Patterns of tree species composition and richness across the principal biomes of lowland tropical South America and their underlying environmental drivers

Pedro Luiz Silva de Miranda

Thesis submitted in fulfilment of the requirements for the degree of Doctor of
Philosophy to the University of Edinburgh – 2018

Declaration

I hereby declare that the work contained in this thesis is my own unless otherwise acknowledged and cited. This thesis has not in whole or in part been previously presented for any degree.

Pedro Luiz Silva de Miranda

16th February, 2019

"When I don't know where I put away an important piece of paper and the search for it proves to be useless, I ask myself: What if I were myself and had an important piece of paper to put away, what place would I choose? Sometimes it works, but most of the time I get so pressed by the phrase "What if I were myself" that the search for the piece of paper becomes secondary and I begin to think, better yet, I begin to ***feel****.*

And I don't feel well...

Try it…"

Clarice Lispector ("Se eu fosse eu" in: **A Descoberta do Mundo**. Rio de Janeiro: Rocco, 1999. Freely translated from Portuguese into English)

Abstract

Lowland tropical South America encompasses some of the most species-rich and threatened ecosystems in the world, spanning across countries such as Brazil, Bolivia, Colombia, Ecuador, Peru and Venezuela, which are known for their biodiversity. However, due to its incredible environmental and ecological complexity and that most of its area has yet to be scientifically studied in any depth, controversy surrounds its biomes' identities, the limits of their geographic and environmental distributions and estimates of their tree species richness. The main objective of this thesis is to study the phytogeography of lowland Tropical South America by delimiting its biomes through a floristic approach, by investigating these biomes' environmental controls and dynamics and by assessing their tree species richness and endemism. In order to fulfil this objective, we have employed a dataset of thoroughly checked tree species checklists, the NeoTropTree (NTT) dataset, which covers more than 8000 locations across South, Central and southern North America and encompasses occurrence records for more than 12000 tree species.

Firstly, I defined and mapped the main biomes in lowland tropical South America (LTSA) through the means of a hierarchical clustering analysis based on tree species composition associated with an *a priori* classification of 4103 NTT sites into vegetation types. I then proceeded to map these biomes geographically and to assess their environmental overlaps (both climatic and edaphic) through a classification tree approach (random forest analysis). I was able to delimit five main biomes in LTSA: Amazon Forest, Atlantic Forest, Chaco, Savanna and Seasonally Dry Tropical Forest (SDTF). I also show that there is an important environmental overlap amongst biomes. Error rates for site classification into biome using solely environmental data ranged from 19-21% when only climate was considered and 16-18% when I also took edaphic variables into account. I conclude that it is viable and advisable to use tree species composition to determine biome identity, at least within individual continents. In the case of LTSA, there is high biome heterogeneity at small spatial scales, which explains why it is so challenging to use climatic and/or interpolation-based edaphic data, or remotely-sensed imagery, to map tropical biomes. Because of this, I then conclude that biome delimitation using floristic information may enable more efficient biome conservation and management efforts.

Secondly, I investigated the environmental controls distinguishing biome limits for two regions of LTSA with high biome heterogeneity – eastern Brazil and Bolivia. To this end, I selected 182 NTT sites in these two regions, collected detailed soil data from the field and extracted climate and fire data from publicly available GIS data layers. I assigned these sites to one of three states based on their tree species composition: moist forest (including both Atlantic and Amazon Forests), SDTF or savanna. Selected environmental variables were organized into three distinct categories describing functional environmental regime: water availability, soil fertility and fire, and their significance as predictors of biome identity was assessed within a structural equation modelling framework. I found that environmental controls behind biome distribution differ between the two studied areas and according to the biomes involved. I concluded that water availability, soil fertility and fire are all important determinants of biome limits. Amongst the three categories, water availability was the most important one in determining biome identity at our study sites, with soil fertility differentiating eastern Brazil SDTFs from the other biomes, and fire representing an important determinant of savanna’s environmental limits.

Thirdly, I estimated and compared tree species richness and endemism levels of LTSA’s main biomes using NTT’s tree species checklists and incidence (i.e., occurrence) data. To do so, I extracted tree species information for 4540 sites registered in NeoTropTree distributed across four biomes: Amazon Forest, Atlantic Forest, Savanna and SDTF. I first compared how tree species accumulated with number of sites sampled for biomes and then estimated biomes’ total tree species richness using non-parametric approaches (species extrapolation curves). I also estimated the number of endemic tree species to these areas with two approaches: indicator species analyses and absolute unique/shared species counts. I was able to show that the Amazon Forest is the most tree species-rich environment in LTSA, followed by the Atlantic Forest, Savannas and then SDTFs. In relation to endemism levels, the Amazon and Atlantic Forests’ tree flora are mainly composed of endemic tree species whereas that is not the case for the savanna and SDTF. The estimation of total tree species richness through extrapolation curves revealed that around 94% of the tree flora of the Amazon forest, the Atlantic forest and the SDTF have already been recorded. According to the same analysis, only around 70% of the savannah tree flora has been recorded. However, this pattern might be related to the high number of biome intrusions into this biome. The differences in richness and endemism between the moist (Amazon and Antlantic forests) and drier biomes (savanna and SDTF) suggest that drought-sensitivity and biogeographic history are drivers of tree species distribution in LTSA.

Finally, by integrating biome delimitation based on floristic composition with knowledge on these environments’ environmental correlates and tree species richness, I was able to describe LTSA’s main phytogeographic features in a way that has never been done before, drawing attention to its complexities and performing novel cross-biome comparisons. My study shows that LTSA’s biomes are interspersed across geographic space, especially in the Dry Diagonal located between the Amazon and Atlantic Forests, and that environmental controls driving these ecosystems’ distributions can vary according to the biomes being considered and the geographic location. I also show that LTSA’s most tree species-rich biomes are the ones with the highest quantity of endemic tree species and that taxonomic expeditions to the Amazon Forest can potentially lead to more species being described in these environments. To summarize, I was able to highlight LTSA’s main floristic patterns and link them to environmental drivers and tree species richness, thereby substantially transforming how these biomes are perceived by biodiversity scientists and conservationists.

Lay Summary

South America is one of the most biodiverse areas of the world and its countries are known for their astonishing species richness. In this study, I used data on tree species composition (i.e. lists of the tree species present) at a variety of localities across lowland tropical South America in order to assess how its vegetation can be divided into major categories/groups – called biomes – and understand how these biomes are related to their surrounding environmental conditions. To this end, I have examined the relationship between biomes and climatic features, such as temperature and precipitation (rainfall), soil features such as texture and fertility, and the occurrence of fire. In addition, I have quantified how many tree species can be encountered in each biome, how many are shared among biomes and how many can be found only in a given biome and nowhere else.

I was able to divide lowland tropical South America's tree flora into five sub-groups or biomes: Amazon Forest, Atlantic Forest, Savanna, Seasonally Dry Tropical Forest (SDTF) and Chaco. The Amazon and Atlantic Forests are moist, largely evergreen forests with closed canopies. Savannas are formed by trees arranged sparsely across the landscape, interspersed with a grass and forb layer. SDTFs are forests formed by trees that can withstand very dry conditions and that do not have leaves for part of the year. The Chaco is similar to SDTF, but it can also tolerate low temperatures. I was also able to show that these biomes are not contiguously distributed across geographic space, especially within the centre of LTSA. Rather, their spatial distributions overlap substantially. In relation to climate, soils and fire, I was able to show that each of these is important to limiting biome distributions, and that their relative importance changes across LTSA. I discovered that climate (water availability) is more important to biome distribution than soil fertility or fire. However, soil fertility can differentiate SDTF from the other biomes, and fire is important for Savanna dynamics.

In relation to how many tree species each biome encompasses, I studied each biome's largest continuous area of occurrence – or core area – and concluded that

the Amazon Forest and the Atlantic Forest are the most tree species-rich biomes in LTSA. I also showed how most of the tree species found in these two areas are unique and not found anywhere else. In relation to the Caatinga (SDTF's core area) and the Cerrado (Savanna's core area), I showed that there are tree species that can only be encountered in these regions, but in a smaller proportion relative to the Amazon and Atlantic Forests. In addition, I demonstrated how more expeditions to collect plants in the Amazon Forest and the Cerrado would likely lead to many more tree species being discovered in these areas. The same can also be observed for the Caatinga and the Atlantic Forest, but to a lesser degree.

These results are of fundamental importance for our ecological understanding of LTSA's biomes and can be used to enhance conservation efforts and management policy in these regions. This is because our results show how complex the relationships between biome distribution and environmental conditions truly are. Also, I was able to provide guidelines for which biomes are most in need of further exploration in order for their tree species diversity to be better known, an important finding during a time in which conservation prioritization is a necessity, due to temporal and financial constraints.

