HAL open science

Monitoring and understanding tropical agricultural frontiers: A remote sensing based approach

Damien Arvor

To cite this version:

Damien Arvor. Monitoring and understanding tropical agricultural frontiers: A remote sensing based approach. Geography. Université Rennes 2, 2022. tel-03894655

HAL Id: tel-03894655

https://shs.hal.science/tel-03894655

Submitted on 12 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

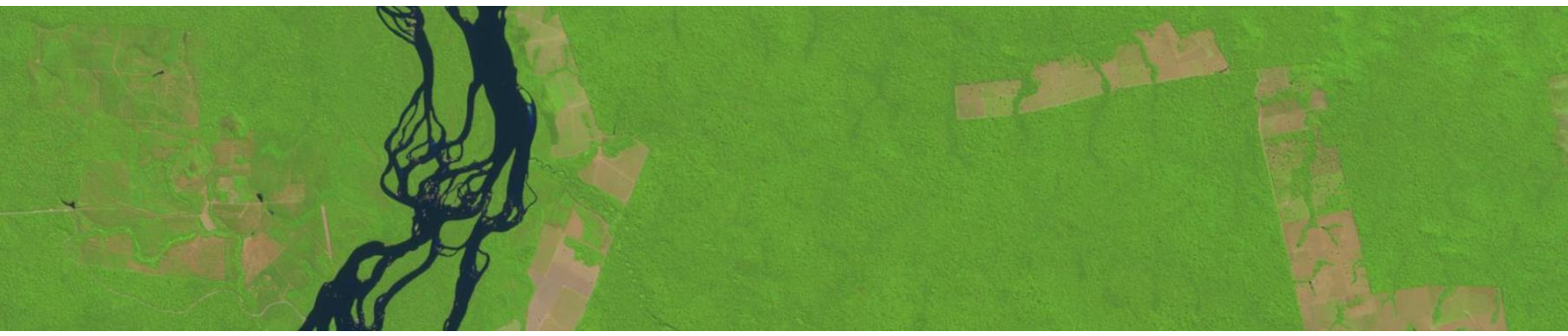
Mémoire d'Habilitation à Diriger Des Recherches

présenté par

Damien ARVOR

Chargé de Recherches
Université Rennes 2 - UMR LETG CNRS 6554
Ecole doctorale Espaces, Sociétés, Civilisations (ESC)

Monitoring and understanding tropical agricultural frontiers: A remote sensing based approach



Volume 1:
Position et projet scientifique

Mémoire présenté et soutenu à Rennes le 25 octobre 2022
devant le jury composé de:

Agnès Bégué	Chercheuse, CIRAD	Garante
Vincent Dubreuil	Professeur, Université Rennes 2	Examinateur
Liana O. Anderson	Chercheuse, CEMADEN	Examinatrice
Marcos H. Costa	Professeur, Universidade Federal de Viçosa	Rapporteur
Valéry Gond	Chercheur, CIRAD	Rapporteur
Christophe Claramunt	Professeur, Ecole Navale	Rapporteur

UNIVERSITÉ RENNES 2

Monitoring and understanding tropical agricultural frontiers : a remote sensing based approach

by

Damien Arvor

Mémoire d'habilitation à diriger les recherches

in the
UFR Sciences sociales
Département de géographie et d'aménagement de l'espace

August 2022

Acknowledgements

I would like to express my very great appreciation to :

- the members of the jury for accepting to revise this thesis.
- my colleagues and friends from LETG research unit : Jean, Thomas C., Samuel, NDVI², Beatriz, Thomas H., Valérie, Anne-Julia, Simon, Hervé, Solène, Johan, Adeline, Yann, Romain and Laurence. I am also greatful to Françoise and Brice for their efforts while leading the unit. I sincerely acknowledge Claudie and Roselyne for their essential support ! Finally, I thank the "young generation" of LETG for their dynamism, unlimited motivation and fruitful exchanges.
- my colleagues and friends from CIRAD Montpellier, especially Agnès (who supervised this dissertation), Julie, Lilian, Renan, Hélène, Valéry, Sylvie and Emilie.
- my colleagues and friends from IRD Montpellier : Laurent (Durieux), Nadine, Jeff, Thibault, Laurent (Demagistri) and Marie-Paule.
- my numerous colleagues and friends from Brazil : Margareth ("minha mamae brasileira") and Rodrigo from Embrapa Rio de Janeiro, Christovam from Friocruz, Carlos from UNEMAT, Vinicius from ICV, Andre from SEMA, Dona Vitoria from FEC, "Dudy e Nando" Paiva from CAT Sorriso, Laurimar from Embrapa Sinop, Margarete from UNESP, Stéphanie and Marcel from Brasilia, Ari from CIFOR, Claudio from INPE, Andre and Alex from Maranhão, Pedro from UFRN. I also wish to express my gratitude to Hervé and Neli from USP who kindly shared their endless experience of the Amazon.
- my colleagues and friends worldwide: Mariana from Netherlands, Prof Tang, Prof Zhang and Prof Le Yu from Chinese Academy of Sciences and Tsinghua University in Beijing, Bill and Prof Clobite from Brazzaville, the R students from Costa Rica and Thailand.
- my partners from Alkante with whom I hope we will manage to strengthen our joint researches.
- my friends from Brittany, Montpellier and other distant locations.
- my family and especially my son, Noan.

"La dernière épreuve que l'homme à la bêche s'apprête à subir de son vivant, c'est la destruction de son jardin. Car les paysans sont les véritables jardiniers de la terre, les seuls qui soient à l'échelle de la géographie. Leur travail patient a minutieusement établi la plupart des paysages sur lesquels nous reposons nos yeux. Ce faisant, certes, ils n'étaient pas menés par un souci artistique, encore qu'ils aient toujours su admirer les beautés de la nature aussi bien que le premier Jean-Jacques Rousseau venu. N'ayant d'autres moyens que leurs bras, quelques outils rudimentaires et quelques animaux domestiques, il leur fallait bien s'acommoder du relief, laisser les hauteurs aux végétations sauvages ou à la nudité, cultiver les pentes en respectant les courbes de niveau, aménager les vallées en fonction des cours d'eau. En somme, la terre et le ciel demeuraient leurs maîtres. Ils alignaient leurs travaux sur la matière première qui leur était donnée. Ils se gardaient bien de la violer, sachant par expérience que toute violation amène des catastrophes. Le résultat de cette sagesse obligée a été de conserver aux paysages l'essentiel de l'ordre et de l'harmonie qu'ils devaient aux grandes forces de la création. C'est à peine si les cultures, sacrifiant de larges pans de forêts, ont causé quelques touches baroques dans la grave diversité des plaines et des versants. Les villes et les villages eux-mêmes choisissaient leur place aux carrefours des voies de terre et d'eau."

"Les jardiniers s'en vont, avec les houes, les paniers, les faucilles et les chevaux paisibles, rejoindre les vieilles lunes tombées d'un ciel où se tracent, désormais, les routes futures. Des machines géantes tranchent directement dans le gâteau des collines et livrent à l'herbe les vieux chemins qui faisaient le tour sans se presser. Des cantons entiers retournent à la sauvagerie pendant qu'ailleurs d'énormes usines font place nette pour le béton, les poteaux électriques et les fumées jaunes. Des lacs apparaissent derrières les barrages imposés aux torrents, des lacs où tinteront bientôt, aux oreilles attentives de la légende, les cloches des villages engloutis. Le visage du monde se farde de plus en plus. L'artifice l'envahit à mesure que l'homme s'en rend maître. Faut-il le déplorer? Non, mais espérer qu'une nouvelle beauté surgira un jour de tout ceci, pourquoi pas! Et tirer le chapeau à l'ancienne paysannerie, les jardiniers du monde."

- Pierre-Jakez Helias, Le cheval d'orgueil - Mémoires d'un Breton du pays bigouden. 1975.

Foreword

"Science is like sex: sometimes something useful comes out,
but that is not the reason we are doing it"
- Richard Feynman

The HDR thesis is an important milestone in a French scientific career. It represents a good opportunity for me to reflect about the evolution of my research studies and my motivations to do science, going beyond the formal scientific considerations discussed in this thesis. This is the reason why this foreword aims to briefly express some thoughts about my research experience and put them into perspective with more personal considerations.

The first point I wish to mention refers to the reasons that led me (and other colleagues) to work on Amazon agricultural frontiers. There are many potential stock answers to such question. One may highlight the importance of the Amazon at global scale, considering biodiversity and climate regulation for example. Other may discuss the international geopolitics dimension of the Amazon agricultural frontier since the agricultural sector is supported by multinational companies that export a large part of the agricultural production to China and Europe. Another answer is to emphasize the importance of international cooperation in science, and especially with regard to Southern countries.

But, to be honest, we also (if not firstly) work in Brazil because we really appreciate this country and its people. In my case, my work in Mato Grosso (since 2005) results from successive opportunities and challenges that appeared during my career and which I benefited. I first graduated in engineering of industrial systems in 2004 and, for this reason, working as a geographer at CNRS was nothing straightforward. During my last year of engineering studies, I met three students from Brazil who encouraged me to do a 6-month internship at the Federal University of Pernambucano in Recife. At that time I used to work on human reliability in a research project for a regional power company. That first Brazilian experience in a research lab convinced me to carry out environmental scientific studies in Brazil. I then had the chance to meet with Vincent Dubreuil, professor at Rennes 2 university, who invited me to enter the Master's degree in geography and remote sensing and offered me the opportunity to discover another face of Brazil, far from my experience in *Nordeste* : the Mato Grosso pioneer frontier.

When I arrived in Rennes in 2005, I knew nothing about geography and even less about remote sensing. Yet, I patiently learned to process and analyze the MODIS Vegetation Index time series of soybean fields and finally got a PhD position at LETG. The PhD,

led in collaboration with Margareth Simões from Embrapa and State university of Rio de Janeiro, was a great experience to better discover Brazil. I bought a car in Cuiaba, ran thousands of kilometres, got it out of the mud with a *Havaianas* sandal, broke the engine and finally sold it in Sorriso in less than one hour. An interesting road-trip to really appreciate the meaning of two Brazilian sentences :

"Tu vai embora, toda vida..."
("You go away, all life long...". So it leaves time to reflect over the notion of distance.)

and then:

"Brazil is not for beginners!"
(An expression that I always heard pronounced in English, and that took sense when my mechanic was assassinated in front of his garage in Sorriso about 15 minutes before I arrived to request the quotation to repair my car.)

After the PhD, I moved to Montpellier to work for CIRAD with Agnès Bégué and for IRD with Laurent Durieux. Laurent contracted me to work on the use of ontologies in remote sensing applications. I had never heard about ontologies and it appeared to me as a tremendous challenge to take up. After a few years of struggle to understand ontologies and how they could advance remote sensing science, I nearly arrived to the conclusion that working on ontologies raised more questions than it solved... and this is actually quite positive since I think this may be the reason why I like science. In a way, it reminds me a famous quote from Richard Feynman (again!) about quantum mechanics and which I would nearly consider valid for ontologies, at least for geographic applications.

"If you think you understand quantum mechanics,
you don't understand quantum mechanics".
- Richard Feynman

Yet, although I still do not understand everything I am working on, at least my studies opened my mind about the importance of sharing experiences on the evolution of rural landscapes, not only in the Amazon but also here in France. As paradoxical as it may seem, I consider my researches in Brazil help me to better understand the evolution of agriculture in my home land, Brittany.

I am fully aware it may appear as a risky exercise to draw a parallel between Brazil and Brittany. To start, I may recall the importance of famous Bretons who contributed to

the Brazilian history. During the 19th century, Auguste Leverger (born in Saint Malo and also known as *"Bretão de Cuiaba"*) defended part of the Pantanal territory for Brazil during the war against Paraguay. He also mapped many river courses in the southern Mato Grosso, was appointed as governor of Mato Grosso and two municipalities were named in his honor: *Santo Antonio do Leverger* and *Barão de Melgaço*. A bit later, during the second half of the 19th century, Auguste Glaziou (born in Rospez, 7 kms from my home town!) was known as the *paisagista do imperador* in Rio de Janeiro and contributed to the discovering of flora biodiversity of the Brazilian Cerrado. His name is now associated to an important virtual Brazilian herbarium (http://glaziou.cria.org.br/project).

Yet, beyond these historical and coincidental considerations, I would prefer to emphasize how the evolution of agriculture in Brittany over the course of the twentieth century actually echoes the evolution of agriculture in the Amazon today (as mentioned in the text from Pier-Jakez Elias introduced at the very beginning of this document). Similar issues (e.g. need to feed the population with little concern for the environment) are repeated again and similar solutions are proposed (e.g. intensifying the production with more inputs and technological innovation). In some way, to speak with Mato Grosso farmers is like speaking with the older generation of Breton farmers. Both have been strongly supported by respective governments to produce food, encouraged to adopt intensive and specialized agricultural practices and even to open new areas in the case of Mato Grosso. Also both have experienced tremendous changes in their daily-existence, surviving very precarious conditions (related to the post-war period in Europe and to the settlement of a very unfriendly environment in the Amazon) before establishing a powerful and modern agricultural sector. And finally, both are currently hardly criticized by public authorities and the civil society because of the environmental impacts induced by the agricultural transition.

A widely agreed consensual discourse for Brazilian farmers to defend their position against charges consists in reminding us that we, Europeans, have already deforested our rural landscapes and so we should not criticize Brazilians for doing exactly the same (actually, they even relatively protect much more forests than we did). But it is worth noting that Mato Grosso farmers are, of course, Brazilians but they are also, for the majority of them, Italian or German descendants who still present a very European vision of agriculture, development and human-environment relationships. In this regard, the Amazon frontier can really be considered as a frontier of the western civilization.

Yet, I consider that what firstly characterizes Mato Grosso and Breton farmers is the sense of initiative and entrepreneurship. For example, both have understood the importance of transport and industrial infrastructures to support agriculture. In Mato

Grosso, farmers contribute to the expansion of the road networks in order to reduce the production costs exactly as Breton farmers avoided road pricing on major roads. Mato Grosso farmers are joining efforts to open a new railway (*a Ferrogrão*) across the Amazon to export soybean whereas Breton farmers joined their efforts to create a maritime company (Brittany Ferries) to export their products to Great Britain. In that sense, the willingness to support a continuous development for the benefit of humanity is fascinating and seems unlimited. And if you mix that determination with some strong religious beliefs about good and evil, you can even organize a parade to present orphan teenagers for adoption in a Cuiaba's mall (as it happened in 2019). And if encouraged by favourable national politics, that desire for growth can be fully realized and recalls the narration of Augusto Boal relating his personal experience in São Paulo prisons during the "Brazilian miracle" period of the early 1970s under the military government:

"- Everything is increasing," replied Copy Desk [a political prisoner in São Paulo's Tiradentes prison] nervously. "Last year, Brazil imported a billion dollars in oil; this year, the country will have to import more than three billion. Last year, it imported $500 million in wheat. This year it will import over a billion. I wonder when this Brazilian miracle is going to end..."
- "The miracle is not going to stop," interrupted Mr. Luis [military officer in charge of the prison]. "Even when it goes down, we can still sell the state oil monopoly... and the dollars will flow again. We still have a lot to sell... Brazil is very big and very rich... We still have a lot to sell"...
- "Except for dignity, which you no longer have..."

- Augusto Boal, Miracle au Brésil. 1979.

But of course there are also important differences between Breton and Mato Grosso farmers. The major one may refer to different relations of farmers with their land. In northern Mato Grosso, where oldest cities were created in the 1970s, few farmers were born in their farms whereas Breton farmers cultivate their land since various generations (e.g. my 17th century ancestors already lived in the same municipality where my parents were born). This difference always appeared to me as a curiosity since the first question to ask to a Mato Grosso farmer is : "where do you come from ?". This point has evident consequences for the human-environment relationships as well. As an example, a farmer from Nova Canãa do Norte (northern Mato Grosso) received us in his office with a large picture of a European "natural" agricultural landscape (likely from Toscana, Italy) on the wall whereas my grand-parents (and parents) display pictures of Brittany's littoral in the living room (... but I display posters of Amazon forests in my son's room). More importantly, whereas Bretons fought hardly to cancel a project of nuclear

power plant, people from Mato Grosso positively accepted the construction of recent hydropower plants (including one in Sinop built by a French electric company) that inundated thousands of hectares of tropical forests and agricultural lands along the *Teles Pires* river. This example at least questions the sense of belonging to a territory which actually may not be experienced in the same way we are used to conceive it in Europe. This reminds me the text from Gabriel Garcia Marques who discussed that issue as follows:

"Only when he began to take down the door of the room did Úrsula dare ask him what he was doing, and he answered with a certain bitterness.
"- Since no one wants to leave, we'll leave all by ourselves."
Úrsula did not become upset.
"- We will not leave," she said. "We will stay here, because we have had a son here."
"- We have still not had a death," he said. "A person does not belong to a place until there is someone dead under the ground."
Úrsula replied with a soft firmness:
"- If I have to die for the rest of you to stay here, I will die."

- Gabriel Garcia Marques, 100 ans de solitude, 1967

At the end, I would like to emphasize the importance of human encounters that underlie my work, a dimension that is too rarely considered in remote sensing studies. This is a pity since the numerous interviews I did with farmers also helped me to analyze the satellite images and understand the ongoing dynamics. For example, the study on the relationship between the adoption of double cropping systems and the duration of the rainy season initiated after discussions with farmers from Sorriso and Lucas do Rio Verde who had already noticed that the rainy season was longer there than in other agricultural regions in Mato Grosso.

The first point to recall about interviews is that farmers always remained respectful and often even cordial. This is for example the case with Dudy and Fernando Paiva who always received us warmly at their farm in Sorriso. I interviewed many farmers with potentially different convictions than mines and it always resulted in formidable moments for sharing experiences and better understand the point-of-view of people leaving on the frontier. This happened, for example, when a farmer considered that François Hollande's French government was communist or when another farmer explained us that he had left 500 ha of standing forest for his son to clear it when he is 18 years old. Actually, the only time I was dismissed by a farmer was in Tangara da Serra because I was wearing a t-shirt advertising for the coming indigenous games in the neighbouring city of Campo

Novo do Parecis. Thas was an oversight of mine but it illustrated a certain disregard for indigenous populations in Mato Grosso. In 2006, Blairo Maggi, at that time governor of Mato Grosso, wrote a letter to Brasilia in order to ask for the freezing of the creation of new indigenous lands in Mato Grosso with the argument that there is 1 *indio* for 650 ha which is much more than sufficient to sustain a family. Considering that Blairo Maggi is the CEO of the Amaggi group that cultivates over 150.000 ha of soybean, this testifies for the consideration of the agricultural sector for the indigenous cause in the Amazon.

Since the beginning of the colonization programs, indigenous populations have always been considered as a constraint on the development and were toughly excluded (either being displaced or even exterminated) and discredited. Still today, these populations remain on the fringes of the society, living in protected areas with limited access to education and health services for example. But things can change rapidly. As a Breton, I may consider myself as an indigenous (finally currently trying to learn the language of my grand-parents...) and it is worth noting that although young Bretons are now proud to claim their origins, things have not always been this way as testified by Jean Rohou:

"Until the 1970s, we [the Bretons] were a usual anti-model for everyone. Around 1972, I heard a Moroccan man on the radio being asked why he had returned to his country, instead of staying in France. "I was despised", he replied. "They treated me like an inferior, a loser, a Breton! At the same time, a car driver was convicted of calling a policeman a "Breton head"."

- Jean Rohou, Fils de ploucs, 2011.

As for Bretons who gained importance in politic, cultural and economic networks, may Amazon indigenous communities also soon gain in visibility (only 2 indigenous federal deputies since the 1980s!) and traditional environmental knowledge be better considered as part of the solution to ensure a sustainable development in the Amazon. But, of course this may sound as a pious hope at a time when the current Brazilian government claims its intentions to open indigenous lands to economic activities such as mining or large-scale agriculture.

Contents

Acknowledgements **ii**

1 Introduction **1**

2 A focus on tropical agricultural frontiers **5**
2.1 Definition and mapping of the pioneer frontier . . . 5
2.2 Public policies and the agricultural frontier in the Brazilian Amazon . . . 9
2.2.1 Ambivalent environmental and development policies on the agricultural frontier . . . 11
2.2.2 Beyond the frontier: the integration stage . . . 13
2.2.3 Back to the future? . . . 14
2.3 The Amazon frontier: from a regional to a global process . . . 14
2.4 Perspectives: Towards the oil palm agricultural frontiers in Indonesia . . . 16
2.4.1 Similarities between the Indonesia and Amazon frontiers . . . 17
2.4.1.1 Two very dynamic frontiers . . . 17
2.4.1.2 The result of political decisions . . . 17
2.4.1.3 The emergence of environmental considerations . . . 19
2.4.2 Differences between the Indonesia and Amazon frontiers . . . 20
2.4.2.1 Different geographic contexts . . . 20
2.4.2.2 Different agricultural systems . . . 20
2.4.2.3 The importance of smallholders plantations in Indonesia 21

3 Land use sustainability in the Brazilian Amazon pioneer frontier **23**
3.1 Proximate causes of environmental degradation . . . 23
3.1.1 Agricultural expansion and intensification . . . 24
3.1.2 Infrastructures: the proliferation of small farm dams . . . 26
3.1.3 Degradation of forest ecosystems . . . 28
3.2 Underlying causes of land use changes . . . 30
3.2.1 Testing the boom-and-bust hypothesis . . . 31
3.2.2 The Environmental Kuznets Curve . . . 34
3.2.3 What is to be expected in 2020? . . . 35
3.3 Perspectives: from pixel to landscape analysis . . . 37

4 Climate variability and climate change in the Amazon: implications for land use/cover dynamics **41**
4.1 Remote sensing-based rainfall estimates . . . 42

4.2 Climate change in the Amazon . . . 45
4.2.1 Trends in the rainy season metrics in the Southern Amazon . . . 45
4.2.2 Precipitation extremes . . . 49
4.3 Implications for agriculture . . . 52
4.4 Perceptions of Climate change and public policies in Mato Grosso . . . 56
4.5 Perspectives: climate-vegetation interactions in the Amazon forests . . . 58
4.5.1 Climate as a driver of phenology in Amazon forests . . . 59
4.5.2 Impacts of climate change on forest phenology/degradation . . . 62

5 Remote sensing: more than a data science 63
5.1 Overview of current research trends in remote sensing . . . 63
5.1.1 Data-driven remote sensing: The advent of Deep Learning? . . . 64
5.1.2 Knowledge-driven remote sensing: GEOBIA as a new paradigm? . . . 66
5.1.3 The challenge of processing big EO data . . . 67
5.1.4 Towards a citizen science for Earth Observation . . . 68
5.2 Ontologies in remote sensing: why, what and how? . . . 70
5.2.1 Why do we need ontologies? . . . 71
5.2.1.1 The complex definition of geographic concepts . . . 72
5.2.1.2 The concept duality of geographic concepts . . . 73
5.2.1.3 The vagueness and ambiguity of geographic concepts . . . 74
5.2.1.4 The sensory gap . . . 75
5.2.1.5 The semantic gap . . . 75
5.2.2 What are ontologies? . . . 76
5.2.2.1 A brief definition of ontologies . . . 76
5.2.2.2 Ontologies as ways of explicitly specifying bodies of knowledge . . . 77
5.2.2.3 Ontologies to manage and share knowledge about the interpretation of remote sensing images . . . 79
5.3 User-adaptive automatic classification of Sentinel-2 time series . . . 81
5.3.1 Ontological conceptual framework . . . 82
5.3.1.1 Purpose and existence . . . 82
5.3.1.2 Entities . . . 84
5.3.1.3 Characteristics: Observables and Similarities . . . 85
5.3.2 Application to Sentinel 2 time series . . . 86
5.3.2.1 System's architecture . . . 86
5.3.2.2 Study area and data . . . 87
5.3.2.3 Image processing . . . 87
5.3.2.4 Image classification . . . 91
5.4 Ontology-based classification of remote sensing images using spectral rules 101
5.4.1 Knowledge management: from conceptualization to image facts . . . 102
5.4.1.1 The reference conceptualization . . . 103
5.4.1.2 The contextual knowledge . . . 104
5.4.1.3 The image facts . . . 108
5.4.2 Implementation and results . . . 108
5.5 Perspectives: Ontologies to advance remote sensing science . . . 109
5.5.1 From pixel- to object-based image analysis . . . 109

5.5.2 Data-driven remote sensing: towards hybrid approaches 111
5.5.3 Connection with end-user expert knowledge 111
5.5.4 Ontologies and big EO data 112
5.5.5 Towards collaborative remote sensing 113

6 Conclusion **115**

Bibliography **119**

Chapter 1

Introduction

This thesis introduces my research work as a geographer interested in the monitoring and understanding of nature-society interactions in tropical regions. I am especially focused on the impacts of societies on the environment and natural resources and the adaptation of societies to environmental conditions, especially considering the climate change context. In order to provide answers to these complex interactions, I rely on remote sensing information to monitor both land use/land cover dynamics and environmental variables (mainly rainfall). Indeed, I consider Earth Observation data as a major source of geographical information with a huge potential for the production of large-scale spatio-temporal indices. Yet, my experience in remote sensing applications also led me to reflect on the evolution of remote sensing science, especially questioning the lack of knowledge sharing and accumulation due to the massive use of data-driven approaches.

I firstly developed my research experience studying the land use dynamics in the Southern Amazon agricultural frontier, especially in the Brazilian state of Mato Grosso (Figure 1.1). This region is characterized by intense and rapid changes led by populations with diverse (not to say opposite) cultures and interests (from indigenous populations to large-scale soy farmers and the international stakeholders behind them) associated to global environmental concerns (e.g. climate change, biodiversity loss). In this regard, Mato Grosso appears to me as a perfect open-air laboratory for geographers... and for many other scientists.

Although I here report my main personal scientific achievements, I wish to emphasize that the results and the thoughts expressed in this thesis ensue from collaborative researches with colleagues from various countries across the world (France, Europe, Brazil, US, China, Congo, etc). All these collaborations were carried out under the frame of research projects funded by French, European, Brazilian or Chinese agencies. In many

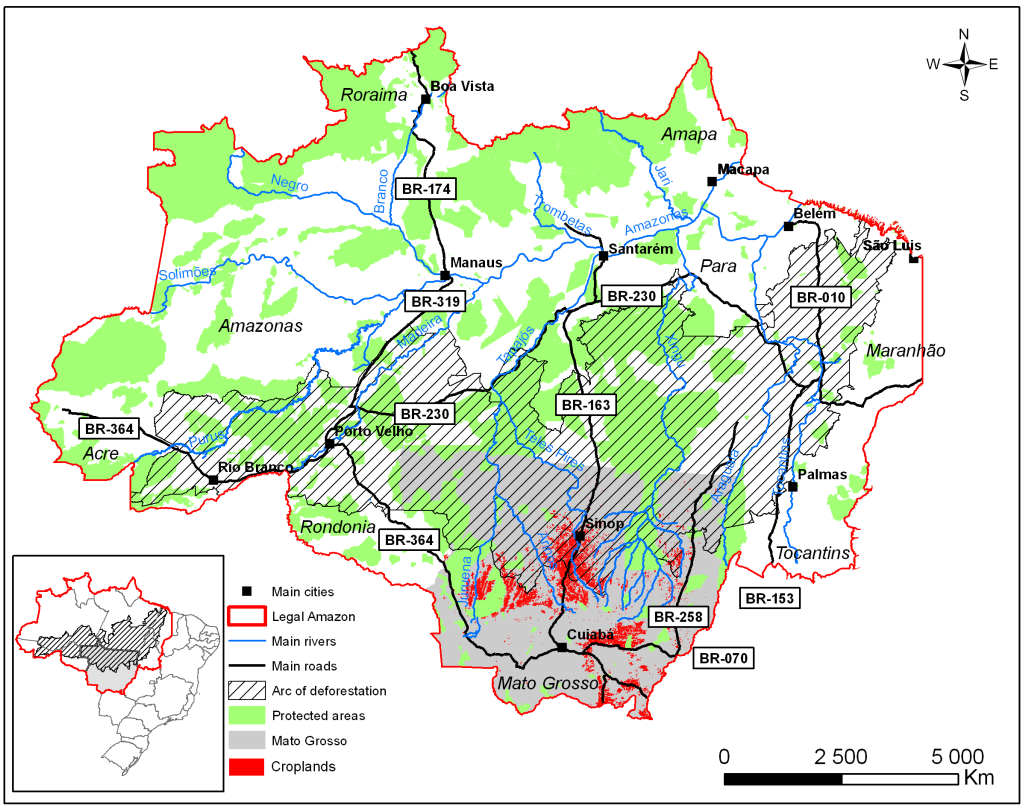


FIGURE 1.1: Location of Mato Grosso in the "Legal Amazon", showing the main roads, rivers, protected areas and the arc of deforestation according to the Brazilian Institute of Environment (Portuguese acronym IBAMA).

projects, I acted as a coordinator or as a Work Package Leader trying to foster scientific exchanges. I also encouraged scientific discussion as a coordinator of scientific teams on remote sensing and geomatics (in my research unit, UMR LETG) and on "ontologies for interdisciplinarity" (at GDR MAGIS). In addition, I initiated a straight collaboration with the State University of Mato Grosso (UNEMAT) in Sinop through a CNRS International Emerging Action to strengthen and formalize the scientific collaborations of the LETG research unit in the Amazon. Finally, I recently got engaged in the approximation of science and civil society through the publication of papers with non-academic partners (NGO, state administration) and the coordination of a Space Climate Observatory (SCO) project to ease the release of scientific results to the civil society.

This thesis is organized in four chapters that correspond to the four research axis I studied during the last fifteen years. These chapters are described below and aim to position my research in a more general scientific context, summarize my major scientific achievements and introduce my future research perspectives.

In chapter 2, I introduce the **tropical agricultural frontier** based on a conceptual model proposed by DeFries et al. [1] and explain how I applied it to the case of the Southern Amazon agricultural frontier (section 2.1). I also discuss the importance of public policies to understand the dynamics of the frontier (section 2.2) and its implications beyond the regional Amazon context (section 2.3). Finally, as a perspective, I express my intentions to analyze other agricultural frontiers in the world, especially the oil palm plantations in Indonesia (section 2.4).

In chapter 3, I discuss the evolution of **land use sustainability in the Amazon frontier**. To do so, I rely on the conceptual model proposed by Geist and Lambin [2] to characterize the direct and proximate causes of land use changes in the Amazon. I introduce how I used remote sensing images and techniques to monitor major land use changes in the Amazon (section 3.1) and understand their underlying causes (section 3.2). As a perspective, I consider I should now orient my research towards more integrated research relying on landscape analysis to assess the environmental impacts of the frontier progress on hydrosystems (section 3.3).

In chapter 4, I explain how I used remote sensing data to analyse **the climate variability and climate change in the Amazon**. I firstly present the quality of remote sensing-based rainfall estimates (section 4.1) and introduce some results on climate change expectations int he Southern Amazon (section 4.2). Then, I discuss the potential impacts of climate variability and climate change on agricultural activities (section 4.3). In this regard, the vulnerability of the agricultural sector should then depend on the severity of climate change but also on the adaptation capacity of societies (section 4.4). Finally, I suggest to combine my experiences in remote sensing-based monitoring of vegetation and rainfall to better analyze the climate-vegetation interactions in the Amazon, with a focus on the impact of climate change on forest phenology (section 4.5).

In chapter 5, I focus on the methodological dimension of my researches to expose my vision on the (recent) past and future evolution of **remote sensing science, which I consider to be much more than a data science** (section 5.1). Indeed, whereas remote sensing science tends to be firstly seen as an image processing science, depending on image characteristics (spectral, spatial and temporal resolution) and advanced techniques (e.g. supervised machine learning classifiers), I wish to emphasize that remote sensing is firstly a science of interpretation where a special attention should be paid to the motivations and intentions of the users, which depend on their knowledge, skills and experiences. In this regard, a major focus should given to the management of remote sensing knowledge to ensure knowledge formalization and reuse. This is why I present my expectations on the use of ontologies as a knowledge management technique that should be further studied to advance remote sensing science. In this regard, I introduce

what are ontologies (section 5.2), two application examples (sections 5.3 and 5.4) and my expectations on ontologies to advance remote sensing science (section 5.5).

Chapter 2

A focus on tropical agricultural frontiers

My studies focus on the monitoring and understanding of land use/cover dynamics in tropical agricultural frontiers. In this chapter, I briefly define the concept of frontier in geography and explain how I mapped its evolution in the southern Amazon (section 2.1). I then linked this model to the evolution of public policies and suggested a new model frontier (section 2.2). Going further, I explain how the Southern Amazon agricultural frontier is actually driven by national and global economic and political interests and, thus, should not be the sole concern of Brazilian authorities (section 2.3). Finally, I introduce some new personal research perspectives, especially to explore land use dynamics in other agricultural frontiers such as in Indonesia (section 2.4).

2.1 Definition and mapping of the pioneer frontier

The pioneer frontier is a well-studied geographical concept, which was first described by Turner [3] who analysed the American frontier in the United States of America. The frontier was defined as a moving limit separating inhabited areas from areas to be settled. In the Brazilian context, Monbeig [4] defined the frontier as a piece of space being incorporated to the ecumene. It is a rural area in contact with wild areas [5], which are subject to drastic transformations (often non-reversible) of the natural space [6]. The incorporation of space is made dynamically, so that it has a spatial depth [7] and cannot be simply thought as a continuous line [8]. Besides, Monbeig [4] preferred talking about a pioneer fringe instead of a pioneer frontier. Moreover, the dynamic is not constant, neither in space, nor in time. It is made of breaks and jumps that express the variability of the progress of the demographic and economic occupation of the territory

[9]. Thus, understanding the pioneer frontier implies analysing societies that make it up [8]. The frontier is a territory where conflicts between actors and social groups prevail and result in a spatial occupation marked by a mosaic of landscapes [7].

In my studies, I was inspired by the work from DeFries et al. [1] who suggested to rely on landscape analysis to monitor the evolution of the pioneer frontier. The landscape-frontier relationship stands on the hypothesis that, within a particular region, land use potentially follows a series of stages, which are intricately related to both economic development and the ecological characteristics of landscapes. Five successive stages of the frontier progress were defined based on different landscape characteristics:

- The **pre-settlement stage** is the initial stage, at which the main land use is represented by wildlands.
- The **occupation stage** begins with the arrival of the first settlers who start clearing wildlands, usually for timber or subsistence agriculture.
- The **consolidation stage** is characterized by a large part of the population dedicated to food production for local consumption.
- The **intensifying stage** starts when small-scale farms are replaced by large-scale farms where commodities are cultivated in order to be sold on national or international markets.
- The **intensive stage** is characterized by a large proportion of the agricultural area being cultivated by intensive practices. In addition, the development of a profitable agriculture attracts new migrants that support a rapid urbanization process. Also, new considerations about the environment and the quality of life appear and result in the emergence of a new model of environmental governance [10], characterized by the creation of protected areas or the reforestation of degraded areas (see the concept of forest transition [11]).

These stages describe the process to convert wildlands into intensive agricultural areas and appear to be valid for the specific case of Mato Grosso, with two main particularities: (1) the magnitude of unprecedented environmental damages caused by the occupation, consolidation and intensification stages and (2) the rapidity of land use conversions and its associated development, since few agricultural regions reached the final intensive stage in less than 40 years, while other areas still remain unoccupied today.

In order to illustrate these two particularities, I published a paper to map the evolution of the pioneer frontier in Mato Grosso [12]. To do do, I defined rules to assign a stage

of the frontier to each rural area. These rules are used to classify rural areas based on the percentage of landscape represented by main land use types per unit area. The final map includes eight stages (instead of the five initial stages) because the occupation, consolidation and intensifying stages were split into croplands and non-croplands classes (croplands here only consider large-scale cropped areas). Indeed, for each stage, I wished to consider if the driver for change was large-scale agriculture or not. For instance, reaching the occupation stage implies that wildlands have been cleared, but it does not give any indication on the motivations for change, i.e. extensive croplands, pasture or small-scale agriculture. The eight evolution classes of the rule-based model are defined in Fig. 2.1:

I applied these rules on the following input land use maps considering unit areas of 0.25*0.25° (about 25 kms):

- the MapBiomas [13] land use maps are provided annually since 1985 based on the supervised classification (Random Forest) of Landsat time series (30 m spatial resolution) on Google Earth Engine. This product was used to compute the proportion of natural vegetation (forest or cerrado), pasture and croplands per unit area.
- the MODIS land use maps produced by Picoli et al. [14] and updated from INPE. It consists in annual land use maps at 250 m spatial resolution from 2000 to 2017. This product was used to compute the proportion of cropland being cultivated in double cropping system (i.e. soybean followed by maize, cotton).

It is worth noting that since MODIS data are available after 2000, the intensive stage could only be mapped after that date.

The evolution of the frontier between 1985 and 2017 is illustrated in Figure 2.2. This figure confirms and completes the original results achieved by comparing the situation in 2001 and 2007 [12]. Indeed, it shows that, in 1985, a major part of the state was at a pre-settlement or occupation stage, confirming that Mato Grosso is a recently colonized state in a process of incorporation to the national space. The southern regions of Mato Grosso are the most affected by deforestation. This is justified by the historical occupation of the state that first began with the colonization of the cerrado biome, located in the southern part of the state and easier to clear than forest. However, the frontier is moving up north into the Amazon forest. Especially, the northern central and northeastern regions have been dramatically affected by deforestation. The most deforested areas are located along the main roads planned by the federal governments (the BR-163 and BR-158 roads in the centre and eastern parts of Mato Grosso, respectively). Creating such roads in wild

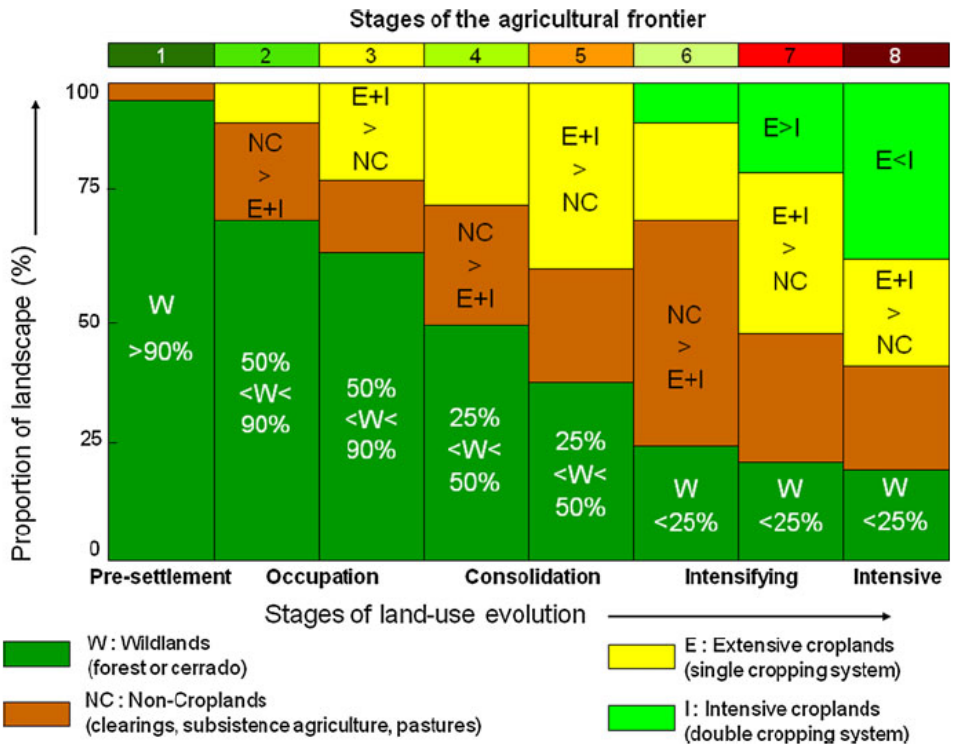


FIGURE 2.1: Rules to adapt the frontier model for mapping the agricultural frontier based on remote sensing data.

areas induced a decrease in transport costs that incentivized deforestation [15]. These roads have led to the appearance of cleared corridors which are now connecting to form super-glades [16]. The main remnants of wildlands areas in 2017 are now located in protected areas (i.e. either conservation units or indigenous lands) in the eastern and central regions of Mato Grosso (Xingu indigenous land) and in the pantanal biome (which is also affected by human impacts).

In addition, the maps also show that the main agricultural regions first appeared in (1) the central region, along the BR-163, (2) the western region of the Chapada dos Parecis and (3) the southeastern area, around Campo Verde and Rondonopolis. The evolution shows that the two main regions (i.e. the central region and the western region) are getting closer and merging rapidly to form a "soybean belt" at the interface between the Amazon and Cerrado biomes. These areas are now also connecting with a fourth agricultural region that appeared in the eastern part of the state between 2001 and 2005. This area then increased constantly, raising important issues for the protection of the Xingu indigenous land that is nearly entirely surrounded by crops in 2017.

Finally, the intensification process first started in the southeastern and central regions at the beginning of the 2000s. In 2017, double cropping systems were applied in the three

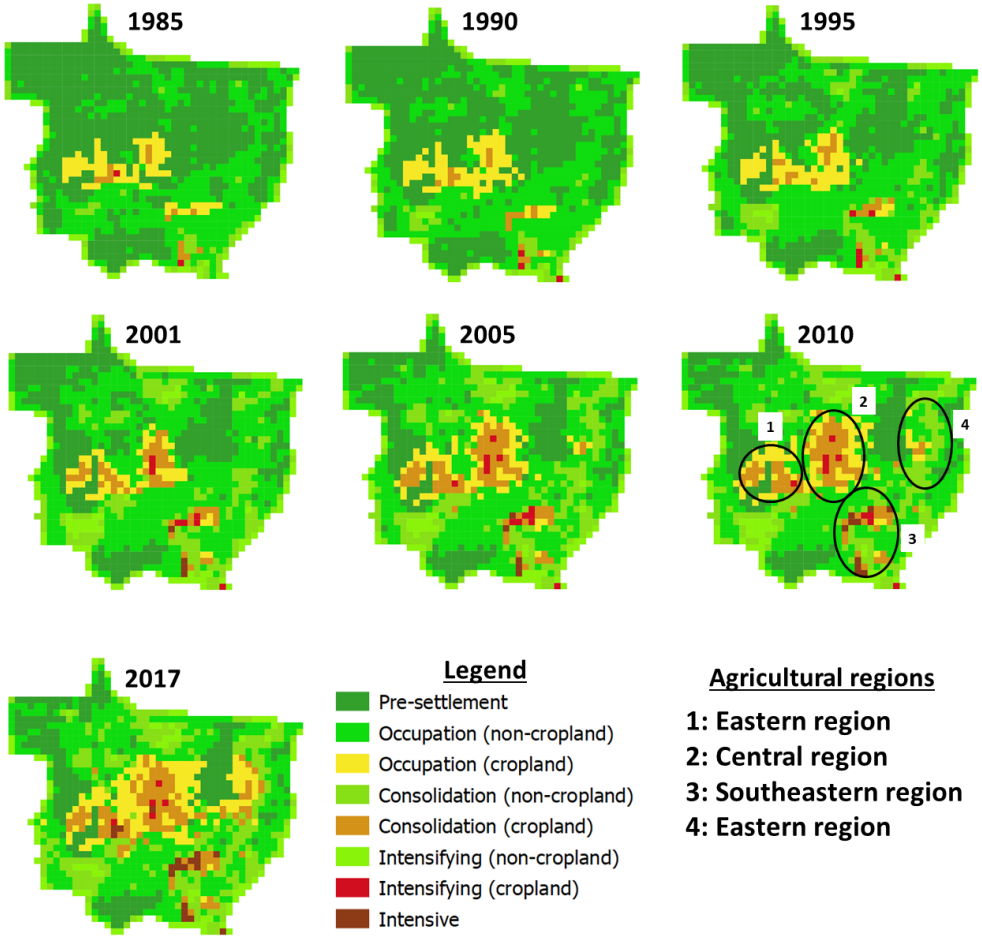


FIGURE 2.2: Map of the evolution of the soybean agricultural frontier in Mato Grosso between 1985 and 2017 according to the land-use transition stages.

main agricultural regions (central, western and southeastern regions). Interestingly, the agricultural increase in the southeastern region mainly relied on the adoption of intensive practices, whereas agricultural areas did not expand so much. On the opposite, the agricultural evolution in the central and western areas was characterized by both rapid expansion and intensification, thus questioning the efficiency of intensification to contain deforestation. To finalize, double cropping systems are not generalized in the eastern region.

2.2 Public policies and the agricultural frontier in the Brazilian Amazon

The rapid progress of the agricultural frontier in the Amazon is the result of many successive government programs to occupy this distant and widely unexplored region. In

order to understand the underlying reasons that drive the frontier expansion, I explored how the historic public policies shaped the present Amazon landscapes and I suggested to improve the frontier model considering recent efforts to take environmental concerns into account [17] (Figure 2.3).

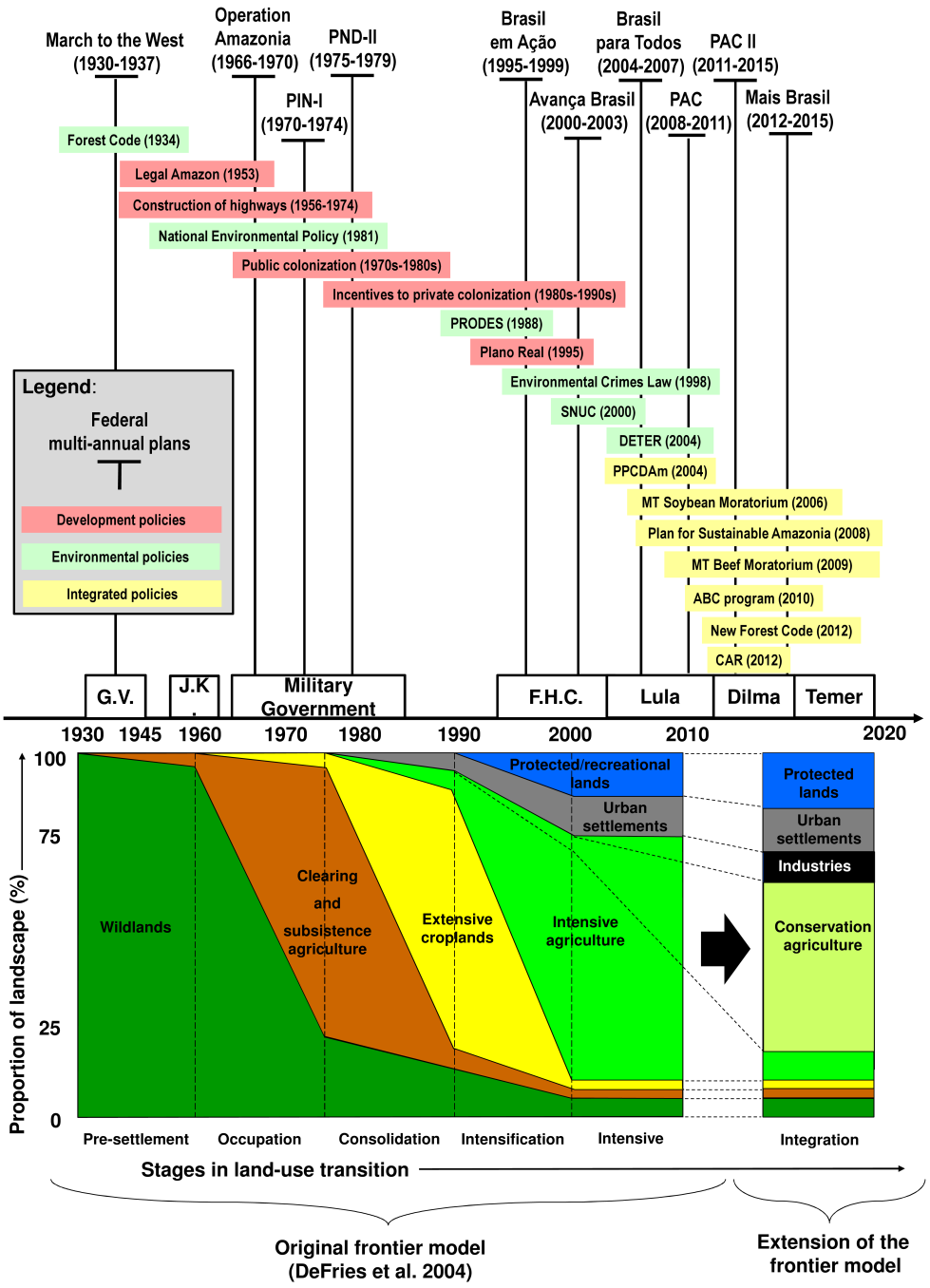


FIGURE 2.3: Conceptual model of the agricultural frontier and links with main development, environmental and integrated policies established in the Amazon and Mato Grosso since the 1930s (G.V. Getulio Vargas, J.K. Juscelino Kubitschek, F.H.C. Fernando Henrique Cardoso, Dilma Dilma Rousseff, Temer Michel Temer) [17]

2.2.1 Ambivalent environmental and development policies on the agricultural frontier

The analysis of the past evolution of the agricultural frontier in the southern Brazilian Amazon emphasizes a long history of ambivalent development and environmental policies [17]. Although the Amazon was inhabited long before the arrival of the first European settlers, we may consider that the process of spatial occupation of Mato Grosso really began in the 1930s with the first governmental colonization programs in the Amazon, e.g. the "March towards the West" (*Marcha para o Oeste*) program. These programs were based on a development-oriented (*desenvolvimentista*) model and intented 1) to ensure the national sovereignty on the Brazilian territory, and 2) to support the United States policy of helping Brazil to expand rubber production in the Amazon during the Second World War [5]. After the Second World War, the new democratic Brazilian government created the "Legal Amazon" in 1953 and financial incentives were offered to attract people. At the same period, President Kubitschek (1956-1961) reinforced the occupation process building major infrastructures such as the new Brazilian capital, Brasilia (1960), and main highways (e.g BR-010 from Brasilia to Belem and the BR-363 from Cuiaba to Porto Velho). Later in the 1970s, this effort was pursued with the construction of new trans-Amazonian roads (especially the BR-163 road linking Cuiabá to Santarém initiated in 1973) to encourage small farmers to migrate to the Amazon [18].

Still in the 1970s, the government decided to open the Legal Amazon to private capital through the Program for Agricultural, Livestock and Mineral poles in Amazonia (Portuguese name *Polamazonia*), which emphasized the development of large-scale export-oriented development poles dedicated to mining, livestock and agriculture [19]. The occupation of Mato Grosso was consolidated through the implementation of 88 isolated and spatially-distributed colonization projects from 33 private colonization companies, confirming that the progress of the pioneer frontier is made of "breaks and jumps". Powerful mechanized agriculture was encouraged by policies (with limited taxes on land and agricultural income and facilitated access to credit) to attract farmers from southern Brazil and with sufficient capital to invest in the agricultural potential of the large plains of Mato Grosso [20]. In 1994, the government introduced the Plano Real, which reduced inflation and attracted multinational companies to promote the cultivation of crop commodities (soybean, corn, cotton) in Mato Grosso. These incentives to promote large-scale agriculture were effective but also encouraged widespread deforestation since efforts to establish efficient environmental policies were of little importance compared with economic priorities. Whereas environmental issues were discussed at the Rio Conference (1992), the unstability of environmental federal and state institutions in

the 1980s-1990s hindered the implementation of an efficient environmental governance model.

As cropped areas expanded, yields also increased rapidly thanks to agricultural intensification. Since the 1980s, the government encouraged the Brazilian Agricultural Research Corporation (Portuguese acronym EMBRAPA) to develop new soybean varieties adapted to tropical conditions [21]. The intensification process continued in the 2000s with the adoption of double cropping systems (soybean followed by corn or cotton) promoted by national and regional research institutes. In addition, investments in huge infrastructure projects under the frame of multi-annual federal plans aimed at reducing the cost of transport and energy (the recent asphalting of the BR-163 in the state of Para is one such strategy).

While the main political preoccupation was focused on the construction of infrastructures to export agricultural commodities, environmental protection efforts were limited to a few international and federal programs (e.g. PPG7 Pilot Program for the Conservation of Brazilian Rainforests). The Lula's government then adopted the "Plan for Sustainable Amazonia" (PAS - Plano Amazonia Sustentavel) as a main project of the Environment Ministry, prioritizing transversal efforts across the Amazon. However, in Mato Grosso, the intensification process remained a priority, and the implementation of the environmental plan consequently increased tensions between producers and environmentalists. As a result, the environmental initiatives made no significant progress and deforestation and agricultural intensification continued to increase until the state of Mato Grosso integrated the PAS in 2006.

Finally, although agricultural expansion and intensification are still increasing in Mato Grosso, recent decades have also been characterized by urbanization and the creation of new protected areas, i.e. indicators of the final intensive stage of the frontier model. The urban population has increased dramatically since the 1970s (from 239,524 inhabitants in 1980 to 2,482,801 inhabitants in 2010) whereas protected areas increased after the creation of the National Conservation Unit system (Portuguese acronym SNUC) (Law 9.985, 2000). There are two statuses for legally protected areas in Amazonia: conservation units protected for environmental reasons (including areas under integral protection and areas for sustainable use) and areas protected for social reasons (indigenous lands, traditional populations, and quilombolas). It is noteworthy that both types of protected areas have proved to be efficient to contain deforestation.

