the ways in which humans have appropriated natural land across the South American continent at staggering rates. Human impact on natural land cover in South America increased by 60% since 1985, resulting in a 20% decrease in unaltered natural land cover in a 34-year time period. Since 1985, pasture area grew by a quarter, cropland area more than doubled, and plantation area nearly quadrupled. Pasture, the single largest land use category across every geographic unit examined in this study, occupied 17% of the continental landmass by 2018. A particularly concerning finding was that 55Mha of land across the continent are in a long-lasting transitional state – land that is not natural nor a land use. An analysis of the most important land cover/land use transitions at the continental level provides evidence of a trend toward the increased intensification of land use, since the largest from-to categories are tree cover to pasture and pasture to cropland. Numerous results presented in this study are entirely new to the scientific literature, as there are no studies covering the geographic and temporal scales presented herein. In particular, there is a dearth of information on the theme of pasture, both in terms of its spatial characterization through mapping and of the quantification of its area. This is partly

due to the difficulties associated with mapping such a heterogeneous and typically low-intensity land use. This obstacle can be overcome through the use of sampling for area estimation, as was done here. Further, an important contribution to the literature is the introduction of the transitional land class in the context of South American land cover/land use change. The concept itself is not new: it fits neatly in the category of semi-natural land as defined in the FAO's Land Cover Classification System (LCCS) (200). Yet it has not been widely applied in the South America land cover/land use change scientific literature. This class should be of particular concern to decision-makers as it represents the worst of all worlds: land that is compromised in terms of its ecosystem functioning, yet provides no economic gains.

Chapter 4 complements Chapter 3: it provides an analysis of the spatial component of the study – the map. Maps of land cover and land use change are useful for providing the necessary spatial context of the change. Further, they are useful in bolstering the area estimation process. A map of continental land cover/land use change was created using 34 years of Landsat data, with a particular emphasis on commodity land uses. The map themes are cropland, plantation, natural tree regrowth, stable land cover, and other land uses. The map can serve as a reference for better understanding the dynamics of change at play in the continent. It could also serve as an input for modeling applications. The stable land cover class may be particularly useful as an input for studies that focus on impacts on natural vegetation in South America beyond forests, since no Landsat-spatial-resolution data exist on themes like shrubland or grassland, for example, for this region. Further, this map provides spatial information on cropland and plantation area for countries outside of Brazil, for which

there is little to no spatial data. The map was also found to significantly contribute to the area estimation process through its use as a stratifier for the random sample. The area estimates computed from the sample correspond strongly with those from a comparable but significantly larger sample, which confirmed the results. Further, the map strata were found to provide significant efficiency gains for some classes and modest losses for others, resulting in an overall reduction in the sample size needed to achieve the uncertainties reported in Chapter 3. The evaluations performed in Chapter 4 led to a set of lessons learned for the purpose of using a stratified random sample to produce area estimates when there are a multitude of target classes. These include the fact that it may not be useful to stratify if the target class is too large. Further, the map accuracy and the sampling efficiency are not directly linked, but errors of omission can have a large impact on efficiency. Finally, stratifying to better target one class may be disadvantageous for other classes, so it is important to have a clear understanding of the study's priority objectives when developing the sampling design. Studies that seek to employ good practice recommendation to quantify areas for a multitude of land cover/land use themes as opposed to a single one are not common in the literature. However, as countries continue to develop mapping and sampling capabilities for important reporting requirements like those of the IPCC National Greenhouse Gas Inventories, for which several classes must be mapped and quantified, the use of sampling to quantify several land cover themes at once will become increasingly common. Lessons learned here can be applied for other studies, but further guidance is needed on which sampling strategies can most effectively and efficiently accomplish the goal of targeting many classes at once.

5.2 **Outstanding issues**

There are outstanding issues with this dissertation research both in terms of the data and methodological approaches employed as well as regarding the themes addressed. Some of the outstanding issues are the following:

- The regrowth map produced for the continental-scale study was inadequate in terms of accuracy and resulted in a loss of efficiency for the area estimation. More and better training data should have been supplied to the classification tree model in order to improve the map. The iterative phase of the classification process ended prematurely because I mistakenly considered the map to be an appropriate representation of tree regrowth in the continent. Lack of proper geographic knowledge of land cover and land use distributions will often lead to poor maps. This mistake is an example that underscores the continued need for and importance of the field of Geography today. My knowledge of the geographic distribution of tree regrowth was substantially improved through the evaluation of the map and the interpretation of the sample pixels, which leads me to believe that I could produce a better regrowth map today.
- Although the area estimates produced here are accompanied by error estimates, there are sources of uncertainty that cannot be quantified. Amongst these is the uncertainty associated with sample interpretation to create reference data. The reference data sources employed in both the Brazil and the continental-scale cases had high temporal and spatial representation which enabled us to label each sample pixel with higher accuracy than through the

map classification. However, a quality assurance protocol would have required that some or all sample units be interpreted by more than one interpreter to ensure consistency in labeling. Although this was partly the case for the Brazil study (four interpreters aided me in labeling reference data, but in case of disagreement, my interpretation overruled the others), I was the single interpreter for the South America reference sample. Having an additional interpreter label a random sample of the sample would provide an improvement in terms of ensuring consistency in labeling and to reduce the uncertainty associated with the reference sample. While other approaches employ many interpreters and tools such as voting to assess confidence of interpretation, our research group often employs a pyramid approach, as in Chapter 2. In this approach, a team of less experienced interpreters assign labels, after which a more experienced expert filters those results, confirming and relabeling as needed. Our strategy is to use consensus as a starting point, but to have a single, best interpreter be responsible for final assignments. For Chapter 3, my interpretations were the only ones used in estimation due to resource limitations and the comparatively fewer samples assessed. When publishing our results, I will share all sample templates so that anyone may perform the interpretation independently to confirm or otherwise, my results. The uncertainty from the interpretations is referred to as measurement error and is not accounted for in these studies, nor generally in the discipline. Measurement error remains a poorly characterized source of error in good practice area estimation methods, and should be a focus of future research.

- The map strata used to select the sample of the South America study resulted in a loss of efficiency for estimating the area of the tree cover loss and pasture gain classes. Using tree cover loss data from the Global Forest Change product to create an additional strata would have likely yielded a gain in efficiency for both of these classes (since tree cover loss most often results in pasture gain). This was a missed opportunity that would have helped target two classes that are of high priority. However, we decided against this as we have tree cover loss data only since 2000, precluding the applicability and expected sampling efficiency of this stratum across the study period.
- Not enough attention was paid to the allocation of sample pixels amongst the strata for the South America study. The cropland strata were prioritized in terms of number of sample units allocated, which resulted in very good area estimates of cropland extent. However, if more sample pixels had been allocated to the "other land uses" class, which was substantially larger and contained a theme of great interest (pasture), we would have likely had smaller uncertainties associated with some of the area estimates, such as those of tree cover loss and pasture gain. As mentioned in the conclusion of Chapter 4, a better approach to sample allocation would have been to allocate an initial sample equally amongst the strata, use it to calculate initial uncertainties of the area estimates, and then allocate more sample pixels in order to decrease the uncertainty of the most uncertain estimates.
- If the sample size of the South American study was increased, not only would we reduce the uncertainty of the reported area estimates, but we could also

report areas for smaller geographic units, including all of the South American countries and all of the ecozones. Further, a larger sample size would allow us to report areas for a whole host of additional themes, including, for example:

- Average time between a degradation event and tree cover clearing
- Average time between tree cover to pasture to crop conversions
- Area of land that is abandoned and reverts back to a natural land cover
- Area of tree encroachment (new trees in natural land on which trees were not previously present)
- Area of pasture that is intensively managed versus low intensity pastures
- Area of double cropping

The sample size was a function of the difficulty and the time requirement associated with interpreting each of the sample pixels. To label each sample pixel, a land cover/land use label had to be assigned for every one of the 34 years of the study period. In contrast, for the Brazil study, a much larger sample size was employed, but each sample pixel had to be labeled only in relation to the cropland class: a sample pixel could be labeled as cropland over the entire period, cropland expansion, or not cropland. Additionally, the study period was only 14 years. Thus, the Brazil cropland gain sample interpretation exercise was much simpler than the multi-themed, continental-scale, 34-year study. Increasing the sample size of the South America study would yield substantial benefits but it would also require substantial additional work.