Acknowledgments

Firstly, I would like to thank my PhD sponsor – CAPES. Thank you for supporting my studies during these four years. CAPES is a Brazilian public institution that provides funding and sponsorship for research through tax-payers' contributions, so I would like to thank the Brazilian people for trusting science can contribute towards Brazil's future. I would also like to thank the Davis Expedition Fund for financing my fieldtrip to Bolivia, which laid the foundation for the third chapter of this thesis.

During my PhD, I learned a lot about what it means to be a scientist and how science is conducted. However, above all else, writing this thesis was an act of belief and trust. Belief in my potential and trust in my capacity to produce the product you, the reader, hold in your hands. It was not easy, but writing theses lines make me realise and appreciate how much I have grown during the past four years.

I wold like to thank Kyle Dexter, my main supervisor. Thank you for being the amazing human being and inspiring scientist that you are. I would like to thank you for all of your teachings and for always being so considerate and thoughtful. I have no words to explain how much I learned from you during the past four years and how much I admire you. Thank you for moving past that first e-mail I sent you (the one which started with a "Hy") and agreeing for me to become one of your students.

I would like to thank Toby Pennington and Caroline Lehmann, my co-supervisors. I guess I never expressed how happy I was when I learned that you two would be in my supervisory team. Thank you so much for all the teachings and ample support you have given me. It was (and is) an honour to learn from you.

I would like to thank Ary Oliveira-Filho, supervisor turned into best friend. Thank you for saying yes to me becoming your student eight years ago. Thank you for being such a good and loyal friend. Also, thank you for keeping your ears and heart open to all I had to say and for teaching me the importance of honesty and kindness to ourselves. When I made the decision of pursuing a PhD degree abroad, you put me in the right direction and placed in my hands the fruit of 25 years of your labour – NeoTropTree. It was my first time seeing such an act of selflessness in academia, so thank you for the example. For all the love, friendship and trust, I thank you.

I would like to thank my family for their unshakable support and love throughout my PhD. Mom, you are one of the most incredible and strong human beings I have known. Thank you for being an endless source of inspiration and the very definition of resilience. Dad, thank you for expressing your love and care in so many ways. Thank you for everything. I also want to thank my two brothers. Flávio and Hudson; thank you for always checking in on me and making sure everything was OK and for all the fun we had together every time I went back home.

I would like to thank all of my friends from home and from Edinburgh. My dear friends from home: Alan, Mateus, Michel, Natália, Marina, Mariana, Fazito, Luara, Nathália, Marcela, Camila, Rafael and Vinícius, thank you for always receiving me back with open arms and hearts (and parties). My dear friends from Edinburgh: Karina, Julieth, Flávia, Andy (academic baby brother), Cristina, Jaclyn, Abby, Vanessa, Polyanne, Sudhansu, Dami, Jack, Abdulaziz, Ebuka, Pippa, Erica, Rosa, Sandra, Mariana, Haydn, Toby and many others, thank you for paving my stay in Edinburgh with love, joy, companionship and laughter. Jaime, Paula and Sara, thank you so much for being the best friends/office buddies one can ask for. Thank you so much for everything, truly.

I would like to thank Carlos and Chrystiann, my brothers in Edinburgh. You two have my deepest gratitude for the friendship and for the love shared. Thank you for being the safe space and my home away from home during so many times.

I would like to thank Hadesh Sobral, my counsellor, loved brother and good friend. You taught me so much about life and its mysteries and empowered me. Thank you for giving me love, friendship, your full attention and precious tools. It is an honour and a privilege to learn from you. When fear took hold of me, I went to you and you helped me stand once more on my feet and reminded me of my strengths and power. For that, I thank you.

I would like to thank my depression. When you first came along, you were enormous and powerful and you showed me what fear, panic and despair truly are. For more than a year you kept me in your embrace. I felt broken and powerless. However, time passed (as it always does) and I took you in, with kindness and courage. You had a message for me and I listened to it. I listened to your teachings about the love one has to have for oneself, about how life is a beautiful thing and far much more than what we envision and how vulnerability is not a weakness, but a strength. More importantly, you taught me that it takes guts to be what we are and to better ourselves. Out of all of my teachers, you were the most ruthless one. However, you made me grow like no one else did. I sat down and listened to you. For the teachings, I am grateful.

I would like to thank my partner, André Leão. It’s been ten years of the most profound love I’ve experienced in my entire life. Thank you for your love, unshakable support, companionship and understanding during my PhD. Me coming to Scotland nearly broke us, but you agreed with keeping us together as much as we could and I have no words to express how grateful I am for that. I love you.

Table of Contents:

List of Figures

List of Tables

1

Introduction to the thesis

1 - Introduction

1.1 - Setting the Scene

The great species diversity of South America has intrigued researchers for centuries, as can be seen in the early works of Humboldt (1816), Darwin (1859) and Wallace (1855), and it continues to inspire and confound researchers to the present day. Recent years have seen major advances in our understanding of large-scale patterns of tree diversity in South America, particularly in the wetter domains (e.g. Gentry 1988; ter Steege et al. 2006, 2013; Oliveira-Filho et al. 2013). Previous large-scale studies have focused on the roles of precipitation regime and soils in determining the distribution of major vegetation types in Lowland Tropical South America. For example, (ter Steege et al. 2006) demonstrated that major floristic gradients in the Amazon are associated with gradients in precipitation seasonality and soil fertility and that these gradients are also associated with functional changes (e.g. changes in wood density). For the Atlantic Forest, (Sanchez et al. 2013) pointed out the overriding influence of altitudinal gradients on tree species distributions.

Drier biomes, such as the savannas, seasonally dry tropical forests (SDTF) and the Chaco have also been the subject of phytogeographical studies (Ratter, Bridgewater, and Ribeiro 2003; Santos et al. 2012; Velloso et al. 2001), especially after the concept of the Diagonal of Dry formations was published (Vanzoline 1963; Prado & Gibbs 1993). For example, Ribeiro and Walter (2008), Bridgewater et al. (2004) and Werneck et al (2012) showed how the Brazilian savannas are influenced by different precipitation regimes and soil types and harbor a variety of ecoregions, endemism and species tolerant to fire. Even though savannas and SDTFs are commonly linked to soil and precipitation, Neves et al. (2015) showed that tree species turnover across SDTF is also related to temperature regime. Morrone (2000) demonstrated that the Chaco is characterized by a lack of endemic species, rather being composed by dry forest and savanna formations and their accompanying plant species.

Presently, with the advent of comprehensive biodiversity databases and online data repositories, such as GBIF, NeoTropTree (NTT) (Oliveira-Filho 2017) and remote sensing technologies, there has been a shift in the geographic scale of research in South American macroecology and biogeography. These new resources have enabled research encompassing much broader areas and, when aligned with previous knowledge on each one of these biomes, has enabled cross-domain/biome comparisons and the tackling of more elaborate questions, such as where most of South America's biodiversity originated (Antonelli et al. 2018). However, this effort is still young in South America and there is yet much to learn on macroecological patterns of species distributions and their environmental correlates.

1.2 - South America

The South American continent comprises over 17,819,000 km^2, which is 12% of the Earth's land area. South America has five of the world's biodiversity hotspots (Myers et al. 2000) and possesses the largest amount of preserved tropical land in the world (Conservation international 2013, http://www.conservation.org.br). Its flora is recognized for its great diversity and high levels of endemism (Gentry 1982); (Govaerts 2001). I have focused my research on lowland tropical South America and have included sub-tropical areas when necessary. The term "lowland" refers to areas lower than 1000m of elevation.

As a way of delimitating and understanding the patterns underlying plant species distribution on this continent, researchers have proposed various vegetation classification systems (e.g. Cabrera and Willink 1973; Ab'Sáber 1977; Oliveira-Filho 2009) dividing this large flora into Biomes, Domains, Ecoregions, woody vegetation types and/or other categories. Biomes are globally convergent vegetation formations (Moncrieff, Bond, and Higgins 2016; see section below for further discussion on this concept). Phytogeographic Domains, according to Ab'Saber (1977), can be defined by affinities in climatic conditions and geomorphological, floristic and ecological similarities. Ecoregions, defined by Olson et al. (2001), can in turn be described as fairly large areas of land containing a characteristic set of environmental conditions and ecological dynamics, besides being also similar in species composition. Vegetation types, as defined by Oliveira-Filho (2009) and IBGE (1992), also consider the ecological dynamics of a given region, but further take into account endemic plant species and the vegetation physiognomy.

Oliveira-Filho and Eisenlohr (2012) split lowland tropical South America's tree vegetation into five Phytogeographic Domains (Atlantic Forest, Amazon Forest, Cerrado, Caatinga and Chaco). This was done considering IBGE's classification system (IBGE 1992; IBGE 2012) and the vegetation physiognomy of these regions. The Atlantic and Amazon Domains are known for their diverse floras with wet climates and high levels of endemism Fiaschi and Pirani (2009). The Caatinga, Cerrado, and Chaco Domains form the area known as the Dry Diagonal (Prado and Gibbs 1993). This area is characterized by the seasonality of its climates and also has high levels of endemism. These Phytogeographic Domains are similar to the equivalent Brazilian biomes as defined by IBGE (1992, 2012).