2.2.2 Beyond the frontier: the integration stage

The Amazon frontier, characterized by expansion and intensification processes,; remains very dynamic to date. Yet, the intensive stage should not be looked on as the last stage of the frontier model. I thus proposed to extend it by adding an integration stage (Figure 2.3) marked by important efforts to turn development and environmental policies complementary and promote the emergence of a new agricultural model based on the adoption of conservation agriculture practices and industrialization [17]. Three major initiatives illustrate that integration stage.

First, NGOs, producers, private companies, and state authorities agreed to establish the Soy and Beef moratorium (signed in 2006 and 2009, respectively), a market-oriented policy to limit the impact of soybean and pasture expansion on deforestation by preventing traders from financing and trading soybean and beef originating from recently cleared areas.

Second, command-and-control policies were reinforced through the Action Plan for The Prevention and Control of Deforestation in the Legal Amazon (2004), which invested in deforestation control systems (e.g. PRODES project implemented by the National Institute for Spatial Research, INPE) and promoted laws for taxing illegal clearings. Third, the adoption of more "environmentally-friendly" agricultural practices are now encouraged to maintain crop production while limiting or reducing impacts on natural ecosystems. In this regard, the ABC plan ("Plano para uma Agricultura de Baixo Carbono", i.e. low carbon agriculture) has been implemented since 2010 by the Ministry of Agriculture as part of the national policy on climate change with the objective to encourage a few cropping practices such as:

- **No-tillage practices** consist in leaving the soil profile undisturbed, sowing successive crops in between the stubble of the previous one. Thus, no-tillage is usually accompanied by the use of a second crop (usually maize, cotton, sorghum or millet) sown after the soybean harvest in order to improve the soil's quality by limiting the loss of chemical products (e.g. improving nitrogen fixation) and organic matter via erosion and by retaining water for a longer period.

- **Integrated systems (IS)** have been encouraged due to their environmental benefits. Four types of IS are usually defined: crop-livestock systems, livestock-forestry systems, crop-forestry systems and crop-livestock-forestry systems. According to [22], IS already covered about 500,000 ha in Mato Grosso, among which crop-livestock systems are largely dominant, accounting for 89% of the state's IS area.

- **Restoration of degraded pastures**. Pastures currently represent the main land use in the Amazon and Cerrado biomes [23], of which about 50% in the Cerrado biome and more than 60% in the Amazon biome are degraded pastures [24]. Integrated systems based on an interannual rotation between pasture and crop are also part of a process to restore degraded pastures.

2.2.3 Back to the future?

Although the frontier model appears to be linear, it is noteworthy that the frontier does not change constantly, neither in space nor in time. One should thus keep in mind that: (1) land use transitions between the stages of the pioneer frontier are not unidirectional, e.g. cultivated areas can be abandoned and regenerate as secondary forests (see the forest transition concept [25–27] or intensive practices may be abandoned due to financial restrictions; (2) not all the stages of the pioneer frontier may affect a given territory, e.g. one would not expect protected areas to ever be cleared; (3) the duration of each stage may vary depending on the location and the period, e.g. the entire process of transition from wildlands to intensive agriculture took thousands of years in Europe, whereas it is expected to last only a few decades in recently colonized areas such as Amazonia.

These considerations are of particular importance in the current Brazilian context. Indeed, most advances (e.g. efficient market-oriented and command-and-control policies) were noticed during the economic boom of the 2000s, when part of the agricultural incomes was reinvested in environmental initiatives. Yet, the actual economic and political crisis, that culminated when Jair Bolsonaro was elected President supported by the agricultural sector, questions the sincere intentions of the agricultural sector to evolve towards increased land use sustainability [28, 29]. Instead, development policies with little consideration for environmental concerns tend to be prioritized as it was the case in the 1970s. The INPE, responsible for monitoring deforestation in the Amazon, has been severely called into question. In this regard, the recent high deforestation and fire rates observed in the Amazon confirm the occupation stage is still ongoing while the increased pressure on protected areas bears witness of a somehow retrogression in the evolution of the agricultural frontier in the Amazon.

2.3 The Amazon frontier: from a regional to a global process

Beyond the regional evolution of the Southern Amazon agricultural frontier, I also wished to contextualize its evolution with regard to national and global-scale issues. To do so,

I discussed [30] the fact that the efforts to reduce the impacts of agriculture on deforestation in the Southern Amazon since the mid 2000s may potentially be correlated to negative indirect land use changes, i.e. changes that arise from displacement effects when one land use (e.g. soybeans) encroaches on another land use (e.g. pasture), causing the impacted land use to encroach on wildlands [31]. In this regard, land use sustainability is closely related to the spatial scale of analysis since "reduced land degradation in one place is also to some extent the result of a displacement of food production and wood extraction to other places through trade" [32].

Recent studies have analyzed the processes of indirect land use changes in the Brazilian Amazon. Richards et al. [33] explained how land use changes in the Amazon resulted from complex interactions occurring in other Brazilian regions and Walker [34] also suggested that deforestation in the Amazon could be related with the forest transition observed in the Brazilian Atlantic Forest. Upscaling the interactions, crop expansion (and consequent deforestation) in the Amazon can be related to global drivers such as international migrations (e.g. European settlers who moved to southern Brazil from the late 1800s to early 1900s) (Figure 2.4), the RS$-US$ exchange rates [35], the US policy on corn-based biofuels [36], the European policy for Bovine Spongiform Encephalopathy (BSE) prevention [37], and the increasing world consumption of beef, especially in Asia [38].

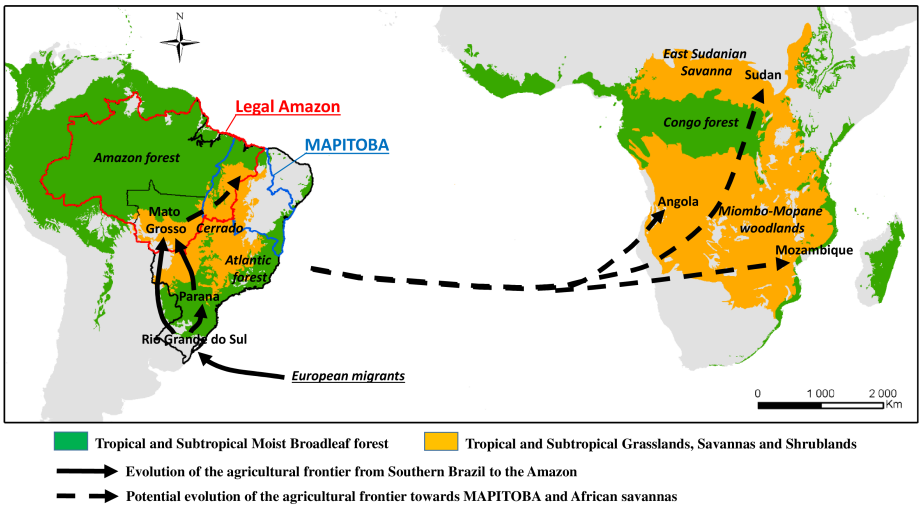


FIGURE 2.4: Evolution of the agricultural frontier from Southern Brazil to the Amazon (Mato Grosso state) and then potentially to MATOPIBA Cerrado areas and African savannas [30].

Yet, few studies question the potential changes that could derive from a slowdown of deforestation in the Amazon. The case of the Brazilian Cerrado biome is especially interesting because it experienced the greatest increase in new cropland in Latin America during the last decade (i.e. a 9.96 Mha increase between 2001 and 2013 [39]. Whereas first waves of soy expansion in Brazil started in the Southern states of Rio Grande do Sul and Parana in the 1960s and 1970s before reaching Mato Grosso in the 1980s, the last frontier of soy expansion (started in the 2000s) appears to be located in the Cerrado states of Maranhão, Tocantins, Piaui and Bahia, also known as Matopiba.

At international scale, the expansion of the Brazilian agricultural frontier to neighboring countries in Latin America (e.g. Bolivia) has already been well documented [40]. Few studies also evidenced the emergence of a soybean frontier in Africa related to the South American agricultural sector [41]. Although soybean expansion there remains low (750,000 ha in 2013 across Southern Africa as compared to the 620,000 ha of soybean cultivated in the sole municipality of Sorriso, Mato Grosso), these authors consider that there were around 247 Mha of unprotected areas (croplands, forests and shrubs) which are highly suitable for soybean cultivation [41]. A few private Brazilian agribusiness actors, among others, have already invested in Africa (e.g. Angola, Ethiopia, Tanzania, Mozambique and Ghana) to benefit from low production costs, the proximity to international export markets such as China and access to new local markets. However, the interest of the Brazilian agricultural sector for Africa is mainly the result of geopolitical concerns. Indeed, the agricultural expansion in Africa is mainly supported by knowledge transfer and cooperation under the frame of South-South telecouplings [41]. A major symbol of such strategy is the Program for the Development of Agriculture in the Tropical Savannas (ProSavana) in Mozambique launched by Brazil, Mozambique and Japan in 2009. This program can be considered as a sister program to the PRODECER (Program for the Development of Cerrados, already supported by Japan) that encouraged agricultural expansion in the Brazilian Cerrados in the late 1980s [42, 43]. Considering that the Brazil agricultural boom of the 1990s-2000s is primarily the result of the clearance of large areas of tropical wildlands, increasing investments to develop the African agricultural sector following the Brazilian model also raise important environmental questions.

2.4 Perspectives: Towards the oil palm agricultural frontiers in Indonesia

Theoretically, the model of the frontier's dynamic can be applied to any location in the world. In the future, I intend to discuss its adaptation to other places and especially in

Indonesia, which represents another major hotspot of tropical deforestation related to the rapid expansion of a very intense agricultural frontier [44]. This will be done through the co-supervision (with Samuel Corgne from LETG and Jérémie Gignoux from INRAE-PSE) of a PhD student (Carl Bethuel) who initiated his research on the monitoring of oil palm plantations in Indonesia in January 2022, under the frame of the ANR PALMEXPAND project. In order to introduce this new scientific issue, I discuss here preliminary thoughts on the interesting similarities and differences between the Indonesia and Amazon agricultural frontiers.

2.4.1 Similarities between the Indonesia and Amazon frontiers

2.4.1.1 Two very dynamic frontiers

As for the soybean cultivation in Brazil, oil palm plantations have been expanding rapidly to reach about 15 Mha in 2020 (Figure 2.5). Yet, recent estimates from Gaveau et al. [45] suggest that the total planted area may actually approximates 16.24 Mha. Such expansion has led to dramatic impacts on forest cover and industrial plantations were associated with three times as much forest conversion as smallholder plantings. As for soybean, the oil price is considered as a major (economic) underlying cause of deforestation [45].

In addition, both soybean and palm oil are major commodities whose expansion relies on interesting advantages. On the one hand, soybean is the most efficient way to produce vegetable protein required to feed livestock worldwide. On the other hand, the expansion of oil palm plantations is due to its extraordinary potential to supply the increasing demand for vegetable oil at global scale. Indeed, oil palm productivity is extremely efficient since one hectare of oil palm plantation can produce 4 tons of oil annually, i.e. at least four times the yield of 1 hectare of rapeseed, sunflowers, or soybeans planted in temperate regions [46].

2.4.1.2 The result of political decisions

Similarly to the soybean expansion in Brazil, oil palm plantations have overall been supported by politic decisions (figure 2.6). The first plantations were cultivated in the second half of the 19th century in Dutch colonies during the colonial era. In the the 1930s, the planted area reached 91,000 hectares, nearly equivalent to the rubber plantations. After the national independence in 1945, the "old order" period was marked by an anti-capitalist regime that did not consider oil palm plantations as a priority. For this reason, the expansion thus effectively began after 1967, when the recently elected

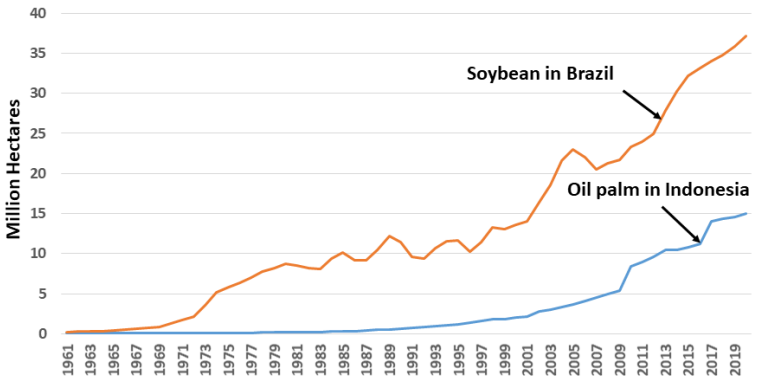


FIGURE 2.5: Comparative expansion of soybean and oil palm areas in Brazil and Indonesia, respectively (source: https://www.fao.org/faostat/fr/#data/QCL)

president Suharto installed an authoritarian regime to develop the country economically. As for the private-public colonization projects implemented in the Amazon, transmigration programs were established in Indonesia to create oil palm plantations managed by public companies (called Perusahaan Negara Perkebunan, PNP). After the 1970s, these companies were then transformed into semi-public enterprises (called Perseroan Terbatas Perkebunan, PTP) and the private sector also developed through the emergence of companies (called Perkebunan Besar Swasta, PBS) relying on foreign investments.

But the expansion really began after the 1998 economic crisis, during the "reformasi era". At that time, the government changed its industrial strategy from heavy intervention to more indirect action focused on liberalization and democratization to gain the confidence of international institutions. In 1999, the government promulgated two laws (No. 22 and No. 25) aimed at decentralizing powers to regional authorities (districts and municipalities) in order to ease local management of natural resources. Yet, the lack of governmental investments to develop infrastructure programs forced many smallholders to sell their land due to lack of market opportunities. This is the reason why the government then established the "Plantation Revitalization Program" to stimulate partnerships between the government, smallholders and companies.

The importance of infrastructures to support the economic development in Indonesia inevitably echoes the Brazilian situation where the construction of roads and harbours played a seminal role in the colonization of the Amazon. In this regard, it is also worth mentioning the present plan to move the Indonesian capital from Jakarta in Java island to a new city named Nusantara in East Kalimantan on Borneo island. Although it happens more than sixty years after the creation of Brasilia as the new Brazilian capital in 1960, it symbolizes the importance of national infrastructure programs to support the process of land occupation and its related (at least expected) economic development.

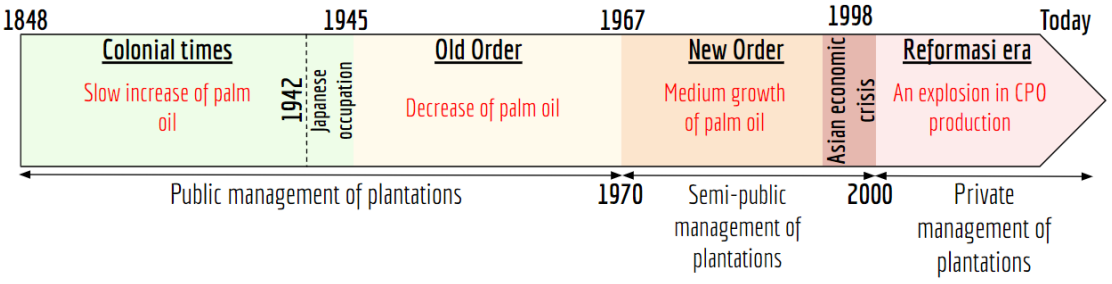


FIGURE 2.6: History of palm oil expansion (by Carl Bethuel, unpublished)

2.4.1.3 The emergence of environmental considerations

Beyond the impacts of both soybean and oil palm expansions on deforestation, the use of very intensive agricultural practices relying on large amounts of agrochemicals have also been criticized. In this regard, the efforts of the Brazilian government (through the ABC plan) and the private sector (through market-oriented policies such as the Soy and Beef moratorium) have already been introduced in this thesis (section 2.2). In Indonesia, the same issues are being raised and similar solutions are being discussed Rochmyaningsih [46]:

- **Moratorium on deforestation :** In 2011, the Indonesian government issued a 2-years moratorium to stop new permits for deforestation, especially in the Papua province. This moratorium has then been renewed but its effects have been limited because of weak command-and-control policies.
- **Expansion on degraded lands only :** In Malaysia, plantations are allowed to expand only onto existing agricultural fields or degraded lands (but the situation is different in Indonesia) [47]. In Brazil, the expansion of soybean areas is also expected to occur on degraded lands.
- **RSPO Roundtable on Sustainable Palm Oil :** Today, about 19% of the oil palm production is certified as "sustainable" by the Roundtable on Sustainable Palm Oil, which includes thousands of growers, traders, manufacturing companies and NGOs. As for the RTRS (Round Table on Responsible Soy), the certification is based on environmental and social considerations.
- **Landscape preservation and restoration :** Since 2015, it is forbidden to plant oil palms in 50-meterwide river banks in order to restore riparian vegetation. In Brazil, these areas are also considered as Permanent Preservation Areas (PPA) by the Forestry Code.

- **Towards conservation agriculture :** New cultivation practices are promoted to ensure minimal use of agrochemical products and support the growth of understory vegetation in oil palm plantations. Such practices can be related with the generalization of no tillage practices in Brazilian soybean fields to ensure soil and water quality.

Finally, it is worth mentioning that these efforts may be criticized because they focus on large companies whereas smallholder farmers who own 40% of Indonesia's oil palm plantations may be less informed about environmental preservation and oil palm management [46]. I also raised such issue in the Amazon context when I criticized the fact that the efforts to reduce deforestation during the 2000s were focused on containing deforestation in large soybean farms whereas little solution was proposed to family agriculture [48].

2.4.2 Differences between the Indonesia and Amazon frontiers

2.4.2.1 Different geographic contexts

Although the Brazilian spatial organization has been compared to an archipelago made of "isolated" settlements (today very large urban areas) with different populations and economic activities [49], Indonesia's geography corresponds to an archipelago in its original meaning. It consists of over 17,000 islands, including the major ones: Sumatra, Sulawesi, Java and parts of Borneo and New Guinea. Such configuration assumes a wide environmental, social, cultural, economic diversity as it also represents a challenge to maintain national territorial unity. Yet, from a frontier viewpoint, it exacerbates the fact that the land occupation is not necessarily spatially continuous but can be made of jumps.

2.4.2.2 Different agricultural systems

Still from a frontier viewpoint, the cultivation of tree plantations at the heart of the agricultural frontier also implies a major difference when compared to annual crops such as soybean. Indeed, oil palm trees are planted for long periods since the productivity starts declining after 25 years. As a consequence, the frontier is unlikely to regress as it happened in the Amazon after the 2004-2005 economic crisis. This is confirmed in figure 2.5 where the evolution of oil palm areas in Indonesia appears much more linear than the soybean evolution in Brazil.

In addition, fresh palm fruit bunches need to be processed in crude oil palm within 24 hours after the harvest. This fact raises important logistical considerations since it implies that fruits are delivered at the mills very rapidly. Overall, it means that, unlike the soybean agricultural frontier, the occupation and intensive stages (including industrialization through mills and roads constructions) needs to be initiated simultaneously to ensure the local transformation of raw production.

2.4.2.3 The importance of smallholders plantations in Indonesia

Another major difference between both frontiers lies in the profile of populations involved in the agricultural activities. In the Brazilian Amazon, the soybean cultivation is made in large farms (usually larger than 1,000 ha) owned by moderately capitalized pioneer families who moved to the Amazon in the 1980s or 1990s. For example, this was the case of Andre Maggi who moved to Mato Grosso in 1977, participated to the foundation of the municipality of Sapezal and then founded the Amaggi company, today considered as the biggest soybean company in Brazil and whose CEO Blairo Maggi used to hold important political responsibilities in Brazil (i.e. Mato Grosso state governor and Minister of agriculture). In Indonesia, the expansion of oil palm plantations requires huge investments and thus depends on highly capitalized national or foreign companies. Yet, besides these industrial plantations, about 40% of the oil palm production is ensured by smallholders who benefit from the installation of mills. In addition, the Indonesian government improved the participation of smallholders in the oilseed sector through the Nucleus Estate and Smallholders (NES) program. These programs promoted the integration of smallholders into industrial plantations through the allocation of land for oil palm cultivation. Thus, industrial plantations are divided in a nucleus (i.e. the core industrial plantation) and the "plasma" located at the periphery of the nucleus. These "plasma" plantations are supervised by the neighbouring company but the lands (generally 2 ha) are entrusted to small planters, mainly transmigrants, within the framework of various partnership contracts. It results in a diversity of planters (i.e. industrial, smallholder and plasma) who share common interests on the development of the agricultural sector whereas large-scale soybean farmers and small-scale family farmers often hold diverging opinions on the agricultural development in the Amazon.

Chapter 3

Land use sustainability in the Brazilian Amazon pioneer frontier

The rapid progress of the agricultural frontier in the state of Mato Grosso raised major concerns with regard to environmental conservation [50–52] and increased considerations for an integrated governance model based on complementary socioeconomic and environmental policies. Such evolution questions the land use sustainability in the Southern Amazon [53]. Yet, assessing land use sustainability of society-environment interactions implies not only to consider (1) how socio-economic activities affect natural ecosystems, i.e. which driving forces for which environmental changes, but also (2) how these changes impact on society and (3) how society can cope with them [54]. In the following sections, I summarize some of my researches on these different questions.

To do so, I adapted the conceptual framework proposed by Geist and Lambin [2] to describe the proximate and underlying causes of deforestation to the case of degradation of natural resources (figure 3.1). I then introduce spatio-temporal indicators I produced to monitor the evolution of proximate causes of land use changes in Mato Grosso (section 3.1) and I discuss underlying drivers of land use changes with a focus on economic factors (section 3.2). Finally, as a future research perspective, I emphasize the importance of considering the landscape scale in studies about land use sustainability (section 3.3).

3.1 Proximate causes of environmental degradation

The figure 3.1 presents three proximate causes of deforestation, i.e. agricultural expansion, infrastructures (e.g. construction of water reservoirs) and timber logging. Here I introduce some of my studies on these different issues.

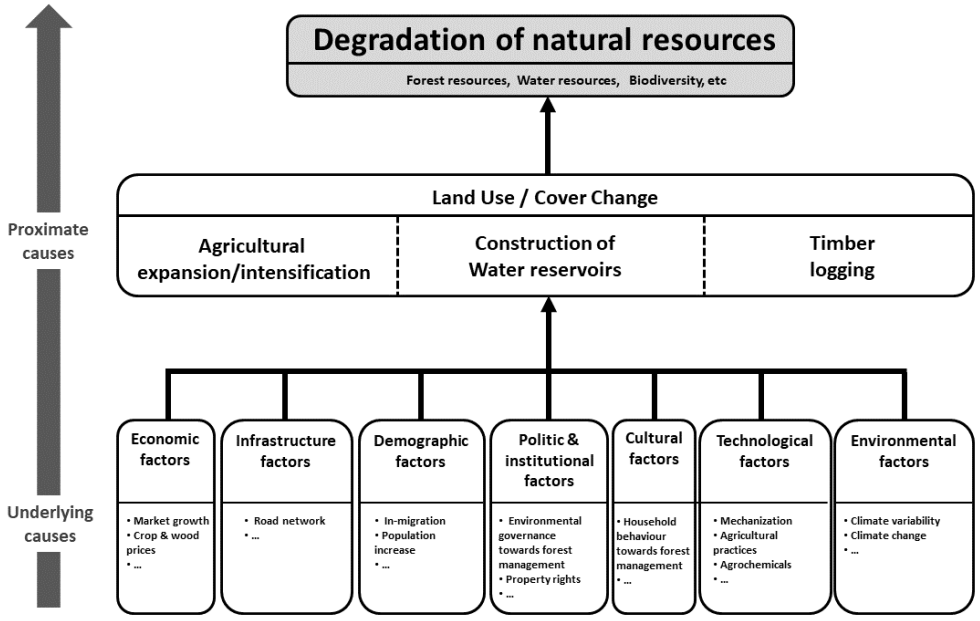


FIGURE 3.1: Drivers of degradation of natural resources (inspired from Geist and Lambin [2]).

3.1.1 Agricultural expansion and intensification

My researches firstly focused on the monitoring of agricultural dynamics in Mato Grosso. During my PhD, I paid a special attention to the characterization of the agricultural transition process, i.e. the evolution of the agricultural model from expansion towards intensification. To this end, I analyzed time series of MODIS EVI satellite images between 2001 and 2007 to map agricultural areas and then define the cropping systems, i.e. crop types and the annual crop frequency. At that time, I evidenced a rapid expansion (+43%) of cultivated areas, to reach a maximum area of 61,140 km^2 in 2005-2006. In addition, I estimated that 12.6% of the deforested area from 2001 to 2004 in Mato Grosso was directly converted into cropland, i.e. within two years after deforestation. Finally, the proportion of the cultivated area sown with double-cropping systems (considering only commercial crops such as soybean maize and cotton) increased from 6% in 2000-2001 to 30% in 2006-2007.

Although these results are outdated today, it is worth noting that many colleagues have then continued to monitor cropped areas in Mato Grosso based on MODIS time series. For example, Kastens et al. [55] estimated that the proportion of double cropping systems increased from 9.3% to 62.3% between 2001 and 2014 (Figure 3.2). More importantly, they point at the importance of the Soy Moratorium (i.e. a market-oriented

policy to contain deforestation due to crop expansion) to decouple the process of agricultural expansion from deforestation. More recently, I contributed to the work of Picoli et al. [14] who also used MODIS VI time series to map crops in Mato Grosso. Their results confirmed the estimates from Kastens et al. [55] since the proportion of double cropping systems was estimated to be 62.5% in 2016. More recently, colleagues from the State University of Mato Grosso (UNEMAT, Prof. Carlos Antonio da Silva Junio) also launched a web platform to monitor the expansion of soybean areas in Mato Grosso based on MODIS data (https://pesquisa.unemat.br/gaaf/plataformas/spaceagro).

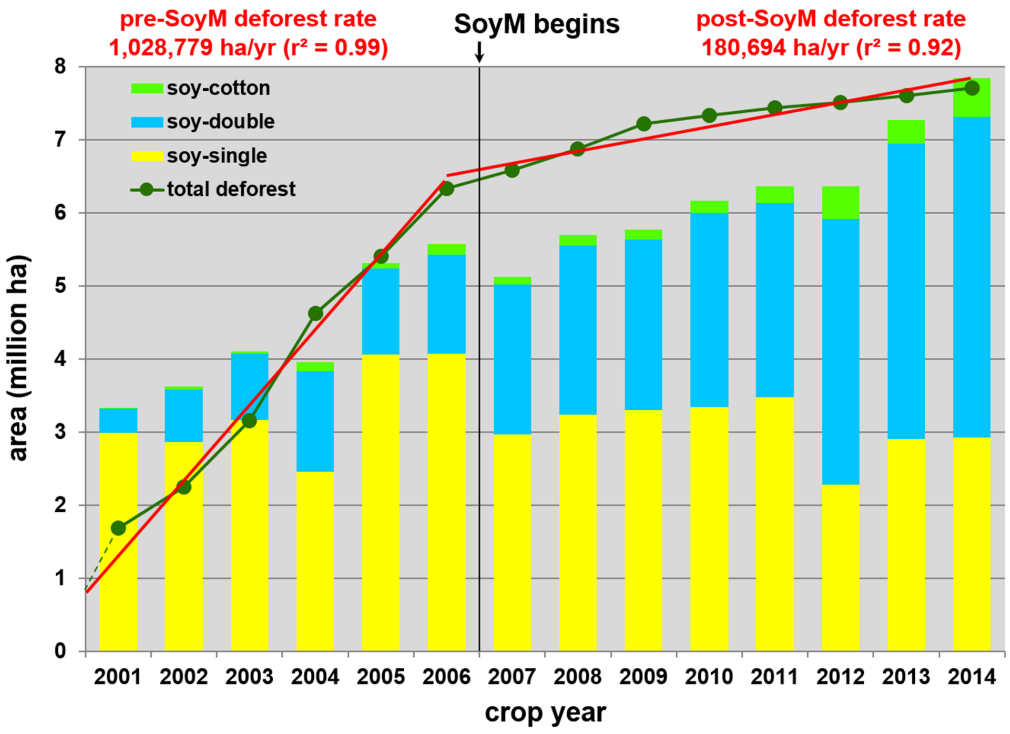


FIGURE 3.2: Evolution of agricultural areas and main copping systems in Mato Grosso [55].

The generalization of double cropping systems is often considered as a relevant strategy to contain deforestation because it implies higher incomes on reduced areas. Yet, it also raises new important environmental issues that may partially counterbalance the benefits from deforestation reduction. First, the production of soybean, maize and cotton requires large amounts of agrochemical products that may affect Amazonian biodiversity [56]. Second, many pivot irrigation systems were installed and numerous artificial ponds were created on river networks to ensure water supply for the second crop (maize or cotton), which is harvested at the beginning of the dry season (June and July).

In this regard, in 2017, I joined a collaborative research project led by Thomas Corpetti from CNRS (UMR LETG) and Chinese colleagues from The Chinese Academy of

Sciences with a huge experience in remote sensing image processing. We jointly implemented a method to detect and delineate center pivot irrigation systems in Sentinel-2 images (Figure 3.3) [57]. In 2017, our estimate of irrigated areas in Mato Grosso achieved 70,654 ha and the radii of the detected center pivot irrigation systems ranged from 31 to 870 meters. These results lend insights into the farmers' adaptation strategies to climate change. Indeed, while most soy producers still consider local climate conditions (i.e. a long and regular rainy season) as a major asset for soybean cultivation in the Southern Amazon and deny climate change, analyzing the medium to long term spatio-temporal proliferation of center pivot irrigation systems may provide very important information on the concrete actions carried out by farmers to mitigate the expected impacts of climate change on agricultural production [58, 59].

Finally, as part of the Brazil-French GeoABC program, I worked with colleagues from Embrapa and CIRAD on the implementation of new methods to map integrated crop-livestock systems. These practices are promoted by the Brazilian government through the Low Carbon Agriculture Plan (ABC Plan) launched in 2010 in order to limit the environmental impacts of the agricultural intensification process. This plan includes six programs to support the adoption of environmentally friendly practices including no-tillage practices, nitrogen fixation, the restoration of degraded pastures and integrated systems. In this context, I especially accompanied the PhD thesis from Patrick Kuchler to map integrated systems in Mato Grosso based on MODIS time series [60]. Four types of IS are usually defined: crop-livestock systems, livestock-forestry systems, crop-forestry systems and crop-livestock-forestry systems. According to [22], IS already covered about 500,000 ha in Mato Grosso, among which crop-livestock systems were largely dominant, accounting for 89% of the state's IS area. Integrated systems based on an interannual rotation between pasture and crop are also part of a process to restore degraded pastures. The results from [61] indicated that areas cultivated in integrated systems increased from 18.6% (1.1 Mha) to 28.9% (2.6 Mha) of cutivated areas between 2012 and 2019.

3.1.2 Infrastructures: the proliferation of small farm dams

The last decade testified of a very dynamic agricultural frontier characterized by the rapid adoption of new agricultural practices such as double cropping, irrigation and integrated systems. Yet, beyond the positive expected impacts of such strategies on deforestation, intensifying and diversifying the production also implies new pressures on natural resources. The generalization of intensive double-cropping systems requiring high amounts of agrochemicals, the adoption of irrigation systems to prevent harvest losses for the second crop and the development of fish farming raise important concerns about the future of soil and water resources in the region [62]. From that point-of-view,

FIGURE 3.3: Automatic detection of center pivot irrigation systems on Sentinel 2 images [57].

the price to pay for controlling deforestation may actually be the degradation of other natural resources such as water resources. In this regard, I paid a special attention on the proliferation of small artificial water reservoirs to ensure crop irrigation, water livestock and farm fishes.

Scientific knowledge about the impacts of water reservoirs is partial since most studies

to date focus on major dams, especially in tropical regions such as the Amazon where the construction of numerous hydropower dams has drawn the attention of the scientific community [63, 64]. Conversely, small reservoirs have received less attention despite their multiplication can generate significant cumulative effects on chemical fluxes (e.g. [65]), fish communities (e.g. [66]) and hydrological cycles (e.g. [67]). In this specific regard, a high number of dams, although small when considered individually, can have similar or even greater impacts than one single large dam [68–70]. In this regard, mapping and counting the small artificial reservoirs at regional scale is of extreme importance since the spatial distribution of the water bodies (i.e. density, position in the basin, etc.) is crucial to assess their cumulated effects [71].

I implemented a fully automatic methodology to map water bodies based on Landsat data and monitor the expansion of small water reservoirs in Mato Grosso. This method is very similar to the one described in section 5.3 but using Sentinel 2 time series. As a case study, I monitored the rapid expansion of water bodies during the last thirty years in the municipality of Sorriso, known as the main soybean producer at national scale and which I study since my PhD [72]. The cumulated area increased from 153 hectares (mean value between the low and high estimates) in 1985 to 1707 hectares in 2015, i.e. the area of a medium-size farm in the region. The total area thus increased more than tenfold in thirty years. The results showed an especially rapid growth after 2000 corresponding to the period of agricultural intensification in the region [73], which was partially based on irrigation thus requiring increased access to water resources. It is worth noting that the continuous increase in water areas after 2000 is opposite to the decrease in deforestation rates observed after the mid-2000s.

The number of water bodies also increased rapidly during the study period, from 86 to 522 (mean values of 1985 and 2015). In addition, this increase accelerated (+50%, from 347 to 522, mean values) between 2010 and 2015 whereas the cumulated area only increased moderately (+14%, from 1494 to 1707 ha). This result may be explained by the recent development of numerous small excavated dams dedicated to fish farming, which could not be accurately delineated at Landsat scale but whose detection may be used as a proxy to observe the process of agricultural diversification in the southern Amazon.

3.1.3 Degradation of forest ecosystems

As the decrease in deforestation rates may be (partially) counterbalanced by the degradation of water resources, one should also notice that forest degradation represents a

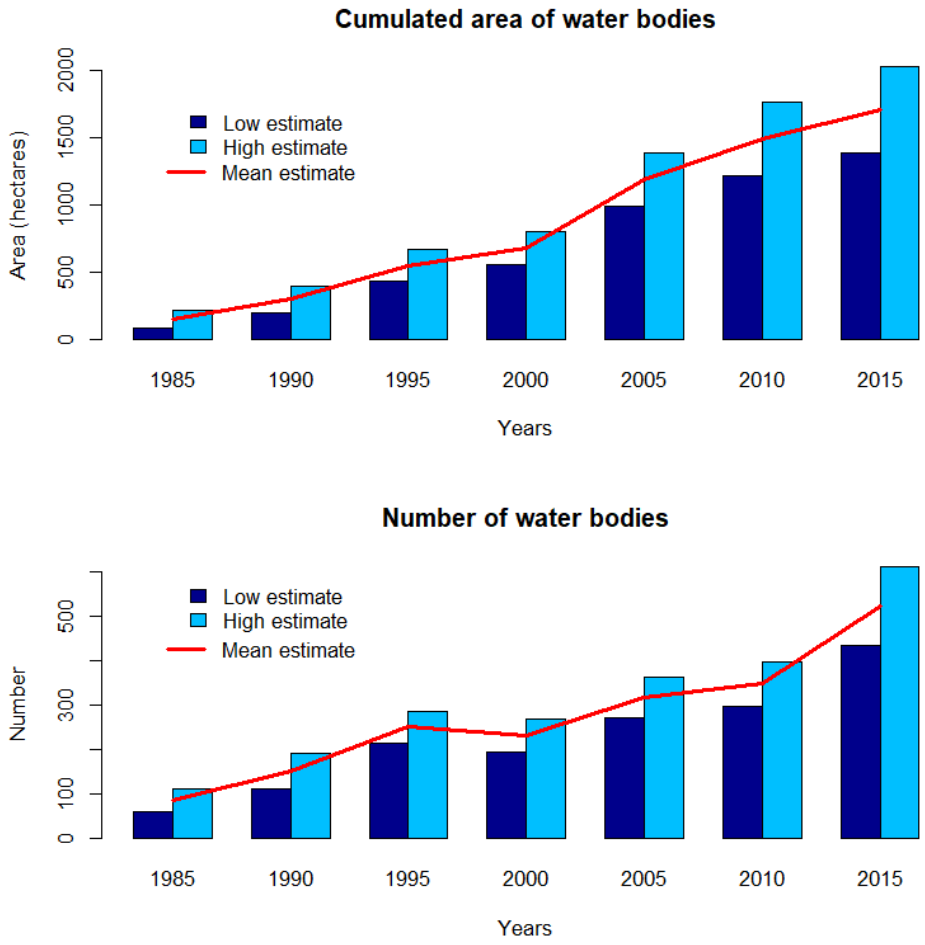


FIGURE 3.4: Temporal evolution of the cumulated area and number of water bodies in Sorriso between 1985 and 2015.

major environmental issue. Forest degradation denotes damages in structure, composition and function with no change in land use [74]. Forest degradation in the world accounts for 68.9% of overall forest carbon losses [75] and affects about 60% of tropical forests. It can be directly related to a loss of ecosystem services [44], leading for example to a loss of biodiversity and carbon, increased soil erosion, or reduced water quality. Degraded forests are characterized by a partial destruction of structure resulting from cumulative disturbances caused by timber extraction (logging activities) and/or fire.

In collaboration with colleagues from CIRAD (i.e. Julie Betbeder, Lilian Blanc, Valéry Gond, Clément Bourgoin and Renan Le Roux), I worked on the implementation of an automatic method based on optical remote sensing data to identify and characterize forest degradation. This was a very challenging task since the partial opening of the canopy is a fleeting phenomenon that is difficult to capture by satellite imagery. Indeed, the vegetation recolonizes these openings more or less quickly according to the soil condition and the surrounding vegetation. We applied the method on an optical time

series of Landsat 5-TM, Landsat 7-ETM+ and Landsat 8-OLI images (30 m resolution) from 2000 to 2017 in the municipality of Paragominas (20,000 km^2, 3 Landsat scenes, in Para state), a regional center for timber production. The 17-year period contained in average 12 surface reflectance images per scene (an average of one image every 1.5 years).

The figure 3.5 shows the detection of logging and fire disturbances in the CIKEL forestry company. The results are in accordance with the dates of forest logging disturbances in 17 forest management units (each one including about 50 plots) provided by the CIKEL company. Then, we derived temporal indicators and landscape metrics from patches of forest disturbances in order to characterize the type of forest disturbance, i.e. FSC (Forest Stewardship Council), conventional and illegal logging. The temporal indicators referred to:

- Cumulative number of disturbances : sum of "logged area" and "sum of burned areas" in a time series;
- Frequency of disturbances: (year of the last disturbance – year of the first disturbance) / sum of disturbances.

Since a conventional or a FSC forest plot in the Brazilian Amazon cannot be logged more than once every 30 years, we reclassified each pixel of the "frequency of disturbances" map in two classes : 1- "=<1 disturbance" and 2- "> 1 disturbances". Finally, for each class of "frequency of disturbances", we computed the "Intensity of disturbance", i.e. the proportion of logged pixels (%).

In the CIKEL forestry company (Figure 3.6), the comparison of the different indicators evidenced the positive effect of the certified forest management techniques. Over the 2000-2017 period, these indicators showed that FSC forests have been disturbed only once with low disturbance intensity (between 7 and 22%, orange areas in figure 3.6) whereas illegal logging areas showed higher frequency of disturbances (average 2.5), a sum of impacts greater than 1 and a higher intensity (>75%, red zones in figure 3.6).

3.2 Underlying causes of land use changes

In order to deepen the land use change analysis, I also studied the underlying causes of land use changes as expressed in figure 3.1. For example, I correlated the evolution of environmental public policies (i.e. politic and institutional factors) with deforestation (see 2.2). I also analysed the relationship between soybean prices driven by international

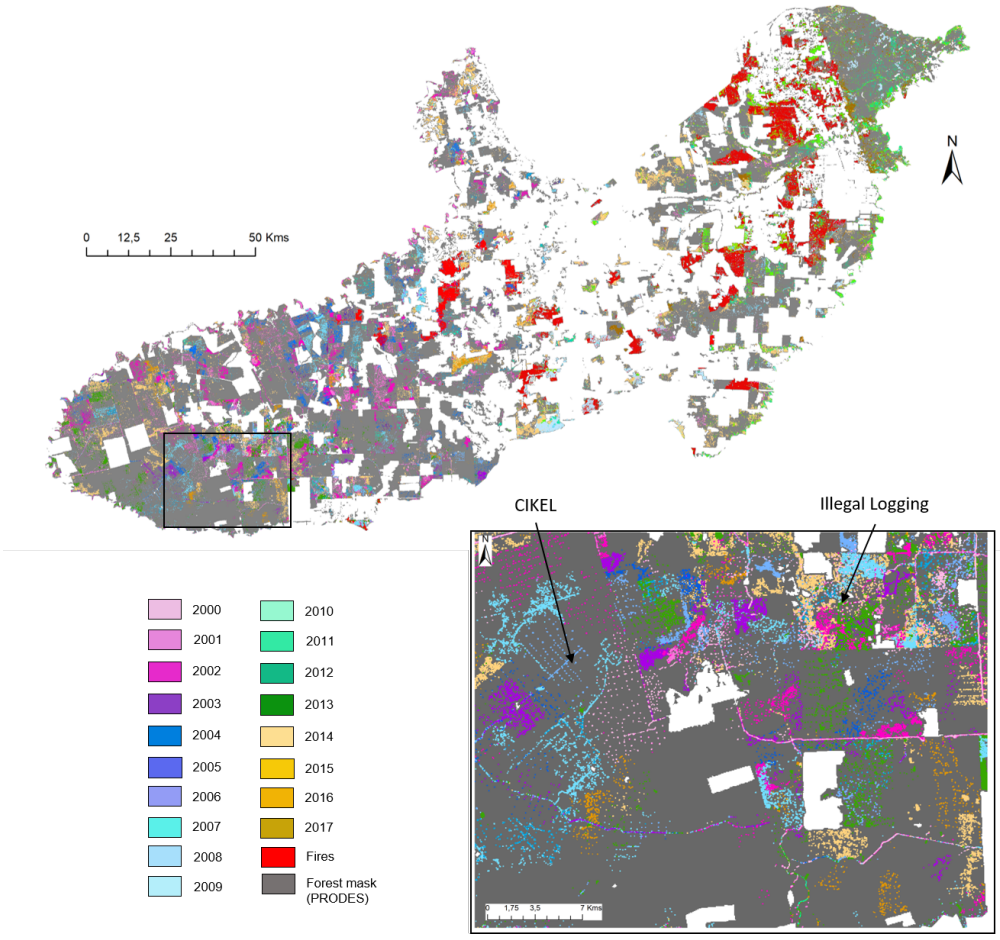


FIGURE 3.5: Detection of annual logging and fire disturbances in Paragominas over the 2000-2017 period. Zoom on the CIKEL Company (western part of the zoom) and on an area of illegal logging (northeastern part of the zoom).

exchange rates (i.e. economic factors) and agricultural expansion [72]. Finally, I studied the spatial variability of the rainy season in Mato Grosso (i.e. Environmental factors) to explain the spatial distribution of double-cropping intensive systems (see section 4.3).

In this section, I introduce a study published with Isabelle Tritsch (from CIRAD) to assess the relationship between socio-economic development in the Amazon and deforestation. This research emerged after seminal studies evidenced a boom-and-bust development pattern in the Brazilian Amazon in the early 2000s [76, 77]. We first tested if the hypothesis was still valid in 2010 and suggested the emergence of a new development pattern named "Environmental Kuznets Curve" (EKC).

3.2.1 Testing the boom-and-bust hypothesis

The boom-and-bust hypothesis states that timber extraction and the conversion of forests into cropland and pasture generate a boom in incomes in the first years of land

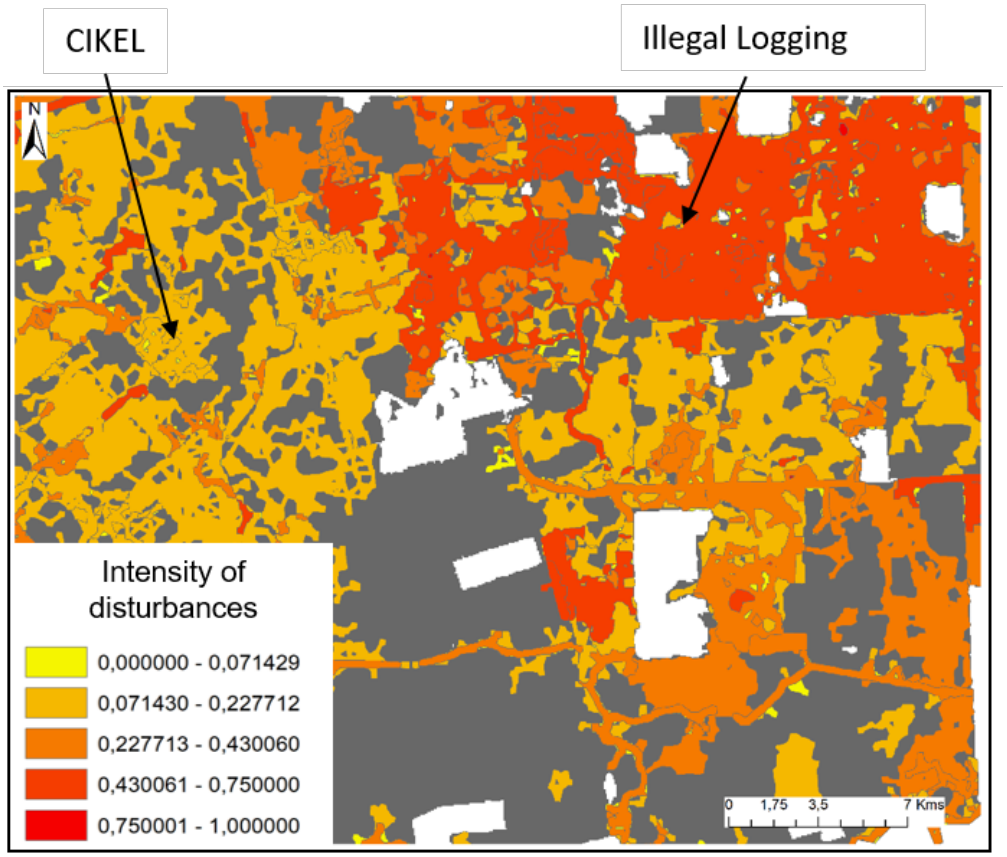


FIGURE 3.6: Intensity of disturbances over the 2000-2017 period: Proportion of logged pixels (%) per "disturbance area".

occupation. This boom is then followed by a collapse as forest resources and soil fertility are depleted [78]. This scenario recalls the historical pattern of economic development in the Amazon when the cycles of extraction of forest products used to follow a boom-and-bust pattern, characterized either by a rapid rise and decline in market prices or by the depletion of resource stocks. Associated with high deforestation rates over the past four decades, the boom-and-bust pattern may be considered as a lose-lose scenario from a long term perspective since it generates high environmental costs and limited socio-economic benefits [79].

In 2009, Rodrigues et al. [76] crossed socio-economic indices (human development index, standard of living, literacy and life expectancy) and remote sensing-based deforestation rates at municipality level to emphasize the existence of a boom-and-bust development pattern in the Amazon. They showed that these indices firstly increase as deforestation begins but then decline as the frontier evolves, so that pre- and post-frontier levels of human development are similarly low. We pursued that seminal work to check the evolution of the boom-and-bust development pattern in the 2000s. To do so, we crossed socio-economic data from full-population censuses in 2000 and 2010 provided by the Brazilian Institute of Geography and Statistics (IBGE) at census sector scale with the

PRODES deforestation maps provided by the Brazilian National Institute for Space Research (INPE). Both datasets were projected on a 10x10 km regular grid [80] and the cells were classified in seven frontier classes (A to G) based on both the extent of deforestation and deforestation activity. The classes ranged from pre-frontier classes (A, B) characterized by a limited area of deforestation and low deforestation activity, through active frontier classes (C, D, E) with increasing deforestation extent due to high deforestation activity, up to post-frontier classes (F, G) where the extent of deforestation is high but deforestation activity is low.

Our results (Figure 3.7) showed that all socio-economic indices (i.e. income, education and sanitation) increased significantly at each frontier stage between 2000 and 2010. Yet, the development pattern changed dramatically. In 2000, a boom-and-bust trend could be observed for all variables since the post-frontier class (G) evidenced a lower level of socio-economic development than the active frontier classes (C, D, E) and a similar level to that of the B pre-frontier class. Conversely, in 2010, the post-frontier areas (F, G) had better levels of socio-economic development than the pre-frontier (A, B) and active frontier (C, D, E) areas. It indicated that the last decade was marked by the end of the boom-and-bust socio-economic development pattern in the Brazilian Amazon.

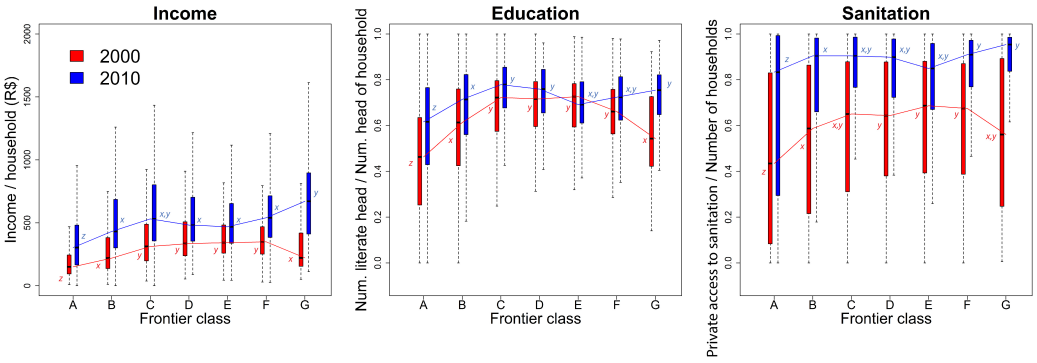


FIGURE 3.7: Variation in income per household, literacy and access to basic sanitation services in 2000 (in red) and 2010 (in blue) across the frontier classes. Dots indicate median values and bars indicate first and third quartiles. Dashed lines represent extreme values. All variables present significant variation across the frontier classes (Kruskal–Wallis:P¡ 0.0001); classes that do not have a letter (u, v, w, x, y, z) in common differ significantly (Tukey's honestly significant difference test: P¡ 0.05).

We explained these results by a spatial bias in the bust observed in 2000. Indeed, 44% of post-frontier cells in 2000 were spatially clustered in the state of Maranhão, which had a significantly low level of development and thus pulled the average socio-economic indices of post-frontier areas downwards. In contrast, ten years later, post-frontier areas expanded to other states, mainly Rondônia (11% of total G cells), Mato

Grosso (22%) and Pará (30%) and better represented the spatial variability of socio-economic conditions across the entire arc of deforestation. In this regard, the end of the boom-and-bust development pattern observed in the cross sectional analysis may appear to be the result of a "business as usual" expansion of the deforestation frontier.

3.2.2 The Environmental Kuznets Curve

While these results provide evidence that deforestation has not been followed by a socio-economic decline and thus reject the boom-and-bust hypothesis in the Brazilian Amazon, they provide little insight into the current scenario of socio-economic development and deforestation across the Amazon. Thus, we suggested that the increased socio-economic development associated with a decrease in deforestation activity emphasized profound changes on the Amazon frontier and attested to the emergence of an environmental Kuznets curve (EKC). The EKC hypothesis predicts that environmental degradation may have an inverted U-shaped relationship with development: environmental degradation goes along with development until primary needs are met; then a turning point is reached when concerns for the environment increase and environmental degradation decreases [81, 82]. In other words, while the boom-and-bust hypothesis analyses a pattern of socio-economic development depending on deforestation, the EKC explores how socio-economic development acts on deforestation.

It is worth noting that the EKC was first conceptualized considering air pollution levels as a proxy for environmental degradation. Yet, its application to tropical deforestation has long been debated and appears to be somewhat controversial [83, 84]. Indeed, deforestation may be considered as irrelevant to monitor a decline in environmental degradation since it is usually seen as an irreversible process. However, deforestation still represents the main threat to the environment in the Amazon and the major source of greenhouse gas emissions [85].

Our results confirmed the emergence of an EKC relationship in the Brazilian Amazon, thus implying the existence of a turning point above which environmental degradation starts declining. The figure 3.8 depicts the deforestation activity as a function of income, and hence the EKC. It emphasizes a positive relationship between deforestation and income in low income areas. In the case of high incomes, the relationship was stable in 2000 and negative in 2010. For 2000, the curve did not illustrate the EKC hypothesis but the inverted U-shaped for 2010 curve was typical of an EKC and illustrated a decoupling between deforestation and socio-economic development in the Brazilian Amazon. Concerning the income threshold, in 2000, we found an inflection point over an income per household equal to 2.5 times the Brazilian minimum wage, above which

the relationship between income and deforestation reached a plateau at high levels of deforestation activity. In 2010, the turning point was observed around one minimum wage per household. Above this income value, deforestation activity decreased significantly.