5.3 **Possible ways forward**

The introduction of this dissertation highlights the ways in which land cover and land use change have repercussions across systems and scales. The human appropriation of natural land cover can have important negative impacts on populations across the world, today and into the future. Further, poverty exacerbates people's vulnerability to these impacts, highlighting the injustices associated with the destruction of natural ecosystems. Simultaneously, the increase of commodity land uses has contributed significantly to the economic development of South American countries. Reliable and up-to-date information on the spatial distribution and the areal extent of land cover and land use and its change is crucial to inform land management and environmental protection policies and decisions. Results from this dissertation research contribute to this task. However, far more can and needs to be done.

First, area estimates should be produced at the country level and even at the state/departmental level. Area estimates for themes that are consistent across countries (such as tree cover loss, pasture gain, and cropland gain) should be regularly quantified with increasingly low uncertainty. A continental-scale assessment of important land cover/land use themes would be advantageous in that the results would be consistent (derived using the same methodology and data) and therefore comparable. This would enable studies examining the possible factors influencing the dynamics of land use change, such as land use policies, environmental regulations, market conditions, climatic factors, issues related to crop diseases, and other potentially consequential factors within and across national borders. However, if good practice methodologies are employed and the data and definitions used are

similar, national-level studies could also yield results that can be compared to each other.

Certain land cover and land use themes need to be researched further. In particular, information (area estimates and maps) on pasture in South America is lacking. This makes it difficult to implement policies regulating pasture expansion. This is of urgent concern given the role of pasture as the dominant fate of forests across the continent. However, pasture is a challenging land use to characterize given the wide range of intensification and land covers associated with it. Additionally, to the extent that it is possible using remote sensing data, we need to move towards mapping and quantifying not only land cover and land use, but also land use management and how it changes through time. Specifically, distinguishing between single and double cropping is important as intensifying cropland land use can have environmental, management, and economic impacts. Another example of a land use management change that could be monitored is the intensification of cattle ranching through planted pastures, rotational practices, feedlots, etc. Increasing production outputs through intensification of cropland and pasture land uses has been hypothesized to be a potential land-sparing strategy. Monitoring land management practices would enable further studies in this research area.

The use of probability sampling protocols to derive area estimators that are unbiased and have known uncertainties needs to become ubiquitous in the field of land cover/land use change monitoring. Aside from being the recommended good practice method for area estimation, sampling has a number of additional benefits: it is low cost, it does not require the data volumes and processing power that mapping

does, it is applicable across geographies and themes, etc. For some land cover/land use themes, field visits can serve as the reference data source, therefore enabling quantification of themes such as crop type that cannot easily be mapped. Commercially available high spatial and temporal resolution imagery is another data source that can be leveraged to improve the accuracy of reference data labels and to increase the thematic richness of area estimates.

Ideally, every country should have an agency dedicated to operational land use and land cover change monitoring. Themes should be clearly defined and standardized. A combination of freely available Landsat and Sentinel-2 data, high resolution commercial data, and field surveys could be employed to map and estimate areas of important land cover and land use themes. The resulting data, which should be openly available to the public, would be useful for a myriad of applications, for example supporting land use management policies and enforcement of environmental regulations, enabling accurate reporting for REDD+, IPCC Greenhouse Gas Inventories and other international commitments, driving climate and economic modeling efforts, monitoring supply chains, tracking economic development, quantifying health outcomes associated with the loss of ecosystem services, and ensuring the maintenance of protected areas, among other use cases.

It is evident that there are tradeoffs associated with converting natural land cover to economically productive land uses. In an ideal world, the decision to convert natural ecosystems for human land use would be made after having taken into account all the costs and benefits, both economic and environmental. Quantifying land cover/land use changes is an urgent and necessary first step towards better

understanding the associated environmental costs. Currently, most land change decisions are not made with such a holistic perspective in mind.

Appendix

Comparison of Brazil cropland expansion data (Chapter 2) and cropland data from the Brazilian Institute of Geography and Statistics (IBGE)

In response to a reviewer during the peer-review process of the publication entitled "Near doubling of Brazil's intensive row crop area since 2000," a thorough analysis and comparison of our results and the data available through the Brazilian Institute of Geography and Statistics (IBGE) was performed. The analysis and comparison, as provided in our response to the reviewers, are detailed here.

The IBGE SIDRA database provides a wealth of data on a variety of metrics related to agricultural activities in Brazil. For the purposes of comparing our results on cropland area and expansion in Brazil to those presented in the SIDRA database, we use the "planted area" metric provided through the Produçao Agricola Municipal (PAM) project, which has historical data on 33 temporary crops.

It is important to note that the aim of our study is to estimate cropland area in the country, which is fundamentally different from a measure of planted area for each crop. Therefore, in order to compare our results to data from the IBGE SIDRA database, we first need to try to extract an estimate of the area of land covered by crops using the database.

To do this, we start by selecting the crops within the database that would fit our definition of "cropland." The definition of the "cropland" class we mapped in our study is "areas of intensive row crop agriculture" (lines 744-745). Intensive row crop agriculture has a particular spectral response that is distinct from non-intensive, non-row crop agriculture. Thus, crops that are produced with low intensity,

generally over small areas, such as tomatoes or onions, were not included in our definition of cropland. Permanent crops such as oranges were also not included.

In order to simplify this comparative analysis, we narrowed the number of crops we examined to the most important crops in Brazil in terms of area occupied for their production. Figure 1 shows that a few crops make up the bulk of total crop area in the country. The planted area of only eight of the 33 crops available through the database (soy, corn, sugarcane, beans, rice, wheat, manioc, and cotton) represent 95% of total crop planted area in Brazil in the year 2000 and 96% in 2014. We will thus consider the area of these eight crops as a starting point to estimate cropland area using the SIDRA database. This corresponds to 43.5 Mha in 2000 and 67.9 Mha in 2014.

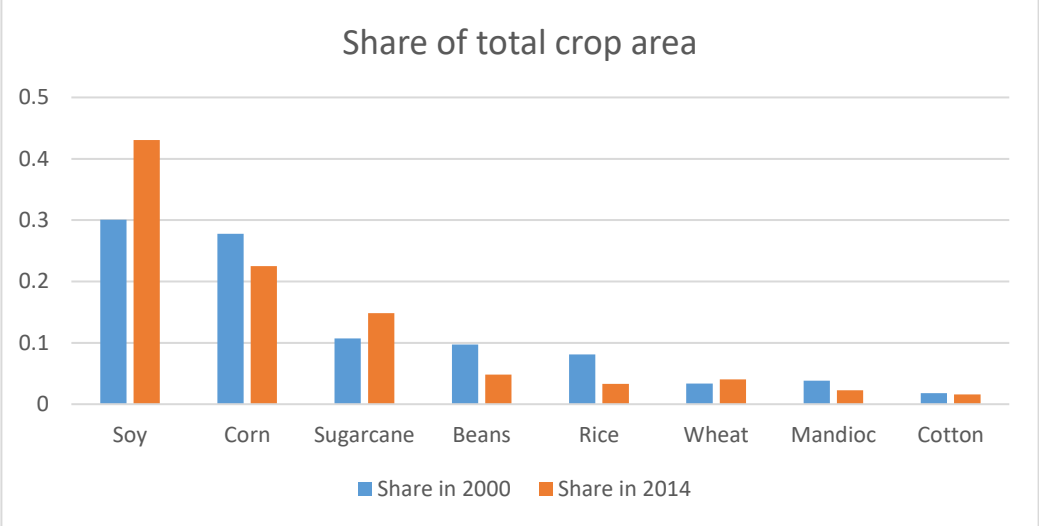


Figure 1. Proportion of total agricultural area of each of the eight largest crops in Brazil. Data from IBGE SIDRA database.

In order to estimate cropland area using data from the SIDRA database, we need to try to account for areas that are double cropped. Wheat, for example, is a winter crop in Brazil and therefore is almost exclusively double cropped. Further,

corn and cotton are often double cropped and beans can be triple cropped. The IBGE SIDRA database does provide data on the different crops for beans and for corn starting in 2003 (Figures 2 and 3), but unfortunately no such data is available for cotton.

Since we aim to remove double cropped areas from the SIDRA database planted area estimate, we will take into account only the area of the first crop of bean and corn. Since data are not available for 2000-2002, we extrapolated the tendencies shown in figures 2 and 3 to estimate the proportion of the total bean and corn area that were planted as the first crop during those years. In the case of beans, 32% of the total planted area corresponded to 2nd and 3rd crops during 2003, so we estimated the areas of 1st vs. 2nd and 3rd crops based on the assumption that this proportion stayed the same for 2000-2002. In the case of corn, 25% of the total planted area corresponded to the 2nd crop for 2003, so we estimated the areas of 1st vs. 2nd crops based on the assumption that this proportion stayed the same for 2000-2002.