1.3 - Ecologically meaningful regionalization – delimiting biomes through tree species composition

To classify entities into different categories is to organize knowledge. For science, the construction of classifications is a necessity and an objective. It is a necessity because it is the way differences and similarities among objects of scientific observation will be formally and objectively described. It is an objective because through classifications a new understanding of the entities being classified can be reached. More importantly, classifications create a system of symbols that has the potential of being understood and used by a wider community.

Classifications in the biological and ecological sciences are created and employed at different levels. Taxonomy is the base of all biological classifications. In biogeography, a plethora of classification systems were created in order to organize assemblages of organisms distributed across geographic space into understandable units (Coutinho

2006; Moncrieff, Bond, and Higgins 2016 and references therein). Along with these classification systems, ranked categories and classes were also created, the most common ones in biogeography and macroecology are: biome, formation, phytophysiognomy, phytogeographic domain and life zone (to name a few). These terms were created at different moments in historic time and they encompassed specific environmental, biological and geographic (scale) requirements in how they were defined by their authors. However, these names are used loosely in the literature and most of these terms have undergone a dramatic change in meaning and usage since their creation. The biome concept is a good example.

The word biome was created by Clements (1917) for a talk presented at a meeting of the Ecological Society of America held in New York City in 1916. In his talk, Clements defined biomes as *"The biotic community regarded as an organic unit comprising all the species of plants and animals at home in a particular habitat"*. He also adds that *"the biotic community, or* biome, *is fundamentally controlled by the habitat, and exhibits a corresponding development and structure".* This definition is well aligned with how Clements perceived ecological communities – *"organic units"* – and his ideas about community succession and climax. Two elements are noticeable in Clements' definition: 1) species composition is fundamental for biome delimitation and 2) biomes are defined at the community scale.

However, Clements' definition is quite different from how biomes are perceived in the present literature (100 years later). According to Moncrieff et al. (2016), biomes *"refer to globally convergent vegetation formations similar in structure and function, explicitly ignoring floristic differences"*. There are two obvious differences between Clements (1917) and Moncrieff et al. (2016): 1) species composition went from being the core aspect of biome definition to being purposefully ignored when delimiting biomes and 2) Biomes are no longer delimited through a community perspective, but through a global perspective based only on similarities in vegetation structure and function. These differences represents a shift in concept applicability, recognition and scale.

Even though Clements (1917) created the biome concept through a community perspective, the way biomes have been conceptualized in recent years is more similar to how floristic formations and phytophysiognomies have been defined in the literature – units of vegetation that are similar in structure and delimited and determined by climatic conditions. Humboldt (1816) was the first to employ the term phytophysigonomy in order to describe how geographically disjunct vegetations can be similar in structure – i.e. have the same phytophysiognomy – if they are under the same climatic conditions. Schimper (1903), Holdridge (1947) and Clements (1949) defined formations in a similar fashion to how the term was envisioned by Humboldt (1816).

Subsequently, the works of Whittaker (1975), Odum (1971), Walter and Box (1976) and Clapham (1973) and Cox & Moore (1993) developed the concept of biome by incorporating climatic and other environmental limits to these units' location and distribution. For example, Odum (1971) defined biomes as the product of the interaction between biota, climatic conditions and substrate. To Odum (1971), this interaction would form large recognizable units, called biomes. With the acknowledgement of the effect of climatic and other environmental conditions on biomes' limits, the only existing difference between these two concepts, as defined by the references cited so far, is a simple one. Formations focus on plants only and biomes focus on both plants and animals (Coutinho, 2006). The idea of placing the biota as a central part of a biome classification system reached its peak with the ecoregions of Olson et al. (2001) (Figure 1.1). More recently, this classification system was revisited and updated by Dinerstein et al. (2017), who organized ecoregions into biomes.

Recent years have seen a shift on how biomes are mapped and circumscribed due to the advancement of remote-sensing technologies. Structure and function have been incorporated into how biomes are defined and delimited since the works of Shimper (1903). Such inclusion is also present in biota-focused biome classification systems (Olson et al., 2001; Dinerstein et al., 2017). However, it was only through remote-sensing that the systematic inclusion of vegetation structure and function became possible at a global scale. Remote-sensing tools enables the quantification of structural traits such as canopy height and percent tree cover of an area. Remote sensing also enables measurement of functional traits such as phenology (leaf-flush) and photosynthesis-related traits such as efficiency and productivity (reflectance and photo-respiration).

Application of remote sensing tecnologies led to the creation of world-wide land cover classification systems and maps (e.g. Friedl et al. 2002; Friedl et al. 2010) and global biomes classification systems and biome maps such as the ones proposed by Woodward, Lomas, and Kelly (2004) and Higgins et al. (2016) (Figure 1.2). However, in order to provide 'wall to wall' mapping or complete terrestrial coverage, these studies had to disregard species composition in their biome circumscriptions and focus on vegetation aspects that can be observed across different land masses – structure (phytophysiognomy) and ecosystem function. Functional traits such as wood density and leaf mass per area have also been used to map biomes and differences in vegetation cover (e.g. Asner et al. 2017).

By considering phytophysiognomy and function, the aforementioned new biome classification systems allowed the study of how these units might have been distributed in the past and how their distribution might be affected by various environmental change scenarios (e.g. Higgins et al., 2016). This is performed by allying structure and function to climate distribution and modelling this relationship for different environmental change scenarios. This shift has allowed much progress to be made on biome and ecosystem function research, such as the monitoring of biomass dynamics, the study of biome boundaries and biome distribution dynamics and their environmental correlates (Staver, Archibald, and Levin 2011; Dantas et al. 2016; Lehmann et al. 2011; Hirota et al. 2011). However, these global schemes often fail to reach an agreement on which biomes exist (for example, Higgins et al., 2016 versus Dinerstein et al., 2017). Even when they agree, a given biome's locations and boundaries can still vary (Särkinen et al. 2011).

More importantly, the degree to which these maps delimit the spatial distribution of ecosystem function is still debatable (Moncrieff et al., 2015). Beuchle et al. (2015) showed that, in South America, remote-sensing tools fail to accurately map the distribution of open-canopy biomes – Cerrado and Caatinga – revealing that there are limitations to the usage of remote-sensed imagery for biome mapping. Also, there are relevant barriers to the use of functional traits to produce biome maps, especially for species rich regions. Knowledge on functional traits is constrained to a small proportion of these biomes' species pools, and trait data have been obtained through a plethora of non-standardized protocols (Violle, Borgy, and Choler 2015; Sandel et al. 2015, Baker et al. 2017). Therefore, even though biome mapping has seen a shift from species composition to function, the construction of accurate biome maps based on function in association with environmental limits is still problematic.

The re-inclusion of species composition in how biomes are defined and delimited can contribute to documenting biome distributions and their environmental correlates, especially where remote-sensing and structure are not fully able to. Another important advantage to this incorporation would be the creation of units that are fully aligned with biogeographic history and that can be used as the basis for biogeographic and phytogeographic studies. It can also serve as the basis for conservation initiatives and efforts. The inclusion of species into biome delimitation can only be done within continents given the few species shared between continents, but it can still be a powerful tool in order to advance our understanding of biomes and their environmental limits. In this thesis, I am proposing that it is possible to use tree species composition to map biomes within lowland tropical South America. However, I do not advocate the dismissal of functional traits, phytophysiognomy and structure in order to do so. What I suggest is that biomes should be defined, delimited and mapped by relying on all of these evidences.

The biome concept we use here is that of a widespread vegetation formation with distinct ecosystem function and species composition.