These results raised important issues about the relative responsibility of low and high income populations in explaining deforestation. Actually, whereas 85.3% of the total deforested area in the Amazon in the 1998-2000 period was carried out in cells where the mean income per household was higher than one minimum wage, this proportion fell to 48.9% in the period 2008-2010. The growing importance of low income populations in total deforestation was first explained by the rapid increase (from 33.9% to 56.7%between 2000 and 2010) in the number of cells with low income values (< one minimum wage) due to the fact that the minimum wage evolved faster than inflation in the 2000s. However, the increase in average deforestation activity (from 0.61%to 0.85% during the same period) in low income cells should not be overlooked. Indeed, it is the result of the establishment of an effective environmental governance aimed at controlling deforestation and the expansion of large-scale commodity-oriented agriculture. In this context, continuing deforestation in the 2000s despite environmental legislation appeared to be a sign of socio-economic precariousness. Then, future social and environmental policies should be more concentrated on better supporting low income farmers in the improvement of their agrarian and economic situation, while achieving legal compliance.

3.2.3 What is to be expected in 2020?

This analysis was led with socio-economic data from 2000 and 2010 and should thus be updated. Unfortunately, the national census scheduled for 2020 has been delayed. However, it appears that the current context of political and economic crisis weakens the Brazilian institutional and financial capacities to ensure environmental preservation in the Amazon and thus questions the validity of the EKC for the coming years. Interestingly, in their seminal paper about EKC, Grossman and Krueger [86] estimated that the EKC turning point where environmental degradation starts declining usually comes when a country reaches a per capita GDP of US$8,000. In Brazil, the figure 3.9 shows that all peaks of Amazon deforestation correspond with drops in GDP per capita (in 1995, early 2000s and late 2010s). The highest levels of GDP per capita, i.e. higher than the US$8,000 threshold, are observed after 2008 and correspond to the lowest levels of deforestation, i.e. below 10,000 km^2 per year from 2009 to 2018. Finally, the GDP per capita decreased dramatically since 2012, corresponding to an increase in deforestation, as evidenced by the correlation between both variables is $R^2 = 0.62$ (p-value = 5.313e-08). Although such analysis would need to be statistically deepen, it illustrates

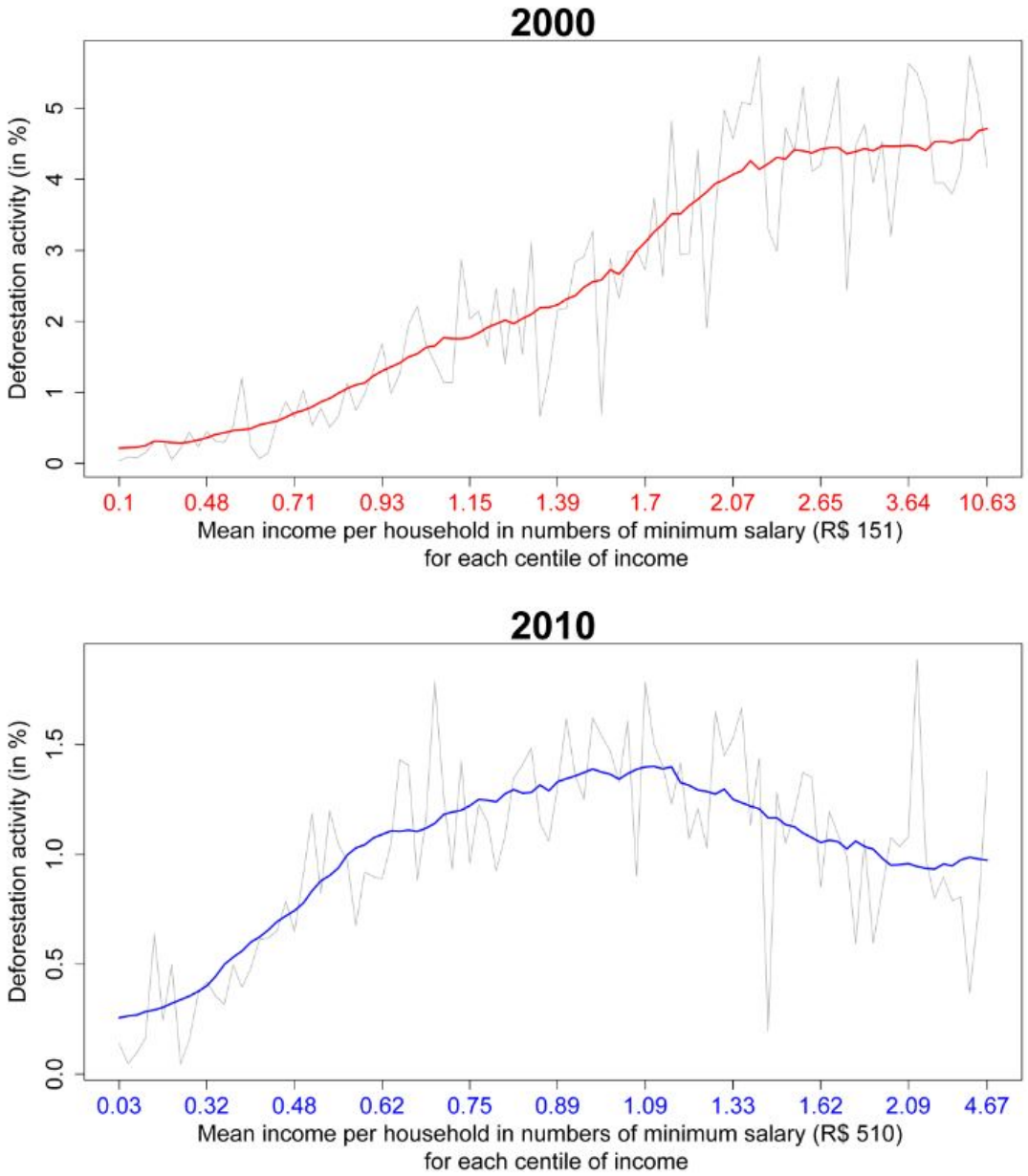


FIGURE 3.8: Deforestation activity as a function of mean income per household in 2000 and 2010. In 2000, the curve illustrates a positive relationship between income and deforestation activity. In 2010, this relationship follows an inverted U-shaped, evidence of the emergence of the EKC: beyond a income threshold of one minimum wage per household, deforestation activity begins to decrease.

the difficult challenge to establish an efficient environmental governance model on the long run, since all positive achievements of the 2000s are currently challenged.

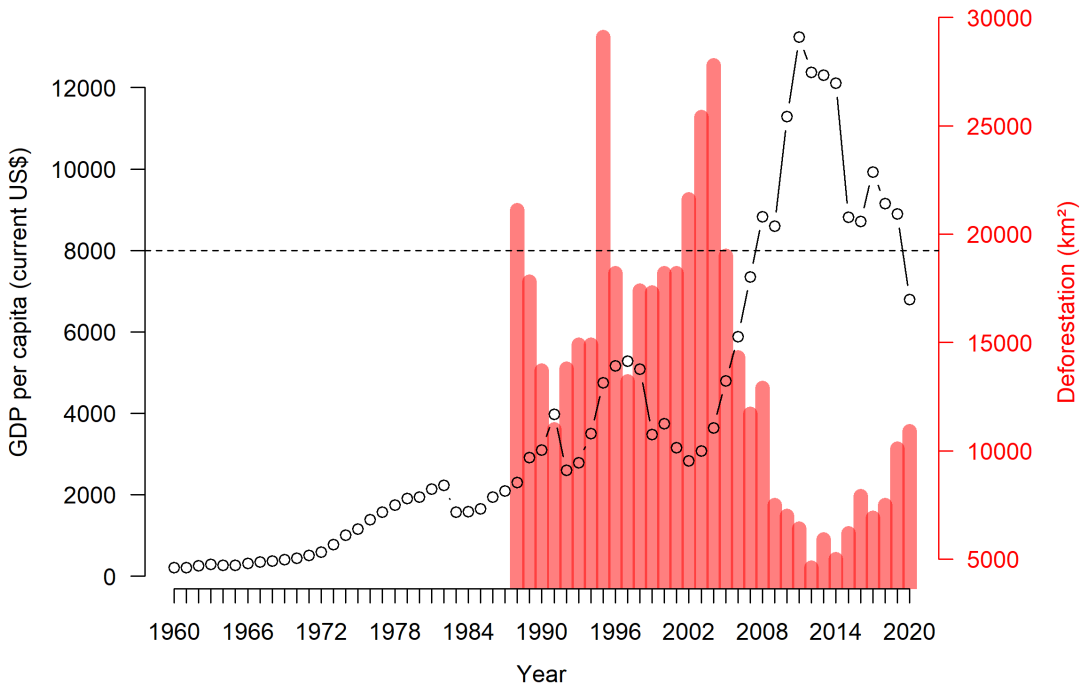


FIGURE 3.9: Long-term evolution of the Gross Domestic Product (GDP) per capita (in current US$) in Brazil and deforestation in the Amazon, in km^2. The black dashed lines shows the US$8,000 threshold identified by Grossman and Krueger [86].

3.3 Perspectives: from pixel to landscape analysis

The spatio-temporal indices I implemented so far to monitor the socio-environmental dynamics in the Amazon were computed at pixel level (or cell level if we consider the study about deforestation and socio-economic development). These indices thus only provide information on the landscape composition but fail to take into account the landscape configuration and temporal dynamics. In this regard, I wish to adopt a whole-landscape approach that incorporates the full spatial, temporal, and socioeconomic domains [87].

I recently contributed to two studies [88, 89] coordinated by colleagues from ecology and forestry that focus on landscape analysis to analyse the impacts of land use changes on biodiversity and forest degradation, respectively. These studies illustrate the various dimensions of the landscape analysis which I consider necessary to take into account:

- Scale, i.e. the definition of a landscape unit (e.g. a buffer around a point of interest, a grid cell, or the result of a segmentation process);

- Thematic resolution, i.e. the definition of land cover/use classes (e.g. the definition of complex classes such as forest or urban affects the computation of landscape metrics);
- Composition, i.e. the distribution of land cover patches (e.g. number of classes in a landscape unit, proportion of a landscape unit being covered by a land cover class);
- Configuration, i.e. the spatial organization of land cover patches (e.g. fragmented or connected);
- Legacy, i.e. the temporal dynamics of land cover patches (e.g. deforestation cycles, number of harvests per year);

I consider that a future major research direction should consist in adapting landscape analysis approaches to assess the impacts of land use changes on water resources. Indeed, there is an urgent need to look beyond the impacts of land use changes on forests, with an important focus on the degradation of water resources. Indeed, as explained by Pelicice et al. [90] regarding the Tocantins-Araguaia basin, capital-concentrating activities such as agribusiness, aquaculture with non-native species, mining, and hydropower promoted extensive changes in land cover, hydrology, and environmental conditions, which are now jeopardizing the basin's biodiversity and ecosystem services.

Yet, monitoring the impacts of land use dynamics on freshwater systems is a very challenging task due to the diversity of land use changes and the complexity of their impacts, especially considering it implies to study both the impacts on water flows and water quality. To date, most methods to assess anthropogenic pressures on freshwater systems are based on place-based metrics (e.g. water quality or species presence/composition) that describe in-stream ecosystems and are thus difficult to apply at large scale [91]. Improvements have then been proposed by developing functional indicators at watershed level. For example, Latrubesse et al. [64] proposed a Dam Environmental Vulnerability Index using information on the location of built, under construction or planned large dams in the Amazon basin. Such approach needs to be adapted to include information on the proliferation of small dams that may have dramatic impacts on hydrosystems. For example, their accumulation can affect river flows because the water collected is often used for irrigation and thus does not contribute to river flow [67]. Additionally, stored water can evaporate or seep, thus leading to a water loss for river network depending on the climate context and the local geological settings. Water stagnation and lamination in impoundments also modify stream water thermal variations due to temperature differences between input and output. Also, reservoirs are commonly nutrient and pollutant sinks at watershed scale [92] and can be affected by eutrophication [93, 94]. Downstream

of the dams, the entire freshwater system is modified due to hydrological and sedimentological changes that can significantly affect the river morphology and habitats [95–97]. Such hydrological and morphological modifications can alter biodiversity of the freshwater systems. Indeed, run-of-river dams can represent obstacles to species migration by limiting mobility, habitat changes, disturbance regime and water quality modification (for reviews see Bunn and Arthington [98] and Nilsson [99]).

Analyzing the impacts of land use changes on water resources implies computing additional landscape metrics than the number and cumulated area of small reservoirs (as already introduced in section 3.1.2). Such new metrics should pay special attention to the location of the reservoirs (e.g. distance to headwaters, impacts on the fragmentation of river streams) at watershed level. In addition, it is necessary to assess how aquatic ecosystems respond both to watershed landscapes and land use/cover patterns within riparian zones and watershed uplands [100, 101]. Finally, it is worth noting that major efforts to contain the environmental impacts of the development model consist in the reforestation of riparian areas and the adoption of integrated crop-livestock systems. In this regard, landscape analysis should not be only considered relevant to estimate the impacts of land use changes on hydrosystems but also to assess the resilience of the systems facing changes.

Chapter 4

Climate variability and climate change in the Amazon: implications for land use/cover dynamics

The intense land use changes related to the rapid advance of a very dynamic agricultural frontier raise concerns about the capacity of the Amazon region to continue providing essential ecosystem services, including food and timber production, biodiversity conservation, carbon storage, hydrology and climate regulation [102]. These concerns are even exacerbated when considering that the Amazon basin environment is now considered to be in transition to a disturbance-dominated regime [103]. In this regard, understanding the interactions between environment and societies is a key issue for assessing the causes and consequences of global climate change, especially in tropical areas. On the one hand, the tropics concentrate nearly two-thirds of the planetary rainfall [104] and are affected by a strong spatio-temporal variability. This variability is expected to be magnified by anthropogenic pressures (e.g. high deforestation rates in the tropical rainforests and savannas [2, 105]), whose impacts on the mid- and long-term climatic trends and water cycle are worrying. On the other hand, the tropics also hold a large proportion of human population and important agricultural activities (including for global commodities markets) which can be severely affected by climate change.

In this context, I focused my researches on 1) the characterization of climate change, 2) the impacts of climate variability on human activities and 3) the adaptation and mitigation strategies to be implemented by local communities. Here, I introduce important results achieved with colleagues from LETG unit using various remote sensing-based

rainfall estimates (section 4.1). We used this data to characterize climate change in the Amazon (section 4.2) and its implications for agricultural activities (section 4.3). Then, I discuss the environmental policies implemented locally to support adaptation/mitigation strategies (section 4.4). Finally, I suggest future research directions on the relationships between climate and natural vegetation, especially through the analysis of phenological cycles (section 4.5).

4.1 Remote sensing-based rainfall estimates

In the Southern Amazon, since temperatures are fairly constant throughout the year (between 22 and 26 ° C) and diurnal temperature range is higher than the annual average temperature range, rainfall appears as the main parameter used for describing the different subequatorial to tropical climates. However, monitoring rainfall in the Amazon is challenging. At regional scale, rainfall regimes in the basin are far from uniform [102] and, at local scale, precipitation is closely linked to the presence of convective cells resulting in marked spatio-temporal variability [106, 107]. A sparse raingauge network (e.g. a bit more than a hundred stations in Mato Grosso) and a lack of maintenance leading to many gaps in the historical series [108] limit the potential of ground data for analysing the spatiotemporal variability of rainfall regimes in Mato Grosso. It has thus been proposed since the 1970s to use remote sensing data to produce homogeneous rainfall estimates.

Since my PhD, I worked with three products which are introduced below:

- **TRMM**: The Tropical Rainfall Measuring Mission (TRMM) products for rainfall estimation combine different types of sources [104, 109]. In my studies, I mainly focused on the daily 3B42 (version 7) TRMM products at a 0.25 ° spatial resolution available since 1998. This product is gauge-adjusted, but the independence between estimated and observed data is assumed because rain gauge data used in 3B42 consists in monthly estimates from the Global Precipitation Climatology Centre product, which considers few rain gauges in the Amazon Basin. The TMPA algorithm combines multiple independent precipitation estimates from multiple instruments on board: the Advanced Microwave Scanning Radiometer for Earth Observing Systems (AMSR-E), Special Sensor Microwave Imager (SSMI), Special Sensor Microwave Imager/Sounder (SSMIS), Advanced Microwave Sounding Unit (AMSU), and Microwave Humidity Sounder (MHS), each calibrated to the TRMM Combined Instrument (TCI) [109].

- **PERSIANN-CDR**: The Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Climate Data Record (PERSIANN-CDR) product was released by the National Oceanic and Atmospheric Administration (NOAA; [110]). It consists of daily precipitation estimates based on infrared and daytime visible imagery from geostationary satellites and built on neural networks classification and approximation techniques [111, 112]. The spatial coverage ranges from 60°S to 60°N and the spatial resolution is of 0.25°. Data are available for the 1983-delayed present period, thus completing long time series relevant for climate change studies.

- **CHIRPS**: Finally, the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) rainfall gridded product combines 0.05 ° high spatial resolution satellite imagery and in-situ station data, providing near-global (50°N-50°S ; all longitudes) daily precipitation estimates [113] since 1981 at a spatial resolution of 0.05 ° and 0.25°.

Although remote sensing rainfall products are promising, their quality needs to be assessed. In this regard, a prior objective of the studies I participated aimed at assessing the accuracy of these estimates in the Amazon. My colleagues and myself conducted quantitative analyses to compare the rainfall measured at rain gauges with the corresponding rainfall estimates from TRMM 3B42, PERSIANN-CDR and CHIRPS data. This comparison was based on continuous and categorical statistical indices as proposed by Tan et al. [114],Miao et al. [115] and Ringard et al. [116]. The continuous indices included correlation coefficient (CC), root mean square error (RMSE) and relative bias (RB) measured at daily and monthly scales. An accurate rainfall estimate is characterized by a high CC combined with low RMSE and RB values. The categorical indices analyzed the capacity of rainfall products to detect rainy days. These indices were accuracy (ACC) (eq. 4.1), probability of detection (POD) (eq. 4.2) and false alarm ratio (FAR) (eq. 4.3). In our case, ACC and POD must be maximized while FAR should be minimized.

$$ACC = \frac{A + D}{n} \tag{4.1}$$

$$POD = \frac{A}{A + C} \tag{4.2}$$

$$FAR = \frac{B}{A + B} \tag{4.3}$$

where A refers to rainfall detected by the rain gauge and the remote sensing estimates; B refers to the detection of rainfall by remote sensing data when no rainfall was measured at the rain gauge; C refers to precipitation missed by remote sensing data when rainfall was recorded by the rain gauge; D refers to the non-detection of rainfall both by the rain gauge and by the remote sensing data and n refers to the total number of observations. In the present study, a rainy day corresponds to a day with estimated rainfall > 0.2 mm, i.e. the minimum rainfall usually measured by a rain gauge.

An example of the application of such indices to validate the quality of PERSIANN-CDR data (but results for other products usually follow the same pattern) in the Southern Amazon is introduced in table 4.1. At daily scale, the coefficient correlation was 0.347 and increased to 0.835 at monthly scale. These values underlined the difficulty to estimate rainfall at a fine temporal scale but confirmed the potential of PERSIANN-CDR data to capture rainfall regimes. The relative bias was positive, corresponding to a slight (around 6%) overestimation of remote sensing based rainfall estimates. Daily statistics computed on a monthly basis emphasized the irregular performance of the PERSIANN-CDR according to the seasons (fig. 4.1). Correlation and RMSE were higher during the dry season (May to September). The relative bias showed rainfall was underestimated during the dry season (especially in June and July) and overestimated in the rainy season. Accuracy (ACC) was also better during the dry season ($> 75\%$) while the months with lowest accuracy (but still greater than 50%) were at the beginning (October) and the end (March-April) of the rainy season. On the contrary, the POD was always greater than 50% but values were lower in the dry season.

Finally, consistent with the previous values, the FAR values were lower in the rainy season and higher in the dry season. In July and August, over 75% of the rainy days detected by PERSIANN-CDR data were in fact not recorded as rainy days by the rain gauges. The contrast between the rainy and dry seasons could be explained by (1) the type of clouds in winter, with sparsely distributed and less frequent rainy convective clouds (which can lead to a miss at the rain gauge, but are detected by remote sensing data which covers a broader area), (2) the difference in spatial resolution between the two datasets and (3) the presence of many outliers, some of them due to suspicious rain gauge data [117].

TABLE 4.1: Validation results for main statistical indices applied to daily and monthly rainfall values and to onset and demise dates of the rainy season.

	CC	RMSE	Relative Bias	ACC	POD	FAR
Daily	0.347	12.122	6.058	0.618	0.913	0.430
Monthly	0.835	72.657	6.476	-	-	-

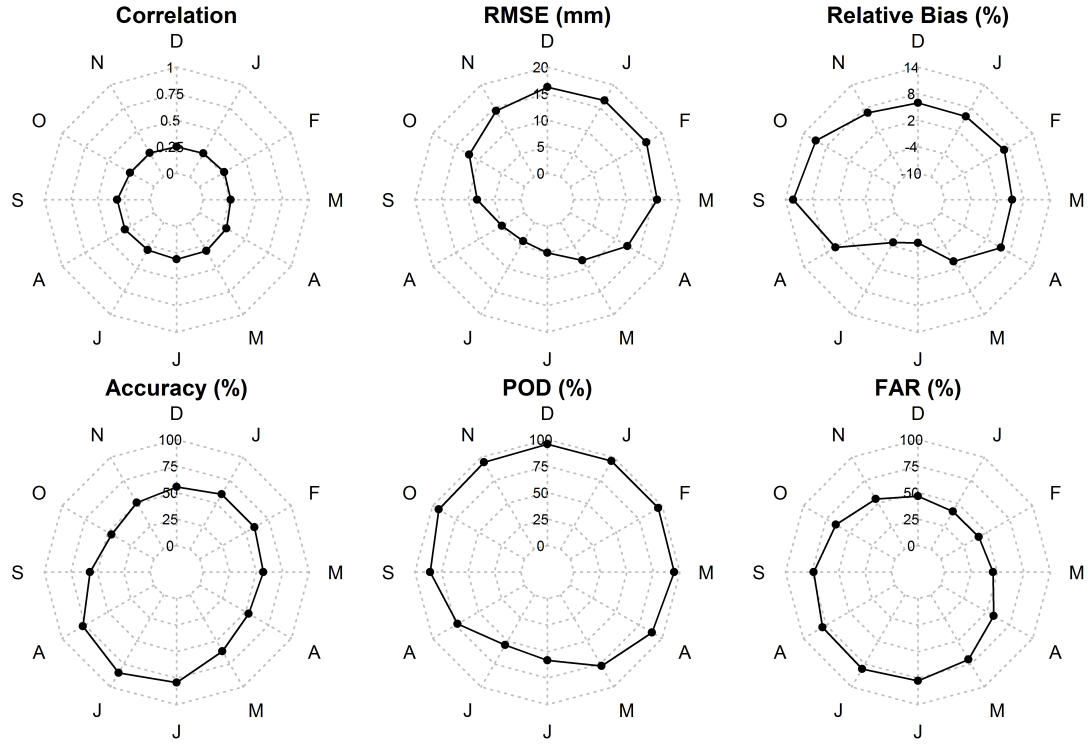


FIGURE 4.1: Daily validation statistics of PERSIANN-CDR data computed per month on 71 rain gauges located in the southern Amazon.

4.2 Climate change in the Amazon

I then used these rainfall estimates to monitor climate change in the Southern Amazon. Indeed, the impacts of deforestation and agricultural development on local and regional climate rainfall regimes are of particular concern [107, 118, 119]. Indeed, Khanna et al. [120] suggested that deforestation is sufficiently advanced to have caused a shift from a thermally to a dynamically driven climatic regime, which may affect the rainy season in the future. In this connection, previous studies have pointed to the intensification of the dry season in the southern Amazon [119, 121, 122], and predicted intensification of rainfall during the rainy season [123] or changes in the wet-day and dry-day frequencies [124].

In this context, I contributed to the monitoring of changes 1) in the temporality of the rainy season and 2) in the occurrences of precipitation extremes.

4.2.1 Trends in the rainy season metrics in the Southern Amazon

I first focused on the temporality of the rainy season. For this reason, besides the annual total rainfall, I considered three metrics of the rainy season, i.e. onset date, demise date and duration of the rainy season. These metrics were computed on a pixel-by-pixel

basis for each hydrological year from 1983 to 2013 (with the end of year 2013 occurring in 2014) using PERSIANN-CDR data. Determining the onset and demise of the rainy season can be rather difficult as there are several definitions of the onset (demise) of the rainy season according to the scientific area of interest, i.e. meteorology, hydrology or agronomy [125]. I used the agronomic definition of the rainy season thus considering that the onset (demise) of the rainy season cannot be followed (preceded) by long dry periods that can affect the crops. In that case the main difficulty consisted in avoiding the detection of false onset (demise) dates. Various authors have proposed methods to overcome this issue. Here, I considered the Anomalous Accumulation (AA) method proposed by Liebmann et al. [126]. This method is based on the definition of an index called AA to calculate the difference between the accumulation of daily precipitation and an average theoretical accumulation considered as an average daily rate (i.e. annual rainfall / number of days). The equation for calculating AA is (Equation 4.4):

$$AA(d) = \sum_{n=1}^{d} [R(n) - \bar{R}] \tag{4.4}$$

where d is the number of days, $R(n)$ is daily rainfall and $\bar{R}$ is average daily rainfall. From this annual time series, the onset (demise) of the rainy season is defined as the day in which AA is minimum (maximum) (fig. 4.2). The duration of the rainy season then consists in the number of days between the onset and the demise dates for each pixel.

After computing the four metrics of the rainy season for all years of the 1983-2013 time period, I finally calculated the means and trends based on the Mann-Kendall test (e.g., [128]). Here I present only the coefficient τ of Mann-Kendall, which varies from -1 to +1. The value of -1 (+1) indicates a trend of continuous decrease (continuous growth) in the study period. The value 0 indicates that there is no trend. Finally, a p-value of less than 0.05 was used to assess whether trends were significant.

The results on spatial patterns of annual rainfall were consistent with those found in other studies on rainfall regimes in the Amazon [128, 129], highlighting specific rainy regions such as the Serra do Cachimbo in southern Pará. The rainy season started earlier in northeastern Mato Grosso (MT) and southeastern Pará (fig. 4.3, top right) and later in western Mato Grosso (Chapada dos Parecis) compared to other areas of the same state.

In addition, the onset date was characterized by a higher interannual variability than the demise date. The onset of the rainy season is associated with a rapid increase in specific humidity prior to the onset and changes in the horizontal and vertical winds (characteristic of the South American monsoon) that follows the period when temperatures reach

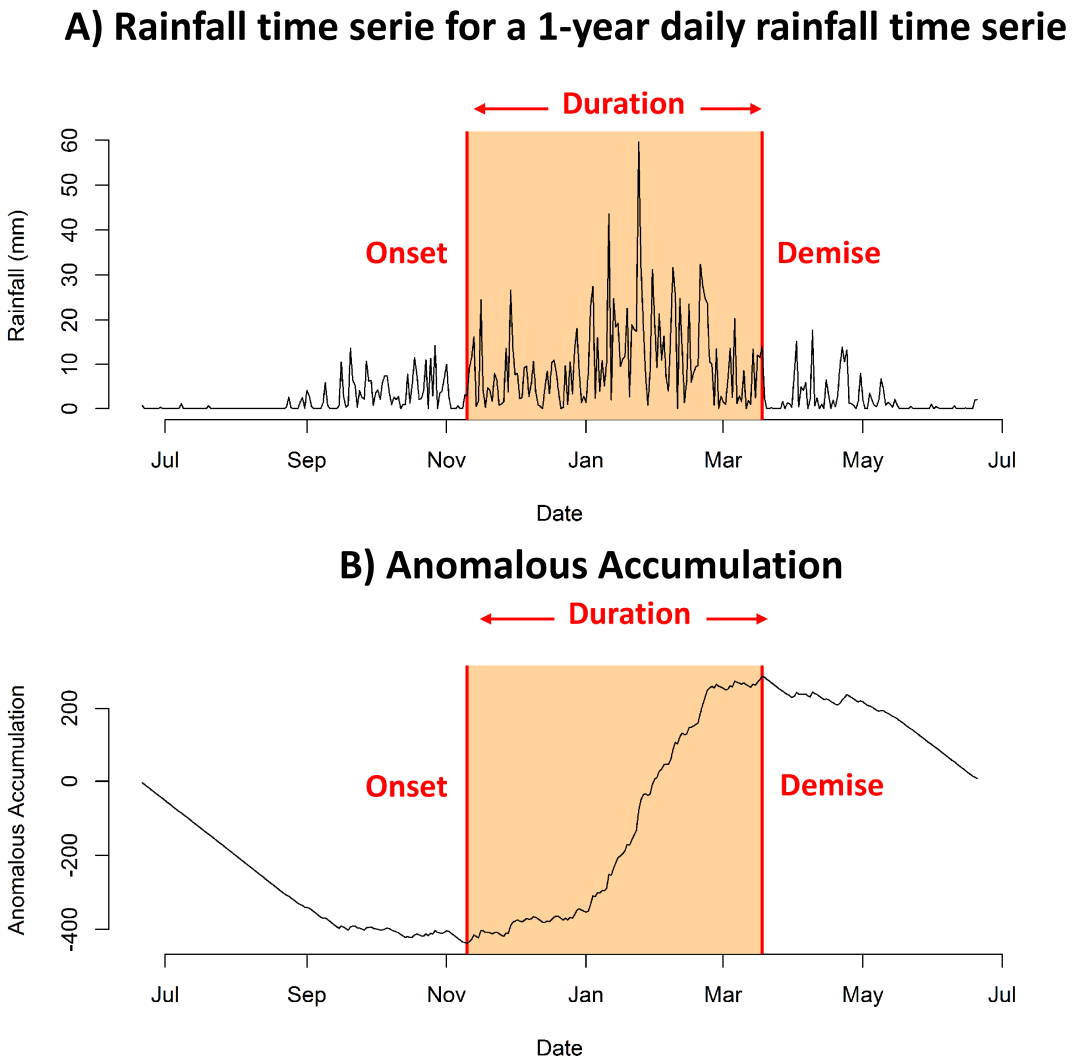


FIGURE 4.2: Illustration of the computation of the rainy season temporal metrics based on the method proposed by Liebmann et al. [126]. It includes: (A) a one year daily rainfall time series for a given pixel and (B) the corresponding anomalous accumulation time series with identified onset and demise dates [127]

their seasonal maximum [130]. In this context, the high variability of onset dates has been directly linked to pre-seasonal conditions during the dry season in the Pacific and South America [131], which modulate moisture availability, atmospheric instability and atmospheric circulation in both low and upper levels.

The demise dates followed a north-south gradient that was more pronounced than for the onset dates (fig. 4.3, bottom left). The end of the South American monsoon was linked to the northward migration of the Atlantic ITCZ and the strengthening of low-level easterly winds that reduce the moisture supply from the Atlantic over the southern Amazon [130, 132]. Consistent with these results, the duration of the rainy season was longer in the north than in the south of the study area (fig. 4.3, bottom right), with

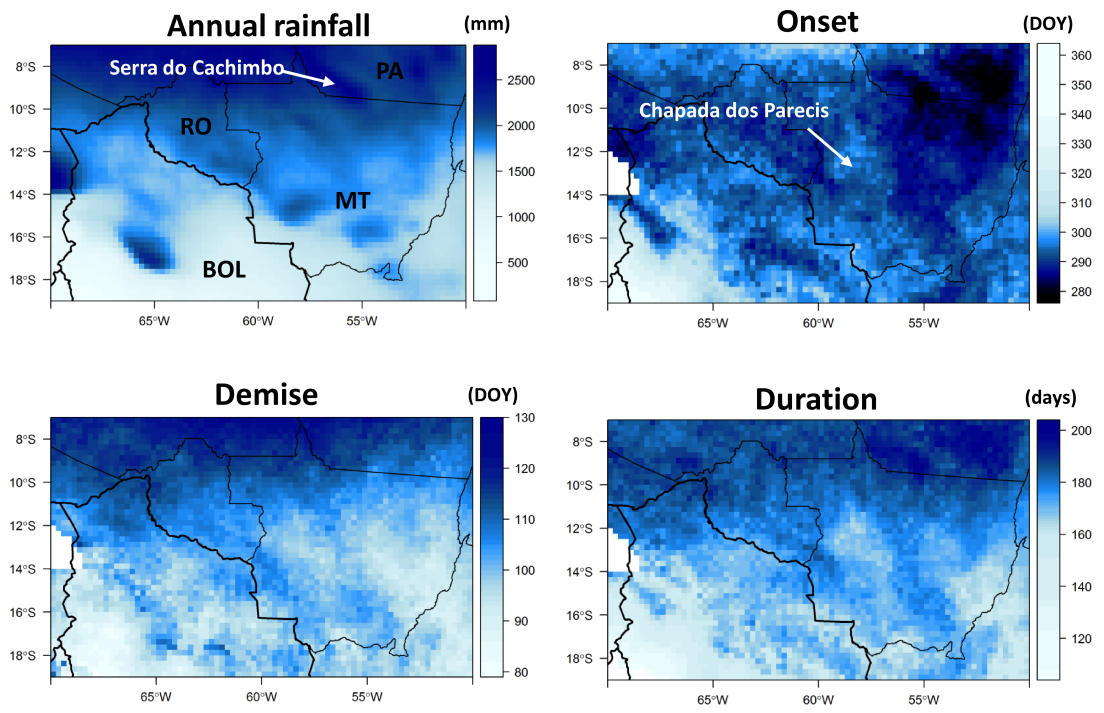


FIGURE 4.3: Mean annual rainfall, onset date, demise date and duration of the rainy season in the southern Amazon estimated from PERSIANN-CDR for the 1983-2013 time period. The onset and demise dates are defined as the number of days after January 1. Areas in white correspond to pixels where the rainy season is not well defined according to the Köppen classification rules. (PA = Pará, MT = Mato Grosso, RO = Rondônia, BOL = Bolivia) [127]

greater variability in Bolivia, along the Andean Cordillera, and in the southern and northern parts of the study area.

Concerning long term changes in rainfall patterns, annual amounts tended to decrease regionally (fig. 4.4, top left), although trends varied strongly and were mainly statistically non significant, except for regions in Bolivia and western Mato Grosso. In addition, we observed positive trends for onset dates (delayed onset) in Bolivia and northern Mato Grosso (fig. 4.4, top right), although they were mainly significant in Bolivia. Conversely, trends were negative (early onset) in the center of Mato Grosso and in Rondônia but not significant. Trends on demise dates were significantly negative for the majority of the study area including southern Amazonas, parts of Acre and Bolivia, Rondônia, and western Mato Grosso (fig. 4.4, bottom left) indicating a trend toward earlier demise of the rainy season in these areas. As a consequence, we also observed a trend towards a reduction of the length of the rainy season, with especially significant trends in Bolivia, Rondônia state and southern Amazonas (fig. 4.4, bottom right).

To summarize, we identified two major cases of shortening of the rainy season that appeared to coexist with a small overlap: regions where the reduction in the length

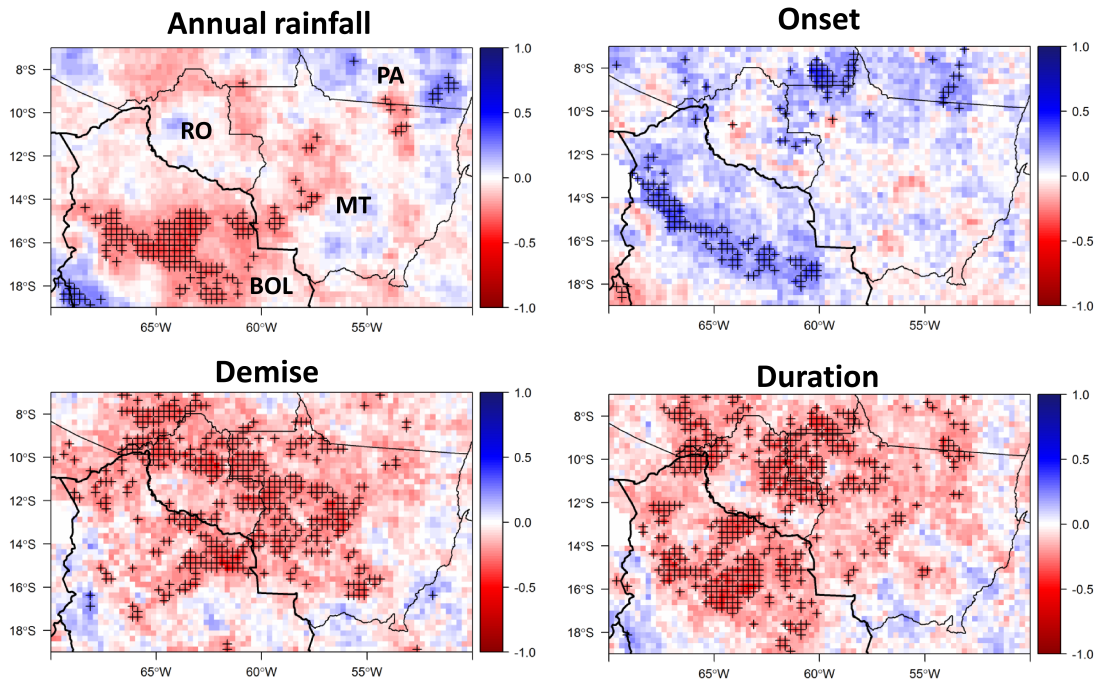


FIGURE 4.4: Trends in annual rainfall, onset date, demise date and duration of the rainy season in the Southern Amazon estimated from PERSIANN-CDR for the 1983-2013 time period. The trends correspond to the τ of Mann-Kendall. Significant trends are indicated with crosses. Areas in white correspond to pixels where the rainy season is not well defined according to the Köppen classification rules. (PA = Pará, MT = Mato Grosso, RO = Rondônia, BOL = Bolivia) [127]

of the rainy season was linked to a delayed onset (mainly central Bolivia) and regions where the reduction was linked to an early demise (northern Bolivia, Rondônia, southern Amazonas and western Mato Grosso). These general conclusions on the shortening of the rainy season in the southern Amazon are in agreement with results of previous studies [119, 121]. However, our study disagreed with the results of the study of Fu et al. [121] who also observed a shortening of the rainy season but related to the delayed onset of the rainy season. These discrepancies may be due to: (1) the large standard deviation of the onset dates that may prevent the identification of statistically robust trends; (2) the use of different datasets; and (3) perhaps most importantly, the different methodologies used to define the rainy season.

4.2.2 Precipitation extremes

In addition to the temporal metrics of the rainy season, I worked with Beatriz Funatsu (also from LETG research unit) to look at the evolution of extreme precipitation events such as droughts and heavy precipitation [133]. Such events are part of the natural variability of the climate system, however the frequency of precipitation extremes is expected to increase in the context of climate change. Several studies have pointed to

a seemingly increase in the frequency of "once-in-a-century"-type events of floods and droughts [134, 135] and communities across the basin perceive changes in climate and precipitation patterns occurring in the last decade [136].

Using the CHIRPS data at a 0.25° spatial resolution for the 1981-2018 time period, we computed extreme precipitation indices defined by the World Climate Research Program (WCRP) Expert Team of Climate Change Detection and Indices [ETCDDI; https://www.wcrp-climate.org/etccdi; see also 137, 138] to characterize precipitation extremes. The ETCCDI indices describe moderate climate extremes with re-occurrence times of a year or shorter. These indices comprise: (1) absolute indices, which describe the annual maximum 1 day or (accumulated) 5 day precipitation amounts, (2) heavy precipitation threshold indices, which count the number of days that a particular daily precipitation value is attained or exceeded (e.g., daily rainfall above 20 mm), (3) duration indices, which describe the longest length of wet (CWD) and dry (CDD) spells, and (4) percentile-based threshold indices, which describe the occurrence of precipitation exceeding a high percentile threshold, usually 95th or 99th percentile. The list and definition of indices is presented in Table 4.2.

The figure 4.5 shows the mean spatial distribution of indices listed in Table 4.2. The mean yearly accumulated precipitation above the 37-yr 95th percentile (P95w, Fig 4.5a) and maximum 5-day accumulated precipitation (RX5day, Fig 4.5c) show a spatial pattern similar to that of the annual simple day intensity index (ASDII; Fig 4.5e), with larger values at the northeastern and northwestern edges of the basin, and over the Amazon-Andes transition region, which has been identified as a rainfall hotspot area [139]. There is also a relative maximum of these indices over southeast Amazonas (AM) and northern Mato Grosso (MT) states in BR, with a northwest-southeast axis, close to the climatological position of the SACZ at the southeastern edge [140]. The frequency of very wet days (R95wp, RR $\geq$ 95th percentile; Fig 4.5b) is higher in the north/northwest and decrease towards the south/southeast and along the eastern and southwestern basin edges. R95wp ranges from around 11 days a year in the NW region, decreasing to 5-8 days in the southernmost areas. The spatial distribution of R20mm shows a contrasted north-south frequency, with nearly 50% more occurrences in the northern region compared to southern Amazon. R20mm has a typical yearly occurrence of 30-45 days in the northern Amazon, whereas in the southern and southeastern Amazon the typical yearly frequency is of 10-20 days. ASDII shows a weaker spatial gradient, indicating that rainfall is spread over a longer rainy season in the north, thus "diluting" the daily rainfall intensity. The maximum number of consecutive dry days (CDD; Fig 4.5f) follows an expected spatial distribution of longer dry periods towards the south/southeast and short duration in the northern region (shortest at NW). In the northwest, CDD typically

TABLE 4.2: List of precipitation indices, definitions, and units. Extreme precipitation indices source: http://etccdi.pacificclimate.org/list_27_indices.shtml, accessed on 5 August 2019. [133]

Index Label	Index name	Definition	Units
RR	Precipitation	Daily convective rainfall amount	mm
Rx5day	Max 5-day precipitation	Let RR_{kj} be the precipitation amount for the 5-day interval ending k, period j. Then maximum 5-day values for period j are: $Rx5day_j = \max(RR_{kj})$	mm
R1mm	Wet days	Number of days where RR $\geq$ 1 mm	days
R20mm	Very heavy precipitation days	Let RR_{ij} be the daily precipitation amount on day i in period j. Count the number of days where $RR_{ij} \geq 20$mm	days
SDII	Simple daily intensity	Let RR_{wj} be the daily precipitation amount on wet days, w (RR $\geq$ 1mm) in period j. If $R1mm_j$ represents number of wet days in j, then: $SDII_j = (\sum_1^{R1mm_j} RR_{wj})/R1mm_j$	$mm.day^{-1}$
CDD	Maximum number of consecutive dry days (RR < 1mm)	Let RR_{ij} be the daily precipitation amount on day i in period j. Count the largest number of consecutive days where $RR_{ij} <$ 1mm	days
CWD	Maximum number of consecutive wet days (RR $\geq$ 1mm)	Let RR_{ij} be the daily precipitation amount on day i in period j. Count the largest number of consecutive days where $RR_{ij} \geq$ 1mm	days
P95w	Very wet days precipitation	Let RR_{wj} be the daily precipitation amount on a wet day w (RR $\geq$ 1.0mm) in period j and let P95w be the 95th percentile of precipitation on wet days in the 1981-2017 period. If R1mm represents the number of wet days in the period, then $P95w_j = \sum_1^{R1mm_j} RR_{wj}$, where $RR_{wj} \geq$ P95w	mm
R95wp	Very wet days	Let RR_{wj} be the daily precipitation amount on a wet day w (RR $\geq$ 1.0mm) in period j and let P95w be the 95th percentile of precipitation on wet days in the 1981-2017 period. If R1mm represents the number of wet days in the period, then $VWD_j = \sum_1^{R1mm_j} RR_{wj}$, where $RR_{wj} \geq$ P95w	days

lasts only 5 days a year, whereas in the southeast (Mato Grosso - MT, Tocantins - TO, and south of Maranhão - MA states) it can reach more than 60 days.

Trends for ASDII, R95pw, Rx5day, and R20mm (Fig 4.6a-d) show regional agreement, with a generally congruent signal with respect to the total precipitation trend: Positive trends at the northern and western parts of the basin, and negative trends in parts of Amazonas and in southern Bolivia. In terms of rainfall, persistence positive trends for CWD (that is, longer periods of rain) are found in the northern part of the domain, and the opposite in the southern Amazon (Fig 4.6e). In addition, significant positive trends for CDD (Fig 4.6f) were found in most of the southern half of the area (northern Rondônia - RO), MT, TO, and MA states in Brazil), with longer dry-day periods. Areas

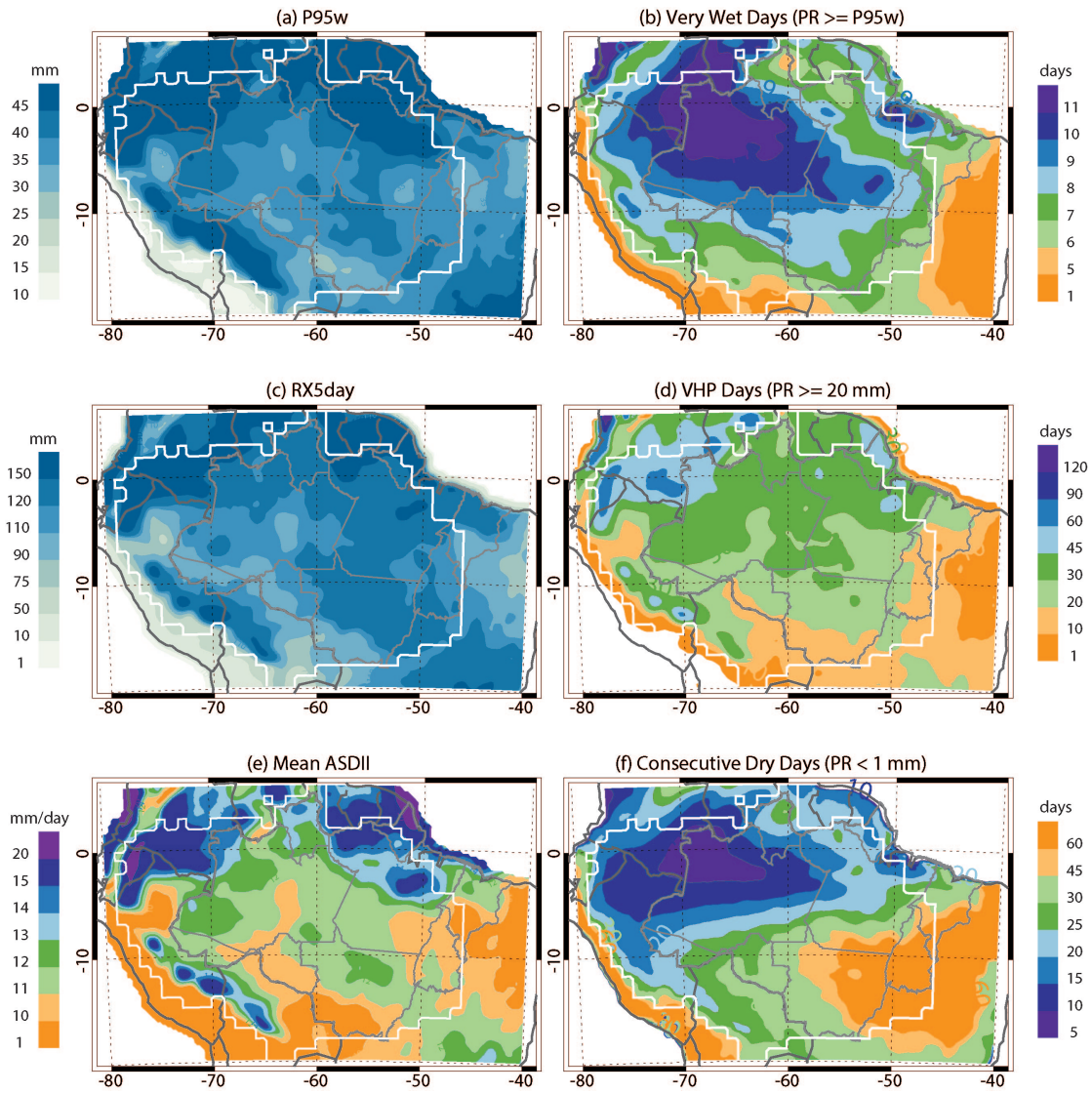


FIGURE 4.5: Spatial distribution of 37-year (1981/82-2017/18) mean (a) P95w (mm), (b) R95w (days), (c) RX5day (mm), (d) R20mm (days), (e) ASDII (mm.day^{-1}), and (f) CDD (days)). White line shows the study domain. A boxcar smoothing filter with a width=5 was used for plotting [133].

with significant negative trends (i.e., decreased persistence of dry-day periods) are found at northwestern Amazon.

4.3 Implications for agriculture

The strong spatio-temporal variability of the rainy season and the observed changes (i.e. shortening of the rainy season) questions the sustainability of the agricultural model in Mato Grosso. This issue actually refers to the first study I led about climate in Mato Grosso, which aimed at relating the regional spatio-temporal variability of rainfall regimes with agricultural dynamics [141]. To do so, I mapped rainfall regimes in Mato

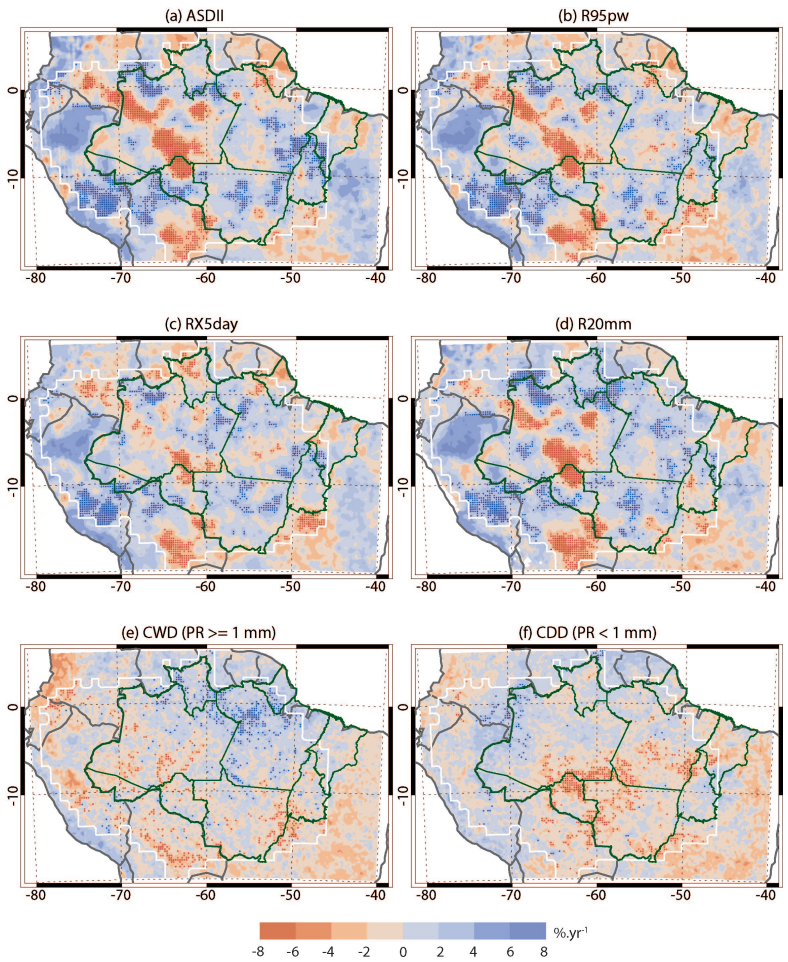


FIGURE 4.6: Linear trends coefficients (in $\%.\mathrm{yr}^{-1}$) for standardized series of (a) ASDII, (b) R95wp, (c) RX5day, (d) R20mm, (e) CWD, and (f) CDD. The color scale for CDD is inverted, so that hues in red (blue) denote increasing (decreasing) CDD anomalies. Values above 95% confidence interval are marked with dark dots [133].

Grosso and assessed how they represent a driver for the adoption of certain agricultural practices such as double cropping systems.

Based on TRMM 3B42 daily product for the 1998-2012 time period and still using the Anomalous Accumulation approach (see section 4.2.1, equation 4.4), I analysed the interannual variability of rainfall regimes in Mato Grosso. As already discussed in section 4.2, the onset date is affected by a larger variability than the end date since the onset date can occur in a time window that lasts from 40 up to 100 days in most regions of Mato Grosso. From an agricultural perspective, it explains why it is sometimes difficult for farmers to plan the optimal seed-time for soybean. Conversely, the window time for the demise date is much more stable, less than 50 days for a major part of the state territory, ensuring the possibility for farmers to harvest maize and cotton during the dry season.