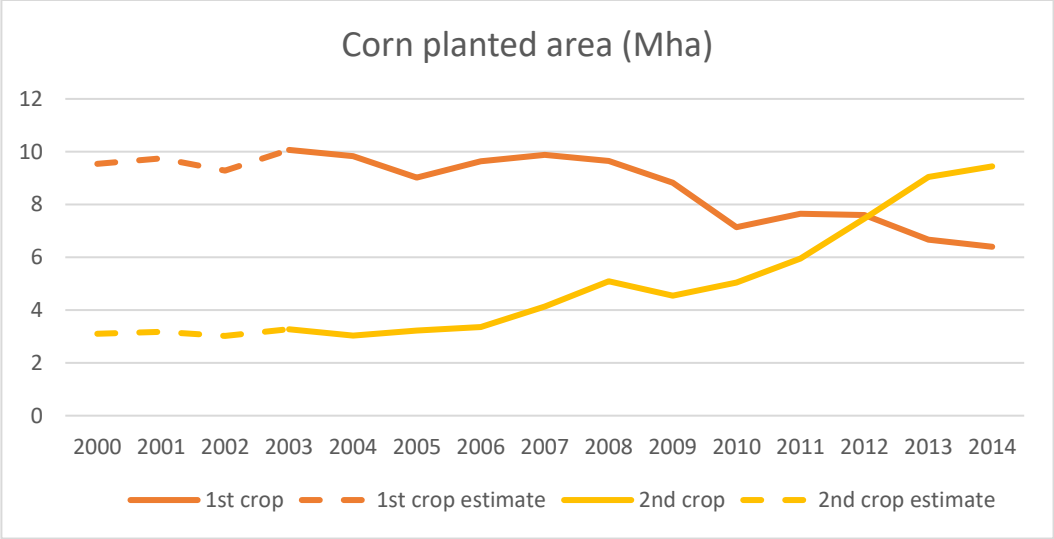


Figure 2. Area of 1st and 2nd season corn crops in Brazil. Dashed line represents estimated area based on the proportions of 1st to 2nd crop in 2003 (the first year for which data is available). Data from IBGE SIDRA database.

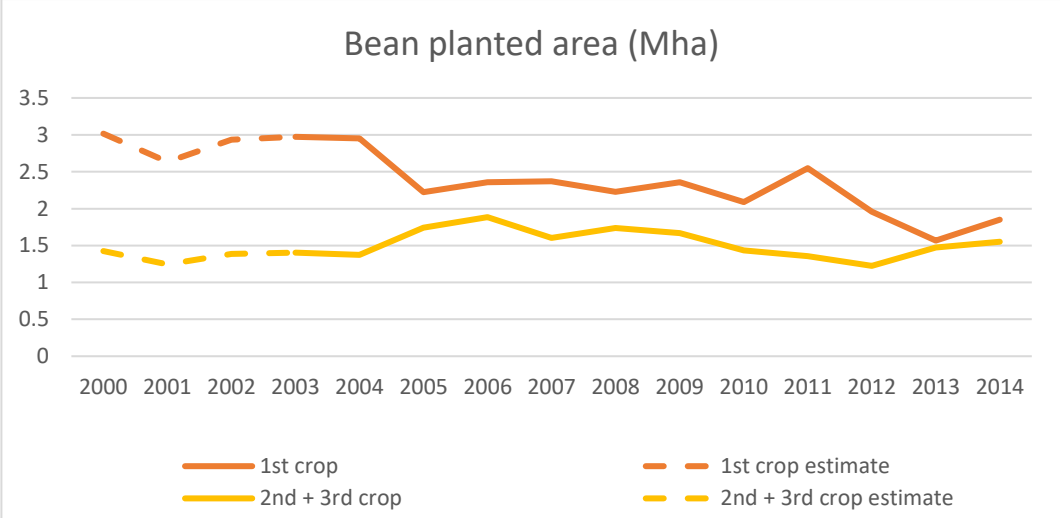


Figure 3. Area of 1st and combined 2nd and 3rd season bean crops in Brazil. Dashed line represents estimated area based on the proportions of 1st to 2nd and 3rd crop in 2003 (the first year for which data is available). Data from IBGE SIDRA database.

Finally, we determined that although manioc planted area is relatively significant at the national level, the way that it is produced in Brazil precludes it

from our definition of "cropland." In Brazil, manioc is produced almost exclusively as a subsistence crop in small fields for personal consumption. Because it is most generally not produced as an intensive row crop, it does not fit within the definition of the land cover class defined in our study. Therefore, manioc planted area was also removed from the total planted area in order to best approximate cropland area from the IBGE SIDRA database.

As mentioned earlier, cotton is at times produced as a second crop in Brazil. This is especially true, for example, in Mato Grosso, where soy-cotton rotations are common. However, the IBGE SIDRA database does not provide any information on the distinction between first and second crop cotton. Thus, we include the totality of cotton planted area in estimating cropland.

To summarize, in order to get an estimate of cropland area from the IBGE SIDRA database that would be comparable to our cropland area estimate, we followed the following steps:

1. Calculated total area of the eight most important crops in the country
2. Removed corn and bean 2nd and 3rd crop areas
    a. This step involved a preliminary step of estimating 1st crop areas for 2000 to 2002 based on the proportion of 1st to 2nd crops in 2003
3. Removed wheat area (it is a winter crop)
4. Removed manioc area (it does not fit our "intensive row crop" definition)

Table 1 shows the numbers corresponding to these calculations for the national level estimate. Numbers that appear in red in Table 1 are numbers that have been estimated.

These steps were followed to estimate cropland area based on the IBGE SIDRA database for the eleven states with the largest cropland areas: Mato Grosso do Sul, Mato Grosso, Goias, Sao Paulo, Para, Parana, Mato Grosso do Sul, Minas Gerais, Bahia, Tocantins, Piaui, and Maranhao. The results can be seen in Figures 4 and 5.

The final cropland area estimates for Brazil derived from the IBGE SIDRA database (henceforth referred to as IBGE LC, because it is the cropland Land Cover estimate) correspond to 35.7 Mha in 2000 and 52.5 Mha.

These estimates are indeed higher than the ones we present in our study (26.0 ± 1.1 Mha in 2000 and 46.5 ± 1.9 Mha in 2014). However, our results are based on unbiased and precise estimates of cropland area and cropland expansion in Brazil for our study period. As noted earlier, our study employs good practice guidance (1–7) on land cover area estimation using a probability sample of reference data. Our estimates are unbiased and reproducible, and associated uncertainties are reported. On the other hand, the IBGE database reports areas based on monthly consensus reports from municipal and regional commissions who gather data from producers. These data are then consolidated at the state level and at the national level. These data are inherently inconsistent through space because different municipal/regional commissions may have differing methodologies, definitions, expertise, capacity, and/or resources to report on cropland area. They may also be inconsistent through time because personnel changes at the commission level may lead to differing area estimates or because producers may incorrectly report planted

area. Lastly, IBGE provides no information on the accuracy or the precision of their area estimates.

When comparing our results to the IBGE LC estimates, we notice that although the national level result is higher than ours, we have similar estimates in many states (Figures 4 and 5). For example, in Mato Grosso, an established agricultural state where crop production is large-scale and intensive, IBGE LC results for 2000 and 2014 fall within one standard error of our area estimates (Figure 4). Other states where IBGE LC results fall within one standard error of our cropland area estimates are Goias (2000 and 2014), Mato Grosso do Sul (2000 and 2014), Sao Paulo (2014), Parana (2014), Bahia (2014) and Tocantins (2014) (Figures 4 and 5). These results only include the eleven states for which the IBGE LC calculation was performed.

In some states, however, we notice that our results vary significantly from the IBGE LC results. This is perhaps most striking in the states of Bahia (2000), Piaui, Maranhao, and Para. Bahia, Piaui, and Maranhao are part of the Matopiba agricultural frontier – an area where croplands have expanded rapidly and significantly in recent years. Because these states have not been agricultural hotspots historically, it is possible that a lack of personnel, funding, and/or capacity at the municipal/regional commission level might have led to inadequate area reports. In Bahia, for example, we report a 143% increase in cropland area whereas IBGE LC indicates a mere 10% increase in area. There is little doubt that IBGE LC is underestimating cropland expansion in Bahia. In Para, another state that has been reported as a newly initiated agricultural frontier, IBGE LC shows a decrease in

cropland area from 2000 to 2014. As a caveat, it is worth mentioning that manioc production is very important in this state, and we have removed manioc planted area from the IBGE LC estimates. However, soy area has grown significantly in this state, so we would still expect the IBGE LC estimate for Para to increase from 2000 to 2014. Instead, we see a 35% decrease which is likely due to a reporting error.