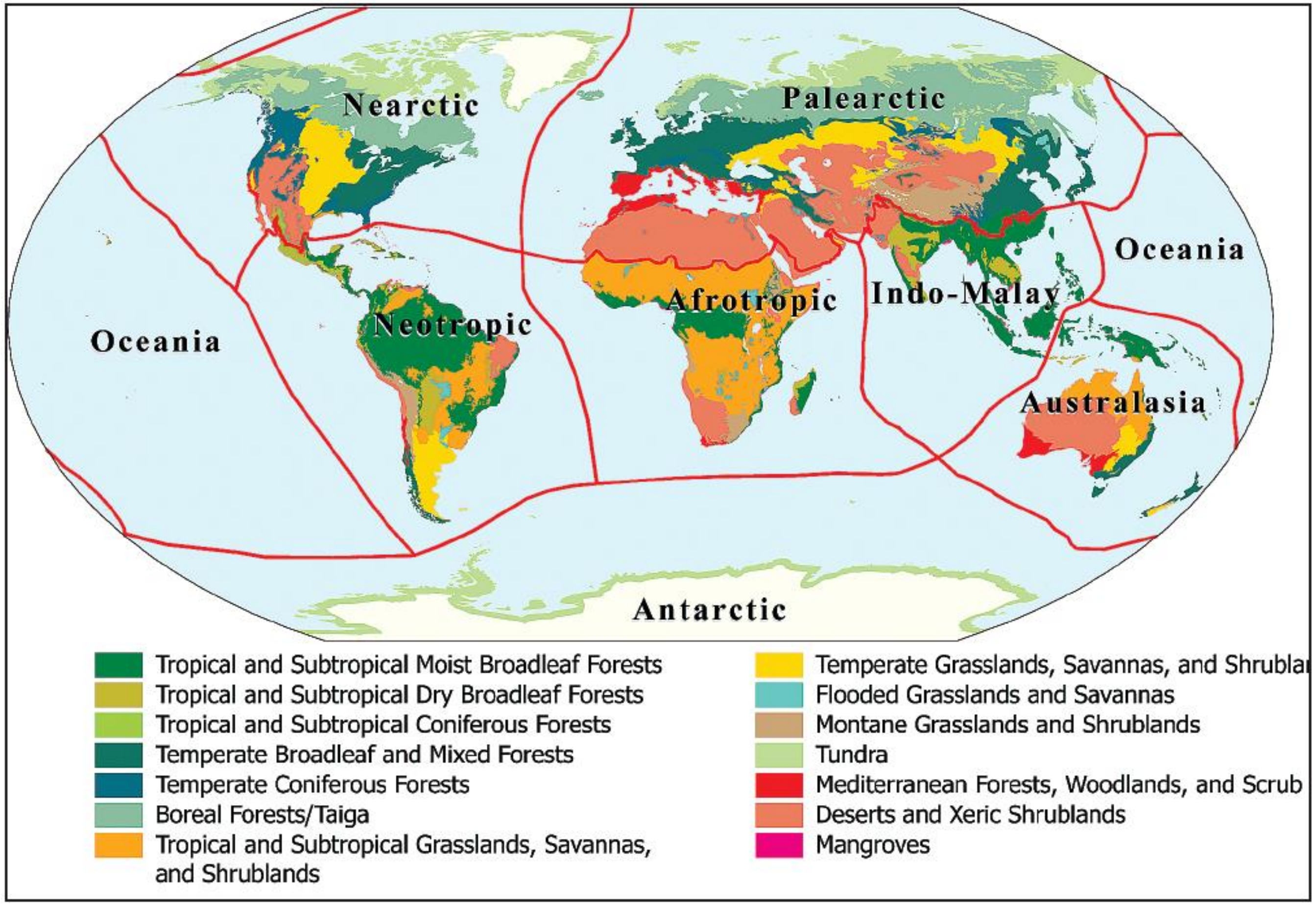

Figure 1.1: World map encompassing all biome classes proposed by Olson et al. (2001). Biomes were delimited and mapped according to previous knowledge on main taxa distribution and specieslists' opinions. Source - Olson, David M., Eric Dinerstein, Eric D. Wikramanayake, Neil D. Burgess, George V. N. Powell, Emma C. Underwood, Jennifer A. D'amico, et al. 2001. "Terrestrial Ecoregions of the World: A New Map of Life on EarthA New Global Map of Terrestrial Ecoregions Provides an Innovative Tool for Conserving Biodiversity." *BioScience* 51 (11): 933–38

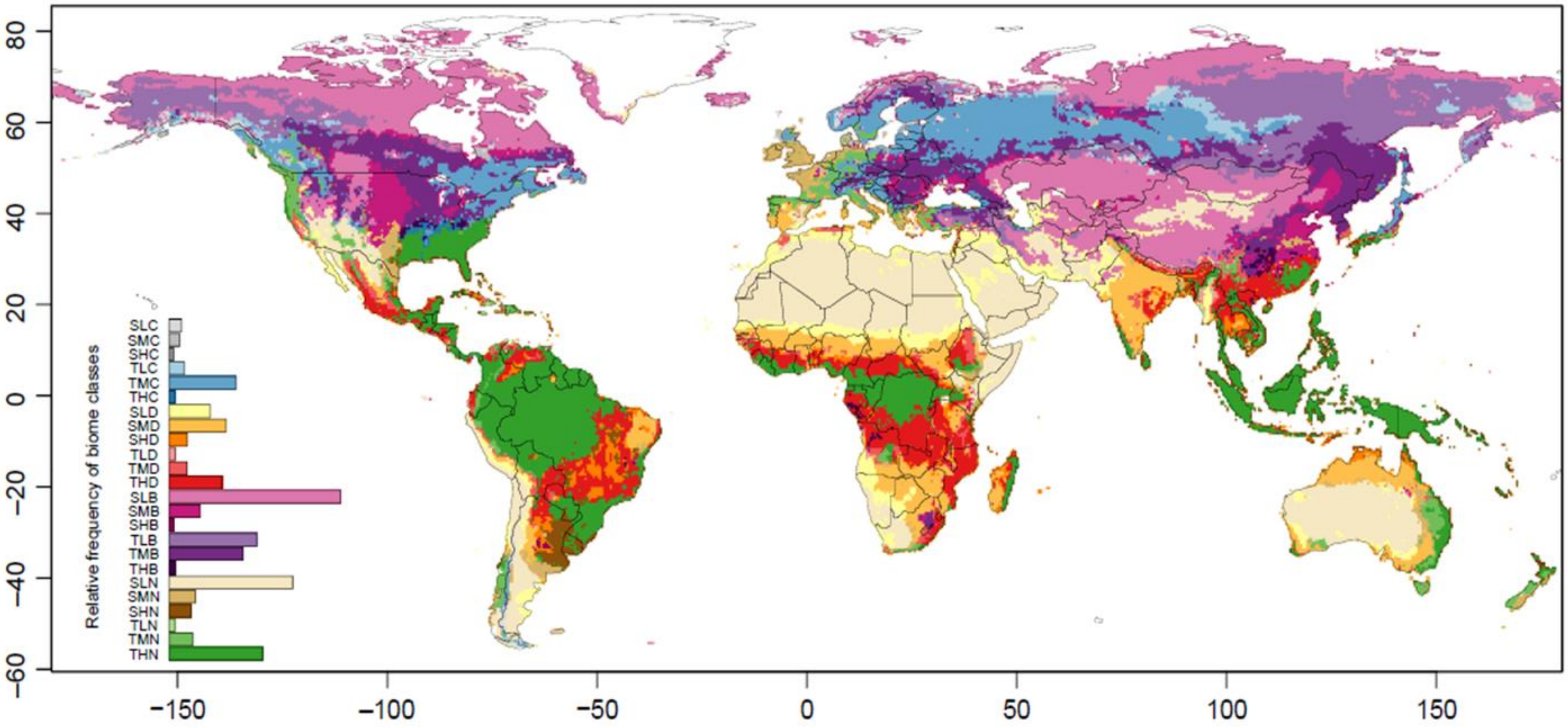

Figure 1.2: World map encompassing all biome classes proposed by Higgins et al. (2016). Biomes were delimited according to canopy height, vegetation productivity index (VPI) and growth limitation index. The biome names are derived from three letters (1) tall vs. short, (2) low, medium and high VPI and (3) cold, dry, both cold and dry, nonseasonal for the growth limitation index.Source - Higgins, Steven I., Robert Buitenwerf, and Glenn R. Moncrieff. 2016. “Defining Functional Biomes and Monitoring Their Change Globally.” *Global Change Biology*, 1–11.

1.4 - Main biomes in Lowland Tropical South America – What we know so far

1.4.1 - Atlantic Forest

The Atlantic Forest ranges from the Rio Grande do Norte state in Brazil to the Rio Grande do Sul state of the same country, forming a strip of forests along the cost. The width of this strip varies according to the region and according to the classification scheme that is being followed (Fiaschi & Pirani 2009). Presently, there is a consensus that the semi-deciduous forests in southeast Brazil are a part of this biome, as well as the Atlantic Auracaria subtropical Forests in the south of Brazil and in Paraguay (Oliveira-filho and Fontes 2000). The Seasonally Dry Tropical Forests (SDTFs) found in Argentina and in the Paranaense Basin were once considered to be a part of the Atlantic Forest, but the works of (Pennington, Lewis, and Ratter 2006) and (Prado 2000) argued that these forests must be considered as a different floristic unit. Indeed, they are considered to be two of the SDTF nuclei spread across South America (Prado & Gibbs 1993), but this is contentious (see Neves et al. 2015).