Then, in order to identify a potential link between agricultural practices and rainfall

regimes at regional scale, I calculated the proportion of the net cropped area being cultivated with double cropping systems for each TRMM 3B42 pixel (0.25 °) and crossed this information with the rainfall metrics. The results (Figure 4.7) clearly confirmed that higher proportion of double cropping systems was observed in areas characterized by high annual rainfall (>1900 mm) and a long rainy season (>180 days). The onset of the rainy season also appeared to be important. Where the onset occurred before October 16th, the proportion of double cropping systems exceeded 45% of cultivated areas whereas this proportion was below 30% in other areas. In the same way, areas where the end of the season occurred before March 30th showed lower proportion of double cropping systems (<20%). These temporal thresholds were in total agreement with information collected during surveys with local farmers who estimated that the second crop cannot be sown after February 18th. Considering that the average soybean crop cycle lasts 120 days, it means that the last day for sowing the first crop of soybean in a perspective of adopting a double cropping system was October 18th.

However, it is worth noting that there exist various adaptation strategies that explain why the double cropping systems can still be adopted (around 30%) even in areas where the onset of the rainy season occurs later. The main strategies relies on the cultivation of soybean varieties with short crop cycles (up to 90 days) and on the installation of central pivot irrigation systems to ensure water supply at the beginning and (overall) at the end of the rainy season (see section 3.1.1).

Finally, these results stressed a major issue regarding the evolution of agriculture in Mato Grosso in a context of climate change. While agricultural expansion has been proven to be a driver of deforestation in the Amazon [51, 73] and while the role of agricultural intensification as a driver of land use change is being discussed [142, 143], if the adoption of intensive practices such as double cropping systems is confirmed to be related to a long rainy season, it is worth discussing the sustainability of the intensive agricultural model in Mato Grosso. Indeed, the expected shortening of the rainy season in the southern Amazon might not be compatible with the adoption of double cropping systems. In other words, we questioned if the generalization of double cropping systems in Mato Grosso, initially made possible thanks to propitious rainy conditions, might not indirectly (as a driver of deforestation) cause the decline of agriculture by contributing to the installation of a shorter rainy season that might precisely prevent the adoption of double cropping systems.

Interestingly, it is worth noting that our concerns on the sustainability of intensive double cropping systems in a climate change context "inspired" (100 citations as for March 4th, 2022) other colleagues who further questioned that important issue, as illustrated in this short list of recent papers published after our research:

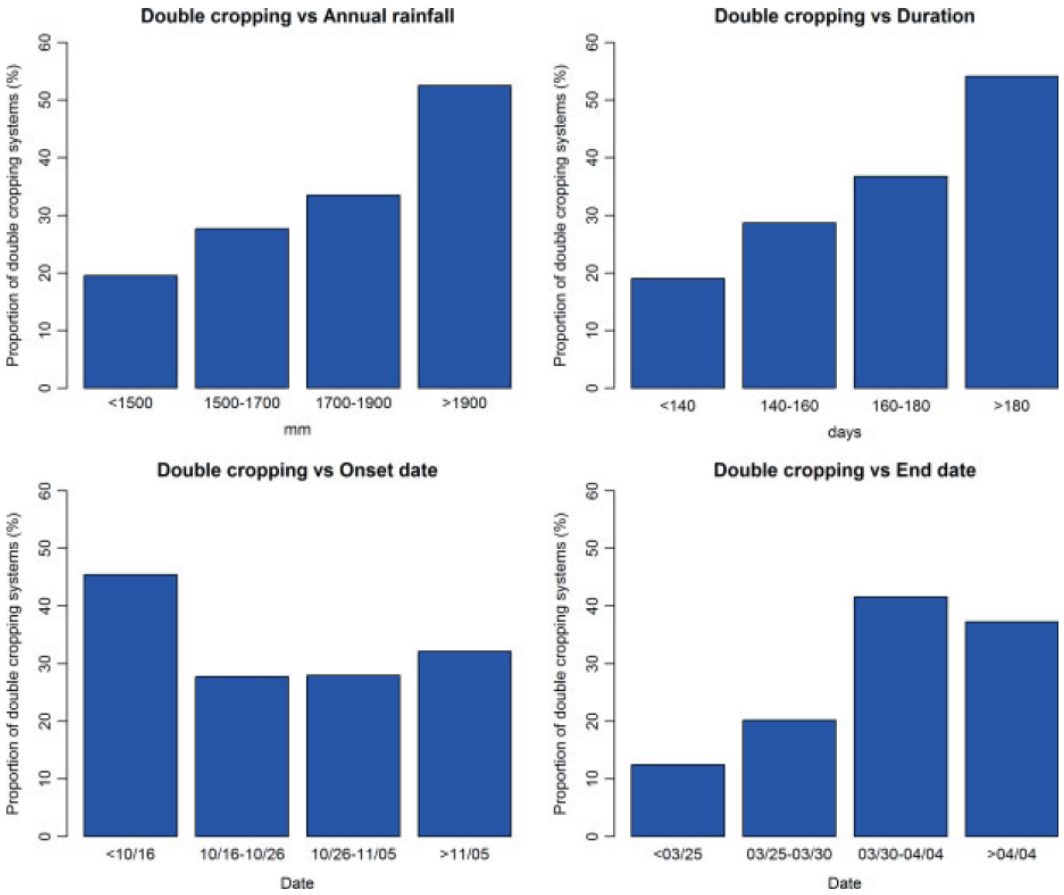


FIGURE 4.7: Relations between rainfall metrics and the proportion of double cropping practices at local scale, i.e. for each 0.25° cell equivalent to a TRMM 3B42 pixel [141]

1. Cropping frequency and area response to climate variability can exceed yield response. Cohn et al., Nature Climate Change, 2016. [58]

2. Increased climate risk in Brazilian double cropping agriculture systems: Implications for land use in Northern Brazil. Pires et al., Agricultural and Forest Meteorology, 2016. [144]

3. Precipitation drivers of cropping frequency in the Brazilian Cerrado: evidence and implications for decision-making. Spangler et al., Weather, Climate, and Society, 2017. [145]

4. Tradeoffs in the quest for climate smart agricultural intensification in Mato Grosso, Brazil. Gil et al., Environmental Research Letters, 2018. [146]

5. Evolution of rain and photoperiod limitations on the soybean growing season in Brazil: The rise (and possible fall) of double-cropping systems. Abrahão and Costa., Agricultural and Forest Meteorology, 2018. [147]

6. Climate risks to Amazon agriculture suggest a rationale to conserve local ecosystems. Costa et al., Frontiers in Ecology and the Environment, 2019. [59]

7. Impacts of Future Climate Predictions on Second Season Maize in an Agrosystem on a Biome Transition Region in Mato Grosso State. Andrea et al., Revista Brasileira de Meteorologia, 2019. [148]

8. Future yields of double-cropping systems in the Southern Amazon, Brazil, under climate change and technological development. Hampf et al., Agricultural Systems, 2020. [149]

9. Challenges to the adaptation of double cropping agricultural systems in Brazil under changes in climate and land cover. Brumatti et al., Atmosphere, 2020. [150]

10. No more double cropping in Mato Grosso, Brazil? Evaluating the potential impact of climate change on the profitability of farm systems. Carauta et al., Agricultural Systems, 2021. [151]

4.4 Perceptions of Climate change and public policies in Mato Grosso

After analyzing climate change in Mato Grosso and its potential implications for agriculture, I worked with colleagues from LETG and especially with Neli Aparecida de Mello-Théry from the University of São Paulo (USP) in order to examine how these issues were comprehended by local populations [152]. Whereas the Brazilian environmental governance model is characterized by a top-down approach in which major decisions are taken at federal and/or state levels [153–155], we wished to analyze the underlying motivations of local populations to implement climate change adaptation/mitigation efforts. Indeed, the recent increase in deforestation rates suggest that economic factors still prevail over environmental considerations to drive the agricultural frontier whereas climate change is still rarely mentioned as a reason to improve the agricultural development model.

To address this issue, we focused on the perception of change and the implementation of adaptation/mitigation efforts at municipality level. We interviewed key stakeholders of the agricultural sector in four municipalities in Mato Grosso (Alta Floresta, Sinop, Sorriso and Lucas do Rio Verde). These stakeholders included federations and associations, trade unions, seed stores, municipal departments and other public organizations, governmental institutions, producers and local researchers, with diverse interests, some being focused on agricultural development while others were more committed to environmental preservation. The interviews consisted in semi-structured questionnaires

applied in 2014 and 2018 and focused on the relation between the perceptions of climate change by different stakeholders, the impacts of climate change on agricultural activities, the vulnerability of current production systems and the relation between climate change and public action. Table 4.3 summarizes the questions of the semi-structured script organized into subtopics.

Subtopics	**Questions**
Seasonality, variation and climate variability	Have you noticed/sensed/perceived that the climate is colder, hotter, rainier, drier? Have you noted climate "markers" in °C, rainfall (mm), at a given time of year?
Consequences of variation and climate variability	Have you noticed/sensed/perceived the need to change the date for sowing or a loss of agricultural products? Any changes to the intensity and/or frequency of extreme weather events? Any change to the start and end of the rainy season and its consequences on sowing and crop productivity?
Vulnerabilities of local production systems	Have you noticed/sensed/perceived any changes to the price of and need for fertilizers, quantity of water, adjustments for production systems and techniques (tillage; increase/decrease of fertilizer or pesticide use, energy consumption)?
Policies and actions	Are you aware of any existing public policies and actions ? Are there any hindrances to production (labor costs, infrastructure, legislation)?

TABLE 4.3: Selection of questions from the semi-structured interview questionnaire [152]

Our first conclusions highlighted the fact that most stakeholders did not notice any changes in precipitation patterns and thus did not see any major concern for agriculture: "The climate here is great for agriculture. Some years are better than others, but it's always good for agriculture". Yet, awareness of environmental issues has begun to impinge on people's consciousness, especially in younger generations who tend to be more aware of these issues and concerned, particularly about conserving water sources. However, it seems that the short-term approach to profit via "adapting" production predominates over long-term environmental sustainability. Indeed, for those stakeholders who have perceived climatic anomalies, their main concern is generally one of targeting, maintenance and expansion of profits of production and productivity. To confirm that point, it is worth noting that most adaptation strategies adopted by producers refer to contouring management practices or installing center pivot irrigation systems, whose first objectives are to ensure agricultural production. As a matter of fact, there is a strong belief that agriculture must rely on technology, both machines and seed varieties in order to adapt to any climate scenario.

Such evolution raises a rhetoric issue about adaptation to climate change. Indeed, the term "adaptation" suffers from a semantic reinterpretation since it refers to the adaptation for the continuation of life, but not necessarily with mitigation of climate change or environmental conservation. More generally, we noticed that, between 2014 and 2018, several of the key stakeholders began using terms and concepts related to the objectives of sustainable development more confidently, and they began to speak out

about suitable transitional public policies that would ensure food security. However, when these issues were more closely scrutinized, many stakeholders indicated that they were still not fully convinced by many of the terms and concepts and continued using expressions upholding traditional practices.

These hesitations about the necessity to promote a sustainable agricultural model with the objective (amongst others) to ensure climate change adaptation and mitigation is partially explained by the fact that international commitments decided at country-level do not always appear relevant to local realities. On the contrary, many farmers even perceive national-scale decisions such as environmental laws (e.g. Forest Code), licensing strategies (e.g. Rural Environmental Registry; CAR) and monitoring efforts (e.g. deforestation monitoring programs) as hindrances and obstacles for the development of agriculture. In this regard, this famous joke heard in Mato Grosso serves as a symbolic example : "What are the two major environmental issues in the Amazon? - SEMA and IBAMA (i.e. the state and federal administrations responsible for the application of environmental laws)".

To conclude, we echoed some authors who consider that education on climate change and closing the gap (perceived or real) between scientists, local populations and decision makers are crucial to develop a trustful relationship that will enable a conjoint action of development and climate change mitigation [152, 156, 157]. There is a need to appease an ongoing feeling that "discussions on climate change are high up in the sky whereas they [public decisions on climate change] happen at ground level. Public policies do not come from the base", and that "Public policies are based on taxing people rather than supporting and educating". In this regard, we emphasized the importance of local leaders, and especially producers, to stimulate debate on climate change at local scale in order to further promote bottom-up approaches in the environmental governance model, as it was the case for the Lucas do Rio Verde Legal project, which was first implemented at municipality level before being transferred to state level. As a researcher, I personally intend to contribute to such initiatives through the implementation of a web platform designed for the diffusion of scientific information on climate and land use dynamics in Mato Grosso (see section 5.1.4).

4.5 Perspectives: climate-vegetation interactions in the Amazon forests

Although my previous researches led me to use remote sensing data to monitor both vegetation (mainly croplands) and climate (mainly rainfall), I did not really integrate

these two research topics so far, with the notable exception of the studies on the relationship between agricultural practices and rainy season metrics (section 4.3). A major perspective of research thus consists in further investigating the relationships between vegetation and climate. In this regard, I rather intend to focus on the influence of climate on forest ecosystems, considering both intact and degraded forests.

To this end, I intend to collaborate with colleagues from CIRAD (UR Forêts et Sociétés) with a great experience in remote sensing-based monitoring of tropical forests. We will especially consider phenology as a significant proxy to assess the role of climate on the functioning of natural ecosystems. Phenology appears essential as it refers to seasonal cycles of canopy development driven by annual and multi-year fluctuations of environmental factors [158]. But monitoring phenology in tropical evergreen forests is challenging since the apparently continuous leaf flush and senescence processes also depend on biological and physiological processes [159]. Indeed, at local scale, forest composition prevails and is firstly characterized by a wide species diversity and variety of phenological patterns [160]. In addition, tropical forests are also vertically complex so that phenological cycles observed at canopy level may differ from lower level and understory phenological patterns [160]. Thus, the influence of climate on phenology must be studied at regional and canopy scale, relying on low to moderate resolution remote sensing data to monitor regional patterns in land surface phenology.

4.5.1 Climate as a driver of phenology in Amazon forests

Amazon tropical forests experience a greening during the dry season attributed to a net leaf-flushing (more new leaves appearing than old leaves being dropped) and an increase in canopy chlorophyll content [161, 162]. During the rainy season, photosynthetic activity declines during the late rainy season, as a a result of senescence, litterfall and envelopment of mature leaves by epiphylls [163].

Yet, this apparently homogeneous situation across the Amazon actually conceals a strong spatio-temporal variability related to major factors that play a role on large seasonal swing leafs, including insect activity and climate variables [160]. The main climate variables that control phenology in tropical ecosystems are rain and solar radiation [158]. To make it simple, humid tropical forests with no constraint on water availability (about 2000 mm annual rainfall or more) experience a rapid leaf-flushing during the dry season in response to an increase in Photosynthetically Active Radiation (PAR). In these forests, the dry season green-up is supported by the deep-rooted trees that ensure water supply. Conversely, the leaf-flushing in drier tropical forests is driven by rainfall occurring at the end of the dry season. Obviously, there is no clear delineation between

water- and light-limited forests so that the distribution of phenological patterns follow a west-to-east moisture gradient related to the dry-season length [160]. This gradient has been mapped by Wagner et al. [164] who represented spatial patterns of climate controls on the leaf growing season in the Amazon (Figure 4.8). Extending the gradient to savannahs and pastures, Huete et al. [163] and Hilker et al. [165] noted a positive relationship with precipitation suggesting an enhancement of photosynthetic activity (i.e. highest LAI and EVI values) during the wet season and a decline during the dry season.

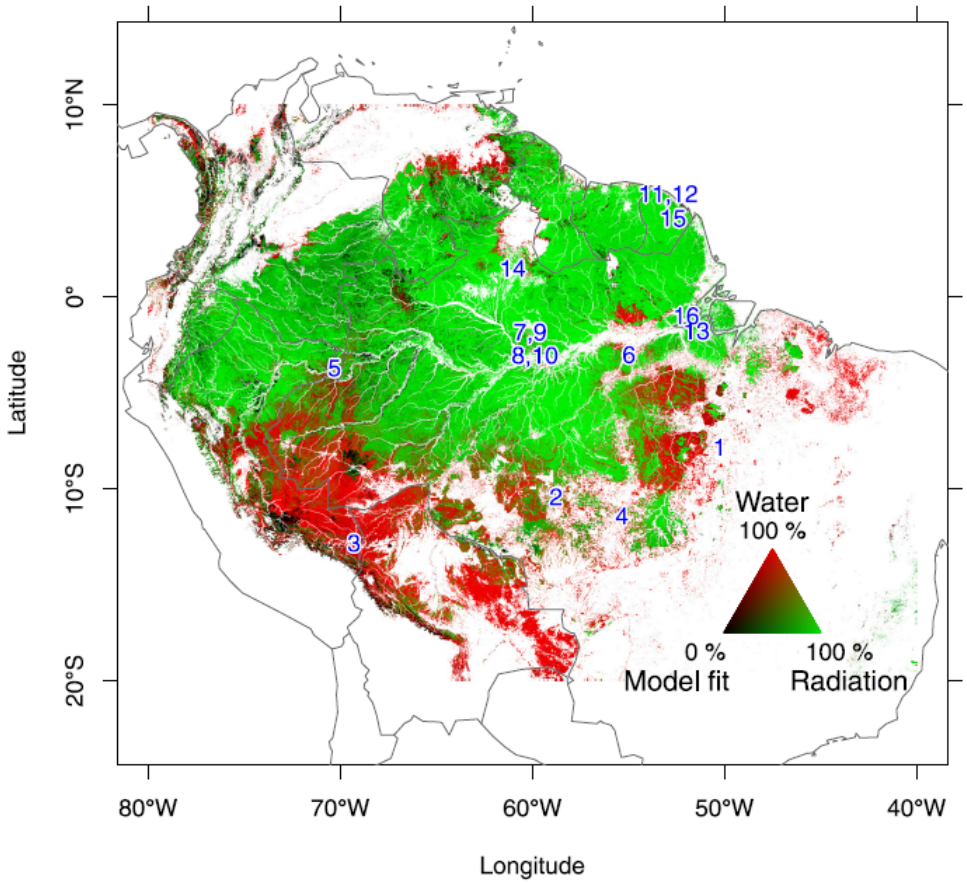


FIGURE 4.8: Spatial pattern of climate controls on leaf growing season initiation in South American tropical forests [164]

In addition, Wagner et al. [164] also evidenced a mean lag of one month between greening and increase in solar radiation in water-limited forests whereas no time lag was measured between greening and rainfall in light-limited forests [164]. Such result is confirmed by [160] who measured phase shifts between Terrestrial Water Storage (TWS), Vegetation Optical Depth (VOD) (i.e. a proxy for canopy biomass), Photosynthetically Acive Radiation (PAR) and Leaf Area Index (LAI) (i.e. proxy for phenology) in the Amazon. They showed that the LAI peak is disconnected with TWS (figure 4.9, row 1) but in phase

with PAR in western light-limited humid Amazon forests. As the dry season length increases and water becomes more limiting (figure 4.9, rows 2 and especially 3), the LAI and TWS signals become ever more synchronous.

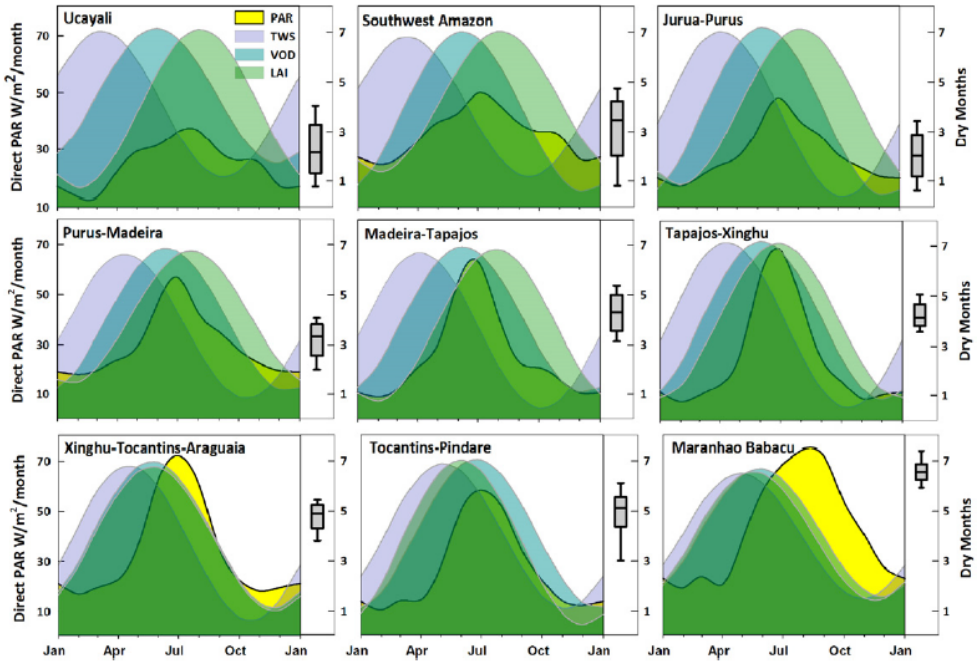


FIGURE 4.9: Sinusoidal model climatology for TWS, VOD, LAI and direct PAR monthly climatology and statistics (box plots) by ecoregion. [160]

Beyond the influence of average climate seasonality on phenology, changes in the range (maximum and minimum) and frequency of temperature and precipitation extremes might have strong effects on tropical phenology [158]. In this regard, the consequences of extreme droughts related to global climate dynamics (e.g. El Ninõ Southern Oscillation, ENSO) has been extensively discussed. Whereas [166] evidenced that intact Amazon forests were drought-sensitive, accumulating less carbon during the dramatic 2005 drought period, Saleska et al. [167] found that intact Amazon forests may actually be more drought resistant than expected after noting significant increases in EVI values following the same 2005 drought. To explain this apparent contradiction between EVI increases after drought events despite tree mortality increases [102], Brando et al. [168] suggested that "drought could increase EVI by synchronizing leaf-flushing via its effects on leaf bud development and tree water status". Yet, [169] consider that the EVI increase may especially reflect structural changes in the forests where emergent and large trees are more affected by droughts, thus reducing the heterogeneity of the canopy reflectance.

4.5.2 Impacts of climate change on forest phenology/degradation

The importance of climate as a driver of forest phenology raises important issues on the potential impacts of climate change on forest ecosystems. As already mentioned in section 4.2, climate projections point to an increasing frequency and severity of drought events [170, 171] and to the lengthening of the dry season [172, 173] in the Southern Amazon. The extent to which long-term droughts (especially related to more intense and frequent El Niño events) or long-term climate change may affect forest ecosystems is a major concern [167, 174, 175].

With regards to droughts, the severe 2005 and 2010 droughts resulted in decreased carbon storage due to the temporary cessation of biomass production and increased biomass loss via tree mortality. Considering that droughts are also accompanied with peaks of fire activity, it calls into question the potential of Amazon forests to act as a carbon sink or as an accelerator of carbon emissions on decadal-scale [176, 177]. Yet, monitoring the impacts of droughts on forests through remote sensing is challenging. As for the understanding of seasonal phenological cycles, studies have led to opposite conclusions. For example, Xu et al. [178] evidenced a widespread browning of Amazon vegetation after the 2005 and 2010 droughts whereas Atkinson et al. [179] concluded from their analysis that it is not possible to detect the response of vegetation to drought from space using VIs.

With regard to the impact of long-term climate change on vegetation, increased photosynthesis due to elevated CO2 concentrations and reduced transpiration related to higher temperatures and lower precipitations may lead to a climate-driven forest dieback [180]. Indeed, drier and warmer climate are expected to support grasses and shrubs over trees [168]. In this regard, Hilker et al. [165] already noted a decline in tropical Amazonian vegetation as a result of reductions in precipitation since 2000. Atkinson et al. [179] also reported an overall declining trend in vegetation greenness (through the observation of MODIS EVI and NDVI time series) over the period 2000 to 2010, but noted that theses trends were only significant at the south and southeastern edges of the intact forest, i.e. on the pioneer frontier.

Chapter 5

Remote sensing: more than a data science

Remote sensing data and associated techniques are very powerful tools to study the interactions between the environment and the anthropogenic activities. Yet, remote sensing is also a science *per se*, which has evolved rapidly during the last decades, accompanying the advent of the digital era. In this context, I spent some time during the last years reflecting on my potential contributions to the development of advanced methodologies in remote sensing science. To do so, I first identified the main research directions in remote sensing (section 5.1) and then decided to pay a special attention to the use of ontologies, which I introduce in section 5.2. Yet, although very promising, the potential contribution of ontologies to advance remote sensing science remains very theoretical to date. In the last years, I led two publications to introduce ontology-based remote sensing applications. The first application (which is actually the most recent [181]) relies on ontology principles to promote user-adaptive classifications of Sentinel-2 time series (section 5.3). The second application [182] introduces a formal ontological framework to classify Landsat images (section 5.4). Finally, I discuss some perspectives to use ontologies to advance remote sensing science (section 5.5).

5.1 Overview of current research trends in remote sensing

I conceive the main research trends in remote sensing science as in figure 5.1. I listed these research directions as data-driven image processing, knowledge-driven image processing, big EO data image processing, citizen science and ontologies. In this section, I review the main strengths and weaknesses of each research topic and I explain how I

contributed (or intend to contribute) in each of these topics. More importantly, I emphasize the reasons why I think we need a paradigm change based on the management of expert knowledge.

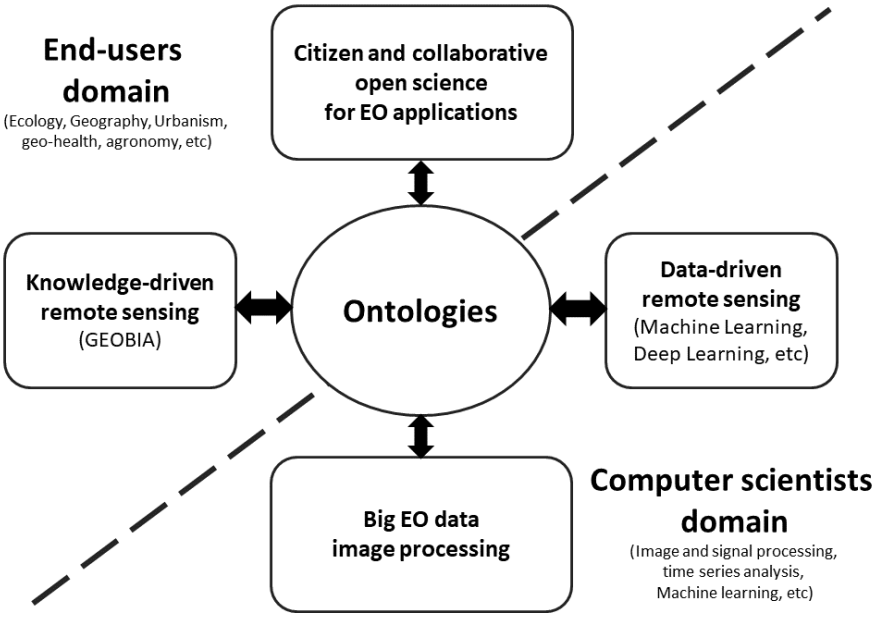


FIGURE 5.1: The current trends in remote sensing science and the pivotal role of ontologies to unify remote sensing and narrow the gap between end-users and computer scientists.

5.1.1 Data-driven remote sensing: The advent of Deep Learning?

A first major research direction in remote sensing consists in the development of data-driven approaches for an optimal use of the quantitative information contained in satellite images. The objectives are then twofold: (1) to improve our understanding of the correlations between the image signal and biophysical variables [183, 184] and (2) to implement efficient algorithms to classify images. In this regard, the last decades have been marked by the rapid adoption of ever more efficient classification algorithms like Maximum Likelihood, Random Forest [185] or Support Vector Machines (SVM) [186] by the remote sensing community. More recently, the advent of Artificial Intelligence, especially through Deep Learning algorithms, appears to be especially promising as it enables to achieve state-of-the-art results regarding map accuracy [187].

To date, I participated to researches relying on all of these techniques. During my PhD, I classified cropping systems based on MODIS EVI time series using a Maximum Likelihood classifier [188]. Later, I participated to a similar work to map land cover types in Mato Grosso based on SVM classifications of MODIS VI time series [14]. More recently,

I contributed to the researches by Kuchler et al. [60] to map crop-livestock integrated systems in Mato Grosso based on a Random Forest classifier (still applied on MODIS VI time series). Finally, I contributed to the work of Tang et al. [57] who implemented a new method that combines a lightweight real-time object detection network (PVANET) based on deep learning, an image classification model (GoogLeNet) and accurate shape detection (Hough transform) to detect and accurately delineate center pivot irrigation systems in Mato Grosso. Beyond their scientific interests, these researches were also important opportunities to establish and/or strengthen collaborations with research institutes such as INPE, EMBRAPA, CIRAD and the Chinese Academy of Sciences (RADI research unit).

However, these experiences also helped me to raise awareness about some major limitations of data-driven approaches for remote sensing applications. First, supervised algorithms require many training data, which is a severe issue since it can induce heterogeneity in the final products depending on the availability and reliability of reference data collected through field campaigns, crowd-sourcing initiatives [189, 190] or very high resolution maps (based on UAVs for example [191]). This assertion is especially valid for Southern countries where major land use changes are expected to occur.

Second, supervised classifiers are often used to produce maps where classes are mutually exclusive with discrete boundaries separating each other. Such approach is not adapted to the monitoring of vague and ambiguous geographic concepts such as vegetation gradients or urban areas. Indeed, the conversion of continuous quantitative information (i.e. spectral signatures) into discrete classes induces a loss of information or leads to an unrealistic representation of reality that may not suit with requirements for ecosystem monitoring [192]. Similarly, Kennedy et al. [193] note that remote sensing approaches based on static representations of geographic entities mainly allow detecting occasional dramatic disruptions (e.g. deforestation) but are limited to capture continuous changes (e.g. degradation).

Third, and most importantly from my perspective, supervised classifiers are inefficient to deal with symbolic systems based on inference and abstraction [194]. In data-driven approaches, the knowledge corresponding to a land cover/land use class is a "numeric knowledge" expressed in classification models (e.g. random forest model) and that suffers from a lack of interpretability [195] whereas most end-users (e.g. geographers, ecologists, agronomists, etc) are used to work with abstract semantic definitions of geographic features [196]. For instance, end-users can easily understand the Land Cover Classification System (LCCS [197]) entirely based on inference rules (e.g. a vegetated area covered by trees whose leaf type is "broadleaved" and leaf phenology is "evergreen" will be assigned to a land cover class named "broadleaved evergreen trees") but they can hardly interpret

a classification model of the same class produced by a random forest or a deep learning classifier. Thus, such algorithms do not match with the way end-users are used to work.

Main strength: Excels at perceptual classification and leads to state-of-the-art mapping results
Main weakness: Is data hungry and is inefficient to deal with inference knowledge

5.1.2 Knowledge-driven remote sensing: GEOBIA as a new paradigm?

In remote sensing, the main effort to promote knowledge-driven approach has been through Geographic Object-Based Image Analysis (GEOBIA [198, 199]), which allows end-users to produce accurate maps by classifying image objects based on end-user expertise [200, 201]. From a technical perspective, GEOBIA relies on two main tasks: (1) image segmentation to delineate image objects (i.e. pixels are grouped according to their measurement values) and (2) image classification to assign the image objects to semantic classes, usually based on expert-based decision rules. These tasks represent the major assets of GEOBIA. First, GEOBIA stands on the fact that partitioning an image into objects is akin to the way humans conceptually organize the landscape to comprehend it [200]. Second, GEOBIA uses semantics based on a descriptive assessment of expert knowledge, i.e. it incorporates the wisdom of the user [202].

Personally, I started working on GEOBIA with Laurent Durieux during my post-doctoral fellowship at IRD (2011-2014), under the frame of the *BIOSOS* FP7 European project. At that time, we tried to implement a robust decision tree to map habitats on very high resolution images [203]. In addition, I started using the InterImage software developed at the Pontifical Catholic University of Rio de Janeiro. Interimage is an open-source solution for object-based remote sensing image analysis. I used to lecture InterImage at Montpellier University and at the Gabonese Agency for Space Studies and Observations (AGEOS) in the early 2010s.

Yet, although GEOBIA has emerged in the 2000s [202] and has been considered as a new paradigm in remote sensing leading to better results than pixel-based classifications [198], its application highlights important scientific issues which can be summarized as follows. First, each GEOBIA expert has his own conceptualization of the reality he intends to represent on the image. Indeed, GEOBIA is characterized by the transformation of knowledge [204] and can be considered as a computer-aided photo-interpretation process [205] in which two experts analyzing the same data will achieve two different results because of their different experiences [206]. Second, the processing chain to produce a classification is not entirely controlled and documented. For example, the segmentation

process is an ill-posed problem [200] based on parameters among which the choice can be scarcely justified and thus hinders the quality assessment of the classification. The classification step itself is usually performed in a laborious way at the end of repeated cycles of "trial-and-error" analysis or using supervised classifiers. In other words, GEOBIA includes a fully data-driven image segmentation and an object classification stage based on rules that can also be considered as data-driven. From that perspective, GEOBIA may actually be considered as a hybrid knowledge- and data-driven classification approach restricted to local studies and difficult to transfer and automate for large-scale applications [203, 207].

Main strength: Integrates expert knowledge in the process of image interpretation
Main weakness: Produces uncontrolled processing chains difficult to be shared with other scientists

5.1.3 The challenge of processing big EO data

Dealing with big data is another current hot topic in remote sensing (as in many other sciences) [208]. The multiplication of remote sensors with high spatio-temporal resolution led to the release of a very high amount of EO data that needs to be processed and analyzed. To date, the remote-sensing community has not fully addressed the challenge to extract useful land cover information from vast volumes of satellite and ancillary data [209]. A major limitation refers to the choice of adequate classification schemes since many classes in existing land cover products (1) do not match with the end-user's requirements for specific applications and (2) are sometimes defined from an ecological or environmental perspective using criteria that may not be directly identifiable from remotely sensed imagery (e.g. vegetation height cannot be measured from optical imagery; [210]). The spatio-temporal resolution is also an issue since global land cover products are usually produced at a coarse resolution (250m to 1km) which is inadequate for many applications such as ecosystem mapping [192]. Recent successful examples of big EO data projects led with Landsat (30 m [13, 211–213]) or Sentinel (10 m [214]) data are important initiatives to tackle that issue. However, the methods used to process remote sensing data are usually based on data-driven supervised or unsupervised classification algorithms [209], which implies the selection of training and validation samples from baseline maps that may induce heterogeneity in the final products.

Overall, and as already mentioned in section 5.1.1, end-users often consider data-driven approaches as "black box" algorithms and their derived land cover products as not sufficiently documented [210]. As a partial answer, recent efforts have been carried out to share methods and algorithms (instead of sharing ready-to-use maps) designed for remote

sensing-based land surface monitoring (e.g. the CLASLite software [215], the EODHaM system [203], the *iota2* processing chain [216]). The application of such algorithms benefits from the emergence of online platforms dedicated to the cloud computing of large-scale geospatial data which allow both the production of high resolution maps and the sharing of image processing algorithms [217].

In my researches, I started working with the R environment to process remote sensing images in 2010. I am now able to process thousands of images in a few hours. Yet, I have to face new issues regarding the data acquisition and storing. Thus, I also intend to use online platforms such as Google Earth Engine (GEE) [217] to produce large-scale land cover maps.

But, whereas I am fully convinced that such technologies improving the transparency and democratization of high-performance technologies contribute to facilitate the processing of big Earth Observation (EO) datasets by scientists from varied scientific domains (i.e. not only remote sensing experts), I feel concerned by the fact it also -and rather paradoxically- supports the evolution of remote sensing from an environmental science to a computer science with a major focus on image and signal processing (e.g. improved classification algorithms, analysis of big EO data time series, etc) rather than on the applications. In this regard, I feel like the expert knowledge mobilized by environmental scientists to interpret remote sensing images tends to be somehow discarded from the image analysis.

Main strength: Automatic processing of big EO data leads to the release of regional and global products
Main weakness: The classification schemes combined with data-driven approaches often results in maps that do not match with end-user requirements

5.1.4 Towards a citizen science for Earth Observation

Another trend in remote sensing science refers to collaborative and open science. For example, many land cover products are freely available online [209, 210] and new image processing methods are also regularly released through libraries in open source software (R or Python for example). In addition, some products are now implemented collectively by joining various research units with different expertise (e.g. the MapBiomas project in Brazil [218]). But even more challenging, recent studies suggest moving towards citizen science for Earth Observation [219]. Citizen science has long existed for geographic studies [220] and is currently gaining importance in remote sensing. Examples of collaborative citizen science applied to Earth Observation include the Open Street Map

collaborative mapping [221] or the geo-wiki project designed to the online sharing of land cover training and validation sample points [222, 223].

Nonetheless, few of these crowd-sourced projects are integrated into operational monitoring systems to date [219]. Whereas collaborative and citizen science for EO remains mainly focused on the sharing of data (e.g. land cover maps or validation/training data) and methods (e.g. image processing chains), efforts are still needed to share knowledge about the theory and practice of remote sensing science. In this regard, precursory initiatives deserve to be spotlighted such as the CEOS data handbook [224] that lists remote sensors and their potential applications or the index database that describes all spectral indices used in remote sensing [225]. However, these remain isolated efforts that need to be better formalized in order to ensure interoperability and thus knowledge sharing. Overall, initiatives to share knowledge about how to use this information should be a priority. Indeed, the interpretation of remote sensing images is a subjective task whose success depends on the experience of the operator [206]. For example, knowing that forests can be mapped using a given remote sensor does not give you any indication on the best way to map forests in a satellite image. The subjective expert knowledge thus also needs to be formalized and shared, not only the objective knowledge about the description of image characteristics.

As for today, my personal experience in citizen/collaborative science is quite limited. However, the recent launch of the CHOVE-CHUVA project, funded by CNES under the frame of the Space Climate Observatory, stands as a first attempt to gather scientists and citizens around remote sensing applications. In that case, the citizens firstly refer to long-term partners, i.e. NGOs (Instituto Centro de Vida in Alta Floresta, Fundação Ecologica Cristalino in Alta Floresta), farmers (through the Clube Amigos da Terra association from Sorriso) and public administrations (Mato Grosso's state Secretary for the Environment). Although we know and work with these partners for a long time, our researches usually relied on a producer-driven approach, where we, as data scientists (e.g. map producers) imposed our own conceptualization of environmental features of interest to the end-users (e.g. we decided to map double cropping or irrigation systems by ourselves). We now consider that we need to develop a user-driven approach to reduce the gap between the data producers and the end-users. End-users needs, intentions and knowledge should be used to guide the data processing steps and to propose multi-uses tools for a wide range of stakeholders. Although maybe not going as far as this for the moment, the objective of the CHOVE-CHUVA project is to develop a tool for disseminating the information produced by scientific teams recently involved in various projects for monitoring territorial dynamics in Mato Grosso. It will then focus on the need to integrate end-users requirements and expert knowledge in the production of spatio-temporal indicators. In addition, the tool will also make it possible to collect

citizen data on the perception of climate change by Amazonian communities and on the location of certain types of land use advocated under the ABC plan, such as reforestation or fields cultivated in integrated systems. This project is in collaboration with a French private company (Alkante) and started in autumn 2021.

Main strength: Opens new perspectives by sharing data and methods within the scientific community
Main weakness: The subjective expert knowledge in EO image analysis is rarely formalized and shared through interoperable systems

5.2 Ontologies in remote sensing: why, what and how?

In my opinion, the major limitation of current research directions in remote sensing science stands on the difficult sharing of experiences in the use of remote sensing data, considering both end-users interested in freely released land cover maps or computer scientists focused on data- or knowledge-driven image processing of (big) EO data. For this reason, knowledge management can be an issue for the future of remote sensing. In this regard, whereas the recent advent of artificial intelligence in remote sensing is mainly based on machine learning for data-driven applications, I consider that the scientific community should pay more attention to Good Old-Fashioned Artificial Intelligence (GOFAI), e.g. artificial intelligence based on mathematics and logic to produce symbolic representations of abstract concepts [194]. More specifically, I think knowledge representation techniques such as ontologies should play a pivotal role at the articulation between knowledge- and data-driven approaches, thus contributing to the unification of remote sensing science by integrating different (not to say opposite) research directions and by reducing the gap between EO end-users and computer scientists (Figure 5.1). I believe ontologies should be used to store expert knowledge so that, in the long run, "big data" in remote sensing should not only be an issue of "big EO data" processing but also of "big knowledge data" management.

Unfortunately, as illustrated in figure 5.2 that shows the evolution of annual publications in these major research directions, the use of ontologies in remote sensing remains widely underexplored to date and their implementation appears very challenging. However, other indices (e.g. > 190 citations for a review paper on ontologies [226]) encourage me to believe that this issue deserves to be further explored by remote sensing scientists. As for today, I contributed to publish four papers on the use of ontologies in remote sensing and the difficulties we (my co-authors and myself) had to face to express our thoughts illustrate the long road ahead before this issue can gain importance. We first published a

review paper to explain "HOW" ontologies could help advancing remote sensing science [226]. Then, we wrote an application paper that was quickly rejected by three remote sensing journals before being rapidly accepted in a computer science journal [182]. The very critical feedbacks we received from editors and reviewers of the remote sensing community led me to consider that the first question to answer was not "HOW to use ontolgies in remote sensing?" but "WHY do we need ontologies in remote sensing?". This is the subject of the third paper [196], which initiated with the responses to the reviewers of the second paper [182]. Finally, a fourth paper was published recently [181] and the globally positive reviews encourage us to pursue in that direction of research. In the following sections, I tried to reorganize (and summarize!) the main findings of these different papers.

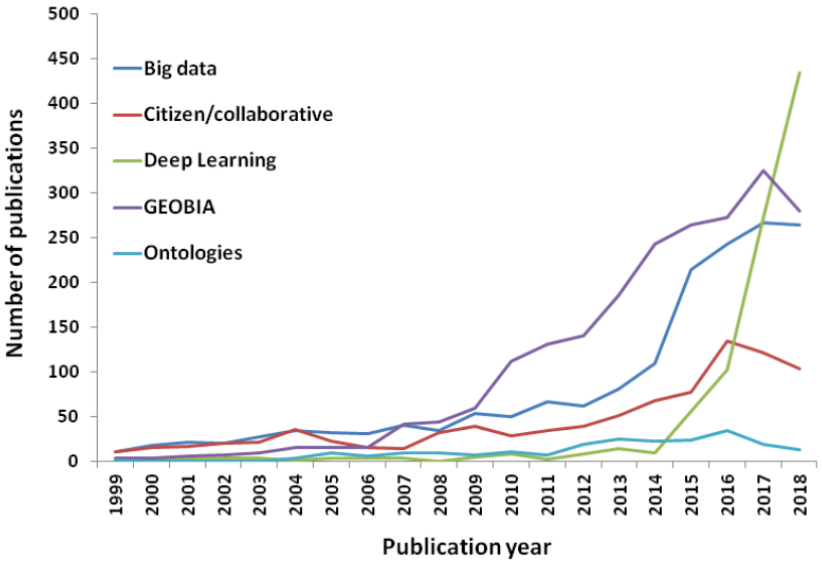


FIGURE 5.2: Number of annual publications (based on Web of Science) for different requests including (Remote Sensing OR Earth Observation) and one of the following terms: (ontology OR ontologies), (deep learning), (citizen science OR collaborative science), (OBIA OR GEOBIA OR Object-based), big data.

5.2.1 Why do we need ontologies?

The interpretation of satellite images is a complex task, from the selection of the relevant data for a given application to the image content analysis, which closely depends on the data characteristics and the users' skills [206]. This task consists in translating a "human perceptual knowledge" (i.e. the visual mechanism that enables a user to identify a class of interest in a remote sensing image) driven by "symbolic knowledge" (e.g. the expert knowledge about a class of interest such as "Forest" and about remote sensing image processing can make him to conclude that "Forest refers to high NDVI values

in optical remote sensing images") into "numeric knowledge" (e.g. the implementation of a classification rule set, e.g. $Forest = (NDVI > 0.6)$). In remote sensing, the symbolic knowledge is usually implicit since the expert generally focuses on the "numeric knowledge" by 1) directly implementing classification rules based on trial-and-error tests or 2) running a supervised classifier (e.g. in a decision tree).

More generally, the entire process to convert a "perceptual knowledge" into a "numeric knowledge" raises important issues that need to be discussed in order to understand some conceptual limitations in the classification of remote sensing images (Figure 5.3).

5.2.1.1 The complex definition of geographic concepts

The first issue concerns the definition of the classes of interest. Indeed, geographic concepts can have many definitions depending on the physical, historical, functional or conventional mode of definition [227, 228] (Figure 5.3, 1 – Modes of definitions). These modes represent alternative points of view of the classification of objects, i.e. a concept can be defined by alternative definitions that are not strictly equivalent, although they are naturally correlated [229]. For example, a forest can be described from a functional viewpoint in which the main interest is the role of forests in storing carbon (by considering NPP), or a forest can also be described based on its historical properties in order to consider only primary forests, or the focus may only be in forests of the Legal Amazon, thus considering a legal definition of forest.

But, certainly more important for remote sensing applications, a land cover class can be defined based on its physical description. In this regard, major efforts are required to standardize the description of land cover classes (e.g. the Land Cover Classification System (LCCS) [197]). Yet, still considering the case of "Forest", various organizations and countries refer to their own definition [230]. For example, in Brazil, a "forest" is defined as "an area greater than 1 hectare, with more than 30 percent canopy cover and a minimum tree height of 5 meters." In China, a forest is a piece of "land with a minimum area of 0.67 ha, minimum 20% crown cover and a minimum tree height of 2 meters." As a consequence, when a remote sensing expert produces a map based on a numeric knowledge (i.e. any classification model), the definitions of the land cover classes usually remain implicit and resulting maps have to be interpreted with caution by end-users or remote sensing colleagues.

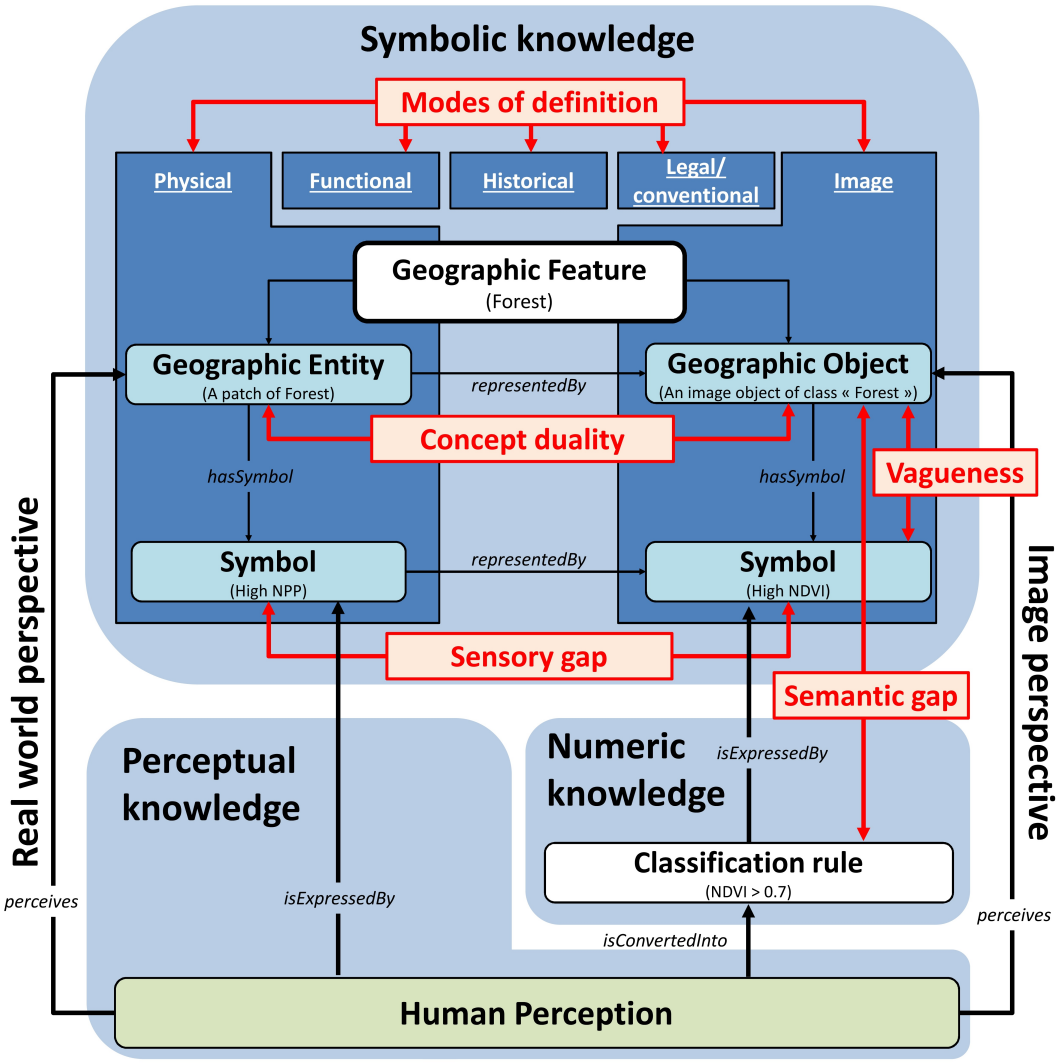


FIGURE 5.3: Schema followed to formalize expert knowledge, according to the real-world and the image perspective. In both cases, the human perceptual knowledge is expressed by a symbolic knowledge (e.g. a "Forest" is a geographic entity with a "high NPP" or "Forest" is a geographic object with a "high NDVI" value). In the image perspective, the symbolic knowledge is then converted into a numeric knowledge (e.g. Forest = (NDVI ¿ 0.7)). In that case, it is worth noting that the symbolic knowledge is actually often implicit since remote sensing experts directly convert their perceptual knowledge into classification rules. The five conceptual issues are then identified in red.
The colour version of this figure is available on the on-line journal paper.

5.2.1.2 The concept duality of geographic concepts

In addition to these modes of definitions, a remote sensing expert has to deal with the image perspective. Strahler et al. [231] distinguish "between the scene, which is real and exists on the ground, and the image, which is a collection of spatially arranged measurements drawn from the scene". More concretely, a land cover class can be defined both from the real world perspective, based on its physical properties (e.g. "a forest has high NPP values") and from the image perspective, based on its image properties (e.g.

"a forest has high NDVI values"). The real-world and image definitions of the same abstract concept are hereafter referred to as concept duality (Figure 5.3, 2 – Concept duality).

The concept duality implies an important vocabulary issue. A geographic feature is an abstraction of a real-world phenomenon (e.g. for example, the concept of "Forest"). A geographic feature is then considered as a superclass including both sub-concepts of a geographic entity and a geographic object [232]. A geographic entity is a real-world entity characterized by attributes (e.g. NPP, leaf type, phenology, tree height, etc.) which are assigned attribute values (e.g. 2000 g/m2/year, broadleaved, evergreen, 15 meters, etc.) and that occupies a position in space (e.g. a patch of forest) [232]. The representation of a geographic entity in an image is called a geographic object where an object is a "thing" in the digital world, which represents real-world phenomena as an instance of a generally recognized category [233]. This geographic object is also characterized by attributes (e.g. NDVI) and its corresponding values (e.g. 0.8). Finally, it is worth noting that the observation of a geographic feature in a remote sensing image depends on the choice of an appropriate scale of analysis, or spatial resolution, for a given application.

5.2.1.3 The vagueness and ambiguity of geographic concepts

Moreover, the definitions of geographic concepts are often vague and ambiguous (Figure 5.3, 3- Vagueness). On the one hand, Bennett [228] identify different kinds of vagueness. Firstly, threshold vagueness arises when qualitative terms (e.g. "high") are used to distinguish objects that in fact exhibit continuous variation in the observables (e.g. NDVI). Asserting that "high NDVI" values represent forests is thus vague since it requires setting a user- and context-dependent threshold in a continuous gradient of NDVI values. Secondly, partiality refers to the fact that "high NDVI" is a partial description of the "forest" concept in remote sensing images since it does not make it possible to either (1) map only forests (other classes such as "crops" may also have high NDVI values), or (2) to map all the different types of forests (some "degraded forests" for example may have lower NDVI values).

On the other hand, ambiguities cover all cases in which natural language expressions can have more than one meaning [234]. The most common ambiguity is lexical ambiguity, which "arises due to homonymy of a natural language expression, i.e. an expression that has several lexical meanings, where each of these meanings can be mapped to one ontology concept unambiguously" [234]. For example, Lund [235] gathered more than 800 definitions of the "forest" concept across the world, each of them being explicitly

defined when considered separately. In addition, deep ambiguity (or open texture) may also exist when "a term has no crisp boundaries and may shift its meaning over time as new technologies appear, as people develop new habits, and in general, as the social and physical context of the term evolves" [236]. For instance, Chazdon et al. [230] considers that commonly used definitions of forest have limited utility for assessing and monitoring new and diverse forms of forest cover such as "reforestation", "degraded forest" or "logged forest". In this regard, it is worth noting that vagueness differs from ambiguity: whereas threshold vagueness can be reduced by defining more precise rules, deep ambiguity cannot [237].