Additionally, we notice that MapBiomas' area estimates approach our estimates in three of the four states where our results differ the most from IBGE LC estimates (Piaui, Maranhao, and Para). In Bahia, MapBiomas presents significantly higher area estimates than ours. This may be due to differing "cropland" land cover definitions. This is difficult to assess because MapBiomas does not provide a formal definition of their "cropland" land cover class.

Figures 4 and 5 support the fact that our results do not systematically underestimate cropland area in Brazil. Additionally, given that our results are consistent (we used the same methodology, the same definitions, and the same data for every area estimate presented), figures 4 and 5 highlight the inconsistencies in the IBGE data. Compounding errors from the municipal to the state to the national level might have led to inflated IBGE LC area estimates for Brazil.

Our manuscript has been revised to include an explanation of the IBGE LC area estimate calculation and a brief comparison with our results. Supplementary material figures 8 and 9 have also been amended to incorporate the IBGE LC estimates.

		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Soy		13.69	13.99	16.38	18.53	21.60	23.43	22.08	20.57	21.25	21.76	23.34	24.03	25.09	27.95	30.31
Corn	**Total**	12.65	12.91	12.30	13.34	12.86	12.25	13.00	14.01	14.75	14.14	12.96	13.61	15.07	15.71	15.84
	1st crop	9.54	9.74	9.28	10.07	9.83	9.02	9.64	9.88	9.65	8.83	7.14	7.65	7.60	6.66	6.40
	2nd crop	3.11	3.17	3.02	3.28	3.03	3.23	3.36	4.13	5.10	4.54	5.05	5.95	7.47	9.05	9.44
Sugarcane		4.88	5.02	5.21	5.38	5.63	5.82	6.39	7.09	8.21	8.85	9.16	9.62	9.75	10.22	10.45
Beans	**Total**	4.44	3.88	4.32	4.38	4.33	3.97	4.24	3.98	3.97	4.28	3.66	3.91	3.18	3.04	3.40
	1st crop	3.02	2.63	2.93	2.97	2.95	2.23	2.36	2.37	2.23	2.36	2.09	2.55	1.96	1.57	1.85
	2nd + 3rd crop	1.43	1.24	1.39	1.40	1.37	1.74	1.88	1.60	1.74	1.67	1.44	1.36	1.22	1.47	1.55
Rice		3.70	3.17	3.17	3.19	3.77	4.00	3.01	2.92	2.87	2.91	2.78	2.86	2.44	2.39	2.35
Wheat		1.54	1.73	2.15	2.56	2.81	2.36	1.77	1.86	2.39	2.44	2.18	2.18	1.94	2.23	2.84
Mandioc		1.74	1.74	1.75	1.65	1.78	1.93	1.97	1.94	2.01	1.80	1.82	1.76	1.76	1.56	1.59
Cotton		0.81	0.89	0.76	0.72	1.16	1.27	0.91	1.13	1.07	0.81	0.83	1.41	1.42	0.95	1.13
Total planted area (top 8 crops)		43.45	43.33	46.04	49.75	53.95	55.01	53.38	53.49	56.51	56.99	56.73	59.36	60.65	64.04	67.92
Proportions of area attributed to first crop (based on 2003 proportions)																
Corn					0.75											
Beans					0.68											
Areas to be removed																
Wheat		1.54	1.73	2.15	2.56	2.81	2.36	1.77	1.86	2.39	2.44	2.18	2.18	1.94	2.23	2.84
Beans	**2nd + 3rd crop**	1.43	1.24	1.39	1.40	1.37	1.74	1.88	1.60	1.74	1.67	1.44	1.36	1.22	1.47	1.55
Corn	**2nd crop**	3.11	3.17	3.02	3.28	3.03	3.23	3.36	4.13	5.10	4.54	5.05	5.95	7.47	9.05	9.44
Mandioc		1.74	1.74	1.75	1.65	1.78	1.93	1.97	1.94	2.01	1.80	1.82	1.76	1.76	1.56	1.59

Table 1. Break-down of IBGE LC calculation for Brazil based on original IBGE data obtained from the SIDRA database. Numbers in red are areas that have been estimated.

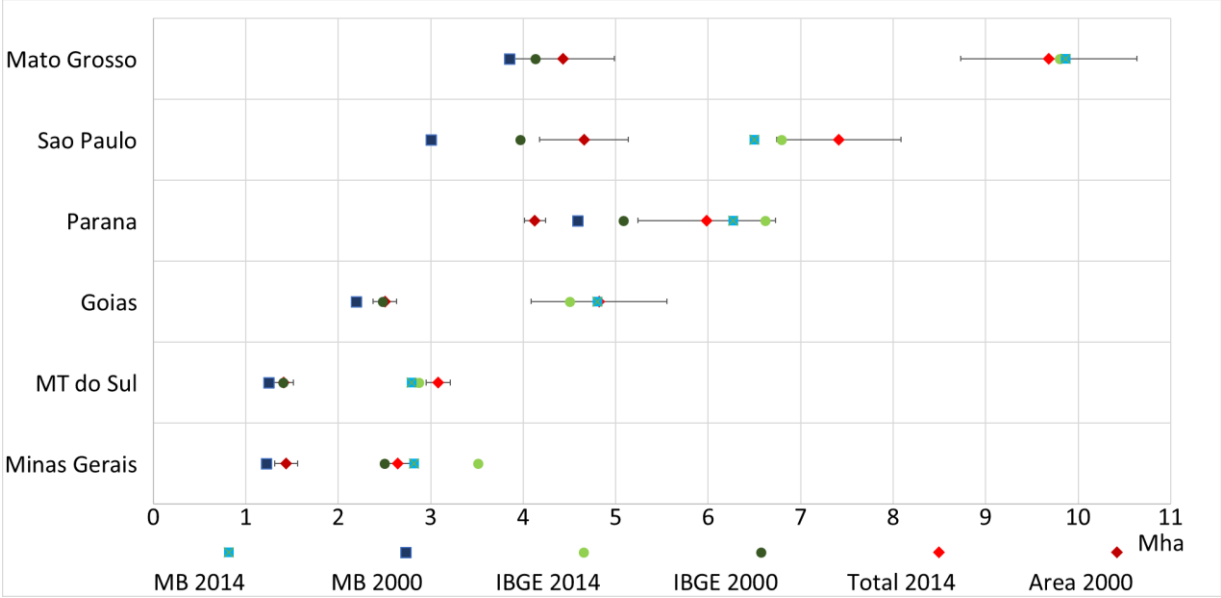


Figure 4. Comparison of area estimates from MapBiomas (MB), IBGE, and our study. Error bars on our estimates correspond to one standard error.

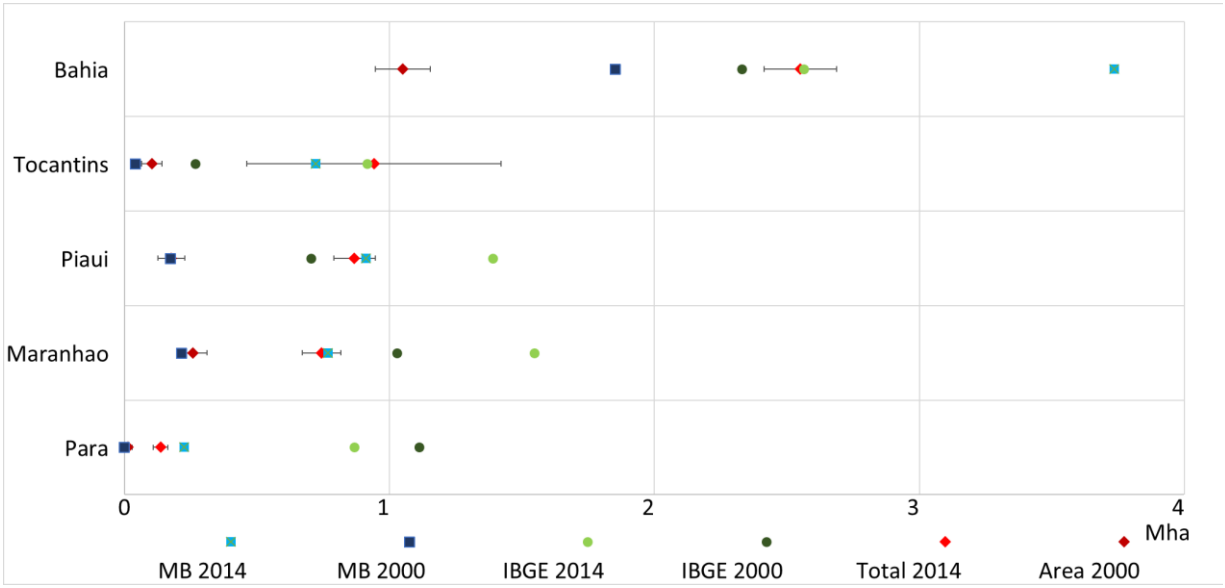


Figure 5. Comparison of area estimates from MapBiomas (MB), IBGE, and our study. Error bars on our estimates correspond to one standard error.