The Atlantic Forest biome is regarded to be one of the most species rich regions in the world in terms of vascular plant diversity (Martini et al. 2007) and is classified as a Biodiversity hotspot – a region with high species richness and levels of endemism, which has also lost a great part of its natural area of occurrence and continues to be under threat by agriculture expansion and economic growth (Myers et al. 2000). The Atlantic forest is the most degraded biome in Brazil and one of the most endangered biomes in South America. This region has approximately 7.5% of its original vegetation left and a large part of this is still not well protected from future deforestation and degradation (Mittermeier et al. 2005).

Sanchez et al. (2013) pointed out the overriding influence of altitudinal gradients on tree species distributions in this region. Separately, Fiaschi and Pirani (2009) in their review, showed that various studies (see the references provided by Fiaschi & Pirani 2009) agree on a historical separation between the northern part of the Atlantic Forest and the southern part, with the boundary lying at the Rio Doce Basin. This separation is based on examples of plant and other taxa restricted to only one of these blocks, and formal analyses of centers of endemism.

The northern part of the Atlantic forest goes from the Rio Grande do Norte state and reaches the Espírito Santo state in southeast Brazil and is basically a strip of forests along the coast bordered by the Caatinga in the west (Thomas and Barbosa 2008). The southern portion of this biome ranges from the state of Espírito Santo till the southern part of Santa Catarina state. The Semi-deciduous forests in the southeast of Brazil, which are composed of tree species from the more humid parts of this biome that are capable of tolerating seasonality of precipitation, and the Auracaria mixed forests in the south, which are forests that are influenced by the occurrence of frost, are also parts of this group (Oliveira-filho and Fontes 2000).

1.4.2 - Cerrado

The Cerrado is located in the center of South America and covers 20% of the Brazilian territory. It is surrounded by the Amazon Forest in the north, by the Caatinga in the northeast, by the Atlantic Forest in the East and South and by the Pantanal wetlands and the Chaco in the west (IBGE 1992). The Cerrado was classified as being a Biodiversity Hotspot by Myers et al. (2000) since it is rich in endemic species and threatened by the expansion of agriculture, cattle and cities.

This biome harbors a variety of vegetation types ranging from grasslands (termed Campos Limpos) to woodlands (termed Cerradão) (Ribeiro and Walter 2008). The number of vegetation types varies from author to author according to what they believe can be considered as Cerrado or not. Authors such as Coutinho (2000) consider only areas with a continuous grassy layer and sparse trees never forming a canopy as Cerrado, while authors such as Ribeiro and Walter (2008) consider the whole vegetational gradient from grasslands to Cerrado Woodlands (Cerradão) as forming the Cerrado biome. Besides these types of vegetation, Seasonally Dry Tropical Forests (SDTFs) patches and gallery forests can also be found within the geographic limits of the Cerrado Domain. The SDTFs are located in areas with rich soils (Toby Pennington, Prado, and Pendry 2000), hence their patchy distribution within the Cerrado, and gallery forests grows by the margins of rivers found in the region (IBGE, 1992).

There is a consensus in the literature that the many vegetation types within Cerrado are the result of different environmental factors acting together, particularly fire, soil fertility, soil aluminium (Al) concentration and precipitation seasonality (Bueno et al. 2018), which create a gradient of vegetation types that vary in their diversity, structure and dynamics (Ribeiro & Walter 2008). Areas that are burned with high frequency and have high concentrations of Al and low soil fertility tend to be grasslands and areas with very low frequency of fire, relatively high soil fertility and/or more access to water tend to become woodlands (Cerradão or SDTF) (Bueno et al, 2018). Basically, in terms of vegetation structure, what changes between these different vegetation types is the continuity of the grass layer and the prevalence and height of trees and shrubs.

Regarding its diversity, the Cerrado is known for its high level of endemism and species richness (Oliveira-Filho and Ratter 1995; Myers et al., 2000). Although the structure may look similar among regions covered by the same vegetation type, the species composition found in these environments can be very different from one site to another, even while there is a small set of species, such as *Curatella americana* L., that can be found almost anywhere within the Cerrado’s distribution (Bridgewater, Ratter, and Ribeiro 2004). According to Ratter, Bridgewater, and Ribeiro (2003), the Cerrado can be divided into seven largely geographic, groups based on tree species composition.

1.4.3 - Caatinga

The Caatinga is located in the northeastern region of Brazil, and it borders the Amazon forest in the northwest, the Cerrado to the west and the Atlantic Forest to the east and south (IBGE, 1992). This biome is the largest area of seasonally dry tropical forests (SDTFs) in all of South America (Prado & Gibbs, 1993; (DRYFLOR et al.

2016). Other areas of SDTF in South America are the enclaves within the Cerrado, in dry inter-Andean valleys, the dry forests in the Misiones and Piedmont regions of Argentina (see below) and along the Caribbean coast of Colombia and Venezuela (Prado & Gibbs 1993). The region where the Caatinga occurs is known for its semiarid climatic regime, with low precipitation concentrated in three to five months of the year (Sampaio 1995). Vegetation in this area can vary from shrub-dominated landscapes to dry forests with low or high canopies (the latter in Caatinga Arborea) and patchy grasses and bromeliads in the understory (Velloso et al. 2001).

Even though Caatinga is not as species-rich as Cerrado or the Atlantic and Amazon Forests, it has a fairly large number of plant species, especially angiosperms, of which 318 species are considered as being endemic (Giulietti et al. 2002). In terms of species composition, this region shows similarity with other SDTF areas in South America described above (Prado & Gibbs 1993; Toby Pennington, Prado, and Pendry 2000), which led to a hypothesis of a Pleistocene Arc of Dry Formations – an arc of continuous SDTFs starting with Caatinga in the northeast of South America, going through Cerrado, reaching the Misiones and Piedmont region in northern Argentina and going north again through the inter-Andean dry valleys and ending with the SDTFs along the northern coasts of Colombia and Venezuela. Essentially, the putative Pleistocene Arc almost completely encircled the Amazon Forest.

In terms of eco-units and sub-regions within the Caatinga, there is little consensus on how this domain can be further divided. A workshop organized by Velloso et al. (2001) with the collaborations of many specialists, defined eight ecoregions for this biome based on vegetation physiognomy, distribution of key taxa, geographical barriers and environmental & soil variables. Queiroz (2006), by using the distribution range of 274 species of legumes, showed that the Caatinga can be divided into two main floristic units separated by the kind of substrate on which they occur. According to this author, there is a unit associated with soils derived from the crystalline shield and a second one related with sandy sedimentary areas spread across this area. In 2011, (Linares-Palomino, Oliveira-Filho, and Pennington 2011) defined the main nuclei of

SDTFs in South America and observed that the Caatinga may comprise more than one floristic unit, even though the Caatinga formed a consistent group in analyses. Further investigation carried out by Santos et al. (2012), using tree species occurrences from Caatinga, Cerrado, coastal dry forests and the SDTFs enclaves within the Cerrado, showed that Caatinga can be divided into four main floristic units: Crystalline Lowland Caatinga, Sand Deposit Caatinga, Rock Outcrop Caatinga and Arboreal Caatinga.

1.4.4 - Seasonally Dry Tropical Forests (SDTFs – Misiones and Piedmont nuclei)

The Misiones and Piedmont SDTF nuclei are both located in mid-south South America, close to the Chaco and near the Chiquitania region of Bolivia. The Misiones Nucleus ranges from the lower course of the Paraná and Paraguay rivers, till the Uruguay River, in Argentina; The Piedmont nucleus is located along the eastern base of the Andes, starting in Bolivia and entering north-western Argentina (Prado & Gibbs 1993).

Specific information about the flora in these two areas is hard to come across, since they tend to be included together in broader studies, along with the dry inter-Andean valleys and the Caatinga. In contrast to the other SDTFs spread across South America, these two areas register the occurrence of regular frost during winter (Neves et al. 2015), which was one of the justifications used by Pennington (2000) (along with differences in soil and plant species composition, considering the work of Prado & Gibbs, 1993), to disregard the Chaco as being one of the SDTFs nuclei. This shows that the floristic identity of the Misiones and Piedmont regions is yet to be determined. Although these areas are considered to be different from the Chaco, they share 8.4% of their plant species diversity with this region (Särkinen et al. 2011).

1.4.5 - Amazon Forest

The Amazon Forest is the largest and most diverse regions in South America and is also one of the least known regarding plant species diversity and distribution (Pitman et al. 1999; Pitman et al. 2001; Hopkins 2007). This region encompasses over six million square kilometers, yet by 2007, only 500,000 plant specimen vouchers could be found for the region, at least in Brazil (Hopkins 2007), which is much less than one individual sampled per square kilometer of forest (Brazilian Amazon has 5.500.000 km^2). Cardoso et al. (2017) registered 6627 tree species for the entire Amazon. However, (ter Steege et al. 2013; ter Steege et al. 2016) estimated a total tree species richness of almost 12,000 for the entire area. This area, according to Antonelli et al. (2018), is of great importance to South America's biogeographic history as the Amazon Forest is a cradle of biodiversity for the other South American biomes.