5.2.1.4 The sensory gap

The sensory gap derives directly from the dual definition of abstract geographic features (Figure 5.3, 4 – Sensory gap). It is defined as the gap between the entity in the world and the information in a (computational) description derived from a recording of that scene [238]. For example, Congalton et al. [210] emphasize that land cover products are often based on classification schemes (such as LCCS) that are defined from an ecological or environmental perspective and include some characteristics (e.g. vegetation height) that cannot be identified from optical remotely sensed imagery used to produce these maps (e.g. Landsat data). According to Smeulders et al. [238], the sensory gap is an ill-posed problem that yields uncertainty in the description of geographic objects.

5.2.1.5 The semantic gap

The semantic gap is closely related to the vagueness and ambiguity of the geographic concepts. The semantic gap is the lack of agreement between the information that one can extract from the visual data and the interpretation made of the same data by a user in a given situation [238]. It describes the gap between the low-level representation of an image (e.g. NDVI values) and the high-level semantic concepts used by users to interpret the image (e.g. forest) [238] (Figure 5.3, 5- Semantic gap). The main explanation for this gap is that humans trigger perceptual mechanisms (i.e. "perceptual knowledge") to draw inferences from what they perceive in remote sensing images [239]. These mechanisms of human observation cannot be easily translated into computational representation, i.e. into "numeric knowledge".

5.2.2 What are ontologies?

Based on the above-mentioned considerations, it appears that remote sensing experts are mainly used to work with numeric knowledge from an image viewpoint (e.g. $Forest = (NDVI > 0.6)$). This approach implies that experts use partial and implicit representations of their expert knowledge, which is then difficult to share with other scientists (e.g. in ecology, agronomy, urban studies, geo-health, etc.) more used to work with symbolic knowledge and real-world concepts (e.g. in ecology, a forest might be defined by concepts such as broadleaved, evergreen or trees and does not need to be linked to image features). In this context, formal ontologies are an especially promising way to enhance remote sensing science since they make it possible to represent both symbolic and numeric knowledge about both real-world and image viewpoints.

5.2.2.1 A brief definition of ontologies

A widely accepted definition of the term "ontology" is the one proposed by Gruber [240], who defines an ontology as a formal, explicit specification of a shared conceptualization. The conceptualization is an abstract, simplified view of the world that we want to represent for a specific purpose [240]. An ontology intends to identify the concepts and their relationships within a scientific domain. In addition, the conceptualization needs to be explicit, shared and formal to be considered an ontology. Here, explicit means that concepts and constraints are precisely defined. A shared conceptualization means that the ontology captures a consensual knowledge. Finally, a formal conceptualization means that the ontology is machine understandable.

When these three qualities are met, the conceptualization can be implemented in dedicated software (e.g., Protégé) that allows exploiting all of the potentialities of ontologies. More specifically, ontologies are implemented according to standard formalisms such as the Web Ontology Language (OWL). OWL is a W3C (World Wide Web Consortium) standard related to description logics (DLs) [241], which is a valuable support for the semantic web as the concept and role hierarchies fit the context of the web particularly well. In DLs, a given subject area is described in terms of (1) concepts (or classes) organized in a hierarchy of concepts (e.g. "Forest" is a sub-class of "Vegetation"), (2) roles (or properties) organized in a hierarchy of roles (e.g. "is Adjacent To" is a sub-role of "has Spatial Relation") and (3) individuals (e.g. an image object classified as "forest" is an instance of the "Forest" concept).

The figure 5.4 introduces a concrete example of the implementation of an Ontology considering a simple classification rule such as $Forest = (NDVI > 0.6)$. Note that,

different from "traditional" remote sensing, both the symbolic knowledge and numeric knowledge are explicitly formalized based on DLs so that they can then be implemented in an ontology editor (e.g. Protégé). Based on OWL and DLs formalisms, the reasoner algorithms then enable new implicit knowledge to be inferred (see section 5.4) or to check for inconsistencies (e.g. estimating vegetation height based on multispectral images may not be reliable).

5.2.2.2 Ontologies as ways of explicitly specifying bodies of knowledge

The formalization of concepts and relations expressed in ontologies is a key to addressing the major issues listed in Section 5.2.1.1. Formal ontologies can be used to explicitly describe an observation (e.g. of a forest) from different perspectives. For example, the Extensible Observation Ontology (OBOE [242, 243]) was specially designed to describe the semantics of scientific observations. In OBOE, an observation of an entity (e.g. a tree) is made through the measurement of a characteristic of this entity (e.g. biomass) based on a measurement standard (e.g. gram). A similar approach can also be used to describe geographic features: a forested area can be described thanks to one characteristic (e.g. NPP) and its associated value (e.g. "HighNPP"). In that case, the description corresponds to the functional definition of a forest, but other characteristics can be included in order to consider other modes of definition. Indeed, measurements of vegetation cover (in %), vegetation type (e.g. trees or shrubs), leaf type (broadleaf or needle), phenology (evergreen or deciduous), vegetation age (in years) may complete the description of a forest using physical and/or historical modes of definition. Further, the observation of an entity (e.g. forest) can be associated with a context to enable the description of the relationships between the entity and other entities at the time of the observation (e.g. the entity of forest is located within the Amazon basin).

Ontologies also allow perspectivalism, that is, separating the field point of view used to describe the properties of the feature in the field reality (e.g., high NPP, leaf type) and the image point of view used to describe the characteristics of the object in the image (e.g. high NDVI, texture, wavelength). This approach takes into account the concept duality of geographic features (including both geographic entities and geographic objects) and improves the semantic interpretation of satellite images. A geographic feature is thus described based on its image characteristics (see sections 5.3 and 5.4). In this case, a pixel (or image object) of a given class (e.g. Forest) is described by an image feature (e.g. NDVI) and its associated value ("HighNDVI") that results from the application of an image processing task (e.g. applying a threshold on an NDVI image). The formalization of the "HighNDVI" concept makes it possible to clarify the relationship between symbolic (e.g. "HighNDVI") and numeric knowledge (e.g. $NDVI > 0.6$) thereby reducing the

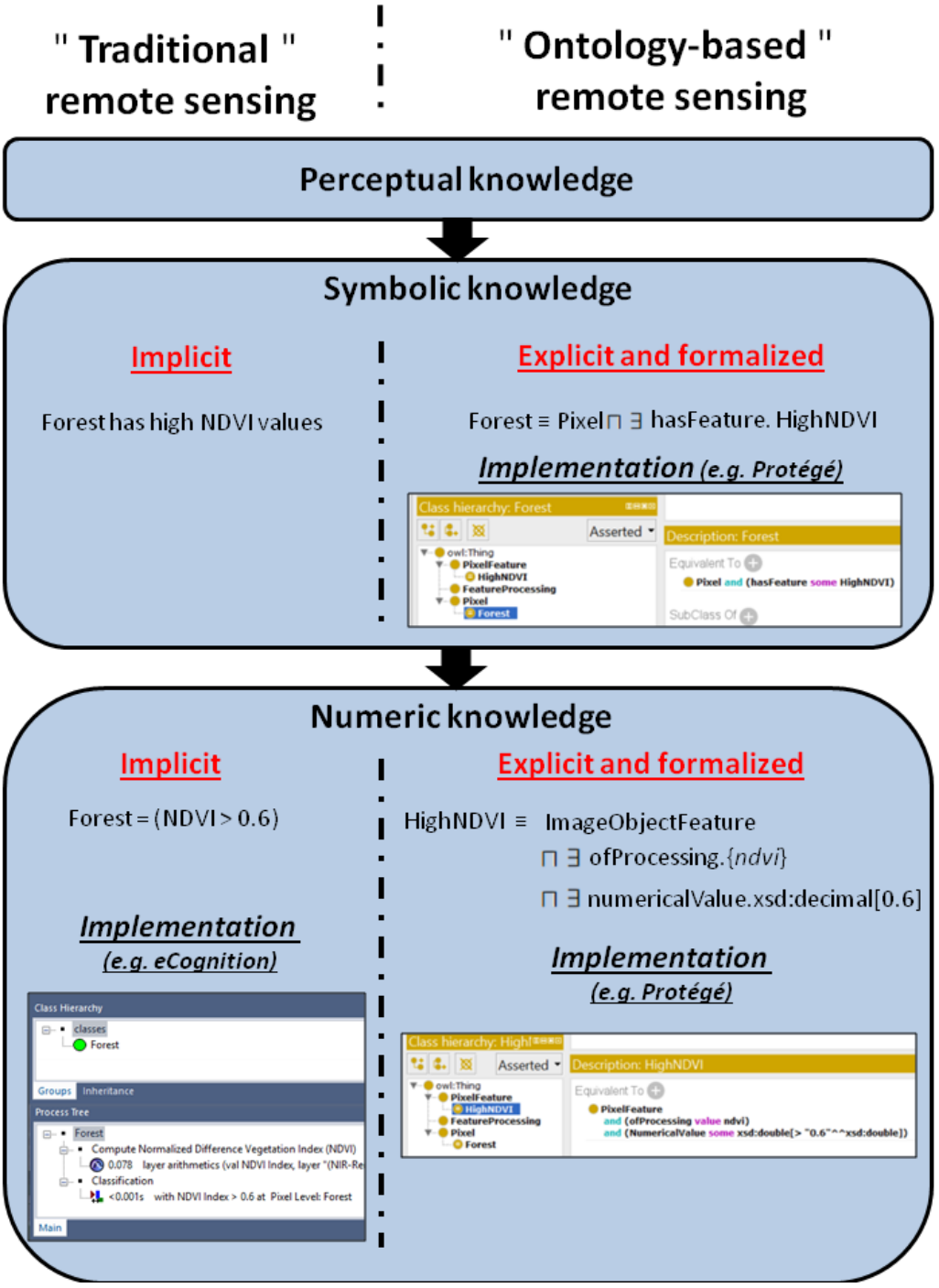


FIGURE 5.4: Implementation of a classification rule in knowledge-driven remote sensing using a "traditional" and an "ontology-based" approach. In "traditional" remote sensing, the perceptual knowledge is implicitly converted into symbolic and numeric knowledge and finally implemented in a remote sensing software (e.g. eCognition). In "ontology-based" remote sensing, both the symbolic and numeric knowledge are explicitly formalized based on Description Logics. They are then implemented in ontology editors such as Protégé. In Description Logics, the symbol ≡ stands for equivalence, the symbol ∃ stands for existential quantifier and can be interpreted as "there exists" and the symbol and the symbol ⊓ stands for intersection. In this example, the conceptual framework to model the classification rule is inspired from Andrés et al. (2017) and can thus be interpreted as follows. At symbolic level, a "Forest" is defined as a pixel that has a characteristic ("hasFeature") which is "HighNDVI". The concept of "HighNDVI" is then defined as a "PixelFeature" resulting from a processing ("ofProcessing") named "ndvi" ("ndvi" being an instance of a "FeatureProcessing" concept) and whose numerical value is greater than 0.6.

semantic gap. Similarly, the description of both real-world geographic entities and their corresponding geographic objects facilitates connections between concepts from both perspectives (for example, between NDVI and NPP), thus reducing the sensory gap (or at least making it explicit).

The explicit formalization of geographic features is also a good way to address the problem of ambiguity. For example, lexically ambiguous concepts can be described by different definitions (e.g. the 800 definitions of the forest concept [235]). However, it is worth noting that formal ontologies do not solve the deep ambiguity issue (e.g. to consider a reforested area as a forest is still ambiguous) but rather force the experts to clarify how they define the concepts used in a given application (e.g. the difference between a forested and a reforested area must be explicitly defined).

Ontologies can also help to address vagueness issues, i.e. threshold vagueness and partiality. On the one hand, there exist fuzzy ontologies that can be used to address threshold uncertainty, since concepts are defined as sets with vague definitions. In that case, qualitative measures like green area, round lake, etc. are given a degree of truth to indicate the degree of greenish, roundness, etc. [244]. On the other hand, formal ontologies are based on the open world assumption (OWA), which makes it easier to handle partial definitions of vague concepts. The OWA works under the assumption that knowledge of the world is incomplete so that anything is true or false unless the contrary can be proved [245]. For example, two concepts overlap unless they are declared as disjoint, or a fact that does not belong to the knowledge base cannot be considered to be false. OWA is quite new to remote sensing experts whose methods are usually based on the closed world assumption (CWA). Indeed, a set of classes of interest is usually defined *a priori* (in case of supervised classification) and all pixels or image objects are assigned to one (and only one) of these classes. While this approach may be appropriate for broad land cover classes (e.g. vegetation, water, bare soil, etc.), it may be conceptually limited to map finer classes (e.g. agricultural practices, flooded areas, forest degradation) and OWA may be relevant.

5.2.2.3 Ontologies to manage and share knowledge about the interpretation of remote sensing images

Although both symbolic and numeric knowledge can be represented in formal ontologies, the focus is clearly on the symbolic knowledge (e.g. "Forest" has "HighNPP" or "HighNDVI" values) rather than numeric knowledge (e.g. $Forest = NDVI > 0.6$). It thus differs markedly from traditional approaches in remote sensing where the knowledge is implicitly contained in image processing tasks, as in an image analysis processing chain.

If formalized in an ontology, the symbolic knowledge turns out to be a new datum, separated from the image and the associated processing tasks, that can be handled and shared easily.

Working with symbolic knowledge independently from numeric knowledge enhances knowledge management. It is possible to formalize a few pieces of knowledge (e.g. concepts and relations) without affecting the entire conceptualization (whereas changing a rule in an image processing chain may require adapting all its child processes). It also facilitates collaborative and interdisciplinary research by connecting concepts from different scientific domains (in ecology, agriculture, geo-health etc.). Improved knowledge management also enhances knowledge sharing. For example, to share classification rules implemented in an image processing chain means sharing a set of processing tasks that have proven to be efficient for a defined remote sensing dataset. Colleagues who wish to reuse the processing chain then will need to adapt it to their own dataset. To do so, they will need to "capture" the symbolic knowledge triggered by the person who initially implemented the processing chain (but did not make his/her symbolic knowledge explicit) in order to build new numeric knowledge (e.g. to set down new rules with thresholds) better suited for their case studies. In this regard, different users may have different interpretations of the implicit symbolic knowledge contained in the same processing chain. With formal ontologies, it is possible to limit the subjectivity of the remote sensing analysis by directly sharing the symbolic knowledge through OWL repositories handled by semantic data stores. For example, Jena TDB allows users to store domain knowledge using OWL formalism. Stored domain knowledge can be queried using the CRUD operations (Create, Retrieve, Update, Delete). These OWL repositories represent a new modular interdisciplinary contextual knowledge base about ever more diverse remote sensing applications and may turn out to be a new object of research *per se*.

Finally, the explicit specification of the knowledge used to extract information from remotely sensed data allows the integration of remote sensing classification products from other information sources. Interoperability is a major advantage of ontologies, especially in a context in which the technical focus has moved from the accuracy of the remote sensing information extraction and inventory to fitness for the purpose of integration of various spatial sources of information in support of specific policy goals (GEO's goal 2016–2025).

5.3 User-adaptive automatic classification of Sentinel-2 time series

In this section, I introduce a recent applied research I conducted that relied on ontologie principles. This research starts form the observation that large-scale land cover maps (regional national or even global) are essential for many applications but the classification schemes (e.g. including a limited a number of land cover classes defined by the map producers) and the methodological approaches (e.g. usually data-driven approach such as supervised classifiers) present important limitations that may deepen the gap between the map producers and the end-users. In this context, I worked on a novel knowledge-driven methodological approach to automatically provide symbolic descriptions of time series of Sentinel-2 spectral signatures. These descriptions are based on low-level characteristics in order to avoid hasty over-interpretation with high-level semantic classes and thus reduce the sensory and semantic gaps. Then, the low-level descriptions can be adapted by the end-users to match their requirements. This work firstly relies on a conceptual framework inspired from ontological modelling and includes a concrete application in the Brazilian state of Mato Grosso. The following sections represent a (incomplete) summary of the very long published paper [181] of this research whereas the table 5.1 defines the main terms used in this study.

Term	Definition
Ontology	Formal, explicit specification of a shared conceptualization [**?**]
Geographic Feature	Abstraction of a real-world geographic phenomenon
Geographic Entity	Real-world entity that occupies a position in space [232]
Geographic Object	Digital representation of a geographic entity [232]
Intentionality	Purposes, intentions, motivations, needs, beliefs of an end-user [246]
Observables	Properties that depend on perceptual experiences, e.g. colour and texture whose description is subjective and difficult to express [247]
Similarities	Properties that refer to measurable and objective characteristics, e.g. vegetation height [247]
OBOE	Formal ontology for capturing the semantics of scientific observation and measurement [242]
Knowledge-driven	Methods that depend on *a priori* knowledge, i.e the classification model exists regardless of the data
Data-driven	Methods that depend on the data, i.e the classification model does not exist if there is no data
EUEK	End-User Expert Knowledge, usually implicit, symbolic and vague
RSEK	Remote Sensing Expert Knowledge, usually numeric and implicit

TABLE 5.1: Definitions of main terms used in this section.

5.3.1 Ontological conceptual framework

The approach relies on a conceptual model that allows describing the way experts observe geographic features of interest in remote sensing images. To do so, we inspired from the Extensible Observation Ontology (OBOE) [242, 243], which provides a formal conceptual framework for capturing the semantics of scientific observations and measurements [242]. To make it short, OBOE asserts that an *Observation* of an *Entity* is carried out through the *Measurement* of one (or more) *Characteristic* using a measurement *Standard*. Finally, the *Measurement* has a *Value*, which can either be numeric (e.g. a reflectance value) or literal (e.g. a class label) (Figure 5.5, A). The OBOE ontology is called "extensible" because key concepts such as *Entity*, *Characteristic* or *Value* can be extended to include sub-concepts from a specific domain such as Earth Observation. For example, the *Forest* concept can be defined as a sub-class of *Entity*, *Leaf phenology* as a sub-class of *Characteristic* and *Evergreen* as a sub-class of *Value*.

In addition, an efficient use of OBOE framework requires investigating the roles of language and cognition in humans' conceptualization of geographic concepts [248]. Couclelis [247] introduced an ontological system to describe the observation of geographic features. This system is based on principles of cognitive geography, which is about the way humans understand, categorize and act in the geographic world [246]. It takes the form of a suitably ordered *semantic contraction* made of seven levels differentiated by their degree of complexity, from highest (level 7) to lowest (level 1) (Table 5.2). According to Lang et al. [249], "the *semantic contraction* is a kind of backtracking concept from rich geospatial information that, stepwise, loses purpose, function, composition, and ultimately its space–time relations, down to a piece of information without spatial or temporal context."

5.3.1.1 Purpose and existence

The main peculiarity of this hierarchy is that any geographic information is understood as a function of at least two kinds of separate considerations: purpose (intentionality) and existence (scale) [246]. These two considerations (intention and scale) represent the two ends of the *semantic contraction*, i.e. level 7 and level 1 respectively. Intentionality refers to the purposes, intentions, motivations, needs, beliefs, and so on, of an observer or a user of the system [247]. For example, an ecologist may consider a hedgerow separating two neighboring cultivated fields as a corridor that allows forest species to move in an agricultural landscape, whereas an agronomist may consider the same hedgerow as an obstacle that impedes disease dispersion between two crop fields. These different point-of-views will drive the intentions of these two scientists at the time of mapping

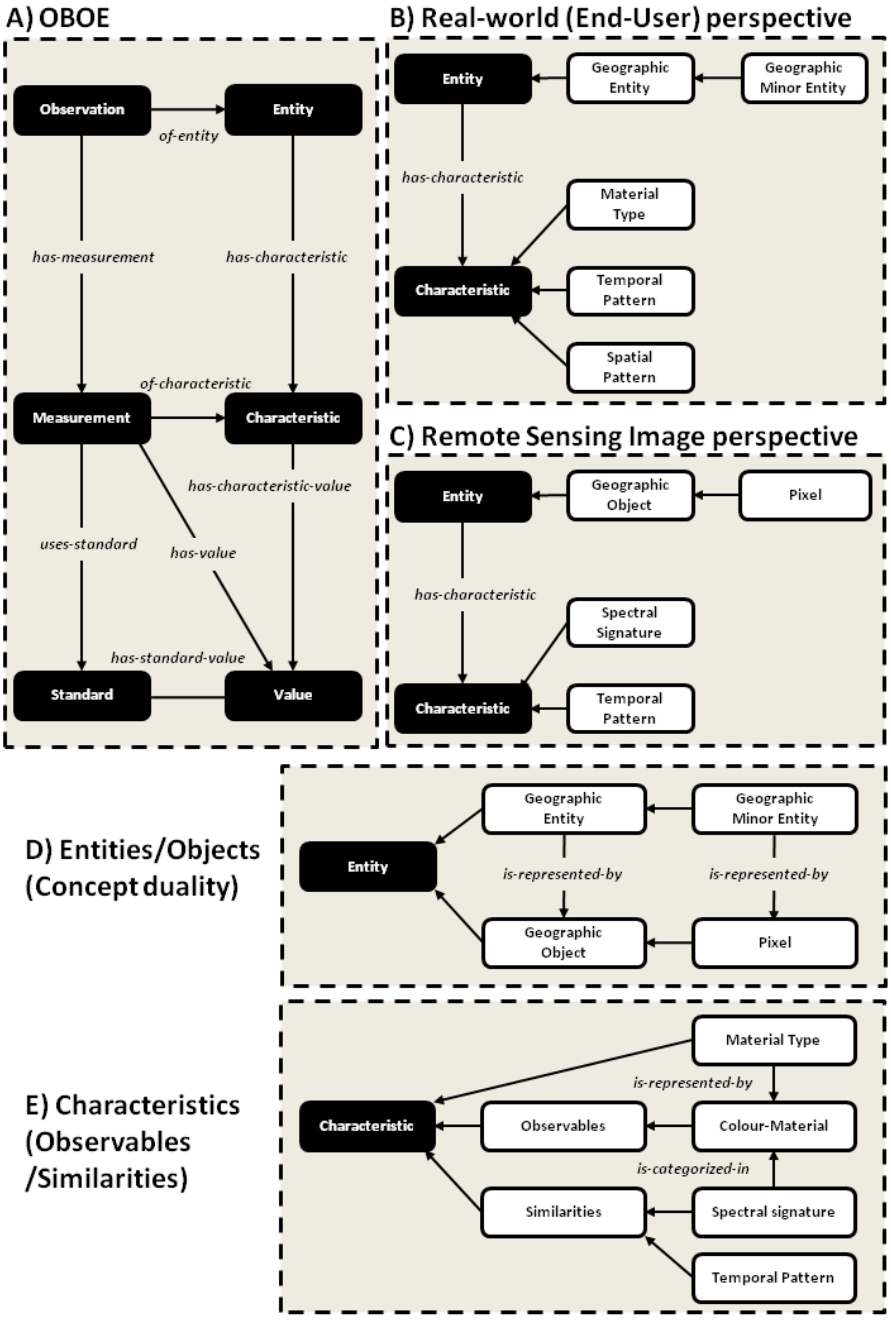


FIGURE 5.5: Description of the ontological conceptual framework including (A) the Extensible Observation Ontology (OBOE), its application for the description of (B) geographic entities and (C) geographic objects, and its extension through the *Entity* (D) and *Characteristic* (E) concept.

hedgerows, and consequently impact on their mapping requirements. The ecologist may require hedgerows to be represented as superficial areas to focus on their characteristics (e.g. width, internal heterogeneity) while the agronomist may represent them as lines.

In consequence, the intention of the end-users impacts the scale of analysis, which consists in identifying an appropriate spatio-temporal framework that enables to observe an entity [247]. It thus refers to the existence of the entity at a particular granule of

Semantic resolution level	Example: A map of Permanent Protected Areas in Mato Grosso, Southern Brazilian Amazon
7 Purpose	Identify Permanent Protected Areas to be restored
6 Function	Represent the areas where natural vegetation has been degraded
5 Composite objects	A network of patches of natural vegetation
4 Simple Objects	Patches of natural vegetation
3 Similarities	Fields of measurable properties exhibiting diverse geometrical patterns (e.g. occurrences)
2 Observables	Field-like properties (e.g. Colours)
1 Existence	The relevant information exists at given spectral, spatial and temporal resolution

TABLE 5.2: Steps of the semantic contraction proposed by Couclelis [247] and an example of its application. Shaded rows identify the steps addressed in the present application. I focused on levels 1-3 of the semantic contraction corresponding to the field representation of geographic entities (pixel-based), in opposition to object representation at levels 4 to 7.

space and time. It is then an issue to determine the geographic features that can be observed in the data to be processed, e.g. 10 m resolution Sentinel-2 images at a 5-10 days temporal resolution. With regards to the temporal resolution, a critical issue to consider is: does a given remote-sensing tool describe change over time in a manner consistent with the ongoing process of interest (e.g. ecological processes, deforestation, urbanization) ? [193]. For example, information on vegetation phenology does not exist in single date remote sensing images and requires a multitemporal analysis. Similarly, considering that the spatial resolution of a sensor system should be less than half the size of the feature measured in its smallest dimension [183], it gives an indication about the kind of geographic features that can (e.g. fields, forests, urban areas, large farm dams, etc) or cannot (e.g. plant crops, individual trees, houses, swimming pools, etc) be observed in 10 m resolution remote sensing images.

5.3.1.2 Entities

Between these two extremities (i.e. purpose and existence), the levels 4-7 of the semantic contraction are focused on the description of geographic entities. The *Entity* concept of the OBOE ontology was then extended to introduce the *Geographic Entities* of interest for the end users (Figure 5.5, B). A *Geographic Entity* is a real-world entity that occupies a position in space (e.g. a forest, a building), i.e. "for every set of coordinates in space, there is corresponding to it or associated with it, an instance of one or more substance terms" [250].

On the other hand, since we are interested in interpreting remote sensing images, we do not actually intend to classify *Geographic Entities* but rather their representation in

satellite images. For this reason, we need to consider the concept duality issue between *Geographic Entities* and *Geographic Objects* [196] (Figure 5.5, D). To this end, let us recall that an object is an *Entity* in the digital world, which represents real-world phenomena as an instance of a generally recognized category [251]. A *Geographic Object* is then defined as a bounded geographic region in the digital world that can be identified for a period of time as the referent of a geographic term [252]. It is noteworthy that there is no semantic identity between geographic entities and corresponding objects since both cannot be described with the same characteristics from the ground and from satellite images. Whereas a *Geographic Entity* (e.g. a forested area) can be characterized by its matter (e.g. vegetation) and form, i.e. its spatial (e.g. patch size) and temporal pattern (e.g. leaf phenology), its representation in an image may be described by its pixel characteristics, i.e. spectral, spatial (e.g. texture) and temporal pattern (e.g. time series analysis) (Figure 5.5, C). This brings us to extend the *Characteristics* concept of the OBOE ontology.

5.3.1.3 Characteristics: Observables and Similarities

Based on Couclelis [247], I differentiated two types of characteristics, namely *Observables* and *Similarities*. *Similarities* are measurable and objective field properties of geographic features (e.g. vegetation height, leaf type, etc) [247]. Such characteristics can be associated to categorical (e.g. "Broadleaved") or numeric (e.g. "20 m") values. In remote sensing, such *Similarities* may stand for spatial (e.g. area of a field of pixels sharing common characteristics) and temporal characteristics (e.g. occurrence of a given class in a time series). But, in the case of a pixel-based classification, the geometrical shape is determined by the pixel's shape (i.e. approximated to a square) and thus cannot be considered to qualify a material type. Consequently, in agreement with the "time-first, space-later" approach [14, 253], I focused on temporal *Similarities*.

The spectral signature can also be considered as a *Similarities* characteristic since it is made of objective numeric reflectance values. However, its classification into spectral categories is subjective and open to debate, depending on the operator's experience and intentions. In that sense, unlike reflectance values that refer to Similarities, spectral categories can be considered as *Observable* characteristics, which refer to Sense Qualia properties, i.e. "certain features of the bodily sensations especially, but also of certain perceptual experiences, which no amount of purely physical information includes" [254]. *Observables* are field-like properties that cannot be measured [247]. Colour is a typical *Sense Qualia* whose perception is subjective and difficult to communicate. For example, the perception of shades of green to describe vegetated surfaces vary among humans (e.g. "green like a rain forest" or "green like a lemon"). According to Jensen [183], *Colour and*

the spectral signature is a fundamental biophysical variable amenable to remote sensing. More specifically, remote sensors measure "spectral signatures of the material within a pixel, thus gaining insight about an object's colour in the visible spectrum and also document its reflectivity beyond our human experience". In other words, classifying spectral signatures means classifying colours of land surface material types. For this reason, I categorized spectral signatures including spectral bands from the non-visible spectrum in classes of *Colour-Material*, which is then an *Observable* characteristic related to a land surface material type (Figure 5.5, E).

5.3.2 Application to Sentinel 2 time series

5.3.2.1 System's architecture

Based on these findings, I implemented a classification system focused on the description of image characteristics defined to describe *Pixels* and relate them to *Geographic Minor Entities* of interest for the end-user. The system is thus primarily designed to describe the information available in the remote sensing images, rather than to describe real-world entities, which is in line with the original approach proposed by Couclelis [247] "whereby it is the information available about the world, and not the world itself, that forms the basis of the ontological system". In addition, it is worth noting that the explicit descriptions of pixels are built using symbolic natural language assertions in order to ease the interpretability of the classification rules by the end-users (Assertion 5.1). By doing so, I intended to reduce the semantic and sensory gaps, allowing the end-user freedom to do the final relation between the geographic entity of interest (e.g. *Natural Vegetation*) and the pixel's description.

Theorem 5.1. *A ["Geographic Object"] represents a ["Geographic Minor Entity"] if that ["Geographic Object"] is [a value to qualify an Occurrence characteristics] classified as [a value to qualify a Colour-Material characteristic].*

For example, a "pixel" represents a "Natural Vegetation" if that "pixel" is "always" classified as "Strong Green Vegetation".

Finally, the system is divided in four successive stages as described below but here I mainly focus on the results obtained at the two last stages (i.e. Modular and Thematic stages):

- **Initial stage** : to discriminate land surfaces, clouds and shadows.
- **Dichotomous stage** : to discriminate land surface material types, e.g. vegetated, mineral and aquatic areas.

- **Modular stage** : to refine the description of land surface material types with additional characteristics in order to provide more complex low-level symbolic descriptions.
- **Thematic stage** : to allow the end-users to relate the modular classes with high-level semantic labels, e.g. natural vegetation and water bodies.

5.3.2.2 Study area and data

Based on this conceptual framework, I implemented an automatic knowledge-driven approach to classify Sentinel-2 time series in Mato Grosso. This work was led in close cooperation with the *Instituto Centro de Vida* (ICV), a Brazilian Non-Governmental Organization (NGO) acting for environmental preservation in the Southern Amazon. Together, we decided to focus on two classes : 1) natural vegetation, with a special emphasis on riparian forests, and 2) water bodies, with a special emphasis on small water bodies. The study area covered three Amazonian watersheds (529.780 km^2), i.e. the Juruena (197.036 km^2), Teles Pires (144.955 km^2) and Xingu (187.789 km^2) watersheds. Only the part of the basins located in the state of Mato Grosso were considered. It thus included a south-north gradient of natural vegetation ranging from Brazilian savannas (*Cerrado*) to dense evergreen rainforests. In order to map this study area, we downloaded 2,069 Level 1-C Sentinel-2 images for the year 2017 and spread across 75 tiles covering a total area of 726,981 km^2. The entire dataset thus consisted in 75 time series including from 13 to 46 images. The dataset consisted in top-of-atmosphere reflectance images of six spectral bands (R, G, B, NIR, SWIR1, SWIR2) resampled at 10 m spatial resolution with a bilinear algorithm in order to benefit from the highest spatial resolution.

5.3.2.3 Image processing

The approach to process Sentinel 2 time series is summarized in Figure 5.6. I first computed the *Observables* and *Similarities* characteristics. For the *Observables*, I relied on the seminal work from Baraldi et al. [255] to process all images of the database. The Satellite Image Automatic Mapper (SIAM) is a rule-based per pixel classifier based on spectral-domain prior expert knowledge taken from the remote sensing literature [255]. It thus does not require any training data and can be considered as a knowledge-driven approach (in opposition to a data-driven approach as explained in section 5.1.1) relying on remote sensing expert knowledge. SIAM assigns pixels to spectral categories exclusively based on relative rules (i.e. applied on spectral indices and on ratios between spectral bands). As such, it aims at describing the pattern of the spectral signature rather than its absolute spectral values and can thus be applied to top-of-atmosphere or

surface reflectance values. Finally, the resulting spectral categories do not correspond to land cover classes but should rather be considered as labels describing spectral signatures. For example, the *SVHNIR* spectral category refers to *Strong Vegetation with High Near-Infrared* values. All SIAM spectral rules are introduced in Baraldi et al. [255] and thus not repeated here. Finally, the SIAM spectral categories were reclassified in six main classes of *colour-material*, whose labels were constructed by associating a colour with the most likely corresponding geographic entity. The six classes of *colour-material* were : "White Cloud" (WhCl), "Black Shadow" (BlSh), "Blue Water" (BlWt), "Strong Green Vegetation" (StGVg), "Weak Green Vegetation" (WkGVg) and "Brown Mineral" (BrMin). All class labels then refer to a field-like properties, e.g. the "White Cloud" label describes a pixel whose spectral signature looks like a cloud spectral signature, but not necessarily representing a cloud due to potential confusions (with buildings for example).

I then computed *Similarities*, which refer to objective and measurable characteristics. Here, I emphasized the importance of the temporal characteristics to describe land surface material types. In this regard, I implemented a scoring approach computed on a per-pixel basis which simply consisted in counting (i.e. summing) the occurrences of a *colour-material* class in a time series. At each date, the counting was weighted by the quality of the corresponding image considering its cloud cover rate. For example, a pixel assigned to the $SVHNIR$ SIAM spectral category in an image with a 30% cloud cover was attributed a score of 0.7 (70%) to the corresponding "Strong Green Vegetation" *colour-material* class. By doing so, the objective was to give more importance to pixels observed in high quality cloud-free images while the information contained in pixels observed in low quality images was more uncertain. Finally, each pixel time series was described by six scores corresponding to the six classes of *colour-material* (Figure 5.7).

These scores were then used to derive various information:

- **Best Possible Score (BPS)** corresponds to the sum of all six scores. It represents the maximum score potentially assigned to a pixel that would always be classified in the same class of *colour-material*. BPS gives a first broad overview of pixel quality since it is related to the number of observations.
- **Best Land Surface Score (BLSS)** is the sum of all scores corresponding to classes of *colour-material* related to a land surface material type, i.e. all classes except "White Cloud" and "Black Shadow".
- **Ocurrence of Colour-Material** ($OCC_{colour-material}$) is the the ratio between the scores of *colour-material* classes related to land surface material types and the

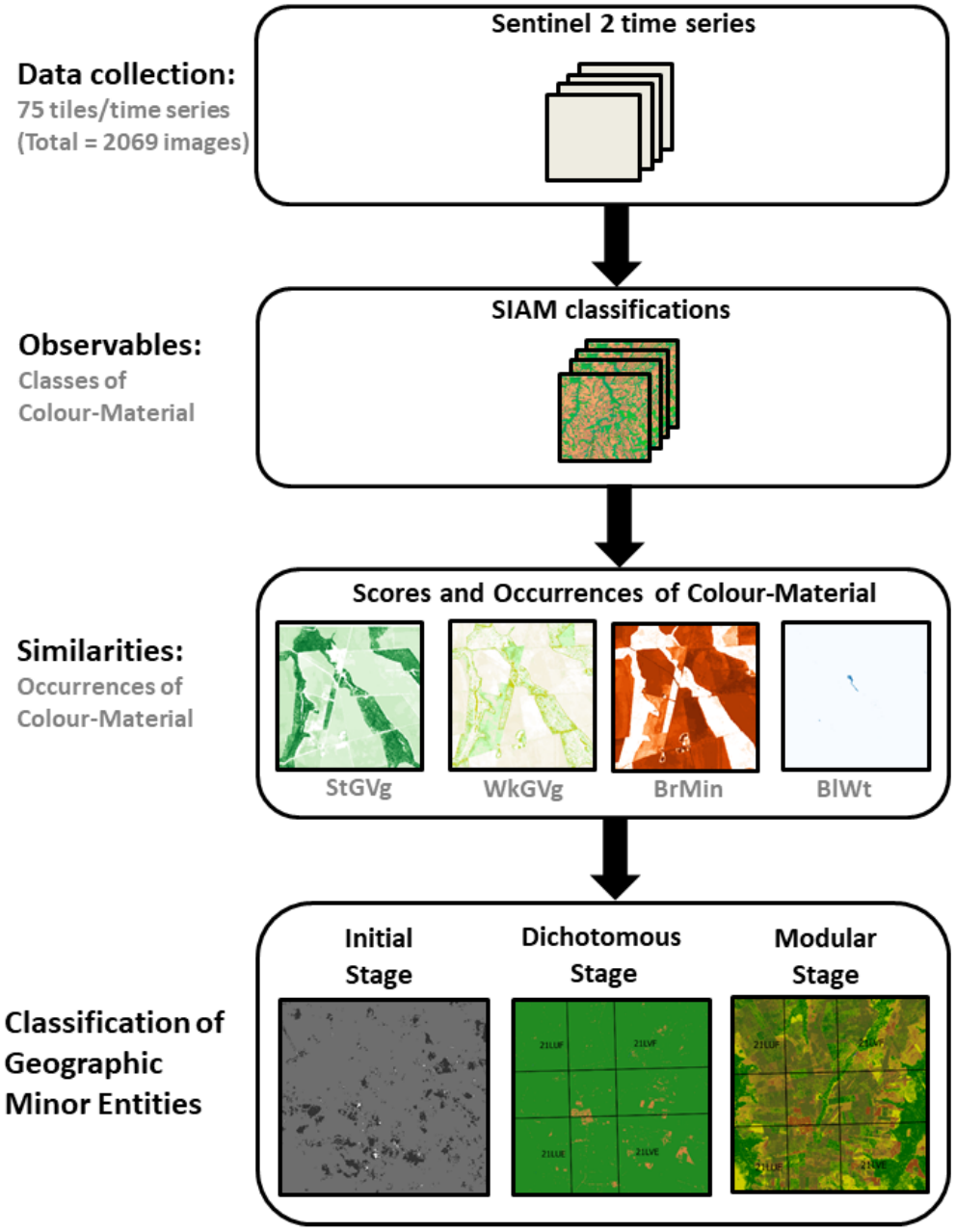


FIGURE 5.6: Flowchart of the proposed methodology to process Sentinel-2 time series.

BLSS, i.e. the proportion (in %) of observations in a given *colour-material* in a time series.

The occurrences of Colour-Material were then partitioned in eight categories labeled with adverbs of time to remain as close as possible to natural language. These adverbs were : "Never", "Rarely", "Sometimes", "Regularly", "Predominantly", "Usually", "Often" and "Always". The thresholds adopted to define each adverb are listed in Table 5.3. It is worth mentioning that the "Never" adverb is chosen when a score is $<= 1$ in a given

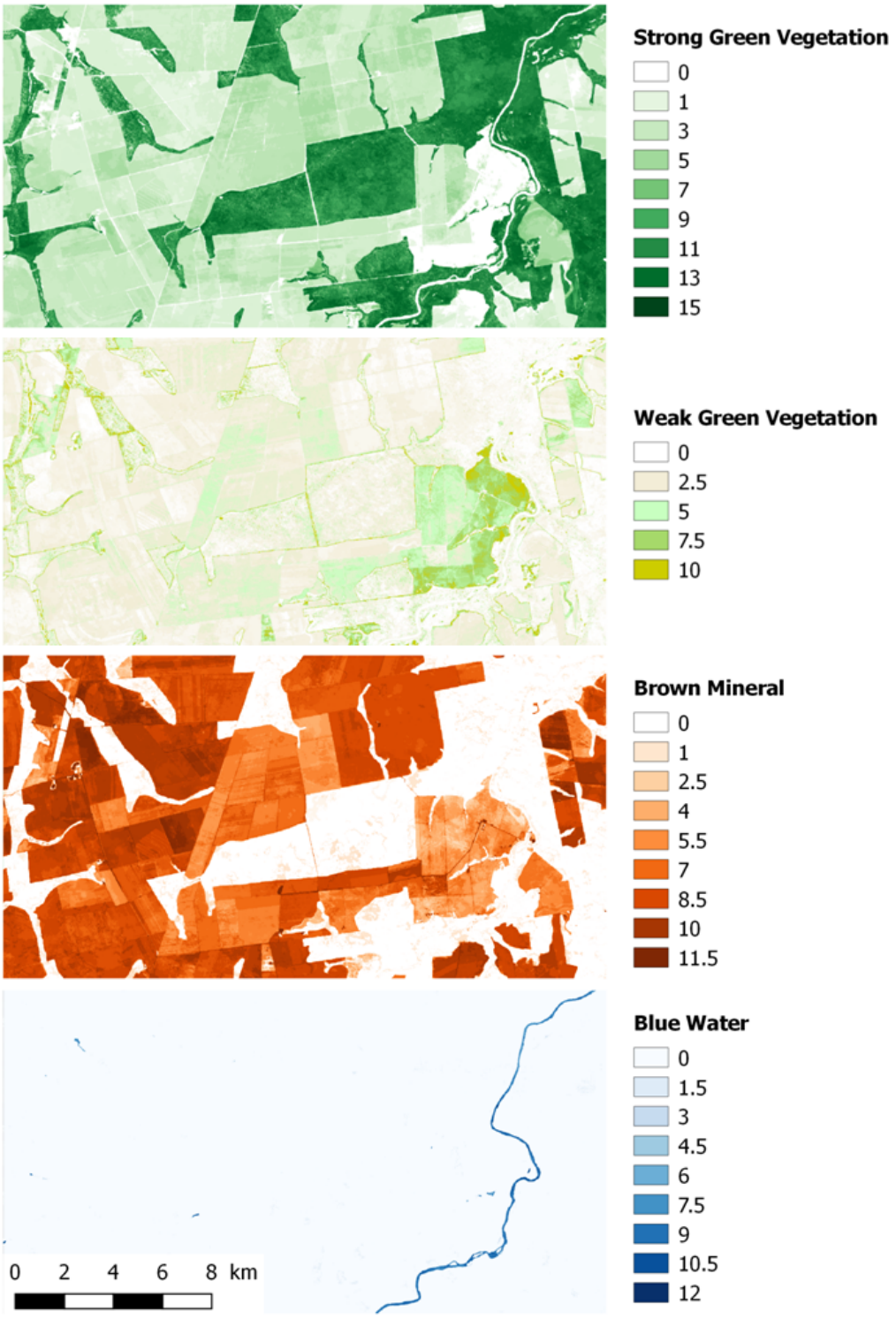


FIGURE 5.7: Examples of scores measured for the main classes of *Colour-Material* (Tile 21LXG).

Colour-Material class in order to avoid giving too much importance to pixels with low absolute score values but high occurrence values.

Adverb of time	Logical Rule	Label	ID
Never	$Score_{colour-material} <= 1$	NEV	1
Rarely	$0\% \leq OCC_{colour-material} < 10\%$	RAR	2
Sometimes	$10\% \leq OCC_{colour-material} < 25\%$	SOM	3
Regularly	$25\% \leq OCC_{colour-material} < 50\%$	REG	4
Predominantly	$50\% \leq OCC_{colour-material} < 75\%$	PRED	5
Usually	$75\% \leq OCC_{colour-material} < 90\%$	USU	6
Often	$90\% \leq OCC_{colour-material} < 99\%$	OFT	7
Always	$OCC_{colour-material} >= 99\%$	ALW	8

TABLE 5.3: Temporal rules to define temporal occurrences. $OCC_{colour-material}$ refers to the occurrence of each class of *colour-material*.

5.3.2.4 Image classification

Initial stage

The initial stage intended to discriminate clouds, shadows and land surfaces. To do so, I set a list of theorems and corresponding classification rules:

Theorem 5.2. *A pixel represents a land surface if that pixel is (at least) sometimes classified in a Colour-Material class that refers to a land surface, i.e. "Blue Water" or "Brown Mineral" or "Strong Green Vegetation" or "Weak Green Vegetation".*

Theorem 5.3. *A pixel represents a cloudy area if that pixel is not classified as land surface but (at least) often classified as either "White Cloud" or "Brown Mineral" Colour-Material.*

Theorem 5.4. *A pixel represents a shadow area if that pixel is neither classified as land surface nor cloudy area.*

These assertions were expressed in equation 5.1:

$$MinorEntity = \begin{cases} LandSurface, & \text{if } (OCC_{BLSS} >= 10\%) \\ Cloud, & \text{if } (OCC_{WhCl+BrMin} \geq 90\%) \\ Shadow, & \text{Otherwise} \end{cases} \tag{5.1}$$

where OCC_{class} refers to the occurrence of the corresponding *class* of *Colour-Material*, i.e. $WhCl$ and $BrMin$ stand for "White Cloud" and "Brown Mineral", respectively.

Dichotomous stage

The dichotomous stage then aimed at discriminating land surface material types. It thus produced a broad classification to separate large chunks of matter considering two binary

classifications to differentiate "Aquatic" and "Terrestrial" areas and then "Vegetation" and "Non-vegetation" areas. The natural language descriptions of these classifications were:.

Theorem 5.5. *A pixel represents an aquatic area if that pixel is (at least) usually classified as "Blue Water" Colour-Material class.*

Theorem 5.6. *A pixel represents a vegetated area if that pixel is not never classified as "Strong Green Vegetation" or "Weak Green Vegetation" in a time series.*

The corresponding equations were:

$$Realm = \begin{cases} Aquatic, & \text{if } (OCC_{BlWt} \geq 50\%) \\ Terrestrial, & \text{Otherwise} \end{cases} \tag{5.2}$$

Where OCC_{BlWt} refers to the occurrence of the "Blue Water" *Colour-Material.*

$$Substance = \begin{cases} Vegetation, & \text{if } (Score_{StGVg+WkGVg}) > 1) \\ Non-Vegetation, & \text{otherwise} \end{cases} \tag{5.3}$$

Where $Score_{SGVg+WGVg}$ refers to the sum of scores in "Strong Green Vegetation" and "Weak Green Vegetation", respectively.

Finally, these two binary classifications were crossed together to produce the dichotomous stage classification including the following four classes of land surface material types : Terrestrial Vegetation, Mineral (i.e. Terrestrial Non-Vegetation), Aquatic Vegetation and Water (i.e. Aquatic Non-Vegetation). In addition, Cloud and Shadow pixels identified at the initial stage are also preserved.

Modular stage

At the modular stage, the descriptions of land surface material types observed at dichotomous stage were extended by considering the occurrences of *Colour-material* classes. The objective was to assign each pixel to a label that explicitly describes its time series of spectral signatures. To date, I only extended the "terrestrial vegetation" class identified at the dichotomous stage to add temporal characteristics. The terminology used at the modular stage was constructed by associating adverbs of time with classes of *colour-material.* All pixels were assigned a composite label describing the temporal signatures based on the possible combinations between the occurrences of "strong Green Vegetation", "Weak Green Vegetation" and "Brown Mineral" *Colour-Material* classes. The "Blue Water" class was not considered here since we were only focusing on areas

of "Terrestrial Vegetation". An example of composite label is "ALW-SV & NEV-WV & NEV-BM", which refers to a pixel *always* classified in "Strong Green Vegetation" and *never* in "Weak Green Vegetation" nor "Brown Mineral". This approach leads to a maximum of $N = 512$ combinations, from which many of them are impossible. For example, a pixel cannot be *always* classified as Strong Green Vegetation and Weak Green Vegetation in the same time.

In addition, each composite label is associated to a 3-digit numeric ID where each digit corresponds to the numeric value (1 to 8) assigned to the adverb of time characterizing the occurrences of the "Strong Green Vegetation", the "Weak Green Vegetation" and the "Brown Mineral" classes of *Colour-Material*, respectively. Consequently, the numeric IDs range from 111 for pixels never classified in any of these three classes of *colour-Material* to 888 for pixels always classified in each of these three classes of *colour-Materials* (which is actually an impossible case).

Finally, since the number of modular classes is potentially very high, I assigned colours to every class automatically in order to ease the map visualisation. For this purpose, I defined one reference colour, i.e. "darkgreen", "yellow3" and "tan4" as defined in the R environment, for each of the three classes of *colour-material* used at the modular stage, i.e. "Strong Green Vegetation", "Weak Green Vegetation" and "Brown Mineral" respectively. These reference colours were then converted to RGB colour space and mixed proportionally to the occurrences of their corresponding classes in order to compute a new colour for each modular class.

In the case study of Mato Grosso, 142 modular classes were encountered thus representing a large diversity of combinations. Nonetheless, the 20 most represented classes covered 87.8% of the total of "Terrestrial Vegetation" pixels (each of these 20 classes accounted for more than 1% of the total number of pixels). The distributions of these 20 classes, their label descriptions and their associated colours are shown in Figure 5.8 and evidence the importance of 4 classes (811, 711, 154 and 145) which accounted for 53.1% of all "Terrestrial Vegetation" pixels. The figure 5.8 serves as a legend to understand the map produced at the modular stage (Figure 5.9). At regional scale, the SE-NW vegetation gradient from Cerrado to rain forests appeared clearly. Pasture areas (tiles 21LUK, 21LXG and 22LCP) were classified with similar colour shades than cerrado areas (tiles 21LXD and 21LUE). This is because they actually both correspond to herbaceous land cover with sparse shrubs or trees. Crops were assigned to brown-green colour shades (e.g. as classes 335 or 415 in Figure 5.8) corresponding to a mix of all classes of *Colour-Material* throughout the year. Finally, forests appeared in dark green. It is also worth noting that other classes discriminated at the dichotomous stage ("Water", "Mineral" and "Aquatic Vegetation") were preserved at the modular stage .

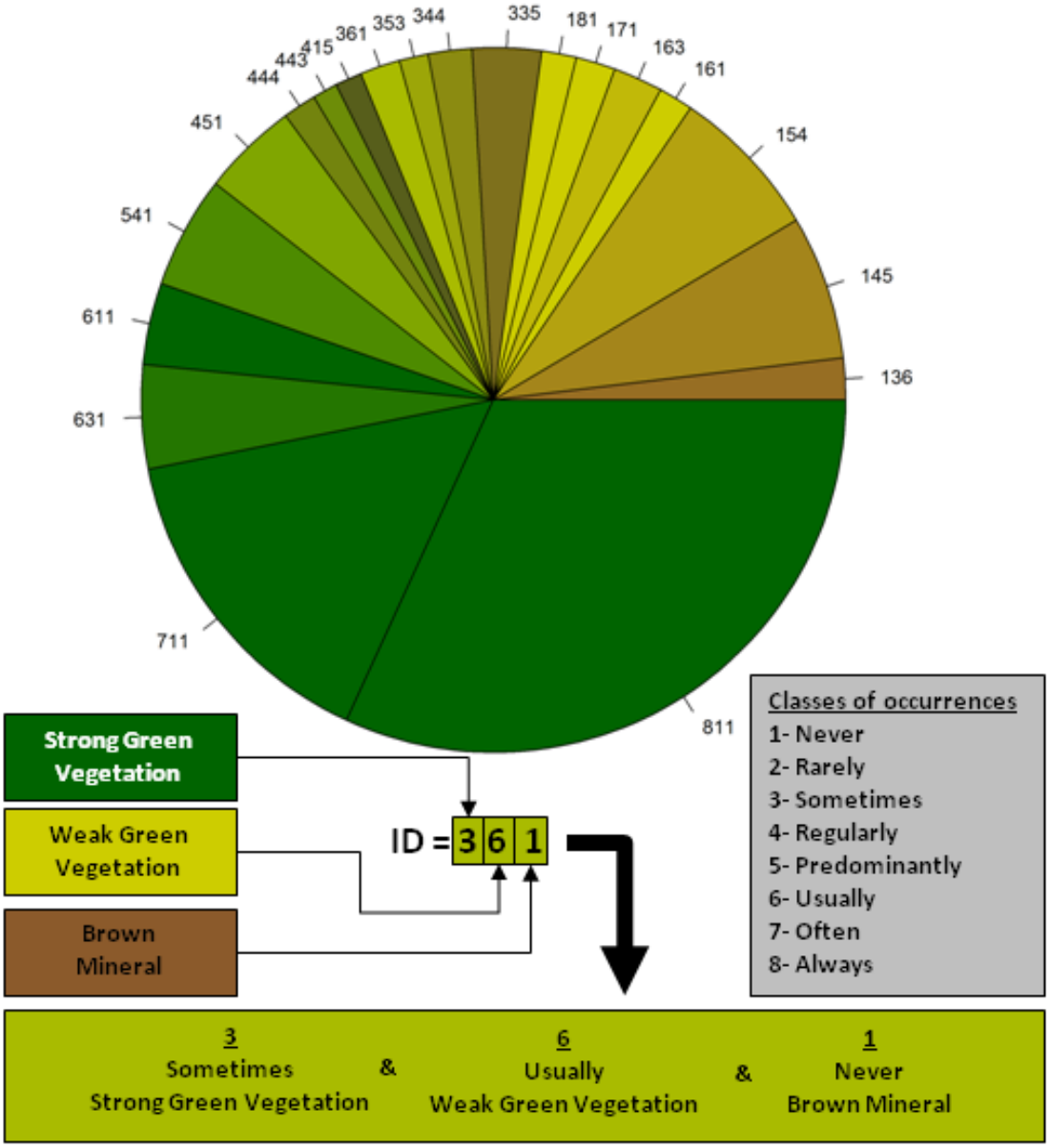


FIGURE 5.8: Distribution of the 20 most represented modular classes with their label descriptions and associated colours.