Bibliography

1. P. M. Vitousek, H. A. Mooney, J. Lubchenco, J. M. Melillo, Human Domination of Earth's Ecosystems. *Science* **277**, 494–499 (1997).

2. R. van Diemen, Ed., IPCC, 2019: Annex I: Glossary. *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*, 803 (2019).

3. V. H. Dale, The Relationship Between Land-Use Change and Climate Change. *Ecological Applications* **7**, 753–769 (1997).

4. Eric F. Lambin, Helmut J. Geist, E. Lepers, Dynamics of Land-Use and Land-Cover Change in Tropical Regions. *Annual Review of Environment and Resources* **28**, 205–241 (2003).

5. J. A. Foley, *et al.*, Global Consequences of Land Use. *Science* **309**, 570–574 (2005).

6. C. Sagan, O. B. Toon, J. B. Pollack, Anthropogenic Albedo Changes and the Earth's Climate. *Science* **206**, 1363–1368 (1979).

7. S. M. Sterling, A. Ducharne, J. Polcher, The impact of global land-cover change on the terrestrial water cycle. *Nature Climate Change* **3**, 385–390 (2013).

8. K. A. Farley, E. G. Jobbágy, R. B. Jackson, Effects of afforestation on water yield: a global synthesis with implications for policy. *Global Change Biology* **11**, 1565–1576 (2005).

9. S. Germer, C. Neill, A. V. Krusche, H. Elsenbeer, Influence of land-use change on near-surface hydrological processes: Undisturbed forest to pasture. *Journal of Hydrology* **380**, 473–480 (2010).

10. S. A. Spera, G. L. Galford, M. T. Coe, M. N. Macedo, J. F. Mustard, Land-use change affects water recycling in Brazil's last agricultural frontier. *Global Change Biology* **22**, 3405–3413 (2016).

11. G. Sampaio, *et al.*, Regional climate change over eastern Amazonia caused by pasture and soybean cropland expansion. *Geophysical Research Letters* **34** (2007).

12. T. E. Lovejoy, C. Nobre, Amazon Tipping Point. *Science Advances* **4**, eaat2340 (2018).

13. S. J. Vermeulen, B. M. Campbell, J. S. I. Ingram, Climate Change and Food Systems. *Annu. Rev. Environ. Resour.* **37**, 195–222 (2012).

14. G. R. van der Werf, *et al.*, CO 2 emissions from forest loss. *Nature Geoscience* **2**, 737–738 (2009).

15. Special Report on Climate Change and Land — IPCC site (April 22, 2020).

16. J. N. Galloway, The global nitrogen cycle: changes and consequences. *Environmental Pollution* **102**, 15–24 (1998).

17. D. Maranguit, T. Guillaume, Y. Kuzyakov, Land-use change affects phosphorus fractions in highly weathered tropical soils. *CATENA* **149**, 385–393 (2017).

18. P. M. Vitousek, *et al.*, Human Alteration of the Global Nitrogen Cycle: Sources and Consequences. *Ecological Applications* **7**, 737–750 (1997).

19. O. E. Sala, *et al.*, Global Biodiversity Scenarios for the Year 2100. *Science* **287**, 1770–1774 (2000).

20. D. U. Hooper, *et al.*, A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *NATURE* **486**, 105–108 (2012).

21. F. S. Chapin Iii, *et al.*, Consequences of changing biodiversity. *Nature* **405**, 234–242 (2000).

22. J. A. Patz, T. K. Graczyk, N. Geller, A. Y. Vittor, Effects of environmental change on emerging parasitic diseases. *International Journal for Parasitology* **30**, 1395–1405 (2000).

23. T. T.-Y. Lam, *et al.*, Identifying SARS-CoV-2 related coronaviruses in Malayan pangolins. *Nature*, 1–6 (2020).

24. Fact sheet about Malaria (April 22, 2020).

25. J. A. Patz, S. H. Olson, Malaria risk and temperature: influences from global climate change and local land use practices. *Proceedings of the National Academy of Sciences* **103**, 5635–5636 (2006).

26. C. Caminade, *et al.*, Impact of climate change on global malaria distribution. *PNAS* **111**, 3286–3291 (2014).

27. T. A. Ghebreyesus, *et al.*, Incidence of malaria among children living near dams in northern Ethiopia: community based incidence survey. *BMJ* **319**, 663–666 (1999).

28. T. Bunnag, S. Sornmani, S. Pinithpongse, C. Harinasuta, Surveillance of water-borne parasitic infections and studies on the impact of ecological changes on vector mosquitoes of malaria after dam construction. *Southeast Asian Journal of Tropical Medicine and Public Health.* **10**, 656–660 (1979).

29. Millennium Ecosystem Assessment (Program), Ed., *Ecosystems and human well-being: synthesis* (Island Press, 2005).

30. S. McGuire, FAO, IFAD, and WFP. The State of Food Insecurity in the World 2015: Meeting the 2015 International Hunger Targets: Taking Stock of Uneven Progress. Rome: FAO, 2015. *Adv Nutr* **6**, 623–624 (2015).

31. E. F. Lambin, P. Meyfroidt, Global land use change, economic globalization, and the looming land scarcity. *PNAS* **108**, 3465–3472 (2011).

32. W. F. Laurance, *et al.*, Reducing the global environmental impacts of rapid infrastructure expansion. *Current Biology* **25**, R259–R262 (2015).

33. M. C. Hansen, *et al.*, The fate of tropical forest fragments. *Science Advances* **6**, eaax8574 (2020).

34. R. J. W. Brienen, *et al.*, Long-term decline of the Amazon carbon sink. *Nature* **519**, 344–348 (2015).

35. M. J. Heckenberger, J. Christian Russell, J. R. Toney, M. J. Schmidt, The legacy of cultural landscapes in the Brazilian Amazon: implications for biodiversity. *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**, 197–208 (2007).

36. N. Myers, R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, J. Kent, Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858 (2000).

37. Y. le P. de Waroux, R. D. Garrett, R. Heilmayr, E. F. Lambin, Land-use policies and corporate investments in agriculture in the Gran Chaco and Chiquitano. *PNAS*, 201602646 (2016).

38. A. J. G. Simoes, C. A. Hidalgo, The economic complexity observatory: An analytical tool for understanding the dynamics of economic development in *Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence*, (2011).

39. R. Müller, D. Müller, F. Schierhorn, G. Gerold, P. Pacheco, Proximate causes of deforestation in the Bolivian lowlands: an analysis of spatial dynamics. *Reg Environ Change* **12**, 445–459 (2012).

40. M. M. Caldas, D. Goodin, S. Sherwood, J. M. Campos Krauer, S. M. Wisely, Land-cover change in the Paraguayan Chaco: 2000–2011. *Journal of Land Use Science* **10**, 1–18 (2015).

41. M. Baraibar, Green Deserts or New Opportunities? : Competing and complementary views on the soybean expansion in Uruguay, 2002-2013 (2014) (April 23, 2020).

42. FAOSTAT (July 6, 2017).

43. P. G. Curtis, C. M. Slay, N. L. Harris, A. Tyukavina, M. C. Hansen, Classifying drivers of global forest loss. *Science* **361**, 1108–1111 (2018).

44. N. Alexandratos, J. Bruinsma, others, "World agriculture towards 2030/2050: the 2012 revision" (ESA Working paper Rome, FAO, 2012) (September 16, 2016).

45. T. Wassenaar, *et al.*, Projecting land use changes in the Neotropics: The geography of pasture expansion into forest. *Global Environmental Change* **17**, 86–104 (2007).