This forest's history is marked by the uplift of the Andes, which caused a great shift in drainage orientation, turning it eastward instead of westward (Hoorn et al. 2010). According to Hoorn et al. (2010), another important event for the biogeographic history of this region was a series of oceanic transgressions entering the continent through the north and covering much of the western Amazon Basin, forming the Pebas System. These introgressions are believed to be one of the main reasons why some studies detect differences in species composition between the northern and southern foreland basins (Nihei and De Carvalho 2007).

While the limits of this biome are not well defined in the literature, especially regarding its borders with the Cerrado (Eisenlohr 2012), most of the literature agrees on four principal vegetation types within the forest: Lowland Rainforests, White-sand Forests, Amazonian Savannas and Tepuís. Lowland Rainforest is the main vegetation type found in the region and is known for its great plant diversity (Tuomisto, Zuquim, and Cárdenas 2014). It tends to form high and closed canopies, and in the southern region of the Amazon Forest, some trees lose their leaves during a rather short dry season. The White-sand forests tend to be shorter and more open than the Lowland Rainforests, even resembling savannas in some cases. This is due to its nutrient-poor soils, believed to be originated from the intense weathering of the Guiana Shield (Wittmann et al. 2006). These forests tend to be less rich in terms of plant species than other formations within the biome.

Amongst the Lowland Forests, there is an important difference between the two types of floodable forests – "Várzea" and "Igapó" – and between these and the "Terra Firme" forests (non-floodable). "Várzea" forests are the most extensive kind of floodable forests in South America and, due to the great seasonal influx of nutrients, they tend to be more productive and species rich than "Igapó" forests, which are seasonally flooded by nutrient poor, organic matter rich waters (usually black in colour) (Haugaasen and Peres, 2006). The difference between floodable and Terra Firme forests is that floodable forests are established along the margins of rivers and, from 50 days up to six months in the year, are completely flooded while Terra Firme Forests rarely ever registers the occurrence of floods (Whittmann et al., 2006). Because of the restrictive anoxic environment created by the recurrent floods, most floodable forests tend to be less species rich than the Terra Firme forests. There is a tree species richness gradient ranging from the rivers' margins until the Terra Firme forests. Located right next to the river is the Low-Várzea forests, which consist of forests with water columns more than three meters high during at least 50 days of the year. These forests are early-successional forests with tree species capable of coping with the anoxic conditions during the flooding period.

Even though the non-floodable "Terra Firme" forests are known to be tree species rich (Gentry 1992), they are also recognized by the presence of 'oligarchs' - a small number of tree species that dominate large areas of the Amazon. This was first proposed by (Pitman et al. 2001), considering a network of permanent plots in Peru and Ecuador, and similar patterns were shown in almost 200 other articles published between 2001 (when the hypothesis was first proposed) and 2011 (Pitman, Silman, and Terborgh 2013). Recently, ter Steege et al. (2013) also showed the existence of 'oligarch' species by using the ATDN network of permanent plots; these authors showed that 227 species account for half of the individual trees in the Amazon Forest. Although the presence of 'oligarch' species is registered for different regions in the Amazon, this biome is also marked by great shifts in tree species composition (including shifts in 'oligarchs'), often related to differences in precipitation and soil fertility (Tuomisto, Zuquim, and Cárdenas 2014). Studies such as the one carried out by Pitman, Cardenas, Spichiger and Valderrama (2008) in a permanent plot sequence ranging from the Andean foothills in Ecuador until the Peruvian-Brazilian border, shows that there are at least two major shifts in tree species composition, both caused by changes in soil texture.

1.4.6 - Sub-Tropical Pampean Forests

The Pampean forests are located in the Pampas region which is situated in the South of Brazil, in Rio Grande do Sul state, and in Paraguay. This region is predominantly covered by grasslands with some woody elements sparsely distributed, according to the phytophysiognomic classification scheme proposed by IBGE (1992). According to IBGE's guidelines, the amount of woody elements in these areas is the main factor used to classify these landscapes. The Pampean forests form enclaves within grasslands, usually along river margins and at the coast line (Oliveira-Filho et al., 2013); these forests are classified by Oliveira-Filho (2009) as being semi-deciduous (riverine) or evergreen dwarf forests (coast line).

Information about the environmental controls of Pampean vegetation types is scarce, especially the forested areas. Oliveira-Filho et al. (2013), in a meta-analysis with all South American sub-tropical woody vegetation types, came to the conclusion that the differences in tree species composition amongst phytophysigonomies and vegetation types in this area is much lower than what can be observed in the tropical vegetations of South America. The environmental controls for the Pampean forests, according to these authors are the distance to the ocean, the annual range of temperature and soil water deficit.

1.4.7 - Chaco

The Chaco represents the most southern biome in the South American Diagonal of Dry Formations (Vanzolini 1963; Prado & Gibbs 1993). Its core area is situated in Paraguay, but it ranges to Brazil, Bolivia and northern Argentina. Regarding its diversity, the Chaco has a complex biota sharing many elements with the surrounding environments, especially the Cerrado and nearby SDTF nuclei (Morrone 2006). Werneck (2011), in her review on the biogeographic history of the Dry Diagonal, do not put forth any evidence regarding endemism in the region. According to Spichiger, Calenge, and Bise (2004), this region can be divided in two sub-groups: *Chaco Húmedo* (Wet Chaco) and *Chaco Seco* (Dry Chaco).

Out of the three ecosystems that make up the Diagonal of Dry Formations, the Chaco is the only one that registers the occurrence of frost, being the furthest from the equator and actually lying largely in a subtropical area (Werneck et al. 2011). Frost is more common in the Dry Chaco, which is located near the Andean Piedmont and is distributed from southern Bolivia to north-central Argentina. This Chaquean region is also known by its saline sandy-loamy xeric soils. Meanwhile, the Wet Chaco, located east of the Dry Chaco, receives more precipitation and more humidity because of the Paraguay River's drainage system; it also has more nutrient rich soils and does not receive as much frost.

1.5 - The role of climatic and edaphic variables in biome distribution

1.5.1 - At the global scale

The relationship between environmental conditions and biome distribution has been acknowledged since the early works of Humboldt (1816). During his voyage, Humboldt noticed that disjunct vegetation formations can present similar phytophysiognomies if they occur under similar climatic conditions. Since then, the works of Schimper (1903), Holdridge (1947), Whittaker (1975) and Odum (1992) have emphasized the strength of this relationship and developed classification systems based on it. For Schimper (1903), each biome was formed by dominant life-forms that presented convergent physiological adaptations in order to cope with the climate they were experiencing. Holdridge (1947) acknowledged that differences in elevation would change climatic conditions which, in turn, would interact with length of the growing season and plant's physiological limits, especially in relation to temperature and precipitation extremes. The works of Whittaker (1975) and Odum (1992) crystalized in the literature the view that biomes are entirely regulated by climatic conditions, especially temperature and precipitation.

While climate is regarded as the most important factor influencing biome distribution, other environmental drivers seem also to be related to biome and vegetation type distributions. Edaphic variables, such as nutrient content and water-holding capacity also have an impact on biome limits (e.g. Ratter and Furley 1992; Bueno et al. 2013). Disturbance regimes, especially herbivory and fire, have also been linked to the distribution of biomes, especially when the environmental limits of a certain area can accommodate more than one biome type (e.g. Staver et al., 2011; Hirota et al., 2011; Dantas et al., 2016). As pointed out by Baldeck et al. (2012), there is a surprisingly small number of studies linking edaphic conditions and biodiversity distribution in tropical communities. However, the studies attempting to understand this link have shown that edaphic conditions can do influence biodiversity distribution.

For example, Lehmann et al. (2011), when studying the environmental limits of savanna systems in relation to forested systems across South America, Africa and Australia, found that soil fertility had little effect at a global scale, but was important at a regional scale. Lehmann et al (2011) also documented that fire impacts these ecosystems’ distributions. By using topographic complexity as a proxy for fire spread and return times, the study showed that this variable has a small effect on biome distribution and is negatively correlated to the distribution of savannas. In a global-scale study encompassing all terrestrial biomes, Moncrieff, Hickler, and Higgins (2015) showed that soil texture and fertility affects biome distribution (represented respectively by % of sand and topsoil C to N ratios extracted from publicly available data layers). Both Lehmann et al. (2011) and Moncrieff, Hickler, and Higgins (2015) observed that the same biome would differ in relation to environmental limits across different continents/biogeographic realms. These are most likely related to differences in floristic composition and biogeographic contingencies

Even though Moncrieff, Hickler, and Higgins (2015) analysed all terrestrial biomes, and Lehmann et al (2011) investigated biome's environmental limits at both wet and arid precipitation extremes, most studies only considers two biome categories: forested (closed-canopy) and non-forested (open-canopy) (e.g. Staver et al., 2011; Hirota et al., 2012; Dantas et al., 2016). In the tropics, studies looking at biome distribution and their environmental correlates are often framed in terms of the Alternative States Theory (AST). AST predicts that different biomes can occupy the same environmental space (climatic and edaphic) as long as they experience different, disturbance regimes e.g. due to herbivory and/or fire (Hanan et al. 2014; Dantas et al., 2016). In order to ascertain the effect of disturbance regimes, such studies focus on canopy cover only, as fire and herbivory will have a detrimental effect on tree establishment and growth. Differences between wet and dry forests are usually disregarded, even though they will respond differently to environmental conditions and disturbances.