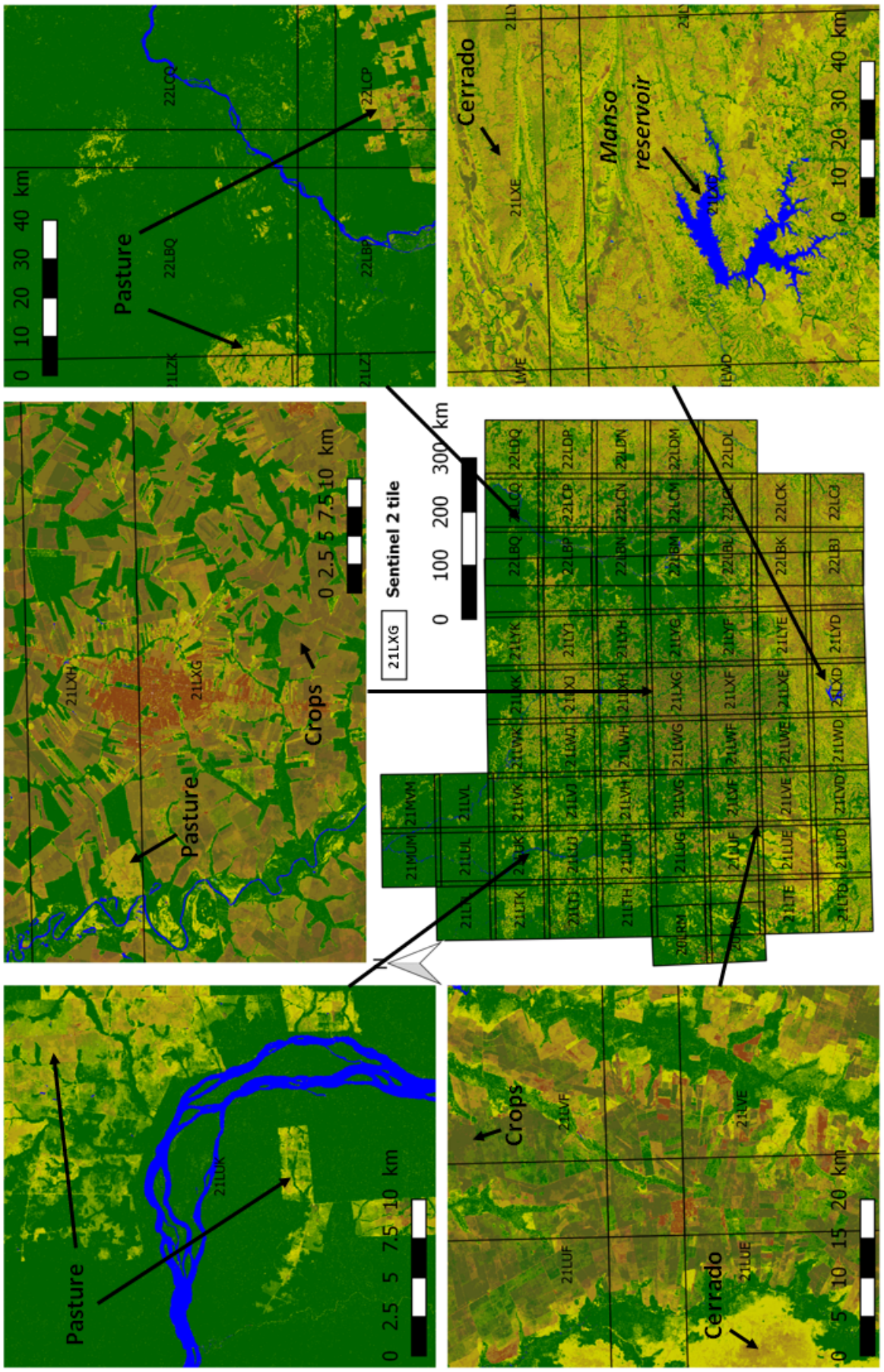


FIGURE 5.9: Map produced at the modular stage.

Thematic stage

The modular classification was finally analysed to assess how it could benefit the end-users with regards to their specific applications. At first, I focused on the classifications of water bodies. After visual assessment, the accuracy of the approach to map water bodies was validated using the F-score statistical index (Table 5.4). Reference samples were extracted from land use maps produced by the end users and compared with the classification. More specifically, two definitions of "Water" were considered, i.e. a restrictive definition considering only the "Water" class and a broad definition considering both classes of "Water" and "Aquatic Vegetation". F-scores were always better when considering both classes of "Water" and "Aquatic Vegetation" as water bodies. As expected, the lowest F-scores (F = 0.802) were obtained when using all reference samples, i.e. including very small water bodies (< 10m width) or hidden by the canopy. Then, the results improved significantly when focusing on water bodies larger than 10 meters (F = 0.874), 20 meters (F = 0.927) or even 100 meters (F = 0.961), i.e. when considering that only water bodies larger than the image spatial resolution (10m) can "exist" in Sentinel-2 images.

	Minimum width of water bodies										
	All	**10m**	**20 m**	**30m**	**40m**	**50 m**	**60m**	**70m**	**80 m**	**90m**	**100m**
Water	0.802	0.867	0.918	0.934	0.938	0.920	0.944	0.947	0.960	0.961	0.960
Water + Aqu. Veg.	0.808	0.874	0.927	0.938	0.942	0.926	0.946	0.951	0.965	0.962	0.961
N	1000	1000	1000	1000	1000	1000	1000	1000	820	626	495

TABLE 5.4: F-score validation statistics for water bodies.

Then, I aimed to assess how efficient the automatic approach was to map areas of natural vegetation. But, beyond the statistical validation with a reference map, I wished to underline to the end-user that various potentially accurate land cover classifications could actually be produced. In other words, various correct interpretations of a unique dataset of satellite images can coexist, thus illustrating the vagueness issue of geographic terms. Here, considering that most tropical native vegetation in the study is evergreen, I focused on the modular classes that were "never" or "rarely" observed as "Brown Mineral" *Colour-Material* and I ordered them in a gradient ranging from classes corresponding to pixels only classified in "Strong Green Vegetation" (i.e. IDs 811, 711, 611, 511, 411, 311, 211) on the one side, to classes referring to pixels only classified in "Weak Green Vegetation" (i.e. IDS 181,171,161,151,141,131,121), on the other side. Between these two extremes, the modular classes were ordered according to their decreasing occurrences of "Strong Green Vegetation" and increasing occurrences of "Weak Green Vegetation". Based on this gradient, I then iteratively computed the F-score comparing the reference labels (retrieved from a reference map provided by the end-user) of 500 randomly selected points in each Sentinel-2 tile with the various potential masks of

natural vegetation computed by integrating a new class of the gradient at each new iteration until all classes of the gradient were considered. By doing so, I intended to set the optimal combination of modular classes that would best discriminate areas of natural vegetation according to the end-user's definition.

Figure 5.10 shows the evolution of the F-score for each combination of modular classes. The high F-score values (> 0.8) achieved with all combinations indicated it did not exist a unique combination of classes to represent the areas of natural vegetation as originally delineated by the end-users. Interestingly, the F-score achieves its maximum value ($F = 0.875$) when considering all modular classes, i.e. pixels that are only classified as "Green Vegetation", either "Strong" or "Weak". This means that including modular classes with occurrences of "Brown Mineral" may potentially improve the mapping of natural vegetation. This is confirmed when analyzing the spatial variability of accuracy metrics across the study area (Figure 5.11) evidencing lower F-score values (e.g. $F = 0.68$ in Tile 21LCJ) in southern areas dominated by open savannas which can be confused with bare soil during the dry season. On the contrary, dense rainforests in northern areas are better classified ($F > 0.9$) when considering less combinations of modular classes such as in tiles 21LWK and 21LYG (Figures 5.10 and 5.11).

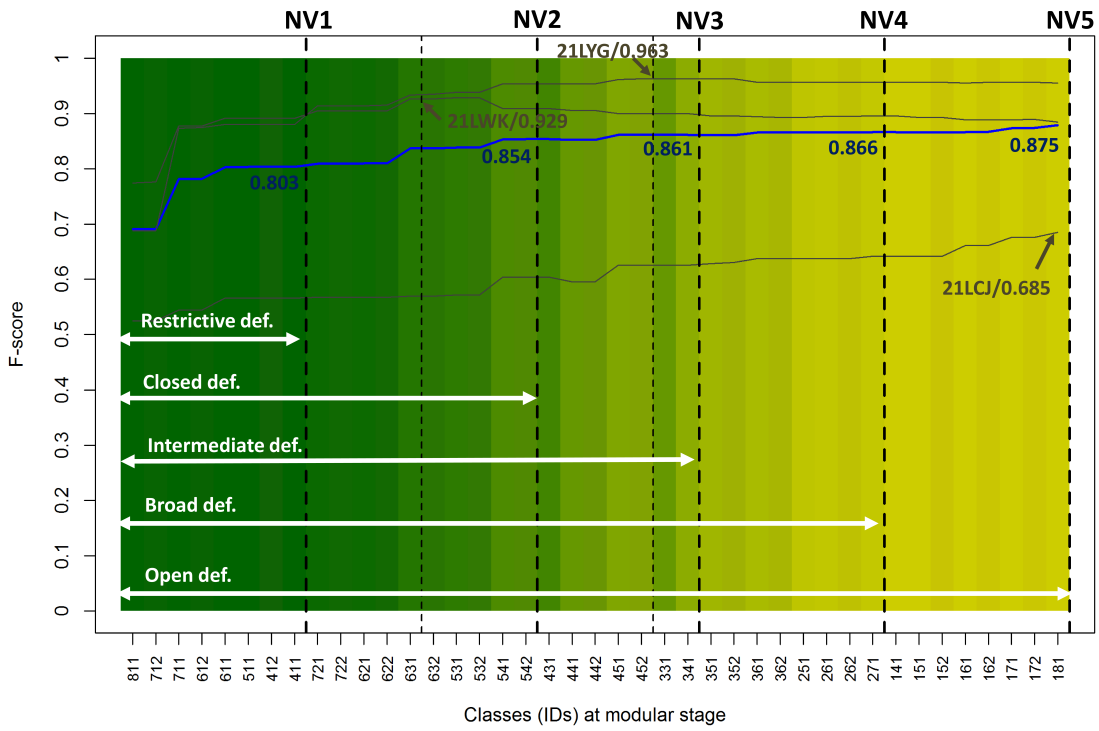


FIGURE 5.10: Validation of the natural vegetation land cover class with F-score statistics computed for various combinations of modular classes.

In order to illustrate the implications of changing the combination of modular classes in the mapping of natural vegetation, I set five thresholds (NV1 to NV5) in the colour gradient of modular classes (Figure 5.10) corresponding to five different definitions, ranging

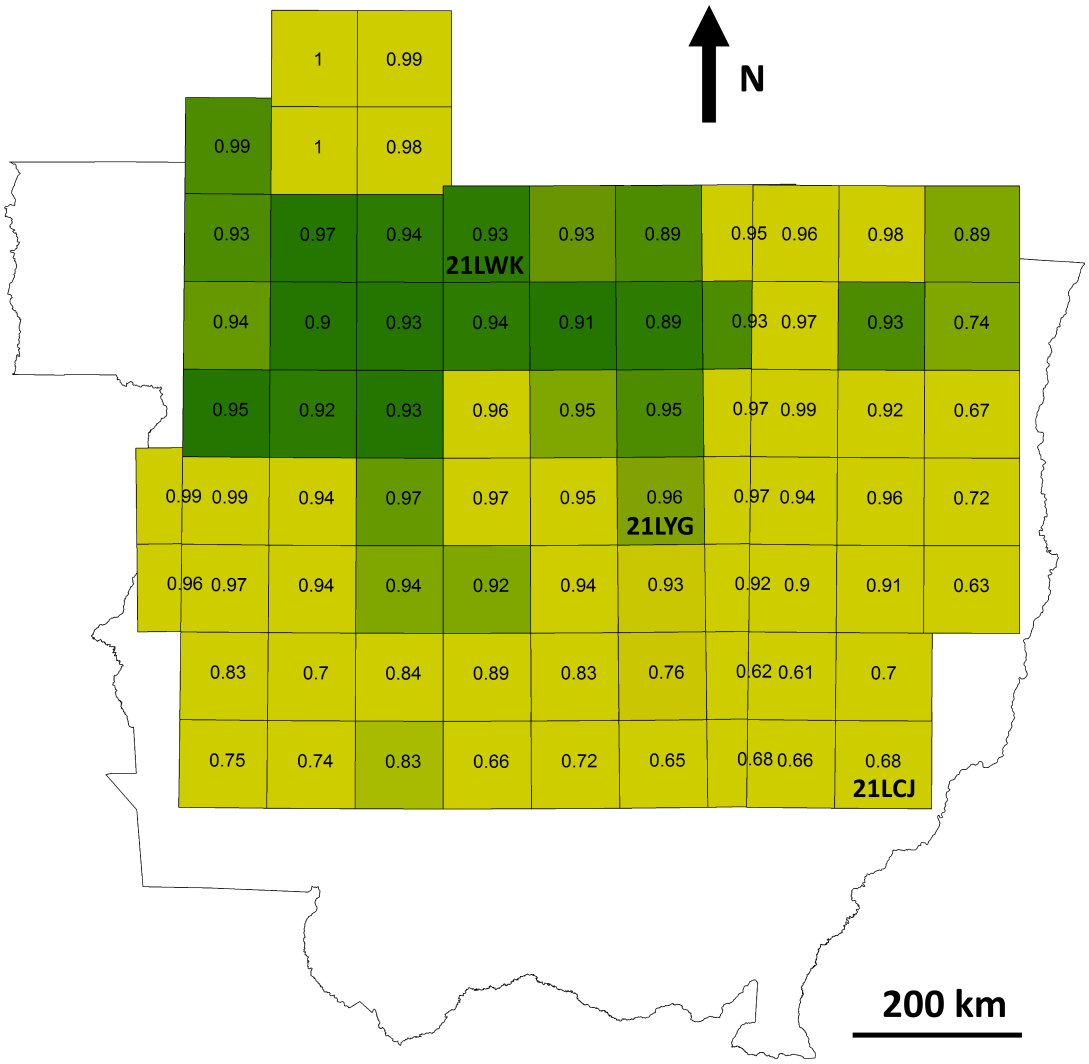


FIGURE 5.11: Validation of the natural vegetation land cover class with F-score statistics for each Sentinel-2 tile. The colour corresponds to the threshold in the colour gradient (see figure 5.10) at which the best F-score value was achieved.

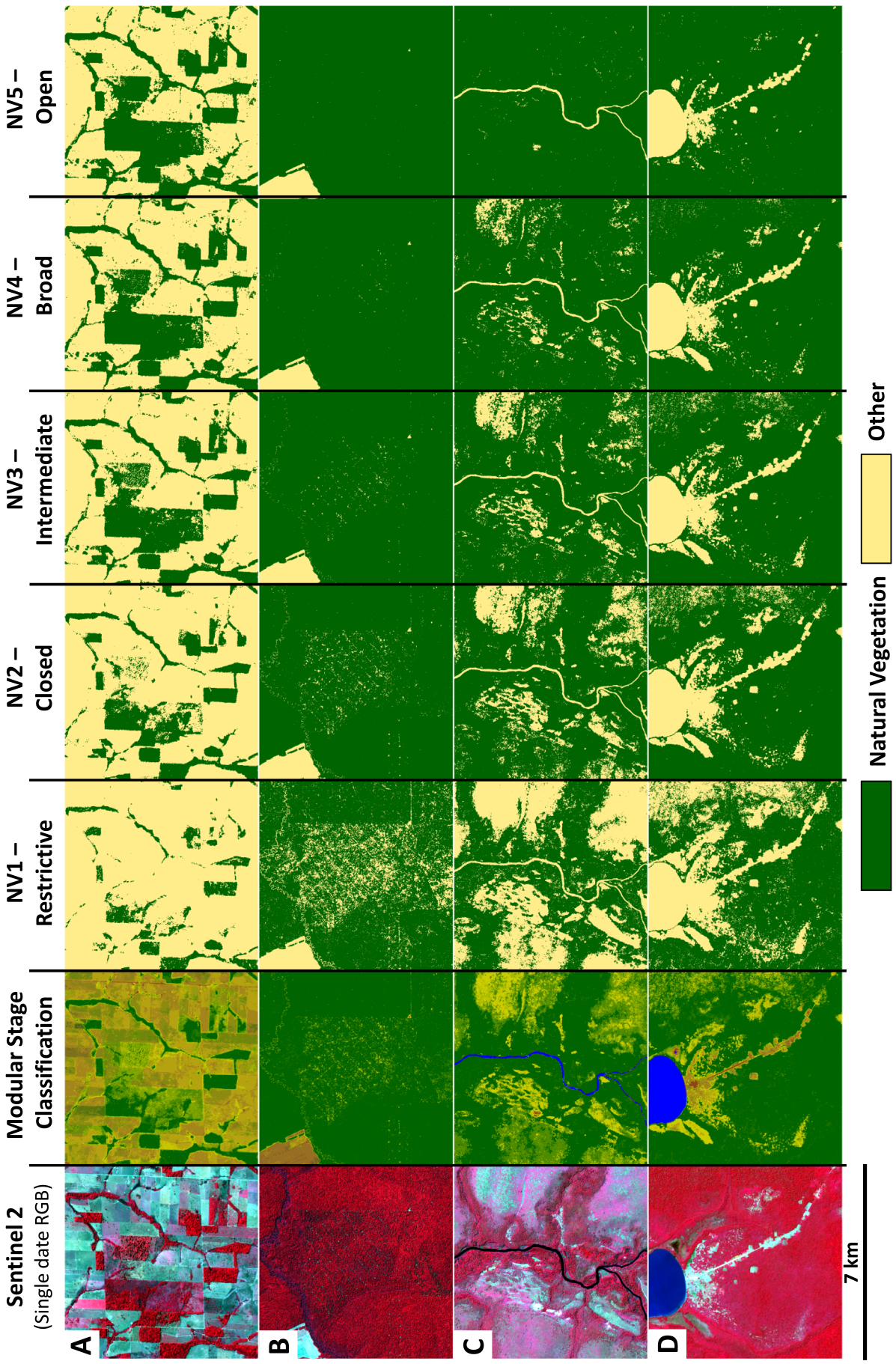
from a restrictive definition of natural vegetation corresponding to pixels always classified as "Strong Green Vegetation" to an open definition of natural vegetation considering all pixels always classified as "Green Vegetation", either "Strong" or "Weak". The natural language definitions corresponding to these five combinations are expressed in Table 5.5. I then identified four examples illustrating the importance to consider vegetation gradients when monitoring fine changes in vegetation cover. The first example (Figure 5.12, A) is centered around an area of degraded forest and shows the difficulty to accurately delineate forest patches. At NV1 (restrictive definition), degraded areas were discarded and forest patches were discontinuous whereas, at NV5 (Open definition), degraded areas were classified as forests and connectivity was preserved. Similar conclusions can be drawn from Figure 5.12 (B) centered on a logged forest. At NV1, forests appeared as highly fragmented and degraded. At NV2 (Closed), NV3 (Intermediate)

and NV4 (Broad), only the human infrastructures (logging roads and log landings) were discriminated. At NV5, the entire area was considered as forest. In Figure 5.12 (C), humid riparian forests weere surrounded by dry forests and Cerrado areas on neighbouring plateaus. Here again, the maps differed dramatically from NV1 to NV5 with Cerrado areas being totally included in the mask of natural vegetation at NV5. Finally, figure 5.12 (D) showed how the extension of an indigenous village (*Camaiura*, inside the Xingu indigenous land) with traditional slash-and-burn agriculture varied depending on the chosen definition of natural vegetation.

Label	**Natural Language Definition**
	A pixel represents a Natural Vegetation if that pixel is...
NV1- Restrictive	...Always Strong Green Vegetation
NV2- Closed	...(At least) Predominantly Strong Green Vegetation and Never or Rarely Brown Mineral
NV3- Intermediate	...(At most) Regularly Strong Green Vegetation and (at most) Regularly Weak Green Vegetation and Never or Rarely Brown Mineral
NV4- Broad	...(At least) Rarely Strong Green Vegetation and Never or Rarely Brown Mineral
NV5- Open	...Always Strong or Weak Green Vegetation

TABLE 5.5: Natural Language definitions of the five potential classes of natural vegetation identified in Figure 5.10.

In addition, I assessed the quality of the maps of natural vegetation to estimate degradation of riparian forests in Permanently Protected Areas (PPA). For this purpose, I compared the proportion of degraded areas measured by the end-users through visual interpretation for $N = 1,239$ watersheds in Northern Mato Grosso with the corresponding estimates produced at modular stage (Figure 5.13). As expected, the restrictive maps of natural vegetation (NV1) tended to overestimate degradation whereas the open forest maps (NV5) underestimated degradation. The highest correlation ($R^2 = 0.841$) was achieved with the closed map (NV2), which means that this definition could be used on a regular basis to produce rapid updates on vegetation degradation in riparian Permanent Preservation Areas.


Sentinel 2
(Single date RGB)
Modular Stage
Classification
NV1 –
Restrictive
NV2 –
Closed
NV3 –
Intermediate
NV4 –
Broad
NV5 –
Open
A
B
C
D
7 km
Natural Vegetation
Other

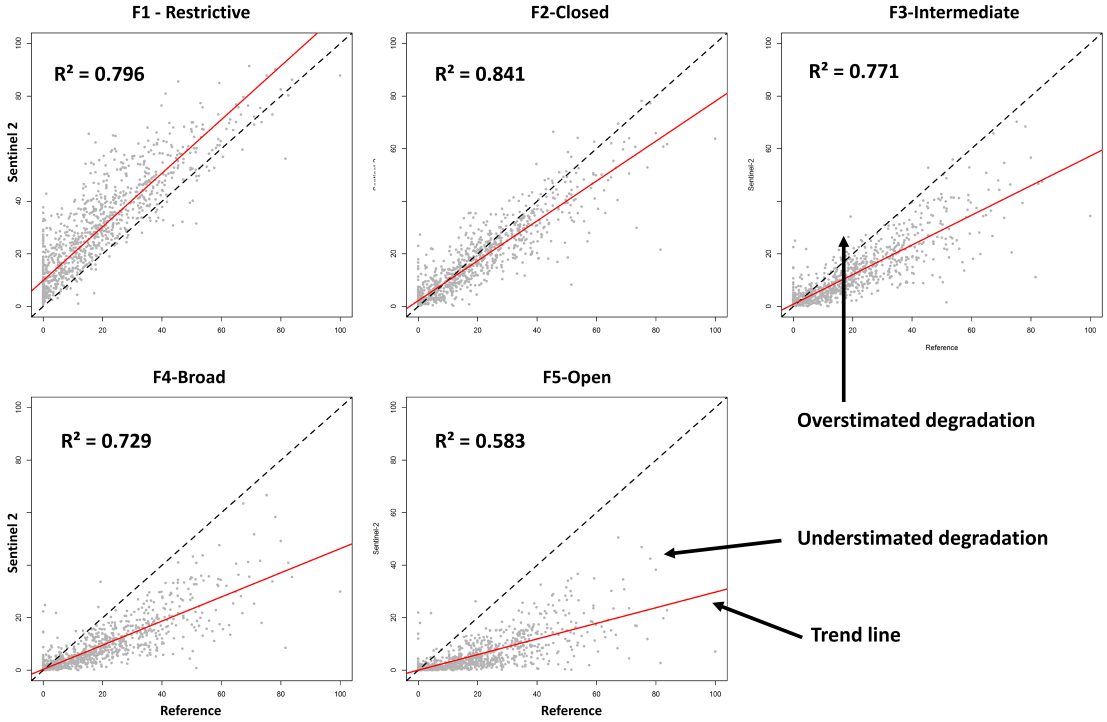


FIGURE 5.13: Correlations between the proportions of degraded riparian forest in $N = 1,239$ watersheds as estimated by visual interpretation (Reference data set) and by the Sentinel-2 classification, using various thresholds at the modular stage.

5.4 Ontology-based classification of remote sensing images using spectral rules

Whereas the first application relied on ontological principles, the use of ontologies actually remained very conceptual since it did not include any formal ontology. As a matter of fact, ontology-based remote sensing applications remain rare and limited to date. The major limitations refer to 1) the number of concepts in the ontologies (i.e. number of classes to be identified in the images) as well as the number of objects to classify are usually restricted and 2) the ontologies are often coupled with an image processing software at the classification step, i.e. to assign image objects to concepts of the ontologies. By doing so, the reasoning potential of the ontologies is often unexplored.

Here, I introduce a novel ontological approach for remote sensing image classification. I worked on this research during my post-doctoral fellowship at IRD (Montpellier) from 2011 to 2014. At that time, I cooperated very closely with Samuel Andrés who was doing his PhD in informatics in the same research unit. Together (and also considering valuable contributions from other colleagues: Laurent Durieux, Isabelle Mougenot, Thérèse Libourel, Marie-Angélique Laporte, Christelle Pierkot), we developed a prototype system through which remote sensing expert knowledge is formalized and analysed to guide

the interpretation of satellite images. The added-value of that system is twofold: i) the ontological approach is coupled with an open source image processing software at the pre-processing step (Orfeo Toolbox, OTB) and ii) the system fully explores the potential of ontologies since the image classification indexing process (i.e. the assignment of an image object to a semantic class) is performed thanks to a reasoner algorithm.

For the purpose of this study, the ontology prototype was built upon the expert remote sensing knowledge expressed in the Satellite Image Automatic Mapper (SIAM) of Baraldi et al. [255], which was already introduced in section 5.3. SIAM consists in rule sets used to classify Landsat images. Our work thus mainly intended to formalize these rules in an ontology, i.e. to first define concepts and relations from the expert knowledge and then set the classification rules. In this regard, it is worth noting that the added-value of an ontology-based approach when compared to the original rule-based approach is entirely methodological and not thematic. The originality of using ontologies stands in its ability to change the paradigm of the image analysis, from a numeric to a symbolic approach. Whereas traditional implementations of classification rules integrate the knowledge directly in the algorithm, our approach extracts the knowledge in order to consider it as a proper object.

5.4.1 Knowledge management: from conceptualization to image facts

In this work, we proposed to use both ontology capabilities (i) to share knowledge between conceptual expertise in remote sensing and image content and (ii) to make some implicit knowledge explicit using inference engines, i.e. assign pixels to semantic descriptors. To do so, we distinguished three levels of knowledge as proposed by Falomir et al. [245]: the reference conceptualization, the contextual knowledge and the image facts. These different parts of knowledge concern the image structure representation, the remote sensing experience and the image content, respectively. All together, they draw a global model (fig. 5.14) with (i) horizontally, the conceptual (concept description) and concrete (instance description) levels, i.e. the TBox and ABox respectively, and, (ii) vertically, the image and image interpretation domains. In this diagram, the bottom right box is the place for image interpretation, i.e. the ontological system calls upon a reasoner to infer implicit knowledge about the individuals (image pixels) of the ontology in order to assign them to concepts (spectral categories) described in the contextual knowledge. Please note that individuals here refer to image pixels but we called them "image objects" because our long-term objective was to integrate the prototype into a GEOBIA approach.

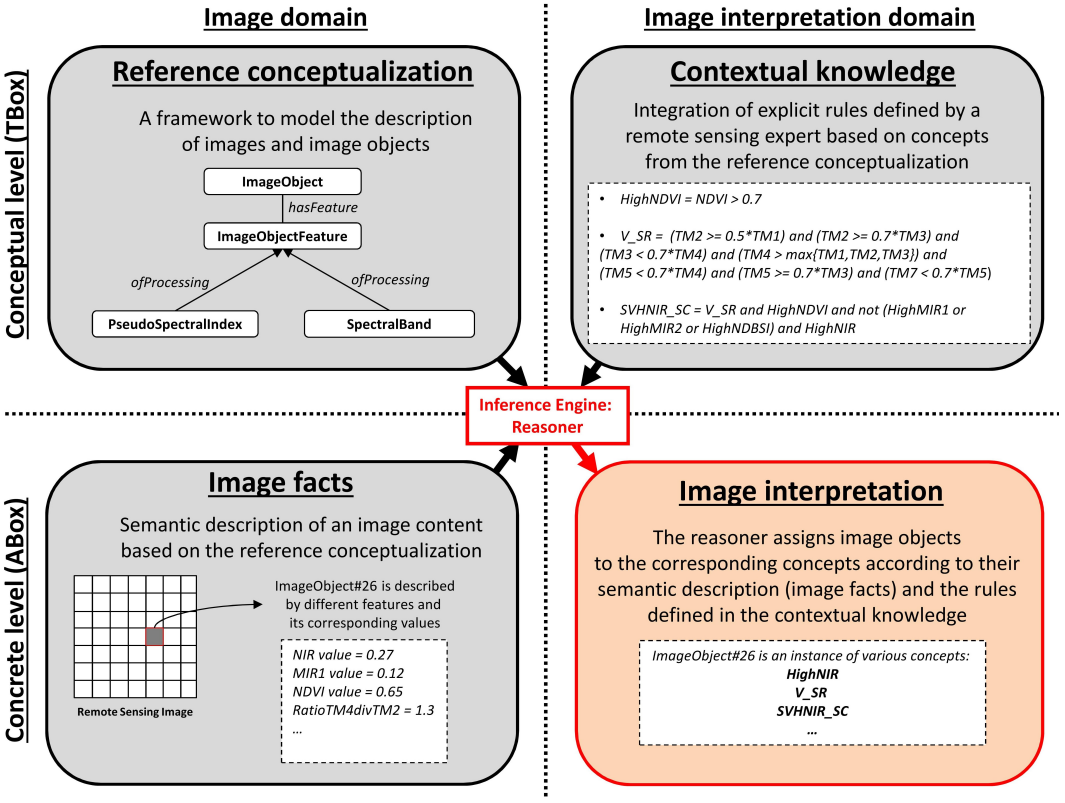


FIGURE 5.14: Structure of the knowledge base comprising two levels (conceptual and concrete) and two domains (image and image interpretation).

5.4.1.1 The reference conceptualization

The reference conceptualization can be viewed as a general model to describe image objects in remote sensing images. It captures a few top-level concepts shared between two packages: the image structure package introduces the *ImageObject* and *ImageObjectFeature* concepts and the image processing package introduces the *PseudoSpectralIndex* and *SpectralBand* concepts (fig. 5.15). These concepts are part of the reference conceptualization because they refer to elementary concepts a remote sensing expert relies on to describe its own contextual knowledge. Concepts related to spectral bands, spectral indices or texture indices are used by experts to interpret a remote sensing image, but they are not defined by these experts. For example, whereas the definition of a vegetation type (e.g. forest) may vary according to the geographical area or the expert's knowledge, the concept of a spectral band is not expected to vary.

The *ImageObject* concept allows to describe image objects according to their characteristics. These characteristics refer to the *ImageObjectFeature* concept so that both concepts are linked by a *hasFeature* relation, e.g. *ImageObject hasFeature ImageObjectFeature*. A characteristic of an image object is the result of a processing task (e.g. a

```
ImageObject ≡ ∃hasFeature.ImageObjectFeature
        ⊓ ∀ hasFeature.ImageObjectFeature (Axiom 1)

ImageObjectFeature ≡ ∃ofProcessing.(PseudoSpectralIndex
 ⊔ SpectralBand) ⊓ ∃ numericalValue.xsd:decimal[operator X]  (Axiom 2)
        where operator is <,>,<=,>=,=,!= and X a quantitative value
```

LISTING 5.1: Concept equivalence for ImageObject and ImageObjectFeature

NDVI value after computing a spectral index) although the processing task is not computed by the ontology. Consequently, the *ImageObjectFeature* is related to concepts of the image processing package (*PseudoSpectralIndex* and *SpectralBand* concepts) through an *ofProcessing* relation. Based on these concepts and relations, the description of any image object (in any type of remote sensing image) can be defined using description logic syntax as in listing 5.1. In axiom 2, the *numericalValue* allows the spectral rules to be included by considering the image object values.

5.4.1.2 The contextual knowledge

The contextual knowledge serves to represent the remote sensing expert knowledge. This knowledge depends on the objectives and the experience of the expert (or of the expert community) and must be coherent with the data to be interpreted. This is why this knowledge is called "contextual knowledge". It differs from the reference conceptualization by the semantics: it is not an "objective" representation of image structure, but a "subjective" description of image interpretation rules. Thus, the contextual knowledge is an extension of the reference conceptualization that contains all rules defined by an expert to interpret a remote sensing image.

In the present study, the remote sensing expert knowledge is derived from SIAM algorithm proposed by Baraldi et al. [255] to automatically interpret Landsat TM and ETM+ images. In this method, as already mentioned previously, pixels are assigned to semantic spectral categories (up to 46 classes) on a per-pixel basis. This method proposes an effective way to describe quantitative data (i.e. remote sensing images) with qualitative (semantic) information (i.e. spectral categories) and it thus appears fully relevant to be converted into ontologies where concepts are semantically described.

The SIAM method is based on three spectral rule sets applied on top-of-atmosphere calibrated Landsat bands (TM_1, TM_2,..., TM_7 bands, where TM stands for Thematic Mapper) and eleven derived spectral indices (*Bright, Vis, NIR, MIR1, MIR2, TIR, MIRTIR, NDSI, NDBBBI, NDVI, NDBSI*). Thus, the contextual knowledge was split into four packages that refer to (i) the description of the input images (e.g. TM_1,*NDVI*,

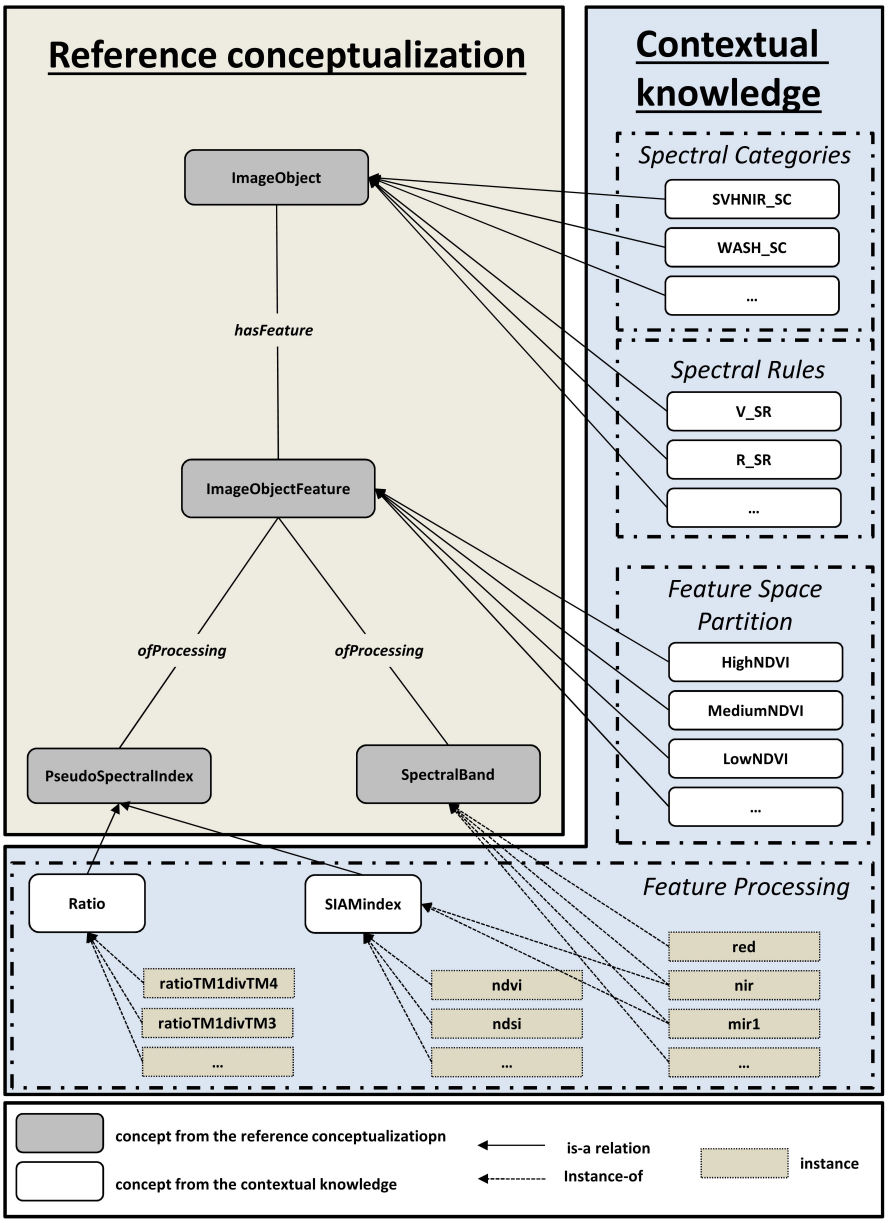


FIGURE 5.15: Concepts, relations and instances included in the ontology at the conceptual level (reference conceptualization and contextual knowledge).

etc.) and (ii) the three rule sets of the methodology (named Feature Space Partition, Spectral Rules and Spectral Categories by Baraldi et al. [255]) (fig. 5.15).

We defined all the spectral bands and spectral indices required to implement the spectral rule sets into the ontology. Since they are the result of image processing tasks we called this package "Feature processing" (although these processes are not computed by the ontology).

First, the seven spectral bands of Landsat TM images were defined as instances of the *SpectralBand* concept and named *blue, green, red, nir, mir1, tir* and *mir2* concepts to

```
HighNDVI ≡ ImageObjectFeature
        ⊓ ∃ ofProcessing.{ndvi}⊓ ∃ numericalValue.xsd:decimal[>0.7] (Axiom 3)
```

LISTING 5.2: Concept equivalence for HighNDVI

refer to TM_1, TM_2,..., TM_7 bands respectively.

Second, we identified two sub-concepts of *PseudoSpectralIndex* concept: *SIAMindex* and *Ratio* (fig. 5.15). The *SIAMindex* concept is used to include the eleven spectral indices initially computed in SIAM. These indices are defined in eleven concepts (named *bright, vis, nir, mir1, mir2, tir, mirtir, ndsi, ndbbbi, ndvi* and *ndbsi*) which are instances of the *SIAMindex* concept. It is worth mentioning that the *nir, mir1, mir2* and *tir* concepts are instances of both SIAMindex and SpectralBand concepts.

The *Ratio* concept is created to handle spectral rules such as "$TM_1 > TM_2$" which cannot be efficiently handled in the ontology. We translated such rules into ratio indices, i.e. $TM_1/TM_2 > 1$. These ratio indices are named *ratioTM1divTM2* (corresponding to TM_1/TM_2), *ratioTM4divMaxTM123* (corresponding to $TM_4/max\{TM_1, TM_2, TM_3\}$), etc.

We then defined the rule sets for feature space partition. The first rule set is applied to the eleven spectral indices and is intended to partition the feature space by assigning a categorical value (low, medium or high) to any pixel depending on its corresponding index value. For example, a high NDVI label (i.e. "HighNDVI") is assigned to any pixel whose Normalized Difference Vegetation Index (NDVI) value is higher than 0.70.

To express such rule in the ontology, the *ImageObjectFeature* concept is used to connect the *ImageObject* with relevant characteristics, i.e. it links an interval of values (e.g. "> 0.7") with a characteristic (e.g. *ndvi*). New concepts referring to the different possible feature partitions (e.g. *HighNDVI, MediumNDVI, LowNDVI, HighBright, MediumBright, LowBright* concepts, etc.) are introduced and described by Description Logics formalism following the same structure as in listing 5.2.

The second rule set refers to logical expressions applied to discriminate fourteen "spectral rules" [255], such as thick clouds, vegetation, water or shadow spectral rules, etc. As an example, the Vegetation Spectral Rule label (V_{SR}) is defined as in equation 5.4:

$$\begin{aligned} V_{SR} = & \\ & (TM_2 >= 0.5 * TM_1) \; and \; (TM_2 >= 0.7 * TM_3) \; and \; (TM_3 < 0.7 * TM_4) \\ & and \; (TM_4 > max\{TM_1, TM_2, TM_3\}) \; and \; (TM_5 < 0.7 * TM_4) \\ & and \; (TM_5 >= 0.7 * TM_3) \; and \; (TM_7 < 0.7 * TM_5) \end{aligned} \quad (5.4)$$

```
V_SR ≡ ImageObject
⊓ ∃ hasFeature.(∃ ofProcessing.{ratioTM2divTM1}⊓ ∃ numericalValue.xsd:decimal[>=0.5])
⊓ ∃ hasFeature.(∃ ofProcessing.{ratioTM2divTM3}⊓ ∃ numericalValue.xsd:decimal[>=0.7])
⊓ ∃ hasFeature.(∃ ofProcessing.{ratioTM3divTM4}⊓ ∃ numericalValue.xsd:decimal[<0.7])
⊓ ∃ hasFeature.(∃ ofProcessing.{ratioTM4divMaxTM123}⊓ ∃ numericalValue.xsd:decimal[>1])
⊓ ∃ hasFeature.(∃ ofProcessing.{ratioTM5divTM3}⊓ ∃ numericalValue.xsd:decimal[>=0.7])
⊓ ∃ hasFeature.(∃ ofProcessing.{ratioTM5divTM4}⊓ ∃ numericalValue.xsd:decimal[<0.7])
⊓ ∃ hasFeature.(∃ ofProcessing.{ratioTM7divTM5}⊓ ∃ numericalValue.xsd:decimal[<0.7])
(Axiom 4)
```

LISTING 5.3: An example of spectral rules definition corresponding to equation 5.4

```
SVHNIR_SC ≡ V_SR ⊓ ∃ hasFeature.HighNDVI ⊓ ∃ hasFeature.HighNIR
          ⊓ ∃ hasFeature.(LowMIR1 ⊔ MediumMIR1)
          ⊓ ∃ hasFeature.(LowNDBSI ⊔ MediumNDBSI)
          ⊓ ∃ hasFeature.(LowMIR2 ⊔ MediumMIR2)
(Axiom 5)
```

LISTING 5.4: An example of spectral categories definition corresponding to equation 5.5.

where TM_i corresponds to the pixel value of the *ith* band of calibrated Landsat TM data.

The "Spectral Rules" are included in the ontology through the *ImageObject* concept. This concept was extended to include fourteen sub-concepts (e.g. V_{SR}) corresponding to the fourteen identified classes. Each class is described by Description Logics formalism following the same structure as in listing 5.3.

Finally, the third rule set identifies final spectral categories based on the combination of linguistic labels resulting from the Feature Space Partition and Spectral Rules rule sets. For instance, the "Strong Vegetation with High Near Infrared Spectral Category" ($SVHNIR_{SC}$) is detected when the following rules are verified (eq. 5.5):

$$\begin{aligned} SVHNIR_{SC} = & \\ & V_{SR} \textit{ and } HighNDVI \textit{ and } HighNIR \\ & \textit{and not } (HighMIR1 \textit{ or } HighMIR2 \textit{ or } HighNDBSI) \end{aligned} \quad (5.5)$$

where V_{SR} is described in equation 5.4 and *HighNDVI, HighMIR1, HighMIR2, HighNDBSI* and *HighNIR* refer to high pixel values (with different thresholds introduced in Baraldi et al. [255]) measured for different spectral indices, namely *NDVI, MIR1, MIR2, NDBSI* and *NIR* respectively.

As for the "Spectral Rules", the *ImageObject* concept is extended to include 46 sub-concepts corresponding to the descriptions of the 46 spectral categories as in listing 5.4.

5.4.1.3 The image facts

Finally, while both the reference conceptualization and the contextual knowledge constitute the TBox of the ontology, the image facts refer to contingent knowledge stored in the ABox. They refer to instances which are automatically computed to provide semantic descriptions of image objects based on concepts from the reference conceptualization, but regardless of any interpretation related to the contextual knowledge. For instance, an image object in the Landsat image is described based on its characteristics, i.e. spectral band and spectral index values. The description is written in Description Logics formalism including the concepts and instances already introduced in the reference conceptualization and contextual knowledge.

5.4.2 Implementation and results

The reference conceptualization, the contextual knowledge and the image facts were implemented in order to form a knowledge base to be analyzed by a reasoner program so as to infer new knowledge, i.e. assign pixels to spectral categories (fig. 5.14). The implementation involves two distinct steps: one is ontological (building the TBox) while the other is procedural (populating the ABox with image facts). At the ontological step, we implemented the reference conceptualization and the contextual knowledge using Protégé , a free ontology editor framework. At the procedural step, we automatically computed the image facts to be integrated as instances of the reference conceptualization. The image facts provide a quantitative description of the image content in order to be as objective as possible. The image processing tasks leading to the image facts were computed with Orfeo Toolbox (OTB). Landsat images were calibrated in top-of-atmosphere reflectance and used to compute eleven spectral indices and thirty-five ratio indices required to apply the three rule sets. We then used the owl_cpp library to integrate both approaches, i.e. to combine OWL-DL ontologies with OTB. This way, all image pixels were semantically described with Description Logics using the concepts of the ontology. Finally, the knowledge base is handled by a reasoner based on Description Logics (FaCT++, [256]). This reasoner makes explicit the instantiation links from the most specific concepts defined in the contextual knowledge, i.e. instances of images objects are assigned to any concept of the ontology whose definition matches the object's description.

We tested our approach on four subsets of Landsat images, including three Landsat 5 TM and one Landsat 8 OLI image (fig. 5.16). The tests were only performed on small image subsets (300 x 300 pixels) due to operational considerations. We then compared the output result of the ontology-based prototype with maps obtained by computing

the spectral rules using the R software [257]. Both results were entirely similar and thus confirmed the relevance of the symbolic ontology-based approach whereby pixels are classified by a reasoner.

Yet, we faced major technical issues because the Fact++ reasoner had to perform very computationally expensive operations. Indeed, the reasoning process applied to assign spectral categories to pixels was quite slow to run due to the large number of concepts and relations combined with the large number of instances to be dealt with (i.e. 90,000 pixels). Actually, we noticed that the processing time required to reason on an ontology grows exponentially with its complexity, i.e. the complexity of the concept descriptions, relationships between concepts and the number of instances to be analyzed. In order to overcome this problem, we had to implement an alternative methodology, which consisted in reasoning on one instance (e.g. one pixel) at a time in order to ensure a linear relation between the processing time and the number of instances to be classified. Although, this approach was efficient to deal with small images, it should be significantly improved in order to process large images.

5.5 Perspectives: Ontologies to advance remote sensing science

The implementation of ontology-based applications raises a wide number of perspectives at the articulation of the main research trends in remote sensing identified at the beginning of this section (Figure 5.1).

5.5.1 From pixel- to object-based image analysis

To date, I focused on the use of ontologies for pixel-based classifications. However, the perspective to integrate additional characteristics (i.e. *Observables* or *similarities*) (e.g. about patch geometry) may require moving from a pixel- to an object-based image analysis, thus adopting the GEOBIA principles [198, 199]. Considering the semantic contraction proposed by Couclelis [247], the "field-to-object" transition occurs between the 3rd and 4th steps of the contraction (Table 5.2), which I did not study yet. One main advantage of the object-based approach is that it allows analyzing a large variety of image object features. Liu et al. [258] consider three levels of features that can be derived from a segmented image: i) Level 1 features are properties of a single image object (e.g. radiometry, texture, geometry), ii) Level 2 features focus on relations between two objects (e.g. spatial or temporal relations) and iii) Level 3 features are spatial patterns in which more than two objects are involved (e.g. aligned objects). Whereas

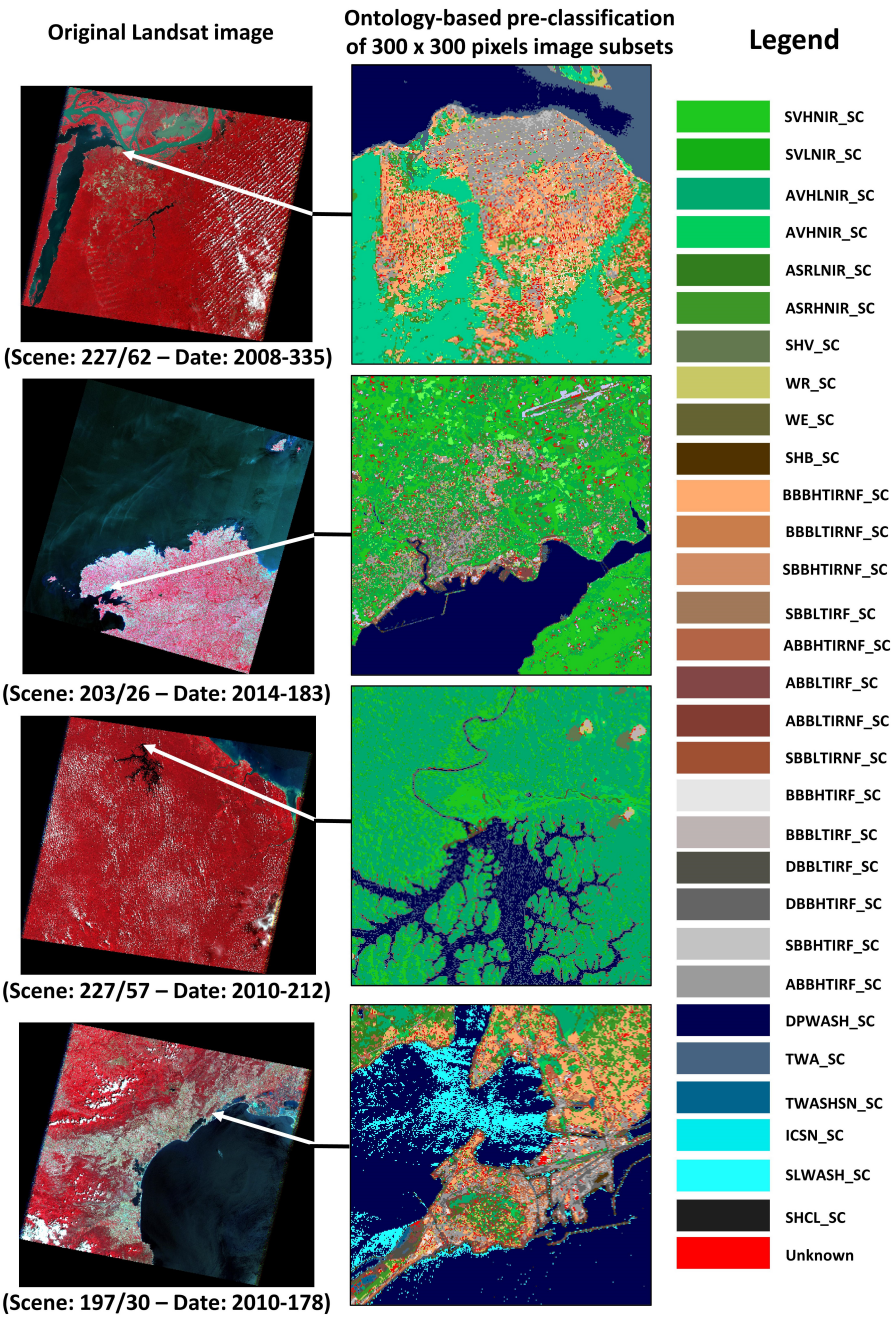


FIGURE 5.16: Ontology-based classifications of four subsets of Landsat images. The complete definitions of the legend acronyms are in Baraldi et al. [255].

the integration of Level 1 features appears quite straightforward (new concepts such as Texture and Geometry can be considered in the same way than spectral information), including Level 2 and Level 3 features in the ontology is quite challenging. The main issue with spatial and temporal relations comes from the fact that i) experts are used to describe them with vague terms (e.g. close to, etc.) which are difficult to formalize and ii) their modelling into ontologies increases the complexity of the reasoning step.

5.5.2 Data-driven remote sensing: towards hybrid approaches

Although I intend to continue working on knowledge-driven approaches, I consider that, at a time when machine learning classifiers tend to prevail in remote sensing science, hybrid approaches combining both knowledge- and data-driven approaches should be investigated. In this regard, I will pay attention to two different research areas: (1) the interpretation of knowledge generated by data-driven approaches and (2) the embedding of expert knowledge in data-driven algorithms.

In machine-learning approaches, the information contained in the numeric knowledge (i.e. resulting thresholds using Random Forests, boundaries using SVM or features using deep networks) can be very meaningful and help end-users to assign semantics on key phenomena that contribute to the recognition of geographic features. Such information should be analyzed from an expert point of view, considering interpretability and meaning in order to serve as a basis to create a "contextual knowledge" to be formalized in ontologies.