46. C. E. Woodcock, *et al.*, Free Access to Landsat Imagery. *Science* **320**, 1011–1011 (2008).

47. M. C. Hansen, T. R. Loveland, A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment* **122**, 66–74 (2012).

48. PRODES — Coordenação-Geral de Observação da Terra (April 14, 2020).

49. Y. Shimabukuro, *et al.*, Near real time detection of deforestation in the Brazilian Amazon using MODIS imagery. *Ambiente e Agua-An Interdisciplinary Journal of Applied Science* **1**, 37–47 (2007).

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

51. M. C. Hansen, *et al.*, Humid tropical forest disturbance alerts using Landsat data. *Environ. Res. Lett.* **11**, 034008 (2016).

52. J. A. Foley, *et al.*, Solutions for a cultivated planet. *Nature* **478**, 337–342 (2011).

53. S. Fritz, *et al.*, Mapping global cropland and field size. *Glob Change Biol* **21**, 1980–1992 (2015).

54. , Famine Early Warning Systems Network (April 29, 2020).

55. , Projeto TerraClass (July 6, 2017).

56. , Projeto TerraClass Cerrado (July 6, 2017).

57. R. Beuchle, *et al.*, Land cover changes in the Brazilian Cerrado and Caatinga biomes from 1990 to 2010 based on a systematic remote sensing sampling approach. *Applied Geography* **58**, 116–127 (2015).

58. M. L. Clark, T. M. Aide, H. R. Grau, G. Riner, A scalable approach to mapping annual land cover at 250 m using MODIS time series data: A case study in the Dry Chaco ecoregion of South America. *Remote Sensing of Environment* **114**, 2816–2832 (2010).

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

60. D. C. Morton, *et al.*, Cropland expansion changes deforestation dynamics in the southern Brazilian Amazon. *PNAS* **103**, 14637–14641 (2006).

61. M. L. Clark, T. M. Aide, G. Riner, Land change for all municipalities in Latin America and the Caribbean assessed from 250-m MODIS imagery (2001–2010). *Remote Sensing of Environment* **126**, 84–103 (2012).

62. V. D. Sy, *et al.*, Land use patterns and related carbon losses following deforestation in South America. *Environ. Res. Lett.* **10**, 124004 (2015).

63. J. Graesser, N. Ramankutty, Detection of cropland field parcels from Landsat imagery. *Remote Sensing of Environment* **201**, 165–180 (2017).

64. B. F. T. Rudorff, *et al.*, Studies on the Rapid Expansion of Sugarcane for Ethanol Production in São Paulo State (Brazil) Using Landsat Data. *Remote Sensing* **2**, 1057–1076 (2010).

65. B. F. T. Rudorff, *et al.*, Remote Sensing Images to Detect Soy Plantations in the Amazon Biome—The Soy Moratorium Initiative. *Sustainability* **4**, 1074–1088 (2012).

66. M. Adami, *et al.*, Remote Sensing Time Series to Evaluate Direct Land Use Change of Recent Expanded Sugarcane Crop in Brazil. *Sustainability* **4**, 574–585 (2012).

67. M. Piquer-Rodríguez, *et al.*, Drivers of agricultural land-use change in the Argentine Pampas and Chaco regions. *Applied Geography* **91**, 111–122 (2018).

68. T. J. Killeen, *et al.*, Thirty years of land-cover change in Bolivia. *AMBIO: A journal of the Human Environment* **36**, 600–606 (2007).

69. T. J. Killeen, *et al.*, Total historical land-use change in eastern Bolivia: Who, where, when, and how much? *Ecology and Society* **13** (2008).

70. M. Baumann, *et al.*, Carbon emissions from agricultural expansion and intensification in the Chaco. *Glob Change Biol*, n/a-n/a (2016).

71. N. I. Gasparri, H. R. Grau, L. V. Sacchi, Determinants of the spatial distribution of cultivated land in the North Argentine Dry Chaco in a multi-decadal study. *Journal of Arid Environments* **123**, 31–39 (2015).

72. G. Baldi, *et al.*, Cultivating the dry forests of South America: Diversity of land users and imprints on ecosystem functioning. *Journal of Arid Environments* **123**, 47–59 (2015).

73. B. Jakimow, P. Griffiths, S. van der Linden, P. Hostert, Mapping pasture management in the Brazilian Amazon from dense Landsat time series. *Remote Sensing of Environment* **205**, 453–468 (2018).

74. P. Arévalo, P. Olofsson, C. E. Woodcock, Continuous monitoring of land change activities and post-disturbance dynamics from Landsat time series: A test methodology for REDD+ reporting. *Remote Sensing of Environment* (2019) https:/doi.org/10.1016/j.rse.2019.01.013 (February 7, 2019).

75. M. Baumann, *et al.*, Deforestation and cattle expansion in the Paraguayan Chaco 1987–2012. *Reg Environ Change* **17**, 1179–1191 (2017).

76. C. A. de M. Scaramuzza, *et al.*, LAND-USE AND LAND-COVER MAPPING OF THE BRAZILIAN CERRADO BASED MAINLY ON LANDSAT-8 SATELLITE IMAGES. *Revista Brasileira de Cartografia* **69** (2017).

77. L. Parente, *et al.*, Monitoring the brazilian pasturelands: A new mapping approach based on the landsat 8 spectral and temporal domains. *International Journal of Applied Earth Observation and Geoinformation* **62**, 135–143 (2017).

78. P. R. Furumo, T. M. Aide, Characterizing commercial oil palm expansion in Latin America: land use change and trade. *Environ. Res. Lett.* **12**, 024008 (2017).

79. H. D. Eva, *et al.*, A land cover map of South America. *Global Change Biology* **10**, 731–744 (2004).

80. C. Giri, J. Long, Land Cover Characterization and Mapping of South America for the Year 2010 Using Landsat 30 m Satellite Data. *Remote Sensing* **6**, 9494–9510 (2014).

81. J. Graesser, T. M. Aide, H. R. Grau, N. Ramankutty, Cropland/pastureland dynamics and the slowdown of deforestation in Latin America. *Environ. Res. Lett.* **10**, 034017 (2015).

82. S. V. Stehman, Estimating area from an accuracy assessment error matrix. *Remote Sensing of Environment* **132**, 202–211 (2013).

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

84. GFOI, "Integration of remote-sensing and ground-based observations for estimation of emissions and removals of greenhouse gases in forests: methods and guidance from the Global Forest Observations Initiative" (2016).

85. Map Accuracy Assessment and Area Estimation: A Practical Guide. *National Forest Monitoring Assessment. Working Paper (FAO)* (2016).

86. "2006 IPCC Guidelines for National Greenhouse Gas Inventories" (National Greenhouse Gas Inventories Programme).

87. F. J. Gallego, *et al.*, Efficiency assessment of using satellite data for crop area estimation in Ukraine. *International Journal of Applied Earth Observation and Geoinformation* **29**, 22–30 (2014).

88. R. E. McRoberts, Satellite image-based maps: Scientific inference or pretty pictures? *Remote Sensing of Environment* **115**, 715–724 (2011).

89. MapBiomas (July 6, 2017).

90. UN Comtrade | International Trade Statistics Database (July 4, 2017).

91. M. N. Macedo, *et al.*, Decoupling of deforestation and soy production in the southern Amazon during the late 2000s. *PNAS* **109**, 1341–1346 (2012).

92. L. C. P. Dias, F. M. Pimenta, A. B. Santos, M. H. Costa, R. J. Ladle, Patterns of land use, extensification, and intensification of Brazilian agriculture. *Glob Change Biol* **22**, 2887–2903 (2016).

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

94. D. Tilman, K. G. Cassman, P. A. Matson, R. Naylor, S. Polasky, Agricultural sustainability and intensive production practices. *Nature* **418**, 671–677 (2002).

95. R. D. Garrett, E. F. Lambin, R. L. Naylor, The new economic geography of land use change: Supply chain configurations and land use in the Brazilian Amazon. *Land Use Policy* **34**, 265–275 (2013).

96. P. M. Fearnside, Soybean cultivation as a threat to the environment in Brazil. *Environmental Conservation* **28**, 23–38 (2001).

97. P. M. Fearnside, Deforestation in Brazilian Amazonia: history, rates, and consequences. *Conservation biology* **19**, 680–688 (2005).

98. D. C. Nepstad, C. M. Stickler, O. T. Almeida, Globalization of the Amazon Soy and Beef Industries: Opportunities for Conservation. *Conservation Biology* **20**, 1595–1603 (2006).