The studies of Staver et al. (2011), Hirota et al. (2011) and Dantas et al. (2016) have studied the transition between forested and non-forested biome states at a pan-tropical scale. They focused on the effects that climatic, edaphic and disturbance-related factors have on biome distribution. Importantly, they showed that boundaries between biomes are dynamic and the three aforementioned drivers influence each other. According to these studies, climate (particularly precipitation and precipitation seasonality) is the chief factor behind the transition between these biome states. High precipitation levels will almost always lead to a forested biome state. At intermediate precipitation levels, disturbance regimes are the chief driver of differentiation between biomes. These studies show that fire and herbivory are capable of maintaining open canopy ecosystems even when precipitation levels can support tree growth, therefore promoting the maintenance of savannas. The effects of edaphic-related conditions is less evident in these works. Staver et al. (2011) reported that soil texture is a weak predictor of tree cover, while Dantas et al (2016) indicated that soil variables (especially soil fertility) will locally influence the effects of disturbance on biome state. However, according to these authors, the effects of soils as a potential driver of biome cover might be underestimated in these studies, since it was obtained through publicly available soil data layers that have poor spatial precision.

The seminal studies of Lehmann et al (2011), Staver et al (2011), Hirota et al. (2011), Dantas et al. (2016) and Moncrieff et al. (2015) have shown the complexities encompassing biomes' environmental limits and biome distribution at the global scale and how climatic, edaphic and disturbance-related variables can function in a synergistic fashion. However, as highlighted by them, soil conditions and disturbance factors need to be studied at the regional/continental scale and more suitable data needs to be obtained in order to truly ascertain the effects of these two factors on biome distribution and identity. Also, critical is the fact that little is still known about the environmental limits of drier biomes since most studies focused on tree cover to define biome states and such technique fails to differentiate biomes properly. Therefore, in this thesis, I propose a large scale study in lowland tropical South America to investigate the influence of climate, soil and fire on biome identity. This will be done considering SDTFs as a distinct biome state, in addition to the moist forest and savanna biomes.

1.5.2 - Within lowland tropical South America

Water availability has been consistently reported as the chief driver of biome identity and distribution (e.g. Lehmann et al., 2011; Staver et al., 2011; Moncrieff et al., 2015). However, edaphic variables, such as nutrient content, water-holding capacity and texture are important for biome distribution as they can influence species' dispersal capacity, distribution ranges (e.g. Borcard, Legendre, and Drapeau 1992; Legendre 1993; John et al. 2007; Baldeck et al. 2012). It is known that these variables have a weak effect on biome distribution at a global scale (e.g. Lehmann et al., 2011). However, the importance of soils in shaping broader patterns across biomes within a continent is yet to be assessed through the means of field collected data. In the case of LTSA, past research has shown how edaphic conditions can drive floristic patterns within biomes. However, how these variables can affect biome distribution and identity in LTSA is yet to be determined.

In the Amazon Forest, Quesada et al. (2011, 2009, 2010), Quesada et al. (2008), through collection and analysis of soil samples and diggings of soil pits, pointed out the existence of at least six different soil categories varying in texture, fertility and capacity to hold water. This is contrary to the belief that the soils in the region tend to be uniformly old, poor and weathered (Quesada et al., 2011). These studies show an east to west gradient in soils in the Amazon Basin, and forest structure and dynamics are related to this gradient. Forest dynamics were found to be more related to soil texture and wood productivity to phosphorous content. They also showed that edaphic factors play an important role in tree species composition and turnover in the Amazon. The transition between the Amazon Forest – a moist forest – and the surrounding savannas and SDTFs is related to a decrease in precipitation (Staver et al., 2011; Lehmann et al., 2011). The role that edaphic factors might play in these transitions remain largely unknown.

The Atlantic Forest, due to its geographic range and complex topography, encompasses great variation in temperature, precipitation and soil features (F. R. Scarano 2002; Neves et al. 2017). Such variation is responsible for species turnover (beta diversity) (Oliveira-Filho & Fontes 2000; Neves et al. 2017). Because this biome spans tropical and sub-tropical areas, mountain ranges and nearly reaches the centre of South America, climate is a significant driver for tree species distribution in this biome (Oliveira-Filho & Fontes 2000; Scarano 2009; Eisenlohr and Oliveira-Filho 2015). However, within the tropical or subtropical portions of the Atlantic Forest, especially at the border with the Cerrado (central Brazilian savannas), edaphic factors affect tree species distribution. Bueno et al. (2018) showed that, within the Cerrado and at its borders with the Amazon and the Atlantic Forest, patches of semideciduous or evergreen Forests are associated with high soil fertility and water availability (e.g. gallery forests). Scarano (2002, 2009) comments on how soil rockiness (shallow soils) can lead to dry vegetations and how soil water-logging can lead to the formation of swamp forests. Even though these studies were conducted within the Atlantic Forest biome, they suggest that soil fertility, temperature, water availability and soil texture are connected to this biome's environmental limits.

For the drier biomes of Lowland Tropical South America, there are fewer large-scale studies discussing the importance of edaphic variables in determining tree species composition and turnover. One of the factors used by Penningnton et al. (2000) to characterize SDTF as being a different biome from Cerrado, besides the occurrence of grass cover and frost, was soil fertility; SDTF soils tend to be more fertile than the soils in Cerrado. This is confirmed by the works of Furley and Ratter (1988) and Furley, Ratter, and Gifford (1988). Soil fertility is also important to define eco-units within the Caatinga area (Velloso et al. 2001; Queiroz 2006; Santos et al. 2012). Importantly, Queiroz (2006) showed that soil depth and texture mark the division between two plant assemblages in the Caatinga – crystalline and sedimentary Caatinga. When identifying the relationships between the Arboreal Caatinga and other SDTF areas within Brazil, Santos et al. (2012) found that soil fertility is highly related to tree species composition patterns observed in the Caatinga and SDTF enclaves spread across the Cerrado. These evidences points to the differentiation between Savannas and SDTFs being marked by edaphic factors, namely soil fertility.

Even though the Cerrado savannas are thought, in general, to grow on old, nutrient-poor, latosols, variation in fertility and water-holding capacity of soils within the Cerrado are two of the factors, along with fire frequency and climatic variables, important to determine different floristic units (eg. Coutinho, 2000; Ribeiro & Walter, 2008, Bueno et al., 2018), especially at a local scale (Furley 1999). Another soil-related variable that has an important role in determining tree species composition amongst the different vegetation types within the Cerrado domain is the content of aluminium (Al) (Arens 1963; Goodland and Pollard 1973). Al is toxic for plants and the concentration of this element in the soil influences which species will be able to grow in that environment (Ruggiero et al. 2002), hence creating a barrier for species from other biomes to colonize Cerrado soils and also creating a gradient in the vegetation. Concerning the biome transitions occurring within the Cerrado, Bueno et al. (2018) showed that high fertility soils associated with low water availability are linked to SDTFs and that semideciduous or evergreen forests were associated with high soil fertility and water availability.

Therefore, in this thesis, I propose a study to investigate the environmental limits driving biome identity in LTSA. Such study will focus on water availability, especially in regards to precipitation seasonality and mean annual precipitation, soil fertility (including soil texture) and fire.

1.6 – Human impact and tree species diversity patterns in LTSA

Humans have transformed natural environments and their plant communities since the beginning of agriculture. By domesticating certain plant species and opening clearings in natural vegetation in order to grow crops and build settlements, humans have favoured particular plant species and changed their distribution and occurrence patterns across the landscape. Nowadays, changes in species distributions are recognized worldwide and across many taxa, a pattern that is now the main feature of the late Pleistocene (Boivin et al. 2016 and references therein). Through changes in patterns of extinction, dispersal and by translocating species, humans can change how the biotas are organized and how communities are assembled. Dornelas et al (2014), when analysing several assemblages' time series, noticed that most of the communities studied underwent changes in species composition, a pattern also partially shown by Newbold et al (2015).