In addition, no prior expert knowledge is usually taken into account in data-driven approaches, especially in Deep Learning applications. This can be prejudicial from an application point-of-view where prior expert knowledge can be of prime importance to have a better conditioning/optimization of the classifier (better initialization, hard constraints to prevent from impossible combinations, etc) and prevent from unexpected/degenerated solution while being faster to optimize. Efficient means to bring domain expertise into deep learning to provide prior knowledge that is tailored to specific remote sensing problems should be explored. To do so, formalized expert knowledge could be used as mathematical constraints in the optimization of neural networks.

5.5.3 Connection with end-user expert knowledge

Furthermore, I consider that the prototype systems introduced in sections 5.3 and 5.4 are still incomplete since they assign image objects (or pixels) to spectral categories, which are not directly useful for domain experts. Thus, we should now connect remote

sensing knowledge with domain expert knowledge to derive land cover/land use classes or ecological habitats. In this regard, it is necessary to develop domain ontologies. For example, I started (but never finished) to implement ontologies of land cover classes and ecological habitats according to the Land Cover Classification System (LCCS; [197]) and General Habitat Categories (GHC; [259]), respectively.

Finally, both systems fall within a top-down approach where a domain ontology representing expert knowledge is first produced and then analysed to interpret the data. Such an approach has long been supported by the ontological community because it is clearly knowledge-driven. However, it has also been criticized because of the difficulty in formalizing expert knowledge and reaching an ontological commitment across a scientific community. A bottom-up approach may thus be preferred to report the diversity of viewpoints in a domain and the heterogeneity of data [260]. I suggest that numeric models derived from supervised classifiers and containing implicit expert knowledge should be converted into ontologies in order to create new knowledge bases that would make up a large set of information to be analyzed thanks to ontological capacities (ontology mapping, reasoning, etc.). In this regard, the reference conceptualization proposed in the present paper could serve as a framework to describe these process trees.

5.5.4 Ontologies and big EO data

The increasing availability of remote sensing images raises new challenges regarding the selection of relevant images in large databases, i.e., data discovery. This process is usually made up two stages: (1) identify relevant EO data type for a given application and (2) select the correct image(s) based on personal experience and knowledge of the data quality. Ontologies can help at both stages because they can be used to explicitly define the data specifications based on end-users' requirements and perform the data discovery. With regards to data identification, the objective is to indicate to the end-users which remote sensing data are relevant according to their requirements. Once the required data type has been identified, it is then an issue of selecting the most relevant data to be used. Usually, this selection process is conducted by performing queries (e.g. on acquisition date, geographic bounding box and acquisition mode) based on personal experience. Ontologies may enhance the capability to discover data sets based on their content rather than on the keywords that are usually considered. For example, an expert in urban planning could only be interested in images containing urban areas larger than a given area, or a hydrologist could select only images containing water bodies.

Similarly, ontologies can also improve the use of EO products (e.g. land cover maps) by end-users. Indeed, ontologies could assist end-users to identify the relevant remote

sensing-based products for their applications. For instance, a forestry scientist may define its own definition of "forest" and then use the reasoning potential of ontologies to look for the land cover maps that contain a "forest" class that better approximates to his definition. In addition, ontologies could also help harmonizing the heterogeneous spatial information produced by the EO community. For example, two classes with different labels (e.g. "crops" and "agriculture") from different land cover maps can be compared if all classes are explicitly and formally described in ontologies. More generally, ontologies could be used to ease the conversion of maps initially produced based on different legends. For this purpose, ontologies of thematic classification systems need to be implemented and connected by ontology mapping.

Finally, ontologies can be used to guide the image processing tasks usually implemented by Earth Observation specialists. In other words, the tasks would be computed in dedicated image processing software but the processing chains would be managed by ontologies. These processing chains actually constitute scientific workflows, i.e., executable descriptions of a scientific process. Ontologies could capture the high-level semantic aspects of a workflow (e.g., first, the user performs atmospheric correction of a Landsat TM image; second, the user performs an image segmentation using a region growing algorithm; and third, the user extracts Haralick textural index values for all segments, etc). These steps would then be performed while hiding the underlying technical details (e.g., required data format conversions, data resampling). This way, ontologies would help to ensure consistency in the scientific workflows by checking necessary interoperability between data and systems.

5.5.5 Towards collaborative remote sensing

The development of ontology-based applications requires the remote sensing community to get more familiar with ontologies. In order to ensure knowledge aggregation, sharing, and subsequently integration among various scientific domains, a highly recommended good practice in ontology building consists in identifying already existing ontologies that could be re-used, interconnected or extended by the remote sensing community. For example, because the focus is on Earth Observation, the Extensible Observation Ontology (OBOE) and the Observations Measurements (OM) ontology appear as relevant ontologies to provide a generic pattern to structure scientific observation and corresponding measurements. Both of them can be used to describe geographic features from a real-world perspective (as an end-user would do) or from an image perspective (as a remote sensing expert would do). As already mentioned in section 5.3, OBOE was first designed for ecological applications [242] but can be adapted for other applications including remote sensing. The O&M [261] ontology commits ISO/OGC standards

and can be connected to the Semantic Sensor Network (SSN) ontology that provides a framework to describe sensors and their observations for a wide range of applications including satellite imagery [262] . With regards to the Earth Environment domain, the Semantic Web for Earth and Environmental Terminology (SWEET) proposed by NASA may be an appropriate framework ontology [263]. This ontology is composed of nine modular ontologies that make it possible to describe any compound concept related to the environment [264]. Ontologies such as OM, SSN, SWEET can be interconnected and extended to include additional concepts more specific to the remote sensing domain, including the elementary concepts remote sensing experts rely on to interpret remote sensing images (e.g. concepts related to spectral bands, spectral or textural indices). Such ontology would then correspond to the "reference conceptualization" knowledge in remote sensing as proposed in section 5.2.

The implementation of such ontology of the remote sensing domain should be built collaboratively following a top-down approach, thus getting scientists involved in the identification of key concepts in remote sensing and their interrelationships. In this regard, the community could rely on interesting successful efforts for the development of web-based tools such as WebProtégé [265] or ThesauForm [266], dedicated to the collaborative construction of ontologies and thesaurus of a scientific domain.

Finally, while a top-down approach appears suitable to design framework ontologies in order to set down a "reference conceptualization" for the description of geographic features, a bottom-up approach is necessary to allow remote sensing scientists formalizing their expert "contextual knowledge". For example, similarly to spectral libraries used to store numeric knowledge about spectral signatures of target objects, I consider that the system introduced in section 5.3 could serve as a basis to build libraries of symbolic definitions of the representation of land cover classes in remote sensing images. Performing additional tests in other contexts would indeed make it possible to accumulate symbolic definitions of land cover classes in order to build a remote sensing expert knowledge database to be stored in ontologies. For example, if N end-users used this system to classify "Forests" in N study areas across the world, we could retrieve the N symbolic definitions of the representation of Forests in remote sensing images and store them into ontologies. Then, any new end-user interested in mapping Forests could reuse these ontologies containing the experience of the N first users to derive N maps of "Forests". Such strategy for ontology-building would follow the recommendations of Janowicz [260] who argued for "a radical paradigm shift in ontology engineering away from a small number of authoritative, global ontologies developed top-down, to a large number of local ontologies that are driven by application needs and developed bottom-up based on observation data."

Chapter 6

Conclusion

My researches focus on the relationships between nature and society in tropical agricultural frontiers. In this regard, I mainly worked in the Brazilian Amazon state of Mato Grosso, which concentrates all the socio-economic and environmental issues related to the conquest of the Amazon. I evidenced the rapid and intense land cover dynamics related to both deforestation and agricultural expansion-intensification processes. Overall, I questioned the land use sustainability of such dynamics, especially considering a context of climate change characterized by dramatic changes in rainfall regimes.

My results emphasize the necessity to consider the tropical agricultural frontiers as part of global dynamics, i.e. frontiers to feed a growing world population (11 billions people expected in 2100) and frontiers driven by global market interests (e.g. soybean prices). In this regard, one should now question the potential impacts of the recent global geopolitic situation. Indeed, the war in Ukraine may set up a new deal leading to huge uncertainties for the future of tropical ecosystems such as the Amazon and Cerrado biomes. On the one hand, as the war may encourage the European Union to limit its dependence on energy from Russia, it may also convince Europeans to limit food imports. Indeed, the lack of agricultural production from Ukraine may increase the commodity prices and convince importing countries to change their agricultural model. But similarly, the increase of prices may also encourage exporting countries to secure their own national production, as Argentina has recently blocked its soybean exports. However, if we consider the business-as-usual dynamics introduced in the present thesis, an increase in commodity prices is unfortunately more likely to drive cropland expansion further in the tropical forests and savannas to increase exports. On the other hand, the extreme sanctions applied to Russia may also reduce the imports of agricultural fertilizers, which are essential to sustain the agricultural production in many countries, including Brazil. If Brazil needs to increase its national production for fertilizers, one should question the

indirect potential environmental impacts of the agricultural sector, especially considering the intentions of the present government to authorize mining activities in protected areas. For this reason, it is necessary to adopt an integrated vision of environmental impacts of the agricultural sector, considering both landscape scale and potentially the life cycle of agricultural production.

For sure, remote sensing science has to play an essential role in the monitoring and understanding of past, present and future land use dynamics. The multiplication of remote sensors and the emergence of advanced techniques to process big Earth Observation data raise promising expectations to produce accurate indices of socio-environmental dynamics. The present thesis introduced a few indices I computed about, for example, the adoption of intensive agricultural practices such as double cropping or irrigation systems, the proliferation of water reservoirs or the level of forest degradation. Yet, the advent of data-driven techniques in remote sensing also calls into question a risk to increase the gap between data producers and end-users from various scientific domains. In this regard, I set out my thoughts about a necessary evolution towards the integration of knowledge management techniques such as ontologies in remote sensing applications. Unfortunately, it is evident that this issue is not a current priority research direction for the remote sensing community but I still strongly believe that it should gain in importance in the coming years. These exciting perspectives of research should now be realized through research projects, gathering interdisciplinary teams interested in remote sensing images analysis and its applications. In this regard, I intend to continue submitting new project proposals at national and international calls. I will pay a special attention to the release of scientific results to the civil society as it is already expected in the Space Climate Observatory CHOVE-CHUVA project which I coordinate and whose aim is to spread most results achieved in this thesis in Mato Grosso through a web infrastructure of spatial data. I consider this project especially relevant since it involves many non-academic partners (administrations, NGOs, associations and a private company) and may open up a new research area related to the integration of citizen science principles in remote sensing applications.

Finally, I will continue paying a special attention to the training of scientists and students to remote sensing techniques. I am already involved in capacity building sessions (which I organized in France and abroad) for the processing of spatial data in the R environment. The supervision of master and overall PhD students also appears very important to me in order to address the numerous scientific challenges ahead and accompany the rapid evolution of methods and data in remote sensing science.

Bibliography

[1] Ruth S. DeFries, Jonathan A. Foley, and Gregory P. Asner. Land-use choices: balancing human needs and ecosystem function. *Frontiers in Ecology and the Environment*, 2(5):249–257, June 2004. doi: 10.1890/1540-9295(2004)002[0249:lcbhna]2.0.co;2. URL https://doi.org/10.1890/1540-9295(2004)002[0249:lcbhna]2.0.co;2.

[2] Helmut J. Geist and Eric F. Lambin. Proximate causes and underlying driving forces of tropical deforestation. *BioScience*, 52(2):143, 2002. doi: 10.1641/0006-3568(2002)052[0143:pcaudf]2.0.co;2. URL https://doi.org/10.1641/0006-3568(2002)052[0143:pcaudf]2.0.co;2.

[3] Frederik J. Turner. *The frontier in American history*. 1893. URL http://xroads.virginia.edu/~Hyper/TURNER/.

[4] Pierre Monbeig. *Pionniers et planteurs de São Paulo*. Cahiers de la Fondation Nationale des Sciences Politiques,A. Colin., Paris, 1952.

[5] Martine Droulers and Francois-Michel Le Tourneau. Amazonie: la fin d'une frontière? *Caravelle*, 75:109–135, 2000. URL http://www.jstor.org/stable/40853836.

[6] Martin Coy. Différenciation et transformation de l'espace au nord du mato grosso. contribution a un modèle dynamique des fronts pionniers en amazonie brésilienne. *L'Harmattan*, pages 103–127, 1996.

[7] Vincent Dubreuil, Anne-Élisabeth Laques, Vincent Nédélec, Damien Arvor, and Helen Gurgel. Paysages et fronts pionniers amazoniens sous le regard des satellites : l'exemple du mato grosso. *Espace géographique*, 37(1):57, 2008. doi: 10.3917/eg.371.0057. URL https://doi.org/10.3917/eg.371.0057.

[8] Hervé Théry. Franges pionnières d'hier et d'aujourd'hui. *La Documentation Française*, pages 113–123, 2006.

[9] Philippe Léna. Aspects de la frontière amazonienne. *Cahiers des Sciences Humaines*, 22:319–343, 1986. ISSN 0768-9829. URL https://www.documentation.ird.fr/hor/fdi:23316.

[10] Christian Brannstrom. South america's neoliberal agricultural frontiers: Places of environmental sacrifice or conservation opportunity. *AMBIO: A Journal of the Human Environment*, 38(3):141–149, May 2009. doi: 10.1579/0044-7447-38.3.141. URL https://doi.org/10.1579/0044-7447-38.3.141.

[11] Thomas K. Rudel, Oliver T. Coomes, Emilio Moran, Frederic Achard, Arild Angelsen, Jianchu Xu, and Eric Lambin. Forest transitions: towards a global understanding of land use change. *Global Environmental Change*, 15(1):23–31, April 2005. doi: 10.1016/j.gloenvcha.2004.11.001. URL https://doi.org/10.1016/j.gloenvcha.2004.11.001.

[12] Damien Arvor, Vincent Dubreuil, Margareth Simões, and Agnès Bégué. Mapping and spatial analysis of the soybean agricultural frontier in Mato Grosso, Brazil, using remote sensing data. *GeoJournal*, 78(5):833–850, October 2013. ISSN 0343-2521, 1572-9893. doi: 10.1007/s10708-012-9469-3. URL http://link.springer.com/10.1007/s10708-012-9469-3.

[13] Carlos M. Souza, Julia Z. Shimbo, Marcos R. Rosa, Leandro L. Parente, Ane A. Alencar, Bernardo F. T. Rudorff, Heinrich Hasenack, Marcelo Matsumoto, Laerte G. Ferreira, Pedro W. M. Souza-Filho, Sergio W. de Oliveira, Washington F. Rocha, Antônio V. Fonseca, Camila B. Marques, Cesar G. Diniz, Diego Costa, Dyeden Monteiro, Eduardo R. Rosa, Eduardo Vélez-Martin, Eliseu J. Weber, Felipe E. B. Lenti, Fernando F. Paternost, Frans G. C. Pareyn, João V. Siqueira, José L. Viera, Luiz C. Ferreira Neto, Marciano M. Saraiva, Marcio H. Sales, Moises P. G. Salgado, Rodrigo Vasconcelos, Soltan Galano, Vinicius V. Mesquita, and Tasso Azevedo. Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine. *Remote Sensing*, 12(17):2735, August 2020. doi: 10.3390/rs12172735. URL https://doi.org/10.3390/rs12172735.

[14] Michelle Cristina Araujo Picoli, Gilberto Camara, Ieda Sanches, Rolf Simões, Alexandre Carvalho, Adeline Maciel, Alexandre Coutinho, Julio Esquerdo, João Antunes, Rodrigo Anzolin Begotti, Damien Arvor, and Claudio Almeida. Big earth observation time series analysis for monitoring brazilian agriculture. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145:328–339, November 2018. doi: 10.1016/j.isprsjprs.2018.08.007. URL https://doi.org/10.1016/j.isprsjprs.2018.08.007.

[15] Diana Weinhold and Eustaquio Reis. Transportation costs and the spatial distribution of land use in the brazilian amazon. *Global Environmental Change*, 18(1):54–68, February 2008. doi: 10.1016/j.gloenvcha.2007.06.004. URL https://doi.org/10.1016/j.gloenvcha.2007.06.004.

[16] Marie Clairay. Structures, composantes et formes spatiales d'un front pionnier situé au mato grosso. *Mappemonde*, 77(1), 2005. URL http://mappemonde.mgm.fr/num5/articles/art05102.html.

[17] Damien Arvor, Marion Daugeard, Isabelle Tritsch, Neli Aparecida De Mello-Thery, Hervé Thery, and Vincent Dubreuil. Combining socioeconomic development with environmental governance in the Brazilian Amazon: the Mato Grosso agricultural frontier at a tipping point. *Environment, Development and Sustainability*, November 2016. ISSN 1387-585X, 1573-2975. doi: 10.1007/s10668-016-9889-1. URL http://link.springer.com/10.1007/s10668-016-9889-1.

[18] Martine Droulers. *L'Amazonie*. Nathan, 1995.

[19] Dennis J. Mahar. Policies affecting land use in the brazilian amazon. *Land Use Policy*, 7(1):59–69, January 1990. doi: 10.1016/0264-8377(90)90055-4. URL https://doi.org/10.1016/0264-8377(90)90055-4.

[20] Wendy Jepson. Private agricultural colonization on a brazilian frontier, 1970–1980. *Journal of Historical Geography*, 32(4):839–863, October 2006. doi: 10.1016/j.jhg.2004.12.019. URL https://doi.org/10.1016/j.jhg.2004.12.019.

[21] Jean-Pierre Bertrand. L'avancée fulgurante du complexe soja dans le mato grosso: facteurs clés et limites. *Tiers-Monde*, 45(179):567–594, 2004.

[22] Juliana Gil, Matthias Siebold, and Thomas Berger. Adoption and development of integrated crop–livestock–forestry systems in mato grosso, brazil. *Agriculture, Ecosystems & Environment*, 199:394–406, January 2015. doi: 10.1016/j.agee.2014.10.008. URL https://doi.org/10.1016/j.agee.2014.10.008.

[23] David M. Lapola, Luiz A. Martinelli, Carlos A. Peres, Jean P. H. B. Ometto, Manuel E. Ferreira, Carlos A. Nobre, Ana Paula D. Aguiar, Mercedes M. C. Bustamante, Manoel F. Cardoso, Marcos H. Costa, Carlos A. Joly, Christiane C. Leite, Paulo Moutinho, Gilvan Sampaio, Bernardo B. N. Strassburg, and Ima C. G. Vieira. Pervasive transition of the brazilian land-use system. *Nature*

Climate Change, 4(1):27–35, December 2013. doi: 10.1038/nclimate2056. URL https://doi.org/10.1038/nclimate2056.

[24] Bernardo B.N. Strassburg, Agnieszka E. Latawiec, Luis G. Barioni, Carlos A. Nobre, Vanderley P. da Silva, Judson F. Valentim, Murilo Vianna, and Eduardo D. Assad. 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, September 2014. doi: 10.1016/j.gloenvcha.2014.06.001. URL https://doi.org/10.1016/j.gloenvcha.2014.06.001.

[25] Edward B. Barbier, Joanne C. Burgess, and Alan Grainger. The forest transition: Towards a more comprehensive theoretical framework. *Land Use Policy*, 27(2):98–107, April 2010. doi: 10.1016/j.landusepol.2009.02.001. URL https://doi.org/10.1016/j.landusepol.2009.02.001.

[26] A S Mather and C L Needle. The forest transition: a theoretical basis. *Area*, 30 (2):117–124, June 1998. doi: 10.1111/j.1475-4762.1998.tb00055.x. URL https://doi.org/10.1111/j.1475-4762.1998.tb00055.x.

[27] Thomas K. Rudel, Laura Schneider, and Maria Uriarte. Forest transitions: An introduction. *Land Use Policy*, 27(2):95–97, April 2010. doi: 10.1016/j.landusepol.2009.09.021. URL https://doi.org/10.1016/j.landusepol.2009.09.021.

[28] Herton Escobar. Scientists, environmentalists brace for brazil's right turn. *Science*, 362(6412):273–274, October 2018. doi: 10.1126/science.362.6412.273. URL https://doi.org/10.1126/science.362.6412.273.

[29] Herton Escobar. Brazilian president attacks deforestation data. *Science*, 365 (6452):419–419, August 2019. doi: 10.1126/science.365.6452.419. URL https://doi.org/10.1126/science.365.6452.419.

[30] D. Arvor, I. Tritsch, C. Barcellos, N. Jégou, and V. Dubreuil. Land use sustainability on the south-eastern amazon agricultural frontier: Recent progress and the challenges ahead. *Applied Geography*, 80:86 – 97, 2017. ISSN 0143-6228. doi: http://dx.doi.org/10.1016/j.apgeog.2017.02.003. URL http://www.sciencedirect.com/science/article/pii/S0143622817301261.

[31] Robert Walker. Sparing land for nature in the brazilian amazon: Implications from location rent theory. *Geographical Analysis*, 46(1):18–36, January 2014. doi: 10.1111/gean.12024. URL https://doi.org/10.1111/gean.12024.

[32] Eric F. Lambin and Patrick Meyfroidt. Land use transitions: Socio-ecological feedback versus socio-economic change. *Land Use Policy*, 27(2):108–118, April 2010. doi: 10.1016/j.landusepol.2009.09.003. URL https://doi.org/10.1016/j.landusepol.2009.09.003.

[33] Peter D. Richards, Robert T. Walker, and Eugenio Y. Arima. Spatially complex land change: The Indirect effect of Brazil's agricultural sector on land use in Amazonia. *Global Environmental Change*, 29:1–9, November 2014. ISSN 09593780. doi: 10.1016/j.gloenvcha.2014.06.011. URL http://linkinghub.elsevier.com/retrieve/pii/S0959378014001162.

[34] Robert Walker. The scale of forest transition: Amazonia and the atlantic forests of brazil. *Applied Geography*, 32(1):12–20, January 2012. doi: 10.1016/j.apgeog.2010.10.010. URL https://doi.org/10.1016/j.apgeog.2010.10.010.

[35] Peter D. Richards, Robert J. Myers, Scott M. Swinton, and Robert T. Walker. Exchange rates, soybean supply response, and deforestation in south america. *Global Environmental Change*, 22(2):454–462, May 2012. doi: 10.1016/j.gloenvcha.2012.01.004. URL https://doi.org/10.1016/j.gloenvcha.2012.01.004.

[36] William F. Laurance. Switch to corn promotes amazon deforestation. *Science*, 318(5857):1721–1721, December 2007. doi: 10.1126/science.318.5857.1721b. URL https://doi.org/10.1126/science.318.5857.1721b.

[37] E.V. Elferink, S. Nonhebel, and A.J.M. Schoot Uiterkamp. Does the amazon suffer from BSE prevention? *Agriculture, Ecosystems & Environment*, 120(2-4): 467–469, May 2007. doi: 10.1016/j.agee.2006.09.009. URL https://doi.org/10.1016/j.agee.2006.09.009.

[38] C.A. McAlpine, A. Etter, P.M. Fearnside, L. Seabrook, and W.F. Laurance. Increasing world consumption of beef as a driver of regional and global change: A call for policy action based on evidence from queensland (australia), colombia and brazil. *Global Environmental Change*, 19(1):21–33, February 2009. doi: 10.1016/j.gloenvcha.2008.10.008. URL https://doi.org/10.1016/j.gloenvcha.2008.10.008.

[39] Jordan Graesser, T Mitchell Aide, H Ricardo Grau, and Navin Ramankutty. Cropland/pastureland dynamics and the slowdown of deforestation in latin america. *Environmental Research Letters*, 10(3):034017, March 2015. doi: 10.1088/1748-9326/10/3/034017. URL https://doi.org/10.1088/1748-9326/10/3/034017.

[40] SUSANNA B. HECHT. From eco-catastrophe to zero deforestation? interdisciplinarities, politics, environmentalisms and reduced clearing in amazonia. *Environmental Conservation*, 39(1):4–19, September 2011. doi: 10.1017/s0376892911000452. URL https://doi.org/10.1017/s0376892911000452.

[41] Nestor Ignacio Gasparri, Tobias Kuemmerle, Patrick Meyfroidt, Yann le Polain de Waroux, and Holger Kreft. The emerging soybean production frontier in southern africa: Conservation challenges and the role of south-south telecouplings. *Conservation Letters*, 9(1):21–31, May 2015. doi: 10.1111/conl.12173. URL https://doi.org/10.1111/conl.12173.

[42] Elizabeth Alice Clements and Bernardo Mançano Fernandes. Land grabbing, agribusiness and the peasantry in brazil and mozambique. *Agrarian South: Journal of Political Economy: A triannual Journal of Agrarian South Network and CARES*, 2(1):41–69, April 2013. doi: 10.1177/2277976013477185. URL https://doi.org/10.1177/2277976013477185.

[43] Gustavo de L. T. Oliveira. The geopolitics of brazilian soybeans. *The Journal of Peasant Studies*, 43(2):348–372, June 2015. doi: 10.1080/03066150.2014.992337. URL https://doi.org/10.1080/03066150.2014.992337.

[44] Peter Potapov, Matthew C. Hansen, Lars Laestadius, Svetlana Turubanova, Alexey Yaroshenko, Christoph Thies, Wynet Smith, Ilona Zhuravleva, Anna Komarova, Susan Minnemeyer, and Elena Esipova. The last frontiers of wilderness: Tracking loss of intact forest landscapes from 2000 to 2013. *Science Advances*, 3(1), January 2017. doi: 10.1126/sciadv.1600821. URL https://doi.org/10.1126/sciadv.1600821.

[45] David Gaveau, Bruno Locatelli, Mohammad Salim, Husnayaen Husnayaen, Timer Manurung, Adrià Descals, Arild Angelsen, Erik Meijaard, and Douglas Sheil. Slowing deforestation in Indonesia follows declining oil palm expansion and lower oil prices. Technical report, January 2022. URL https://www.researchsquare.com/article/rs-143515/v1. ISSN: 2693-5015 Type: article.

[46] Dyna Rochmyaningsih. Making peace with oil palm. *Science*, 365(6449):112–115, July 2019. doi: 10.1126/science.365.6449.112. URL https://doi.org/10.1126/science.365.6449.112.

[47] Richard Stone. Can palm oil plantations come clean? *Science*, 317(5844): 1491–1491, September 2007. doi: 10.1126/science.317.5844.1491. URL https://doi.org/10.1126/science.317.5844.1491.

[48] Isabelle Tritsch and Damien Arvor. Transition in environmental governance in the Brazilian Amazon: emergence of a new pattern of socio-economic development and deforestation. *Land Use Policy*, 59:446–455, December 2016. ISSN 02648377. doi: 10.1016/j.landusepol.2016.09.018. URL http://linkinghub.elsevier.com/retrieve/pii/S0264837716305774.

[49] Hervé Théry. L'archipel brésilien. In *Le Brésil*, pages 23–46. Armand Colin, 2012. doi: 10.3917/arco.ther.2012.01.0023. URL https://doi.org/10.3917/arco.ther.2012.01.0023.

[50] Philip M. Fearnside. Environmental impacts of brazil's tucuru dam: Unlearned lessons for hydroelectric development in amazonia. *Environmental Management*, 27(3):377–396, 2001. ISSN 0364-152X, 1432-1009. doi: 10.1007/s002670010156. URL http://link.springer.com/10.1007/s002670010156.

[51] Douglas C. Morton, Ruth S. DeFries, Yosio E. Shimabukuro, Liana O. Anderson, Egidio Arai, Fernando del Bon Espirito-Santo, Ramon Freitas, and Jeff Morisette. Cropland expansion changes deforestation dynamics in the southern Brazilian Amazon. *Proceedings of the National Academy of Sciences*, 103(39): 14637–14641, September 2006. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0606377103. URL http://www.pnas.org/content/103/39/14637.

[52] Ellen Jasinski, Douglas Morton, Ruth DeFries, Yosio Shimabukuro, Liana Anderson, and Matthew Hansen. Physical landscape correlates of the expansion of mechanized agriculture in mato grosso, brazil. *Earth Interactions*, 9(16):1–18, October 2005. doi: 10.1175/ei143.1. URL https://doi.org/10.1175/ei143.1.

[53] G. L. Galford, B. Soares-Filho, and C. E. P. Cerri. Prospects for land-use sustainability on the agricultural frontier of the Brazilian Amazon. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1619): 20120171–20120171, April 2013. ISSN 0962-8436, 1471-2970. doi: 10.1098/rstb.2012.0171. URL http://rstb.royalsocietypublishing.org/cgi/doi/10.1098/rstb.2012.0171.

[54] Helmut Haberl, Marina Fischer-Kowalski, Fridolin Krausmann, Helga Weisz, and Verena Winiwarter. Progress towards sustainability? what the conceptual framework of material and energy flow accounting (MEFA) can offer. *Land Use Policy*, 21(3):199–213, July 2004. doi: 10.1016/j.landusepol.2003.10.013. URL https://doi.org/10.1016/j.landusepol.2003.10.013.

[55] Jude H. Kastens, J. Christopher Brown, Alexandre Camargo Coutinho, Christopher R. Bishop, and Júlio César D. M. Esquerdo. Soy moratorium

impacts on soybean and deforestation dynamics in mato grosso, brazil. *PLOS ONE*, 12(4):e0176168, April 2017. doi: 10.1371/journal.pone.0176168. URL https://doi.org/10.1371/journal.pone.0176168.

[56] Luis Schiesari, Andrea Waichman, Theo Brock, Cristina Adams, and Britta Grillitsch. Pesticide use and biodiversity conservation in the amazonian agricultural frontier. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1619):20120378, June 2013. doi: 10.1098/rstb.2012.0378. URL https://doi.org/10.1098/rstb.2012.0378.

[57] Jiwen Tang, Damien Arvor, Thomas Corpetti, and Ping Tang. Mapping center pivot irrigation systems in the southern amazon from sentinel-2 images. *Water*, 13(3):298, January 2021. doi: 10.3390/w13030298. URL https://doi.org/10.3390/w13030298.

[58] Avery S. Cohn, Leah K. VanWey, Stephanie A. Spera, and John F. Mustard. Cropping frequency and area response to climate variability can exceed yield response. *Nature Climate Change*, 6(6):601–604, March 2016. doi: 10.1038/nclimate2934. URL https://doi.org/10.1038/nclimate2934.

[59] Marcos H Costa, Leonardo C Fleck, Avery S Cohn, Gabriel M Abrahão, Paulo M Brando, Michael T Coe, Rong Fu, Deborah Lawrence, Gabrielle F Pires, Raphael Pousa, and Britaldo S Soares-Filho. Climate risks to amazon agriculture suggest a rationale to conserve local ecosystems. *Frontiers in Ecology and the Environment*, 17(10):584–590, November 2019. doi: 10.1002/fee.2124. URL https://doi.org/10.1002/fee.2124.

[60] Patrick Calvano Kuchler, Agnès Bégué, Margareth Simões, Raffaele Gaetano, Damien Arvor, and Rodrigo P.D. Ferraz. Assessing the optimal preprocessing steps of MODIS time series to map cropping systems in mato grosso, brazil. *International Journal of Applied Earth Observation and Geoinformation*, 92: 102150, October 2020. doi: 10.1016/j.jag.2020.102150. URL https://doi.org/10.1016/j.jag.2020.102150.

[61] Patrick Kuchler. *Utilisation de la télédétection pour la cartographie des systèmes de culture intégrés : contribution au suivi de l'agriculture bas carbone dans l'Etat du Mato Grosso, Brésil*. PhD thesis, 2021.

[62] C. Neill, M. T. Coe, S. H. Riskin, A. V. Krusche, H. Elsenbeer, M. N. Macedo, R. McHorney, P. Lefebvre, E. A. Davidson, R. Scheffler, A. M. e. S. Figueira, S. Porder, and L. A. Deegan. Watershed responses to amazon soya bean cropland expansion and intensification. *Philosophical Transactions of the Royal*

Society B: Biological Sciences, 368(1619):20120425–20120425, 2013. ISSN 0962-8436, 1471-2970. doi: 10.1098/rstb.2012.0425. URL http://rstb.royalsocietypublishing.org/cgi/doi/10.1098/rstb.2012.0425.

[63] Phillip M. Fearnside. Dams in the amazon: Belo monte and brazil's hydroelectric development of the xingu river basin. *Environmental Management*, 38(1):16–27, 2006. ISSN 0364-152X, 1432-1009. doi: 10.1007/s00267-005-0113-6. URL http://link.springer.com/10.1007/s00267-005-0113-6.

[64] Edgardo M. Latrubesse, Eugenio Y. Arima, Thomas Dunne, Edward Park, Victor R. Baker, Fernando M. d'Horta, Charles Wight, Florian Wittmann, Jansen Zuanon, Paul A. Baker, Camila C. Ribas, Richard B. Norgaard, Naziano Filizola, Atif Ansar, Bent Flyvbjerg, and Jose C. Stevaux. Damming the rivers of the amazon basin. *Nature*, 546(7658):363–369, 2017. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature22333. URL http://www.nature.com/doifinder/10.1038/nature22333.

[65] S. M. Powers, D. M. Robertson, and E. H. Stanley. Effects of lakes and reservoirs on annual river nitrogen, phosphorus, and sediment export in agricultural and forested landscapes: EFFECTS OF LAKES AND RESERVOIRS ON RIVER NUTRIENTS AND SEDIMENT. *Hydrological Processes*, 28(24):5919–5937, 2014. ISSN 08856087. doi: 10.1002/hyp.10083. URL http://doi.wiley.com/10.1002/hyp.10083.

[66] Victor J. Santucci, Stephen R. Gephard, and Stephen M. Pescitelli. Effects of multiple low-head dams on fish, macroinvertebrates, habitat, and water quality in the fox river, illinois. *North American Journal of Fisheries Management*, 25 (3):975–992, 2005. ISSN 0275-5947, 1548-8675. doi: 10.1577/M03-216.1. URL http://www.tandfonline.com/doi/abs/10.1577/M03-216.1.

[67] F. Habets, E. Philippe, E. Martin, C. H. David, and F. Leseur. Small farm dams: impact on river flows and sustainability in a context of climate change. *Hydrology and Earth System Sciences*, 18(10):4207–4222, 2014. ISSN 1607-7938. doi: 10.5194/hess-18-4207-2014. URL http://www.hydrol-earth-syst-sci.net/18/4207/2014/.

[68] J.N. Callow and K.R.J. Smettem. The effect of farm dams and constructed banks on hydrologic connectivity and runoff estimation in agricultural landscapes. *Environmental Modelling & Software*, 24(8):959–968, August 2009. ISSN 13648152. doi: 10.1016/j.envsoft.2009.02.003. URL http://linkinghub.elsevier.com/retrieve/pii/S1364815209000358.

[69] Vanda Tereza Costa Malveira, José Carlos de Araújo, and Andreas Güntner. Hydrological impact of a high-density reservoir network in semiarid northeastern brazil. *Journal of Hydrologic Engineering*, 17(1):109–117, 2012. doi: 10.1061/(ASCE)HE.1943-5584.0000404.

[70] Andreas Gütner, Maarten S. Krol, José Carlos De Araújo, and Axel Bronstert. Simple water balance modelling of surface reservoir systems in a large data-scarce semiarid region / Modélisation simple du bilan hydrologique de systèmes de réservoirs de surface dans une grande région semi-aride pauvre en données. *Hydrological Sciences Journal*, 49(5), October 2004. ISSN 0262-6667, 2150-3435. doi: 10.1623/hysj.49.5.901.55139. URL http://www.tandfonline.com/doi/abs/10.1623/hysj.49.5.901.55139.

[71] Matthew J. Deitch, Adina M. Merenlender, and Shane Feirer. Cumulative effects of small reservoirs on streamflow in northern coastal california catchments. *Water Resources Management*, 2013. ISSN 0920-4741, 1573-1650. doi: 10.1007/s11269-013-0455-4. URL http://link.springer.com/10.1007/s11269-013-0455-4.

[72] Damien Arvor, Vincent Dubreuil, Patricio Mendez del Villar, Carlos Magri Ferreira, and Margareth Simões Penello Meirelles. Développement, crises et adaptation des territoires du soja au mato grosso: l'exemple de sorriso. *Confins*, (6), June 2009. doi: 10.4000/confins.5934. URL https://doi.org/10.4000/confins.5934.

[73] Damien Arvor, Margareth Meirelles, Vincent Dubreuil, Agnes Bégué, and Yosio E. Shimabukuro. Analyzing the agricultural transition in Mato Grosso, Brazil, using satellite-derived indices. *Applied Geography*, 32(2):702–713, March 2012. ISSN 01436228. doi: 10.1016/j.apgeog.2011.08.007. URL http://linkinghub.elsevier.com/retrieve/pii/S0143622811001603.

[74] Francis E. Putz and Kent H. Redford. The importance of defining 'forest': Tropical forest degradation, deforestation, long-term phase shifts, and further transitions. *Biotropica*, 42(1):10–20, September 2009. doi: 10.1111/j.1744-7429.2009.00567.x. URL https://doi.org/10.1111/j.1744-7429.2009.00567.x.

[75] A. Baccini, W. Walker, L. Carvalho, M. Farina, D. Sulla-Menashe, and R. A. Houghton. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science*, 358(6360):230–234, October 2017. doi: 10.1126/science.aam5962. URL https://doi.org/10.1126/science.aam5962.

[76] Ana S. L. Rodrigues, Robert M. Ewers, Luke Parry, Carlos Souza, Adalberto Veríssimo, and Andrew Balmford. Boom-and-bust development patterns across the amazon deforestation frontier. *Science*, 324(5933):1435–1437, June 2009. doi: 10.1126/science.1174002. URL https://doi.org/10.1126/science.1174002.

[77] Danielle Celentano, Erin Sills, Marcio Sales, and Adalberto Veríssimo. Welfare outcomes and the advance of the deforestation frontier in the brazilian amazon. *World Development*, 40(4):850–864, April 2012. doi: 10.1016/j.worlddev.2011.09.002. URL https://doi.org/10.1016/j.worlddev.2011.09.002.

[78] Robin R. Sears, Christine Padoch, and Miguel Pinedo-Vasquez. Amazon forestry tranformed: Integrating knowledge for smallholder timber managemet in eastern brazil. *Human Ecology*, 35(6):697–707, January 2007. doi: 10.1007/s10745-006-9109-y. URL https://doi.org/10.1007/s10745-006-9109-y.

[79] Diana Weinhold, Eustáquio J. Reis, and Petterson Molina Vale. Boom-bust patterns in the Brazilian Amazon. *Global Environmental Change*, 35:391–399, November 2015. ISSN 09593780. doi: 10.1016/j.gloenvcha.2015.09.013. URL http://linkinghub.elsevier.com/retrieve/pii/S0959378015300492.

[80] Isabelle Tritsch and François-Michel Le Tourneau. Population densities and deforestation in the brazilian amazon: New insights on the current human settlement patterns. *Applied Geography*, 76:163–172, November 2016. doi: 10.1016/j.apgeog.2016.09.022. URL https://doi.org/10.1016/j.apgeog.2016.09.022.

[81] Soumyananda Dinda. Environmental kuznets curve hypothesis: A survey. *Ecological Economics*, 49(4):431–455, August 2004. doi: 10.1016/j.ecolecon.2004.02.011. URL https://doi.org/10.1016/j.ecolecon.2004.02.011.

[82] Alexander Pfaff and Robert Walker. Regional interdependence and forest "transitions": Substitute deforestation limits the relevance of local reversals. *Land Use Policy*, 27(2):119–129, April 2010. doi: 10.1016/j.landusepol.2009.07.010. URL https://doi.org/10.1016/j.landusepol.2009.07.010.

[83] Gary Koop and Lise Tole. Is there an environmental kuznets curve for deforestation? *Journal of Development Economics*, 58(1):231–244, February 1999. doi: 10.1016/s0304-3878(98)00110-2. URL https://doi.org/10.1016/s0304-3878(98)00110-2.

[84] Johanna Choumert, Pascale Combes Motel, and Hervé K. Dakpo. Is the environmental kuznets curve for deforestation a threatened theory? a meta-analysis of the literature. *Ecological Economics*, 90:19–28, June 2013. doi: 10.1016/j.ecolecon.2013.02.016. URL https://doi.org/10.1016/j.ecolecon.2013.02.016.

[85] A. Baccini, S. J. Goetz, W. S. Walker, N. T. Laporte, M. Sun, D. Sulla-Menashe, J. Hackler, P. S. A. Beck, R. Dubayah, M. A. Friedl, S. Samanta, and R. A. Houghton. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*, 2(3):182–185, January 2012. doi: 10.1038/nclimate1354. URL https://doi.org/10.1038/nclimate1354.

[86] G. M. Grossman and A. B. Krueger. Economic growth and the environment. *The Quarterly Journal of Economics*, 110(2):353–377, May 1995. doi: 10.2307/2118443. URL https://doi.org/10.2307/2118443.

[87] R. DeFries and C. Rosenzweig. Toward a whole-landscape approach for sustainable land use in the tropics. *Proceedings of the National Academy of Sciences*, 107(46):19627–19632, November 2010. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1011163107. URL http://www.pnas.org/cgi/doi/10.1073/pnas.1011163107.

[88] Christophe Amiot, Cyntia Cavalcante Santos, Damien Arvor, Beatriz Bellón, Hervé Fritz, Clément Harmange, Jeffrey D. Holland, Isabel Melo, Jean-Paul Metzger, Pierre-Cyril Renaud, Fabio de Oliveira Roque, Franco Leandro Souza, and Olivier Pays. The scale of effect depends on operational definition of forest cover—evidence from terrestrial mammals of the brazilian savanna. *Landscape Ecology*, 36(4):973–987, February 2021. doi: 10.1007/s10980-021-01196-9. URL https://doi.org/10.1007/s10980-021-01196-9.

[89] Clément Bourgoin, Julie Betbeder, Renan Le Roux, Valéry Gond, Johan Oszwald, Damien Arvor, Jacques Baudry, Hugues Boussard, Solen Le Clech, Lucas Mazzei, Hélène Dessard, Peter Läderach, Louis Reymondin, and Lilian Blanc. Looking beyond forest cover: an analysis of landscape-scale predictors of forest degradation in the brazilian amazon. *Environmental Research Letters*, 16 (11):114045, November 2021. doi: 10.1088/1748-9326/ac31eb. URL https://doi.org/10.1088/1748-9326/ac31eb.

[90] Fernando Mayer Pelicice, Angelo Antonio Agostinho, Alberto Akama, José Dilermando Andrade Filho, Valter M. Azevedo-Santos, Marcus Vinicius Moreira Barbosa, Luis Mauricio Bini, Marcelo Fulgêncio Guedes Brito, Carlos Roberto

dos Anjos Candeiro, Érica Pellegrini Caramaschi, Priscilla Carvalho, Rodrigo Assis de Carvalho, Leandro Castello, Davi Borges das Chagas, Carine Cavalcante Chamon, Guarino Rinaldi Colli, Vanessa Salete Daga, Murilo Sversut Dias, José Alexandre Felizola Diniz Filho, Philip Fearnside, Wagner de Melo Ferreira, Diego Azevedo Zoccal Garcia, Tiago Kutter Krolow, Rodrigo Ferreira Kruger, Edgardo Manuel Latrubesse, Dilermando Pereira Lima Junior, Solange de Fátima Lolis, Fabyano Alvares Cardoso Lopes, Rafael Dias Loyola, André Lincoln Barroso Magalhães, Adriana Malvasio, Paulo De Marco, Pedro Ribeiro Martins, Rosana Mazzoni, João Carlos Nabout, Mário Luis Orsi, Andre Andrian Padial, Hasley Rodrigo Pereira, Thiago Nilton Alves Pereira, Phamela Bernardes Perônico, Miguel Petrere, Renato Torres Pinheiro, Etiene Fabbrin Pires, Paulo Santos Pompeu, Thiago Costa Gonçalves Portelinha, Edson Eyji Sano, Vagner Leonardo Macedo dos Santos, Paloma Helena Fernandes Shimabukuro, Idelina Gomes da Silva, Lucas Barbosa e Souza, Francisco Leonardo Tejerina-Garro, Mariana Pires de Campos Telles, Fabrício Barreto Teresa, Sidinei Magela Thomaz, Livia Helena Tonella, Ludgero Cardoso Galli Vieira, Jean Ricardo Simões Vitule, and Jansen Zuanon. Large-scale degradation of the tocantins-araguaia river basin. *Environmental Management*, 68(4):445–452, August 2021. doi: 10.1007/s00267-021-01513-7. URL https://doi.org/10.1007/s00267-021-01513-7.

[91] Marianne Evju, Dagmar Hagen, Magni O. Kyrkjeeide, and Berit Köhler. Learning from scientific literature: Can indicators for measuring success be standardized in "on the ground" restoration? *Restoration Ecology*, 28(3): 519–531, March 2020. doi: 10.1111/rec.13149. URL https://doi.org/10.1111/rec.13149.

[92] Paul Passy, Josette Garnier, Gilles Billen, Corinne Fesneau, and Julien Tournebize. Restoration of ponds in rural landscapes: Modelling the effect on nitrate contamination of surface water (the seine river basin, france). *Science of The Total Environment*, 430:280–290, 2012. ISSN 00489697. doi: 10.1016/j.scitotenv.2012.04.035. URL http://linkinghub.elsevier.com/retrieve/pii/S0048969712005670.

[93] Allison A. Oliver, Randy A. Dahlgren, and Michael L. Deas. The upside-down river: Reservoirs, algal blooms, and tributaries affect temporal and spatial patterns in nitrogen and phosphorus in the klamath river, USA. *Journal of Hydrology*, 519:164–176, 2014. ISSN 00221694. doi: 10.1016/j.jhydrol.2014.06.025. URL http://linkinghub.elsevier.com/retrieve/pii/S0022169414004855.

[94] Alain Dauta, Yvon Lapaquellerie, and Noële Maillet. Role of the dams on the river lot on two types of pollution:point-sources (cadmium) and non-point sources (phosphorus). In Josette Garnier and Jean-Marie Mouchel, editors, *Man and River Systems*, pages 325–329. Springer Netherlands, 1999. ISBN 978-90-481-5393-0 978-94-017-2163-9. URL http://link.springer.com/10.1007/978-94-017-2163-9_34. DOI: 10.1007/978-94-017-2163-9_34.

[95] J. R. Williams. SEDIMENT ROUTING FOR AGRICULTURAL WATERSHEDS. *Journal of the American Water Resources Association*, 11(5): 965–974, 1975. ISSN 1093-474X, 1752-1688. doi: 10.1111/j.1752-1688.1975.tb01817.x. URL http://doi.wiley.com/10.1111/j.1752-1688.1975.tb01817.x.

[96] Gert Verstraeten and Jean Poesen. Estimating trap efficiency of small reservoirs and ponds: methods and implications for the assessment of sediment yield. *Progress in Physical Geography*, 24(2):219–251, 2000. ISSN 0309-1333, 1477-0296. doi: 10.1177/030913330002400204. URL http://journals.sagepub.com/doi/10.1177/030913330002400204.

[97] Francis J. Magilligan and Keith H. Nislow. Changes in hydrologic regime by dams. *Geomorphology*, 71(1):61–78, 2005. ISSN 0169555X. doi: 10.1016/j.geomorph.2004.08.017. URL http://linkinghub.elsevier.com/retrieve/pii/S0169555X05000899.

[98] Stuart E. Bunn and Angela H. Arthington. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management*, 30(4):492–507, 2002. ISSN 0364-152X, 1432-1009. doi: 10.1007/s00267-002-2737-0. URL http://link.springer.com/10.1007/s00267-002-2737-0.

[99] C. Nilsson. Fragmentation and flow regulation of the world's large river systems. *Science*, 308(5720):405–408, 2005. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1107887. URL http://www.sciencemag.org/cgi/doi/10.1126/science.1107887.

[100] Zhenyao Shen, Xiaoshu Hou, Wen Li, Guzhanuer Aini, Lei Chen, and Yongwei Gong. Impact of landscape pattern at multiple spatial scales on water quality: A case study in a typical urbanised watershed in china. *Ecological Indicators*, 48: 417–427, January 2015. doi: 10.1016/j.ecolind.2014.08.019. URL https://doi.org/10.1016/j.ecolind.2014.08.019.

[101] Ting Zhou, Jianguo Wu, and Shaolin Peng. Assessing the effects of landscape pattern on river water quality at multiple scales: A case study of the dongjiang river watershed, china. *Ecological Indicators*, 23:166–175, December 2012. doi: 10.1016/j.ecolind.2012.03.013. URL https://doi.org/10.1016/j.ecolind.2012.03.013.

[102] William F. Laurance and G. Bruce Williamson. Positive Feedbacks among Forest Fragmentation, Drought, and Climate Change in the Amazon. *Conservation Biology*, 15(6):1529–1535, December 2001. ISSN 08888892, 15231739. doi: 10.1046/j.1523-1739.2001.01093.x. URL http://doi.wiley.com/10.1046/j.1523-1739.2001.01093.x.

[103] Eric A. Davidson, Alessandro C. de Araújo, Paulo Artaxo, Jennifer K. Balch, I. Foster Brown, Mercedes M. C. Bustamante, Michael T. Coe, Ruth S. DeFries, Michael Keller, Marcos Longo, J. William Munger, Wilfrid Schroeder, Britaldo S. Soares-Filho, Carlos M. Souza, and Steven C. Wofsy. The amazon basin in transition. *Nature*, 481(7381):321–328, 2012. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature10717. URL http://www.nature.com/doifinder/10.1038/nature10717.

[104] J. Simpson, C. Kummerow, W. K. Tao, and R. F. Adler. On the tropical rainfall measuring mission (TRMM). *Meteorology and Atmospheric Physics*, 60(1-3): 19–36, 1996. doi: 10.1007/bf01029783. URL https://doi.org/10.1007/bf01029783.

[105] ERIKA LEPERS, ERIC F. LAMBIN, ANTHONY C. JANETOS, RUTH DeFRIES, FREDERIC ACHARD, NAVIN RAMANKUTTY, and ROBERT J. SCHOLES. A synthesis of information on rapid land-cover change for the period 1981–2000. *BioScience*, 55(2):115, 2005. doi: 10.1641/0006-3568(2005)055[0115:asoior]2.0.co;2. URL https://doi.org/10.1641/0006-3568(2005)055[0115:asoior]2.0.co;2.

[106] L Durieux. The impact of deforestation on cloud cover over the Amazon arc of deforestation. *Remote Sensing of Environment*, 86(1):132–140, 2003. ISSN 00344257. doi: 10.1016/S0034-4257(03)00095-6. URL http://linkinghub.elsevier.com/retrieve/pii/S0034425703000956.

[107] Beatriz M. Funatsu, Vincent Dubreuil, Chantal Claud, Damien Arvor, and Manoel A. Gan. Convective activity in Mato Grosso state (Brazil) from microwave satellite observations: Comparisons between AMSU and TRMM data sets. *Journal of Geophysical Research: Atmospheres*, 117(D16):D16109, August

2012. ISSN 01480227. doi: 10.1029/2011JD017259. URL http://doi.wiley.com/10.1029/2011JD017259.

[108] V. Dubreuil, A. Jallet, J. Ronchail, and G. Maitelli. Estimation des précipitations par télédétection au mato grosso (brésil). *Annales de l'Association Internationale de Climatologie*, 1:133–156, 2004. doi: 10.4267/climatologie.1058. URL https://doi.org/10.4267/climatologie.1058.

[109] George J. Huffman, David T. Bolvin, Eric J. Nelkin, David B. Wolff, Robert F. Adler, Guojun Gu, Yang Hong, Kenneth P. Bowman, and Erich F. Stocker. The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8(1):38–55, February 2007. doi: 10.1175/jhm560.1. URL https://doi.org/10.1175/jhm560.1.

[110] Hamed Ashouri, Kuo-Lin Hsu, Soroosh Sorooshian, Dan K. Braithwaite, Kenneth R. Knapp, L. Dewayne Cecil, Brian R. Nelson, and Olivier P. Prat. PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bulletin of the American Meteorological Society*, 96(1):69–83, January 2015. doi: 10.1175/bams-d-13-00068.1. URL https://doi.org/10.1175/bams-d-13-00068.1.

[111] Kou lin Hsu, Xiaogang Gao, Soroosh Sorooshian, and Hoshin V. Gupta. Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology*, 36(9):1176–1190, September 1997. doi: 10.1175/1520-0450(1997)036⟨1176:pefrsi⟩2.0.co;2. URL https://doi.org/10.1175/1520-0450(1997)036<1176:pefrsi>2.0.co;2.

[112] Kuo lin Hsu, Hoshin V. Gupta, Xiaogang Gao, and Soroosh Sorooshian. Estimation of physical variables from multichannel remotely sensed imagery using a neural network: Application to rainfall estimation. *Water Resources Research*, 35(5):1605–1618, May 1999. doi: 10.1029/1999wr900032. URL https://doi.org/10.1029/1999wr900032.