99. D. Kaimowitz, J. Smith, "Soybean technology and the loss of natural vegetation in Brazil and Bolivia" in *Agricultural Technologies and Tropical Deforestation*, (CABI Publishing, Wallingford, Oxon, UK, 2001), pp. 195–211.

100. E. Barona, N. Ramankutty, G. Hyman, O. T. Coomes, The role of pasture and soybean in deforestation of the Brazilian Amazon. *Environ. Res. Lett.* **5**, 024002 (2010).

101. B. F. T. Rudorff, *et al.*, The Soy Moratorium in the Amazon Biome Monitored by Remote Sensing Images. *Remote Sensing* **3**, 185–202 (2011).

102. S. A. Spera, *et al.*, Recent cropping frequency, expansion, and abandonment in Mato Grosso, Brazil had selective land characteristics. *Environ. Res. Lett.* **9**, 064010 (2014).

103. G. L. Galford, *et al.*, Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sensing of Environment* **112**, 576–587 (2008).

104. J. M. Cardoso Da Silva, J. M. Bates, Biogeographic Patterns and Conservation in the South American Cerrado: A Tropical Savanna Hotspot. *BioScience* **52**, 225–234 (2002).

105. C. Brannstrom, *et al.*, Land change in the Brazilian Savanna (Cerrado), 1986–2002: Comparative analysis and implications for land-use policy. *Land Use Policy* **25**, 579–595 (2008).

106. P. Noojipady, *et al.*, Forest carbon emissions from cropland expansion in the Brazilian Cerrado biome. *Environ. Res. Lett.* **12**, 025004 (2017).

107. H. K. Gibbs, *et al.*, Brazil's Soy Moratorium. *Science* **347**, 377–378 (2015).

108. S. Spera, L. VanWey, J. Mustard, The drivers of sugarcane expansion in Goiás, Brazil. *Land Use Policy* **66**, 111–119 (2017).

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

110. C. A. de Almeida, *et al.*, 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**, 291–302 (2016).

111. Canasat - INPE (July 6, 2017).

112. P. Zuurbier, J. Van de Vooren, Introduction to sugarcane ethanol contributions to climate change mitigation and the environment. *Sugarcane ethanol*, 19 (2008).

113. P. D. Richards, R. J. Myers, S. M. Swinton, R. T. Walker, Exchange rates, soybean supply response, and deforestation in South America. *Global Environmental Change* **22**, 454–462 (2012).

114. Mapeamento Sistemático do Uso da Terra. *Instituto Brasileiro de Geografia e Estatística* (September 27, 2018).

115. D. Arvor, M. Jonathan, M. S. P. Meirelles, V. Dubreuil, L. Durieux, Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. *International Journal of Remote Sensing* **32**, 7847–7871 (2011).

116. IBGE SIDRA - Produção Agrícola Municipal.

117. D. Nepstad, *et al.*, Slowing Amazon deforestation through public policy and interventions in beef and soy supply chains. *Science* **344**, 1118–1123 (2014).

118. D. Nepstad, *et al.*, The End of Deforestation in the Brazilian Amazon. *Science* **326**, 1350–1351 (2009).

119. D. Arvor, *et al.*, Combining socioeconomic development with environmental governance in the Brazilian Amazon: the Mato Grosso agricultural frontier at a tipping point. *Environ Dev Sustain*, 1–22 (2016).

120. E. Mello, GAIN report - Oilseeds and Products Annual 2005 (2005) (July 6, 2017).

121. C. V. Godoy, *et al.*, Asian soybean rust in Brazil: past, present, and future. *Pesquisa Agropecuária Brasileira* **51**, 407–421 (2016).

122. Conab - Companhia Nacional de Abastecimento (July 6, 2017).

123. A. Tyukavina, *et al.*, Aboveground carbon loss in natural and managed tropical forests from 2000 to 2012. *Environmental Research Letters* **10**, 074002 (2015).

124. J. Assunção, C. Gandour, R. Rocha, "DETERring deforestation in the Brazilian Amazon: environmental monitoring and law enforcement" (Climate Policy Initiative, 2013) (July 6, 2017).

125. B. Soares-Filho, *et al.*, Role of Brazilian Amazon protected areas in climate change mitigation. *PNAS* **107**, 10821–10826 (2010).

126. D. M. Lapola, *et al.*, Indirect land-use changes can overcome carbon savings from biofuels in Brazil. *PNAS* **107**, 3388–3393 (2010).

127. S. Andrade de Sá, C. Palmer, S. di Falco, Dynamics of indirect land-use change: Empirical evidence from Brazil. *Journal of Environmental Economics and Management* **65**, 377–393 (2013).

128. Robert K. Hoff, Laura J. Geller, Priscila Ming, "Advances in Agricultural Infrastructure in the North of Brazil" (USDA Foreign Agricultural Service, 2014).

129. E. F. Vermote, N. Z. El Saleous, C. O. Justice, Atmospheric correction of MODIS data in the visible to middle infrared: first results. *Remote Sensing of Environment* **83**, 97–111 (2002).

130. B. Rabus, M. Eineder, A. Roth, R. Bamler, The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. *ISPRS Journal of Photogrammetry and Remote Sensing* **57**, 241–262 (2003).

131. C. O. Justice, *et al.*, An overview of MODIS Land data processing and product status. *Remote Sensing of Environment* **83**, 3–15 (2002).

132. P. V. Potapov, *et al.*, Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with Landsat ETM+ data. *Remote Sensing of Environment* **122**, 106–116 (2012).

133. G. Chander, B. L. Markham, D. L. Helder, Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment* **113**, 893–903 (2009).

134. P. V. Potapov, *et al.*, National satellite-based humid tropical forest change assessment in Peru in support of REDD+ implementation. *Environmental Research Letters* **9**, 124012 (2014).

135. R. DeFries, M. Hansen, J. Townshend, Global discrimination of land cover types from metrics derived from AVHRR pathfinder data. *Remote Sensing of Environment* **54**, 209–222 (1995).

136. L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, *Classification and regression trees* (1984) (July 6, 2017).

137. Q. Ying, *et al.*, Global bare ground gain from 2000 to 2012 using Landsat imagery. *Remote Sensing of Environment* **194**, 161–176 (2017).

138. A. Tyukavina, *et al.*, Types and rates of forest disturbance in Brazilian Legal Amazon, 2000–2013. *Science Advances* **3**, e1601047 (2017).

139. J. Chang, M. C. Hansen, K. Pittman, M. Carroll, C. DiMiceli, Corn and Soybean Mapping in the United States Using MODIS Time-Series Data Sets. *Agronomy Journal* **99**, 1654–1664 (2007).

140. K. Pittman, M. C. Hansen, I. Becker-Reshef, P. V. Potapov, C. O. Justice, Estimating global cropland extent with multi-year MODIS data. *Remote Sensing* **2**, 1844–1863 (2010).

141. J. Dempewolf, *et al.*, Wheat Yield Forecasting for Punjab Province from Vegetation Index Time Series and Historic Crop Statistics. *Remote Sensing* **6**, 9653–9675 (2014).

142. X.-P. Song, *et al.*, National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey. *Remote Sensing of Environment* **190**, 383–395 (2017).

143. L. King, *et al.*, A multi-resolution approach to national-scale cultivated area estimation of soybean. *Remote Sensing of Environment* **195**, 13–29 (2017).

144. A. Khan, M. C. Hansen, P. Potapov, S. V. Stehman, A. A. Chatta, Landsat-based wheat mapping in the heterogeneous cropping system of Punjab, Pakistan. *International Journal of Remote Sensing* **37**, 1391–1410 (2016).

145. D. C. Morton, Forest carbon fluxes: A satellite perspective. *Nature Clim. Change* **6**, 346–348 (2016).

146. P. Olofsson, G. M. Foody, S. V. Stehman, C. E. Woodcock, 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**, 122–131 (2013).

147. F. Krausmann, *et al.*, Global human appropriation of net primary production doubled in the 20th century. *PNAS* **110**, 10324–10329 (2013).

148. X.-P. Song, *et al.*, Global land change from 1982 to 2016. *Nature* **560**, 639–643 (2018).

149. Summary for Policymakers — Special Report on Climate Change and Land (April 6, 2020).

150. UNEP, Ed., *Mainstreaming the economics of nature: a synthesis of the approach, conclusions and recommendations of teeb* (UNEP, 2010).

151. A. P. Kinzig, *et al.*, Paying for Ecosystem Services—Promise and Peril. *Science* **334**, 603–604 (2011).

152. M. M. C. Bustamante, *et al.*, Toward an integrated monitoring framework to assess the effects of tropical forest degradation and recovery on carbon stocks and biodiversity. *Global Change Biology* **22**, 92–109 (2016).

153. J. Penman, *et al.*, Good practice guidance for land use, land-use change and forestry. *Good practice guidance for land use, land-use change and forestry*. (2003).

154. FAO, Global ecological zones for FAO forest reporting: 2010 Update. *FAO: Rome, Italy* (2012).

155. G. B. Martha, E. Alves, E. Contini, Land-saving approaches and beef production growth in Brazil. *Agricultural Systems* **110**, 173–177 (2012).

156. E. F. Lambin, *et al.*, Estimating the world's potentially available cropland using a bottom-up approach. *Global Environmental Change* **23**, 892–901 (2013).

157. R. Müller, D. Müller, F. Schierhorn, G. Gerold, Spatiotemporal modeling of the expansion of mechanized agriculture in the Bolivian lowland forests. *Applied Geography* **31**, 631–640 (2011).

158. D. C. Morton, *et al.*, Reevaluating suitability estimates based on dynamics of cropland expansion in the Brazilian Amazon. *Global Environmental Change* **37**, 92–101 (2016).

159. F. Zabel, *et al.*, Global impacts of future cropland expansion and intensification on agricultural markets and biodiversity. *Nature Communications* **10**, 1–10 (2019).

160. F. M. Resende, J. Cimon-Morin, M. Poulin, L. Meyer, R. Loyola, Consequences of delaying actions for safeguarding ecosystem services in the Brazilian Cerrado. *Biological Conservation* **234**, 90–99 (2019).

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

162. W. F. Laurance, J. Sayer, K. G. Cassman, Agricultural expansion and its impacts on tropical nature. *Trends in Ecology & Evolution* **29**, 107–116 (2014).

163. P. Balvanera, A. Castillo, M. J. MartÍnez-Harms, "Ecosystem Services in Seasonally Dry Tropical Forests" in *Seasonally Dry Tropical Forests: Ecology and Conservation*, R. Dirzo, H. S. Young, H. A. Mooney, G.

Ceballos, Eds. (Island Press/Center for Resource Economics, 2011), pp. 259–277.

164. A. S. Cohn, *et al.*, Cattle ranching intensification in Brazil can reduce global greenhouse gas emissions by sparing land from deforestation. *PNAS* **111**, 7236–7241 (2014).

165. B. B. N. Strassburg, *et al.*, When enough should be enough: Improving the use of current agricultural lands could meet production demands and spare natural habitats in Brazil. *Global Environmental Change* **28**, 84–97 (2014).

166. J. M. Dros, Managing the Soy Boom: Two scenarios of soy production. *Amsterdam: AIDEnvironment* (2004).

167. A. J. Cattelan, A. Dall'Agnol, The rapid soybean growth in Brazil. *Embrapa Soja-Artigo em periódico indexado (ALICE)* (2018).

168. H. K. Gibbs, *et al.*, Brazil's Soy Moratorium. *Science* **347**, 377–378 (2015).

169. E. K. H. J. zu Ermgassen, *et al.*, Using supply chain data to monitor zero deforestation commitments: an assessment of progress in the Brazilian soy sector. *Environ. Res. Lett.* (2019) https:/doi.org/10.1088/1748-9326/ab6497 (December 24, 2019).

170. V. Zalles, *et al.*, Near doubling of Brazil's intensive row crop area since 2000. *PNAS* **116**, 428–435 (2019).

171. J. R. G. Townshend, C. O. Justice, C. O., V. Kalb, Characterization and classification of South American land cover types using satellite data. *International Journal of Remote Sensing* **8**, 1189–1207 (1987).

172. M. A. Friedl, *et al.*, Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment* **83**, 287–302 (2002).

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

174. J.-F. Pekel, A. Cottam, N. Gorelick, A. S. Belward, High-resolution mapping of global surface water and its long-term changes. *Nature* **540**, 418–422 (2016).

175. P. Potapov, *et al.*, Annual continuous fields of woody vegetation structure in the Lower Mekong region from 2000-2017 Landsat time-series. *Remote Sensing of Environment* **232**, 111278 (2019).

176. P. Potapov, *et al.*, Landsat Analysis Ready Data for Global Land Cover and Land Cover Change Mapping. *Remote Sensing* **12**, 426 (2020).

177. M. C. Hansen, R. S. Defries, J. R. G. Townshend, R. Sohlberg, Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing* **21**, 1331–1364 (2000).

178. A. H. Pickens, *et al.*, Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series. *Remote Sensing of Environment* **243**, 111792 (2020).

179. T. Lumley, *Package 'survey'* (2020).

180. S. V. Stehman, Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes. *International Journal of Remote Sensing* **35**, 4923–4939 (2014).

181. E. L. Bullock, C. E. Woodcock, P. Olofsson, Monitoring tropical forest degradation using spectral unmixing and Landsat time series analysis. *Remote Sensing of Environment* (2018) https:/doi.org/10.1016/j.rse.2018.11.011 (December 5, 2018).

182. C. M. Souza, D. A. Roberts, A. Monteiro, Multitemporal Analysis of Degraded Forests in the Southern Brazilian Amazon. *Earth Interact.* **9**, 1–25 (2005).

183. E. W. Sanderson, *et al.*, The Human Footprint and the Last of the WildThe human footprint is a global map of human influence on the land surface, which suggests that human beings are stewards of nature, whether we like it or not. *BioScience* **52**, 891–904 (2002).

184. O. Venter, *et al.*, Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nat Commun* **7**, 1–11 (2016).

185. H. Haberl, *et al.*, Quantifying and mapping the human appropriation of net primary production in earth's terrestrial ecosystems. *PNAS* **104**, 12942–12947 (2007).

186. E. L. Bullock, C. E. Woodcock, C. Souza Jr, P. Olofsson, Satellite-based estimates reveal widespread forest degradation in the Amazon. *Global Change Biology* (2020).

187. World Development Indicators. *World Bank* (May 5, 2020).

188. Chile Case Study Prepared for FAO as part of the State of the World's Forests 2016 (SOFO). 74.

189. Agricultural exports on the rise in Peru. *Oxford Business Group* (2016) (April 7, 2020).

190. M. Carroll, *et al.*, “MODIS Vegetative Cover Conversion and Vegetation Continuous Fields” in *Land Remote Sensing and Global Environmental Change: NASA’s Earth Observing System and the Science of ASTER and MODIS*, Remote Sensing and Digital Image Processing., B. Ramachandran, C. O. Justice, M. J. Abrams, Eds. (Springer, 2010), pp. 725–745.

191. M. C. Hansen, R. S. DeFries, Detecting Long-term Global Forest Change Using Continuous Fields of Tree-Cover Maps from 8-km Advanced Very High Resolution Radiometer (AVHRR) Data for the Years 1982–99. *Ecosystems* **7**, 695–716 (2004).

192. M. C. Hansen, J. R. G. Townshend, R. S. DeFries, M. Carroll, Estimation of tree cover using MODIS data at global, continental and regional/local scales. *International Journal of Remote Sensing* **26**, 4359–4380 (2005).

193. C. Gómez, J. C. White, M. A. Wulder, Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* **116**, 55–72 (2016).

194. L. Breiman, Bagging predictors. *Machine learning* **24**, 123–140 (1996).

195. Chave de Interpretação MapBiomas (April 12, 2020).

196. M. Broich, S. V. Stehman, M. C. Hansen, P. Potapov, Y. E. Shimabukuro, A comparison of sampling designs for estimating deforestation from Landsat imagery: A case study of the Brazilian Legal Amazon. *Remote Sensing of Environment* **113**, 2448–2454 (2009).

197. S. V. Stehman, Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing* **30**, 5243–5272 (2009).

198. M. Hansen, R. Dubayah, R. DeFries, Classification trees: an alternative to traditional land cover classifiers. *International Journal of Remote Sensing* **17**, 1075–1081 (1996).

199. M. C. Hansen, *et al.*, Mapping tree height distributions in Sub-Saharan Africa using Landsat 7 and 8 data. *Remote Sensing of Environment* https:/doi.org/10.1016/j.rse.2016.02.023 (March 8, 2016).

200. A. Di Gregorio, *Land cover classification system: classification concepts and user manual: LCCS* (Food & Agriculture Org., 2005).