[113] Chis Funk, P. Peterson, D. Pedreros, J. Verdin, S. Shukla, G. Husak, J. Rowland, L. Harrison, A. Howell, and J. Michaelsen. The Climate Hazards infrared precipitation with stations - new environmental record for monitoring extremes. *Scientific Data*, 2:150066, 2015. doi: 10.1038/sdata.2015.66.

[114] Mou Leong Tan, Philip W. Gassman, and Arthur P. Cracknell. Assessment of Three Long-Term Gridded Climate Products for Hydro-Climatic Simulations in

Tropical River Basins. *Water*, 9(3):229, March 2017. ISSN 2073-4441. doi: 10.3390/w9030229. URL http://www.mdpi.com/2073-4441/9/3/229.

[115] Chiyuan Miao, Hamed Ashouri, Kuo-Lin Hsu, Soroosh Sorooshian, and Qingyun Duan. Evaluation of the PERSIANN-CDR Daily Rainfall Estimates in Capturing the Behavior of Extreme Precipitation Events over China. *Journal of Hydrometeorology*, 16(3):1387–1396, June 2015. ISSN 1525-755X, 1525-7541. doi: 10.1175/JHM-D-14-0174.1. URL http://journals.ametsoc.org/doi/10.1175/JHM-D-14-0174.1.

[116] Justine Ringard, Melanie Becker, Frederique Seyler, and Laurent Linguet. Temporal and Spatial Assessment of Four Satellite Rainfall Estimates over French Guiana and North Brazil. *Remote Sensing*, 7(12):16441–16459, December 2015. ISSN 2072-4292. doi: 10.3390/rs71215831. URL http://www.mdpi.com/2072-4292/7/12/15831.

[117] Vera Thiemig, Rodrigo Rojas, Mauricio Zambrano-Bigiarini, Vincenzo Levizzani, and Ad De Roo. Validation of Satellite-Based Precipitation Products over Sparsely Gauged African River Basins. *Journal of Hydrometeorology*, 13(6): 1760–1783, 2012. ISSN 1525-755X, 1525-7541. doi: 10.1175/JHM-D-12-032.1. URL http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-12-032.1.

[118] V. Dubreuil, N. Debortoli, B.M. Funatsu, V. Nedelec, and L. Durieux. Impact of land-cover change in the Southern Amazonia Climate: A case study for the region of Alta Floresta, Mato Grosso, Brazil. *Environ. Mon. Assess.*, 184: 877–891, 2012. doi: 10.1007/s10661-011-2006-x.

[119] Carlos A. Nobre, Gilvan Sampaio, Laura S. Borma, Juan Carlos Castilla-Rubio, José S. Silva, and Manoel Cardoso. Land-use and climate change risks in the Amazon and the need of a novel sustainable development paradigm. *Proceedings of the National Academy of Sciences*, 113(39):10759–10768, 2016. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1605516113. URL http://www.pnas.org/lookup/doi/10.1073/pnas.1605516113.

[120] Jaya Khanna, David Medvigy, Stephan Fueglistaler, and Robert Walko. Regional dry-season climate changes due to three decades of Amazonian deforestation. *Nature Climate Change*, 7(3):200–204, February 2017. ISSN 1758-678X, 1758-6798. doi: 10.1038/nclimate3226. URL http://www.nature.com/doifinder/10.1038/nclimate3226.

[121] Rong Fu, Lei Yin, Wenhong Li, Paola A. Arias, Robert E. Dickinson, Lei Huang, Sudip Chakraborty, Katia Fernandes, Brant Liebmann, Rosie Fisher, and

Ranga B. Myneni. Increased dry-season length over southern Amazonia in recent decades and its implication for future climate projection. *PNAS*, 110: 18110–18115, 2013. doi: 10.1073/pnas.1302584110.

[122] Juan P. Boisier, Philippe Ciais, Agnès Ducharne, and Matthieu Guimberteau. Projected strengthening of Amazonian dry season by constrained climate model simulations. *Nature Climate Change*, 5(7):656–660, June 2015. ISSN 1758-678X, 1758-6798. doi: 10.1038/nclimate2658. URL http://www.nature.com/doifinder/10.1038/nclimate2658.

[123] B. Cook, N. Zeng, and Jin-Ho Yoon. Will Amazonia dry out? Magnitude and causes of change from IPCC climate model projections. *Earth Interactions*, 16: 1–27, 2012. doi: 10.1175/2011EI398.1.

[124] Jhan Carlo Espinoza, Hans Segura, Josyane Ronchail, Guillaume Drapeau, and Omar Gutierrez-Cori. Evolution of wet-day and dry-day frequency in the western Amazon basin: Relationship with atmospheric circulation and impacts on vegetation: WET-DAY & DRY-DAY FREQUENCY IN AMAZON AND IMPACTS ON VEGETATION. *Water Resources Research*, 52(11):8546–8560, November 2016. ISSN 00431397. doi: 10.1002/2016WR019305. URL http://doi.wiley.com/10.1002/2016WR019305.

[125] M. Balme, S. Galle, and T. Lebel. Démarrage de la saison des pluies au sahel : variabilité aux échelles hydrologique et agronomique, analysée à partir des données epsat-niger. *Sécheresse*, 16(1):15–22, June 2005.

[126] B. Liebmann, S. J. Camargo, A. Seth, J. A. Marengo, L. M. V. Carvalho, D. Allured, R. Fu, and C. S. Vera. Onset and end of the rainy season in South America in observations and the ECHAM 4.5 Atmospheric General Circulation Model. *J. Clim.*, 20:2037–2050, 2007. doi: 10.1175/JCLI4122.1.

[127] Damien Arvor, Beatriz Funatsu, Véronique Michot, and Vincent Dubreuil. Monitoring Rainfall Patterns in the Southern Amazon with PERSIANN-CDR Data: Long-Term Characteristics and Trends. *Remote Sensing*, 9(9):889, August 2017. ISSN 2072-4292. doi: 10.3390/rs9090889. URL http://www.mdpi.com/2072-4292/9/9/889.

[128] N. S. Debortoli, V. Dubreuil, B. Funatsu, F. Delahaye, C.H. de Oliveira, S. Rodrigues-Filho, C. H. Saito, and R. Fetter. Rainfall patterns in the Southern Amazon: A chronological perspective (1971–2010). *Climatic Change*, 132:1–20, 2015. doi: 10.1007/s10584-015-1415-1.

[129] Jhan Carlo Espinoza Villar, Josyane Ronchail, Jean Loup Guyot, Gerard Cochonneau, Filizola Naziano, Waldo Lavado, Eurides De Oliveira, Rodrigo Pombosa, and Philippe Vauchel. Spatio-temporal rainfall variability in the Amazon basin countries (Brazil, Peru, Bolivia, Colombia, and Ecuador). *International Journal of Climatology*, 29(11):1574–1594, 2009. ISSN 08998418, 10970088. doi: 10.1002/joc.1791. URL http://doi.wiley.com/10.1002/joc.1791.

[130] M. A. Gan, V. E. Kousky, and C. F. Ropelewski. The South America Monsoon Circulation and Its Relationship to Rainfall over West-Central Brazil. *Journal of Climate*, 17(1):47–66, January 2004. ISSN 0894-8755, 1520-0442. doi: 10.1175/1520-0442(2004)017⟨0047:TSAMCA⟩2.0.CO;2. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%282004%29017%3C0047%3ATSAMCA%3E2.0.CO%3B2.

[131] Lei Yin, Rong Fu, Yong-Fei Zhang, Paola A. Arias, D. Nelun Fernando, Wenhong Li, Katia Fernandes, and Adam R. Bowerman. What controls the interannual variation of the wet season onsets over the Amazon? *Journal of Geophysical Research: Atmospheres*, 119(5):2314–2328, March 2014. ISSN 2169897X. doi: 10.1002/2013JD021349. URL http://doi.wiley.com/10.1002/2013JD021349.

[132] Jiayu Zhou and K-M. Lau. Does a Monsoon Climate Exist over South America? *Journal of Climate*, 11(5):1020–1040, May 1998. ISSN 0894-8755, 1520-0442. doi: 10.1175/1520-0442(1998)011⟨1020:DAMCEO⟩2.0.CO;2. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%281998%29011%3C1020%3ADAMCEO%3E2.0.CO%3B2.

[133] Beatriz M. Funatsu, Renan Le Roux, Damien Arvor, Jhan Carlo Espinoza, Chantal Claud, Josyane Ronchail, Véronique Michot, and Vincent Dubreuil. Assessing precipitation extremes (1981–2018) and deep convective activity (2002–2018) in the amazon region with CHIRPS and AMSU data. *Climate Dynamics*, 57(3-4):827–849, April 2021. doi: 10.1007/s00382-021-05742-8. URL https://doi.org/10.1007/s00382-021-05742-8.

[134] J. A. Marengo and J. C. Espinoza. Review: Extreme seasonal droughts and floods in Amazonia: causes, trends and impacts. 36:1033–1050, 2016. doi: 10.1002/joc.4420.

[135] J. C. Espinoza, Josyane Ronchail, José Antonio Marengo, and Hans Segura. Contrasting north-south changes in Amazon wet-day and dry-day frequency and related atmospheric features (1981-2017). *Climate Dynamics*, 52:5413–5430, 2019a. doi: 10.1007/s00382-018-4462-2.

[136] Beatriz M. Funatsu, Vincent Dubreuil, Amandine Racapé, Nathan S. Debortoli, Stéphanie Nasuti, and François-Michel Le Tourneau. Perceptions of climate and climate change by Amazonian communities. *Global Environmental Change*, 57: 101923, 2019. ISSN 0959-3780. doi: 10.1016/j.gloenvcha.2019.05.007.

[137] L. V. Alexander, X. Zhang, T. C. Peterson, J. Caesar, B. Gleason, A. M. G. Klein Tank, M. Haylock, D. Collins, B. Trewin, F. Rahimzadeh, A. Tagipour, K. Rupa Kumar, J. Revadekar, G. Griffiths, L. Vincent, D. B. Stephenson, J. Burn, E. Aguilar, M. Brunet, M. Taylor, M. New, P. Zhai, M. Rusticucci, and J. L. Vazquez-Aguirre. Global observed changes in daily climate extremes of temperature and precipitation. *JGR Atmosphere*, 111(D5), 2006. doi: 10.1029/2005JD006290.

[138] J. Sillmann, V. V. Kharin, X. Zhang, F.W W. Zwiers, and D. Bronaugh. Climate extremes indices in the CMIP5 multimodel ensemble: Part1. Model evaluation in the present climate. *JGR Atmosphere*, 118:1716–1733, 2013. doi: 10.1002/jgrd.50203.

[139] Jhan Carlo Espinoza, Steven Chavez, Josyane Ronchail, C. Junquas, Ken Takahashi, and Waldo Lavado. Rainfall hotspots over the southern tropical Andes: Spatial distribution, rainfall intensity, and relations with large-scale atmospheric circulation. *Water Resources Research*, 51(5):3459–3475, 2015. doi: 10.1002/2014WR016273.

[140] Leila M. V. Carvalho, Charles Jones, and Brant Liebmann. The South Atlantic Convergence Zone: Intensity, form, persistence, and relationships with intraseasonal to interannual activity and extreme rainfall. *Journal of Climate*, 17 (1):88–108, 2004. doi: 10.1175/1520-0442(2004)017⟨0088:TSACZI⟩2.0.CO;2.

[141] Damien Arvor, Vincent Dubreuil, Josyane Ronchail, Margareth Simões, and Beatriz M. Funatsu. Spatial patterns of rainfall regimes related to levels of double cropping agriculture systems in Mato Grosso (Brazil): Spatial patterns of rainfall regimes in Mato Grosso. *International Journal of Climatology*, 34(8): 2622–2633, June 2014. ISSN 08998418. doi: 10.1002/joc.3863. URL http://doi.wiley.com/10.1002/joc.3863.

[142] A. Angelsen and D. Kaimowitz. When does technological change in agriculture promote deforestation? In *Tradeoffs or synergies? Agricultural intensification, economic development and the environment*, pages 89–114. CABI, 2001. doi: 10.1079/9780851994352.0089. URL https://doi.org/10.1079/9780851994352.0089.

[143] A. Angelsen and D. Kaimowitz. Introduction: the role of agricultural technologies in tropical deforestation. In *Agricultural technologies and tropical deforestation*, pages 1–17. CABI, 2001. doi: 10.1079/9780851994512.0001. URL https://doi.org/10.1079/9780851994512.0001.

[144] Gabrielle F. Pires, Gabriel M. Abrahão, Livia M. Brumatti, Leydimere J.C. Oliveira, Marcos H. Costa, Spencer Liddicoat, Etsushi Kato, and Richard J. Ladle. Increased climate risk in brazilian double cropping agriculture systems: Implications for land use in northern brazil. *Agricultural and Forest Meteorology*, 228-229:286–298, November 2016. doi: 10.1016/j.agrformet.2016.07.005. URL https://doi.org/10.1016/j.agrformet.2016.07.005.

[145] Keith R. Spangler, Amanda H. Lynch, and Stephanie A. Spera. Precipitation drivers of cropping frequency in the brazilian cerrado: Evidence and implications for decision-making. *Weather, Climate, and Society*, 9(2):201–213, March 2017. doi: 10.1175/wcas-d-16-0024.1. URL https://doi.org/10.1175/wcas-d-16-0024.1.

[146] Juliana D B Gil, Rachael D Garrett, Alan Rotz, Vassilis Daioglou, Judson Valentim, Gabrielle F Pires, Marcos H Costa, Luciano Lopes, and Julio C Reis. Tradeoffs in the quest for climate smart agricultural intensification in mato grosso, brazil. *Environmental Research Letters*, 13(6):064025, jun 2018. doi: 10.1088/1748-9326/aac4d1. URL https://doi.org/10.1088/1748-9326/aac4d1.

[147] Gabriel M. Abrahão and Marcos H. Costa. Evolution of rain and photoperiod limitations on the soybean growing season in brazil: The rise (and possible fall) of double-cropping systems. *Agricultural and Forest Meteorology*, 256-257:32–45, June 2018. doi: 10.1016/j.agrformet.2018.02.031. URL https://doi.org/10.1016/j.agrformet.2018.02.031.

[148] Maria Carolina da Silva Andrea, Rivanildo Dallacort, João Danilo Barbieri, and Rafael Cesar Tieppo. Impacts of future climate predictions on second season maize in an agrosystem on a biome transition region in mato grosso state. *Revista Brasileira de Meteorologia*, 34(2):335–347, June 2019. doi: 10.1590/0102-77863340241. URL https://doi.org/10.1590/0102-77863340241.

[149] Anna C. Hampf, Tommaso Stella, Michael Berg-Mohnicke, Tobias Kawohl, Markus Kilian, and Claas Nendel. Future yields of double-cropping systems in the southern amazon, brazil, under climate change and technological development. *Agricultural Systems*, 177:102707, January 2020. doi:

10.1016/j.agsy.2019.102707. URL https://doi.org/10.1016/j.agsy.2019.102707.

[150] Livia Maria Brumatti, Gabrielle Ferreira Pires, and Ana Beatriz Santos. Challenges to the adaptation of double cropping agricultural systems in brazil under changes in climate and land cover. *Atmosphere*, 11(12):1310, December 2020. doi: 10.3390/atmos11121310. URL https://doi.org/10.3390/atmos11121310.

[151] Marcelo Carauta, Julia Parussis, Anna Hampf, Affonso Libera, and Thomas Berger. No more double cropping in mato grosso, brazil? evaluating the potential impact of climate change on the profitability of farm systems. *Agricultural Systems*, 190:103104, May 2021. doi: 10.1016/j.agsy.2021.103104. URL https://doi.org/10.1016/j.agsy.2021.103104.

[152] Neli Aparecida de Mello-Théry. O campo é um laboratório para a gestão ambiental. *Confins*, (45), May 2020. doi: 10.4000/confins.27859. URL https://doi.org/10.4000/confins.27859.

[153] Neli Aparecida de Mello-Théry. L'amazonie entre les myriades d'expériences et les politiques publiques dominantes. *Confins*, (26), February 2016. doi: 10.4000/confins.10749. URL https://doi.org/10.4000/confins.10749.

[154] Neli Aparecida de Mello-Théry. Politiques environnementales brésiliennes : intentions et réalités. *EchoGéo*, (41), September 2017. doi: 10.4000/echogeo.15011. URL https://doi.org/10.4000/echogeo.15011.

[155] Carolina Milhorance, Marcel Bursztyn, and Eric Sabourin. From policy mix to policy networks: assessing climate and land use policy interactions in mato grosso, brazil. *Journal of Environmental Policy & Planning*, 22(3):381–396, March 2020. doi: 10.1080/1523908x.2020.1740658. URL https://doi.org/10.1080/1523908x.2020.1740658.

[156] Elke U. Weber. What shapes perceptions of climate change? new research since 2010. *WIREs Climate Change*, 7(1):125–134, November 2015. doi: 10.1002/wcc.377. URL https://doi.org/10.1002/wcc.377.

[157] Carlos A. Nobre, Gilvan Sampaio, Laura S. Borma, Juan Carlos Castilla-Rubio, José S. Silva, and Manoel Cardoso. Land-use and climate change risks in the amazon and the need of a novel sustainable development paradigm. *Proceedings of the National Academy of Sciences*, 113(39):10759–10768, September 2016. doi: 10.1073/pnas.1605516113. URL https://doi.org/10.1073/pnas.1605516113.

[158] Shoko Sakai and Kaoru Kitajima. Tropical phenology: Recent advances and perspectives. *Ecological Research*, 34(1):50–54, January 2019. doi: 10.1111/1440-1703.1131. URL https://doi.org/10.1111/1440-1703.1131.

[159] Alexander Koltunov, Susan L. Ustin, Gregory P. Asner, and Inez Fung. Selective logging changes forest phenology in the brazilian amazon: Evidence from MODIS image time series analysis. *Remote Sensing of Environment*, 113(11):2431–2440, November 2009. doi: 10.1016/j.rse.2009.07.005. URL https://doi.org/10.1016/j.rse.2009.07.005.

[160] Matthew O Jones, John S Kimball, and Ramakrishna R Nemani. Asynchronous amazon forest canopy phenology indicates adaptation to both water and light availability. *Environmental Research Letters*, 9(12):124021, December 2014. doi: 10.1088/1748-9326/9/12/124021. URL https://doi.org/10.1088/1748-9326/9/12/124021.

[161] Jian Bi, Yuri Knyazikhin, Sungho Choi, Taejin Park, Jonathan Barichivich, Philippe Ciais, Rong Fu, Sangram Ganguly, Forrest Hall, Thomas Hilker, Alfredo Huete, Matthew Jones, John Kimball, Alexei I Lyapustin, Matti Mõttus, Ramakrishna R Nemani, Shilong Piao, Benjamin Poulter, Scott R Saleska, Sassan S Saatchi, Liang Xu, Liming Zhou, and Ranga B Myneni. Sunlight mediated seasonality in canopy structure and photosynthetic activity of amazonian rainforests. *Environmental Research Letters*, 10(6):064014, June 2015. doi: 10.1088/1748-9326/10/6/064014. URL https://doi.org/10.1088/1748-9326/10/6/064014.

[162] Russell Doughty, Philipp Köhler, Christian Frankenberg, Troy S. Magney, Xiangming Xiao, Yuanwei Qin, Xiaocui Wu, and Berrien Moore. TROPOMI reveals dry-season increase of solar-induced chlorophyll fluorescence in the amazon forest. *Proceedings of the National Academy of Sciences*, 116(44): 22393–22398, October 2019. doi: 10.1073/pnas.1908157116. URL https://doi.org/10.1073/pnas.1908157116.

[163] Alfredo R. Huete, Kamel Didan, Yosio E. Shimabukuro, Piyachat Ratana, Scott R. Saleska, Lucy R. Hutyra, Wenze Yang, Ramakrishna R. Nemani, and Ranga Myneni. Amazon rainforests green-up with sunlight in dry season. *Geophysical Research Letters*, 33(6), 2006. doi: 10.1029/2005gl025583. URL https://doi.org/10.1029/2005gl025583.

[164] Fabien Hubert Wagner, Bruno Hérault, Vivien Rossi, Thomas Hilker, Eduardo Eiji Maeda, Alber Sanchez, Alexei I. Lyapustin, Lênio Soares Galvão, Yujie Wang, and Luiz E. O. C. Aragão. Climate drivers of the amazon forest

greening. *PLOS ONE*, 12(7):e0180932, July 2017. doi: 10.1371/journal.pone.0180932. URL https://doi.org/10.1371/journal.pone.0180932.

[165] Thomas Hilker, Alexei I. Lyapustin, Compton J. Tucker, Forrest G. Hall, Ranga B. Myneni, Yujie Wang, Jian Bi, Yhasmin Mendes de Moura, and Piers J. Sellers. Vegetation dynamics and rainfall sensitivity of the amazon. *Proceedings of the National Academy of Sciences*, 111(45):16041–16046, 2014. ISSN 0027-8424. doi: 10.1073/pnas.1404870111. URL https://www.pnas.org/content/111/45/16041.

[166] Oliver L. Phillips et al. Drought sensitivity of the amazon rainforest. *Science*, 323(5919):1344–1347, March 2009. doi: 10.1126/science.1164033. URL https://doi.org/10.1126/science.1164033.

[167] Scott R. Saleska, Kamel Didan, Alfredo R. Huete, and Humberto R. da Rocha. Amazon forests green-up during 2005 drought. *Science*, 318(5850):612–612, September 2007. doi: 10.1126/science.1146663. URL https://doi.org/10.1126/science.1146663.

[168] P. M. Brando, S. J. Goetz, A. Baccini, D. C. Nepstad, P. S. A. Beck, and M. C. Christman. Seasonal and interannual variability of climate and vegetation indices across the amazon. *Proceedings of the National Academy of Sciences*, 107(33): 14685–14690, August 2010. doi: 10.1073/pnas.0908741107. URL https://doi.org/10.1073/pnas.0908741107.

[169] Liana O. Anderson, Yadvinder Malhi, Luiz E. O. C. Aragão, Richard Ladle, Egidio Arai, Nicolas Barbier, and Oliver Phillips. Remote sensing detection of droughts in amazonian forest canopies. *New Phytologist*, 187(3):733–750, July 2010. doi: 10.1111/j.1469-8137.2010.03355.x. URL https://doi.org/10.1111/j.1469-8137.2010.03355.x.

[170] Y. Malhi, J. T. Roberts, R. A. Betts, T. J. Killeen, W. Li, and C. A. Nobre. Climate change, deforestation, and the fate of the amazon. *Science*, 319(5860): 169–172, January 2008. doi: 10.1126/science.1146961. URL https://doi.org/10.1126/science.1146961.

[171] J. A. Marengo and J. C. Espinoza. Extreme seasonal droughts and floods in amazonia: causes, trends and impacts. *International Journal of Climatology*, 36 (3):1033–1050, July 2015. doi: 10.1002/joc.4420. URL https://doi.org/10.1002/joc.4420.

[172] R. Fu, L. Yin, W. Li, P. A. Arias, R. E. Dickinson, L. Huang, S. Chakraborty, K. Fernandes, B. Liebmann, R. Fisher, and R. B. Myneni. Increased dry-season length over southern amazonia in recent decades and its implication for future climate projection. *Proceedings of the National Academy of Sciences*, 110(45): 18110–18115, October 2013. doi: 10.1073/pnas.1302584110. URL https://doi.org/10.1073/pnas.1302584110.

[173] Damien Arvor, Beatriz Funatsu, Véronique Michot, and Vincent Dubreuil. Monitoring rainfall patterns in the southern amazon with PERSIANN-CDR data: Long-term characteristics and trends. *Remote Sensing*, 9(9):889, August 2017. doi: 10.3390/rs9090889. URL https://doi.org/10.3390/rs9090889.

[174] Sassan Saatchi, Salvi Asefi-Najafabady, Yadvinder Malhi, Luiz E. O. C. Aragão, Liana O. Anderson, Ranga B. Myneni, and Ramakrishna Nemani. Persistent effects of a severe drought on amazonian forest canopy. *Proceedings of the National Academy of Sciences*, 110(2):565–570, December 2012. doi: 10.1073/pnas.1204651110. URL https://doi.org/10.1073/pnas.1204651110.

[175] Liana Oighenstein Anderson, Germano Ribeiro Neto, Ana Paula Cunha, Marisa Gesteira Fonseca, Yhasmin Mendes de Moura, Ricardo Dalagnol, Fabien Hubert Wagner, and Luiz Eduardo Oliveira e Cruz de Aragão. Vulnerability of amazonian forests to repeated droughts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1760):20170411, October 2018. doi: 10.1098/rstb.2017.0411. URL https://doi.org/10.1098/rstb.2017.0411.

[176] Luiz E. O. C. Aragão, Liana O. Anderson, Marisa G. Fonseca, Thais M. Rosan, Laura B. Vedovato, Fabien H. Wagner, Camila V. J. Silva, Celso H. L. Silva Junior, Egidio Arai, Ana P. Aguiar, Jos Barlow, Erika Berenguer, Merritt N. Deeter, Lucas G. Domingues, Luciana Gatti, Manuel Gloor, Yadvinder Malhi, Jose A. Marengo, John B. Miller, Oliver L. Phillips, and Sassan Saatchi. 21st century drought-related fires counteract the decline of amazon deforestation carbon emissions. *Nature Communications*, 9(1), February 2018. doi: 10.1038/s41467-017-02771-y. URL https://doi.org/10.1038/s41467-017-02771-y.

[177] S. L. Lewis, P. M. Brando, O. L. Phillips, G. M. F. van der Heijden, and D. Nepstad. The 2010 amazon drought. *Science*, 331(6017):554–554, February 2011. doi: 10.1126/science.1200807. URL https://doi.org/10.1126/science.1200807.

[178] Liang Xu, Arindam Samanta, Marcos H. Costa, Sangram Ganguly, Ramakrishna R. Nemani, and Ranga B. Myneni. Widespread decline in greenness of amazonian vegetation due to the 2010 drought. *Geophysical Research Letters*, 38(7):n/a–n/a, April 2011. doi: 10.1029/2011gl046824. URL https://doi.org/10.1029/2011gl046824.

[179] P. M. Atkinson, J. Dash, and C. Jeganathan. Amazon vegetation greenness as measured by satellite sensors over the last decade. *Geophysical Research Letters*, 38(19):n/a–n/a, October 2011. doi: 10.1029/2011gl049118. URL https://doi.org/10.1029/2011gl049118.

[180] P. M. Cox, R. A. Betts, M. Collins, P. P. Harris, C. Huntingford, and C. D. Jones. Amazonian forest dieback under climate-carbon cycle projections for the 21st century. *Theoretical and Applied Climatology*, 78(1-3), April 2004. doi: 10.1007/s00704-004-0049-4. URL https://doi.org/10.1007/s00704-004-0049-4.

[181] Damien Arvor, Julie Betbeder, Felipe R.G. Daher, Tim Blossier, Renan Le Roux, Samuel Corgne, Thomas Corpetti, Vinicius de Freitas Silgueiro, and Carlos Antonio da Silva Junior. Towards user-adaptive remote sensing: Knowledge-driven automatic classification of sentinel-2 time series. *Remote Sensing of Environment*, 264:112615, October 2021. doi: 10.1016/j.rse.2021.112615. URL https://doi.org/10.1016/j.rse.2021.112615.

[182] Samuel Andrés, Damien Arvor, Isabelle Mougenot, Thérèse Libourel, and Laurent Durieux. Ontology-based classification of remote sensing images using spectral rules. *Computers & Geosciences*, 102:158–166, May 2017. ISSN 00983004. doi: 10.1016/j.cageo.2017.02.018. URL http://linkinghub.elsevier.com/retrieve/pii/S0098300417302078.

[183] John R. Jensen. Biophysical remote sensing. *Annals of the Association of American Geographers*, 73(1):111–132, mar 1983. doi: 10.1111/j.1467-8306.1983.tb01399.x. URL https://doi.org/10.1111/j.1467-8306.1983.tb01399.x.

[184] Statistical learning for earth observation data analysis (erc project), 2014. URL http://isp.uv.es/sedal.html.

[185] Mariana Belgiu and Lucian DrÄfguÅ£. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114(Supplement C):24 – 31, 2016. ISSN 0924-2716. doi:

https://doi.org/10.1016/j.isprsjprs.2016.01.011. URL http://www.sciencedirect.com/science/article/pii/S0924271616000265.

[186] Giorgos Mountrakis, Jungho Im, and Caesar Ogole. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3):247 – 259, 2011. ISSN 0924-2716. doi: https://doi.org/10.1016/j.isprsjprs.2010.11.001. URL http://www.sciencedirect.com/science/article/pii/S0924271610001140.

[187] Xiao Xiang Zhu, Devis Tuia, Lichao Mou, Gui-Song Xia, Liangpei Zhang, Feng Xu, and Friedrich Fraundorfer. Deep learning in remote sensing: a review. *CoRR*, abs/1710.03959, 2017. URL http://arxiv.org/abs/1710.03959.

[188] Damien Arvor, Milton Jonathan, Margareth Simões Penello Meirelles, Vincent Dubreuil, and Laurent Durieux. Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. *International Journal of Remote Sensing*, 32(22):7847–7871, November 2011. ISSN 0143-1161. doi: 10.1080/01431161.2010.531783. URL http://dx.doi.org/10.1080/01431161.2010.531783.

[189] Steffen Fritz, Ian McCallum, Christian Schill, Christoph Perger, Linda See, Dmitry Schepaschenko, Marijn van der Velde, Florian Kraxner, and Michael Obersteiner. Geo-wiki: An online platform for improving global land cover. *Environmental Modelling & Software*, 31:110–123, May 2012. doi: 10.1016/j.envsoft.2011.11.015. URL https://doi.org/10.1016/j.envsoft.2011.11.015.

[190] Linda See, Peter Mooney, Giles Foody, Lucy Bastin, Alexis Comber, Jacinto Estima, Steffen Fritz, Norman Kerle, Bin Jiang, Mari Laakso, Hai-Ying Liu, Grega Milčinski, Matej Nikšič, Marco Painho, Andrea Pődör, Ana-Maria Olteanu-Raimond, and Martin Rutzinger. Crowdsourcing, citizen science or volunteered geographic information? the current state of crowdsourced geographic information. *ISPRS International Journal of Geo-Information*, 5(5): 55, April 2016. doi: 10.3390/ijgi5050055. URL https://doi.org/10.3390/ijgi5050055.

[191] Emilien Alvarez-Vanhard, Thomas Houet, Cendrine Mony, Lucie Lecoq, and Thomas Corpetti. Can uavs fill the gap between in situ surveys and satellites for habitat mapping? *Remote Sensing of Environment*, 243:111780, 2020. ISSN 0034-4257. doi: https://doi.org/10.1016/j.rse.2020.111780. URL http://www.sciencedirect.com/science/article/pii/S0034425720301504.

[192] Duccio Rocchini, Giles M. Foody, Harini Nagendra, Carlo Ricotta, Madhur Anand, Kate S. He, Valerio Amici, Birgit Kleinschmit, Michael Förster, Sebastian Schmidtlein, Hannes Feilhauer, Anne Ghisla, Markus Metz, and Markus Neteler. Uncertainty in ecosystem mapping by remote sensing. *Computers & Geosciences*, 50:128–135, January 2013. ISSN 00983004. doi: 10.1016/j.cageo.2012.05.022. URL http://linkinghub.elsevier.com/retrieve/pii/S0098300412001781.

[193] Robert E Kennedy, Serge Andréfouët, Warren B Cohen, Cristina Gómez, Patrick Griffiths, Martin Hais, Sean P Healey, Eileen H Helmer, Patrick Hostert, Mitchell B Lyons, Garrett W Meigs, Dirk Pflugmacher, Stuart R Phinn, Scott L Powell, Peter Scarth, Susmita Sen, Todd A Schroeder, Annemarie Schneider, Ruth Sonnenschein, James E Vogelmann, Michael A Wulder, and Zhe Zhu. Bringing an ecological view of change to landsat-based remote sensing. *Frontiers in Ecology and the Environment*, 12(6):339–346, August 2014. doi: 10.1890/130066. URL https://doi.org/10.1890/130066.

[194] Gary Marcus. Deep learning: A critical appraisal, 2018.

[195] Christopher Small. Grand challenges in remote sensing image analysis and classification. *Frontiers in Remote Sensing*, September 2020. doi: 10.3389/frsen.2020.605220.

[196] Damien Arvor, Mariana Belgiu, Zoe Falomir, Isabelle Mougenot, and Laurent Durieux. Ontologies to interpret remote sensing images: why do we need them? *GIScience & Remote Sensing*, 56(6):911–939, March 2019. doi: 10.1080/15481603.2019.1587890. URL https://doi.org/10.1080/15481603.2019.1587890.

[197] Antonio Di Gregorio, Food and Agriculture Organization of the United Nations, and United Nations Environment Programme. *Land cover classification system: classification concepts and user manual: LCCS*. Number 8 in Environment and natural resources series. Food and Agriculture Organization of the United Nations, Rome, software version 2 edition, 2005. ISBN 92-5-105327-8.

[198] Thomas Blaschke, Geoffrey J. Hay, Maggi Kelly, Stefan Lang, Peter Hofmann, Elisabeth Addink, Raul Queiroz Feitosa, Freek van der Meer, Harald van der Werff, Frieke van Coillie, and Dirk Tiede. Geographic object-based image analysis – towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87:180–191, jan 2014. doi: 10.1016/j.isprsjprs.2013.09.014. URL https://doi.org/10.1016/j.isprsjprs.2013.09.014.

[199] T. Blaschke. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1):2–16, January 2010. doi:

10.1016/j.isprsjprs.2009.06.004. URL https://doi.org/10.1016/j.isprsjprs.2009.06.004.

[200] G. J. Hay and G. Castilla. Geographic object-based image analysis (GEOBIA): A new name for a new discipline. In *Lecture Notes in Geoinformation and Cartography*, pages 75–89. Springer Berlin Heidelberg. doi: 10.1007/978-3-540-77058-9_4. URL https://doi.org/10.1007/978-3-540-77058-9_4.

[201] Thomas Blaschke. Object based image analysis for remote sensing. *ISPRS journal of photogrammetry and remote sensing*, 65(1):2–16, 2010.

[202] Thomas Blaschke and Josef Strobl. What's wrong with pixels? some recent developments interfacing remote sensing and gis. *GeoBIT/GIS*, 6:12–17, 2001. URL https://ci.nii.ac.jp/naid/10027626746/en/.

[203] Richard Lucas, Palma Blonda, Peter Bunting, Gwawr Jones, Jordi Inglada, Marcela Arias, Vasiliki Kosmidou, Zisis I. Petrou, Ioannis Manakos, Maria Adamo, Rebecca Charnock, Cristina Tarantino, Caspar A. Mücher, Rob H.G. Jongman, Henk Kramer, Damien Arvor, João Pradinho Honrado, and Paola Mairota. The earth observation data for habitat monitoring (EODHaM) system. *International Journal of Applied Earth Observation and Geoinformation*, 37: 17–28, May 2015. doi: 10.1016/j.jag.2014.10.011. URL https://doi.org/10.1016/j.jag.2014.10.011.

[204] S. Lang. Object-based image analysis for remote sensing applications: modeling reality – dealing with complexity. In *Lecture Notes in Geoinformation and Cartography*, pages 3–27. Springer Berlin Heidelberg, 2008. doi: 10.1007/978-3-540-77058-9_1. URL https://doi.org/10.1007/978-3-540-77058-9_1.

[205] Damien Arvor, Laurent Durieux, Samuel Andrés, and Marie-Angélique Laporte. Advances in geographic object-based image analysis with ontologies: A review of main contributions and limitations from a remote sensing perspective. *ISPRS Journal of Photogrammetry and Remote Sensing*, 82:125–137, August 2013. doi: 10.1016/j.isprsjprs.2013.05.003. URL https://doi.org/10.1016/j.isprsjprs.2013.05.003.

[206] Mariana Belgiu, Lucian Drăguţ, and Josef Strobl. Quantitative evaluation of variations in rule-based classifications of land cover in urban neighbourhoods using worldview-2 imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87:205–215, 2014.

[207] Maria Adamo, Valeria Tomaselli, Cristina Tarantino, Saverio Vicario, Giuseppe Veronico, Richard Lucas, and Palma Blonda. Knowledge-based classification of grassland ecosystem based on multi-temporal WorldView-2 data and FAO-LCCS taxonomy. *Remote Sensing*, 12(9):1447, May 2020. doi: 10.3390/rs12091447. URL https://doi.org/10.3390/rs12091447.

[208] Accurate and scalable processing of big data in earth observation (erc project), 2017. URL http://bigearth.eu/.

[209] George Grekousis, Giorgos Mountrakis, and Marinos Kavouras. An overview of 21 global and 43 regional land-cover mapping products. *International Journal of Remote Sensing*, pages 1–27, October 2015. ISSN 0143-1161, 1366-5901. doi: 10.1080/01431161.2015.1093195. URL http://www.tandfonline.com/doi/full/10.1080/01431161.2015.1093195.

[210] Russell Congalton, Jianyu Gu, Kamini Yadav, Prasad Thenkabail, and Mutlu Ozdogan. Global Land Cover Mapping: A Review and Uncertainty Analysis. *Remote Sensing*, 6(12):12070–12093, December 2014. ISSN 2072-4292. doi: 10.3390/rs61212070. URL http://www.mdpi.com/2072-4292/6/12/12070/.

[211] M. C. Hansen, P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 342 (6160):850–853, November 2013. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1244693. URL http://www.sciencemag.org/cgi/doi/10.1126/science.1244693.

[212] Chen Jun, Yifang Ban, and Songnian Li. Open access to Earth land-cover map. *Nature*, 514:434, October 2014. URL https://doi.org/10.1038/514434c.

[213] Jean-François Pekel, Andrew Cottam, Noel Gorelick, and Alan S. Belward. High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633):418–422, December 2016. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature20584. URL http://www.nature.com/doifinder/10.1038/nature20584.

[214] Jordi Inglada, Arthur Vincent, Marcela Arias, Benjamin Tardy, David Morin, and Isabel Rodes. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. *Remote Sensing*, 9(1):95, January 2017. ISSN 2072-4292. doi: 10.3390/rs9010095. URL http://www.mdpi.com/2072-4292/9/1/95.

[215] Gregory P. Asner. Automated mapping of tropical deforestation and forest degradation: CLASlite. *Journal of Applied Remote Sensing*, 3(1):033543, August 2009. ISSN 1931-3195. doi: 10.1117/1.3223675. URL http://remotesensing.spiedigitallibrary.org/article.aspx?doi=10.1117/1.3223675.

[216] Jordi Inglada, Arthur Vincent, Marcela Arias, Benjamin Tardy, David Morin, and Isabel Rodes. Operational high resolution land cover map production at the country scale using satellite image time series. *Remote Sensing*, 9(1):95, January 2017. doi: 10.3390/rs9010095. URL https://doi.org/10.3390/rs9010095.

[217] Noel Gorelick, Matt Hancher, Mike Dixon, Simon Ilyushchenko, David Thau, and Rebecca Moore. Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202:18–27, December 2017. doi: 10.1016/j.rse.2017.06.031. URL https://doi.org/10.1016/j.rse.2017.06.031.

[218] Mapbiomas project, 2018. URL http://mapbiomas.org/.

[219] Steffen Fritz, Cidália Fonte, and Linda See. The role of citizen science in earth observation. *Remote Sensing*, 9(4):357, apr 2017. doi: 10.3390/rs9040357. URL https://doi.org/10.3390/rs9040357.

[220] Michael F. Goodchild, May Yuan, and Thomas J. Cova. Towards a general theory of geographic representation in GIS. *International Journal of Geographical Information Science*, 21(3):239–260, March 2007. ISSN 1365-8816, 1362-3087. doi: 10.1080/13658810600965271. URL http://www.tandfonline.com/doi/abs/10.1080/13658810600965271.

[221] M. Haklay and P. Weber. OpenStreetMap: User-generated street maps. *IEEE Pervasive Computing*, 7(4):12–18, oct 2008. doi: 10.1109/mprv.2008.80. URL https://doi.org/10.1109/mprv.2008.80.

[222] Steffen Fritz, Ian McCallum, Christian Schill, Christoph Perger, Roland Grillmayer, Frédéric Achard, Florian Kraxner, and Michael Obersteiner. Geo-wiki.org: The use of crowdsourcing to improve global land cover. *Remote Sensing*, 1(3):345–354, aug 2009. doi: 10.3390/rs1030345. URL https://doi.org/10.3390/rs1030345.

[223] Harnessing the power of crowdsourcing to improve land cover and land-use information (erc project), 2013. URL http://www.iiasa.ac.at/web/home/research/researchPrograms/EcosystemsServicesandManagement/Crowdland.html.

[224] The ceos database, 2019. URL http://database.eohandbook.com/.

[225] Index database. a database for remote sensing indices, 2019. URL https://www.indexdatabase.de/.

[226] Damien Arvor, Laurent Durieux, Samuel Andrés, and Marie-Angélique Laporte. Advances in Geographic Object-Based Image Analysis with ontologies: A review of main contributions and limitations from a remote sensing perspective. *ISPRS Journal of Photogrammetry and Remote Sensing*, 82:125–137, August 2013. ISSN 09242716. doi: 10.1016/j.isprsjprs.2013.05.003. URL http://linkinghub.elsevier.com/retrieve/pii/S092427161300124X.

[227] Brandon Bennett. What is a forest? on the vagueness of certain geographic concepts. *Topoi*, 20:189–201, 2002.

[228] Brandon Bennett. Modes of concept definition and varieties of vagueness. *Applied Ontology*, 1(1):17–26, 2005.

[229] Brandon Bennett. Foundations for an ontology of environment and habitat. In *Proceedings of the 2010 Conference on Formal Ontology in Information Systems: Proceedings of the Sixth International Conference (FOIS 2010)*, pages 31–44, Amsterdam, The Netherlands, The Netherlands, 2010. IOS Press. ISBN 978-1-60750-534-1. URL http://dl.acm.org/citation.cfm?id=1804715.1804722.

[230] Robin L. Chazdon, Pedro H. S. Brancalion, Lars Laestadius, Aoife Bennett-Curry, Kathleen Buckingham, Chetan Kumar, Julian Moll-Rocek, Ima Célia Guimarães Vieira, and Sarah Jane Wilson. When is a forest a forest? forest concepts and definitions in the era of forest and landscape restoration. *Ambio*, 45(5):538–550, 2016. ISSN 1654-7209. doi: 10.1007/s13280-016-0772-y. URL http://dx.doi.org/10.1007/s13280-016-0772-y.

[231] Alan H. Strahler, Curtis E. Woodcock, and James A. Smith. On the nature of models in remote sensing. *Remote Sensing of Environment*, 20(2):121–139, October 1986. doi: 10.1016/0034-4257(86)90018-0. URL https://doi.org/10.1016/0034-4257(86)90018-0.

[232] David M. Mark. Toward a theoretical framework for geographic entity types. In Andrew U. Frank and Irene Campari, editors, *Spatial Information Theory A Theoretical Basis for GIS: European Conference, COSIT'93 Marciana Marina, Elba Island, Italy September 19–22, 1993 Proceedings*, pages 270–283. Springer Berlin Heidelberg, Berlin, Heidelberg, 1993. ISBN 978-3-540-47966-6. doi: 10.1007/3-540-57207-4_18. URL http://dx.doi.org/10.1007/3-540-57207-4_18.

[233] Vlasios Voudouris. Towards a unifying formalisation of geographic representation: the object–field model with uncertainty and semantics. *International Journal of Geographical Information Science*, 24(12):1811–1828, 2010.

[234] Christina Unger and Philipp Cimiano. Representing and resolving ambiguities in ontology-based question answering. In *Proceedings of the TextInfer 2011 Workshop on Textual Entailment*, TIWTE '11, pages 40–49, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics. ISBN 978-1-937284-15-2. URL http://dl.acm.org/citation.cfm?id=2140473.2140479.

[235] H Gyde Lund. *Definitions of forest, deforestation, afforestation, and reforestation*. Forest Information Services Gainesville, VA, 2006.

[236] Roel Wieringa. *Design methods for reactive systems: Yourdan, Statemate, and the UML*. Morgan Kaufmann Publishers, Boston, 2003. ISBN 978-1-55860-755-2.

[237] Stewart Shapiro. *Vagueness in context*. Oxford University Press : Clarendon Press, Oxford ; New York, 2006. ISBN 978-0-19-928039-1.

[238] Arnold W. M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12): 1349–1380, 2000.

[239] Irving Biederman. Recognition-by-components: a theory of human image understanding. *Psychological review*, 94(2):115, 1987.

[240] Thomas R. Gruber. Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies*, 43(5-6): 907–928, November 1995. ISSN 10715819. doi: 10.1006/ijhc.1995.1081. URL http://linkinghub.elsevier.com/retrieve/pii/S1071581985710816.

[241] Franz Baader. *The description logic handbook: Theory, implementation and applications*. Cambridge university press, 2003.

[242] Joshua Madin, Shawn Bowers, Mark Schildhauer, Sergeui Krivov, Deana Pennington, and Ferdinando Villa. An ontology for describing and synthesizing ecological observation data. *Ecological informatics*, 2(3):279–296, 2007.

[243] Joshua S Madin, Shawn Bowers, Mark P Schildhauer, and Matthew B Jones. Advancing ecological research with ontologies. *Trends in ecology & evolution*, 23 (3):159–168, 2008.

[244] Fernando Bobillo and Umberto Straccia. Fuzzy ontology representation using owl 2. *International Journal of Approximate Reasoning*, 52(7):1073 – 1094, 2011. ISSN 0888-613X. doi: https://doi.org/10.1016/j.ijar.2011.05.003. Selected Papers - Uncertain Reasoning Track - FLAIRS 2009.

[245] Zoe Falomir, Ernesto Jiménez-Ruiz, M. Teresa Escrig, and Lledó Museros. Describing Images Using Qualitative Models and Description Logics. *Spatial Cognition & Computation*, 11(1):45–74, February 2011. ISSN 1387-5868, 1542-7633. doi: 10.1080/13875868.2010.545611. URL http://www.tandfonline.com/doi/abs/10.1080/13875868.2010.545611.

[246] Helen Couclelis. People manipulate objects (but cultivate fields): Beyond the raster-vector debate in GIS. In *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, pages 65–77. Springer Berlin Heidelberg, 1992. doi: 10.1007/3-540-55966-3_3. URL https://doi.org/10.1007/3-540-55966-3_3.

[247] Helen Couclelis. Ontologies of geographic information. *International Journal of Geographical Information Science*, 24(12):1785–1809, November 2010. doi: 10.1080/13658816.2010.484392. URL https://doi.org/10.1080/13658816.2010.484392.

[248] Christophe Claramunt. Ontologies for geospatial information: Progress and challenges ahead. *Journal of Spatial Information Science*, (20), June 2020. doi: 10.5311/josis.2020.20.666. URL https://doi.org/10.5311/josis.2020.20.666.

[249] Stefan Lang, Stefan Kienberger, Dirk Tiede, Michael Hagenlocher, and Lena Pernkopf. Geons – domain-specific regionalization of space. *Cartography and Geographic Information Science*, 41(3):214–226, April 2014. doi: 10.1080/15230406.2014.902755. URL https://doi.org/10.1080/15230406.2014.902755.

[250] Robert David Sack. A concept of physical space in geography. *Geographical Analysis*, 5(1):16–34, September 2010. doi: 10.1111/j.1538-4632.1973.tb00994.x. URL https://doi.org/10.1111/j.1538-4632.1973.tb00994.x.

[251] Vlasios Voudouris. Towards a unifying formalisation of geographic representation: the object–field model with uncertainty and semantics. *International Journal of Geographical Information Science*, 24(12):1811–1828, November 2010. doi: 10.1080/13658816.2010.488237. URL https://doi.org/10.1080/13658816.2010.488237.

[252] G. Castilla and G. J. Hay. Image objects and geographic objects. In *Lecture Notes in Geoinformation and Cartography*, pages 91–110. Springer Berlin Heidelberg. doi: 10.1007/978-3-540-77058-9_5. URL https://doi.org/10.1007/978-3-540-77058-9_5.

[253] Gilberto Camara, Luiz Fernando Assis, Gilberto Ribeiro, Karine Reis Ferreira, Eduardo Llapa, and Lubia Vinhas. Big earth observation data analytics. In *Proceedings of the 5th ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data - BigSpatial '16*. ACM Press, 2016. doi: 10.1145/3006386.3006393. URL https://doi.org/10.1145/3006386.3006393.

[254] Frank Jackson. Epiphenomenal qualia. *The Philosophical Quarterly*, 32(127):127, April 1982. doi: 10.2307/2960077. URL https://doi.org/10.2307/2960077.

[255] A. Baraldi, V. Puzzolo, P. Blonda, L. Bruzzone, and C. Tarantino. Automatic Spectral Rule-Based Preliminary Mapping of Calibrated Landsat TM and ETM+ Images. *IEEE Transactions on Geoscience and Remote Sensing*, 44(9): 2563–2586, September 2006. ISSN 0196-2892. doi: 10.1109/TGRS.2006.874140. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1677766.

[256] Dmitry Tsarkov and Ian Horrocks. Fact++ description logic reasoner: System description. In Ulrich Furbach and Natarajan Shankar, editors, *Automated Reasoning*, pages 292–297, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg. ISBN 978-3-540-37188-5.

[257] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. URL https://www.R-project.org.

[258] Yu Liu, Qinghua Guo, and Maggi Kelly. A framework of region-based spatial relations for non-overlapping features and its application in object based image analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(4):461–475, July 2008. doi: 10.1016/j.isprsjprs.2008.01.007. URL https://doi.org/10.1016/j.isprsjprs.2008.01.007.

[259] R. G. H. Bunce, M. J. Metzger, R. H. G. Jongman, J. Brandt, G. de Blust, R. Elena-Rossello, G. B. Groom, L. Halada, G. Hofer, D. C. Howard, P. Kovář, C. A. Mücher, E. Padoa-Schioppa, D. Paelinx, A. Palo, M. Perez-Soba, I. L. Ramos, P. Roche, H. Skånes, and T. Wrbka. A standardized procedure for surveillance and monitoring european habitats and provision of spatial data. *Landscape Ecology*, 23(1):11–25, November 2007. doi:

10.1007/s10980-007-9173-8. URL https://doi.org/10.1007/s10980-007-9173-8.

[260] Krzysztof Janowicz. Observation-Driven Geo-Ontology Engineering: Observation-Driven Geo-Ontology Engineering. *Transactions in GIS*, 16(3): 351–374, June 2012. ISSN 13611682. doi: 10.1111/j.1467-9671.2012.01342.x. URL http://doi.wiley.com/10.1111/j.1467-9671.2012.01342.x.

[261] Simon J. D. Cox. An explicit owl representation of iso/ogc observations and measurements. In *Proceedings of the 6th International Conference on Semantic Sensor Networks - Volume 1063*, SSN'13, pages 1–18, Aachen, Germany, Germany, 2013. CEUR-WS.org. URL http://dl.acm.org/citation.cfm?id=2874543.2874544.

[262] Michael Compton, Payam Barnaghi, Luis Bermudez, Raúl García-Castro, Oscar Corcho, Simon Cox, John Graybeal, Manfred Hauswirth, Cory Henson, Arthur Herzog, Vincent Huang, Krzysztof Janowicz, W. David Kelsey, Danh Le Phuoc, Laurent Lefort, Myriam Leggieri, Holger Neuhaus, Andriy Nikolov, Kevin Page, Alexandre Passant, Amit Sheth, and Kerry Taylor. The SSN ontology of the w3c semantic sensor network incubator group. *Web Semantics: Science, Services and Agents on the World Wide Web*, 17:25–32, dec 2012. doi: 10.1016/j.websem.2012.05.003. URL https://doi.org/10.1016/j.websem.2012.05.003.

[263] Robert G. Raskin and Michael J. Pan. Knowledge representation in the semantic web for earth and environmental terminology (SWEET). *Computers & Geosciences*, 31(9):1119–1125, nov 2005. doi: 10.1016/j.cageo.2004.12.004. URL https://doi.org/10.1016/j.cageo.2004.12.004.

[264] Ajay Tripathi and Hassan A. Babaie. Developing a modular hydrogeology ontology by extending the SWEET upper-level ontologies. *Computers & Geosciences*, 34(9):1022–1033, sep 2008. doi: 10.1016/j.cageo.2007.08.009. URL https://doi.org/10.1016/j.cageo.2007.08.009.

[265] Tania Tudorache, Csongor Nyulas, Natalya F. Noy, and Mark A. Musen. Webprotégé: A collaborative ontology editor and knowledge acquisition tool for the web. *Semantic Web*, 4(1):89–99, 2013. ISSN 15700844. doi: 10.3233/SW-2012-0057. URL https://doi.org/10.3233/SW-2012-0057.

[266] Marie-Angélique Laporte, Isabelle Mougenot, and Eric Garnier. ThesauForm—traits: A web based collaborative tool to develop a thesaurus for plant functional diversity research. *Ecological Informatics*, 11:34–44, sep 2012.

doi: 10.1016/j.ecoinf.2012.04.004. URL https://doi.org/10.1016/j.ecoinf.2012.04.004.