In South America, the works of Levis et al. (2017) and Junqueira et al. (2017) showed that pre-colombian settlements in the Amazon are responsible for some of the tree species diversity patterns observed in the region nowadays. According to the authors, 20 of the hyperdominant tree species (abundant and widely distributed trees) mentioned by ter Steege et al. (2013) were domesticated during pre-colombian times and present higher abundances close to archeolofical sites. An example of this would be the cocoa plant. Zarrillo et al. (2018) showed that pre-colombian civilizations domesticated cocoa (*Theobroma cacao* L.) and other plants from the genera *Theobroma* and *Herrania* and were able to change their natural distribution range, which is now also found in Mesoamerica.

The effects of pre-colombian settlements on the other South American biomes remain largely unknown. However, for the Atlantic Forest, there are studies showing that human impacts, especially fragmentation, can reduce species richness (Silva and Taberelli et al. 2000) tree species composition, erode the functional attributes (Santos et al., 2008) and change guild structure (Taberelli et al. 1999). Also in this biome, fragmentation leads to more areas being exposed to edge effect, which are known for having empovirished tree species assemblages in relation to non-edge areas (Oliveira et al. 2004). For the Cerrado, Chaco and the SDTFs, little is known about past human impacts and their effects on these biomes' tree species diversity.

1.7 - The sampling bias and knowledge gap

A lack of knowledge about tree species composition and distributions limits the development of effective plant diversity conservation measures (Heywood and Iriondo 2003). Additionally, while some South American biomes, such as the Atlantic Forest, are relatively well known (e.g. Stehmann et al. 2009), others, such as SDTF and the Savanna, have received less attention because they are considered to be species poor (Werneck et al. 2011). Nevertheless, some authors, such as Velloso et al. (2001), Acebes et al. (2010), Velloso et al. (2001), Santos et al. (2012) and DRYFLOR (2016) made valuable contributions to understand these dry environments.

The lack of species-richness related information for some biomes is an example of the obstacles that vegetation science in South America still has to overcome. Due to sampling bias and the remoteness and inaccessibility of some sites, knowledge on species diversity is geographically patchy and sampling effort is unevenly distributed across regions (Sousa-Baena, Garcia, and Peterson 2014; Oliveira et al. 2016). These issues block the possibility of assessing species richness levels across geographic space and of drawing systematic comparisons among different biomes and other biodiversity units.

Recent years have seen a large increase in efforts to construct robust biodiversity databases. In Brazil, the SpeciesLink project (http://splink.cria.org.br/), created by the CRIA organization (“Centro de Referência em Informação Ambiental” - a Brazilian civil society of public interest – similar to an NGO) is digitizing herbaria vouchers (concerning Brazilian flora) from herbaria all across the world. Researchers, such as Oliveira-Filho (2017), have assembled datasets of tree species checklists for most of the Americas. At the global scale, initiatives like GBIF are collecting biodiversity information for millions of taxa and making them publicly available. These resources allow the investigation of biodiversity and ecological patterns across large scales. More importantly, they allow educated assessments of differences in species/taxa richness across ecological units.

In South America, some efforts to point out poorly known regions, in terms of which exact species are present, have been made, considering information retrieved from taxonomic monographs (Hopkins 2007), herbaria records (Schulman, Toivonen, and Ruokolainen 2007), or inventory plots (Milliken et al. 2010). For example, (Rezende, Eisenlohr, et al. 2014) studied the sampling effort in vegetation types of the Atlantic Forest that occur in the Brazilian state of Santa Catarina through a network of permanent plots. Also relying on permanent plot data, ter Steege et al. (2013) estimated that there are approximately 8000 tree species yet to be discovered in the Amazon Forest. However, these studies were able to focus only on one biome or limited geographic regions, usually due to the lack of consistent data compilation or alternatively, the absence of data.

For biodiversity groups such as birds (e.g. Hasui et al. 2018; avibase.bsc-eoc.org) and plants/trees (e.g. NeoTropTree), there are good quality datasets available for analyses. However, in such cases, the problem then shifts from a lack of data to the data being compiled for specific countries/political units (e.g countries, states or provinces) or biogeographic regions only and, most of the time, through different methodologies. For example, Sousa-Baena, Garcia, and Peterson (2014) and Oliveira et al. (2016) showed that are significant differences in sampling effort for birds, plants, mammals and invertebrates among Brazil's ecological and political units. Through the efforts of CRIA, GBIF and other initiatives, this is changing and enabling the construction of illuminating biodiversity panoramas for specific groups over large scales. These panoramas provide key information for conservation initiatives (such as the one suggested by Myers et al., 2000) and for further ecological research.

For example, by using plant species records gathered from herbaria, CRIA was able to show that we have a reasonable quantity of information about plant diversity in Brazil and that it is now possible to assign a conservation status for a large number of plant species that were classified as "data deficient" in the past (Sousa-Baena, Garcia, and Peterson 2014). In a subsequent research paper Sousa-Baena, Garcia, andPeterson (2014), showed what was long suspected - we know more about plant diversity in the southeast region of Brazil and near the country's main roads and rivers than in the interior of the country, a result also supported by Oliveira et al. (2017). They point out that the states of Santa Catarina and Bahia have reached high levels of inventory completeness, while states in the north and north-eastern Brazil, especially Amapá, Amazonas, Pará, Mato Grosso, Maranhão and Piauí, still lack much information about their plant diversity. Although the results they presented will help guide sampling and conservation efforts in Brazil, we still lack consistent information and analyses like this for the entirety of South America.

Another advantage of drawing such broad-scale synthesis is the computation of endemism levels. For example, the compilation of species lists at the country level, such as the ones compiled for plant species in Brazil (Flora do Brasil 2010 - http://floradobrasil.jbrj.gov.br), Colombia (Rangel - Ch. 2015) and other South American countries enables the determination of endemic species for particular political and ecological regions. Even though this is being done at the country level in most LTSA countries, the number of endemic species at the biome level, considering the full biome span, remains largely unknown even for well-known biodiversity groups such as trees.

In the present thesis, I aim to conduct a tree species richness synthesis for LTSA and perform cross-biome comparisons of diversity and endemism. These comparisons will provide a background for further ecological research and conservation initiatives for trees in LTSA.

1.8 - The NeoTropTree dataset

The NeoTropTree (NTT) dataset is the main source of biodiversity information employed in the present thesis (Oliveira-Filho 2014, 2017). NTT contains information regarding tree species composition for more than 8000 sites spread across South America, Central America and the south of North America (Figure 1.3). Due to the difficulty of defining what a tree is, NTT was built based on the following broad ("liberal") definition – "tree species are plants able to grow taller than 3 m in stature without climbing or leaning against other plants (free standing stems).

Information on tree species composition was gathered and the dataset assembled using published and unpublished studies, such as floras, floristic or forestry inventories and phytosociological surveys (plots or transects) that could be retrieved from the online literature or from libraries. Academic work (thesis and dissertations) from different universities and research centres were also employed as sources of tree species checklists. Herbaria records that could be retrieved via visits to herbaria in South America, USA and Europe were also added to NTT. This type of data was also retrived by consulting information made publicly available online by herbaria, especially via the species-link tool (CRIA 2012), a Brazilian initiative to make information present in Brazilian herbaria publicly available. The dataset's taxonomy was routinely updated, checked and standardized. This was performed by constantly reviewing the relevant taxonomic literature and by consulting with specialists in various plant groups.

Each site in NeoTropTree is a circle with a 5 km radius for which we have information about tree species composition. Every site has a main source of information: the first tree species list found for the area, which originated the site. After the site is created, other tree species records found for that area (e.g. in herbaria) are also added. Alongside the information on tree species composition, geographical, spatial, topographical and environmental variables (climatic and edaphic) are associated with each site. In addition, each one of these areas is classified into vegetation types following the classification proposed by Oliveira-Filho (2009). This classification system is based on vegetation physiognomy in association with edaphic and topographic features that may have an impact on the vegetation (vegetation types exist at a small scale). At present, the NeoTropTree dataset has information on occurrence for 156 plant families, 1190 plant genera and 11759 tree species.

The sources of environmental data used in NeoTropTree include the 19 variables of WorldClim which describe the precipitation and temperature regime (Hijmans et al. 2005) and the aridity indices (aridity index, real evapotranspiration and potential evapotranspiration) generated by Zomer et al. (2007, 2008). As the principal spatial variable, NTT registers coordinates (latitude and longitude) of the site's center. Altitude is the only topographic variable registered in NeoTropTree, since both aspect and slope varies too much within a given 10 km diameter area. Even though NeoTropTree brings its own set of environmental variables, I have opted to work with other sources of environmental data. By doing so, I was able to work with the best environmental data sources available and better align my research with the current literature. Each data chapter countains a thorough description of the environmental information employed in their analytical workflow and how they were obtained and processed.

NeoTropTree has provided information for a variety of research in the past and is now being used by students and collaborators to cover subjects such as phytogeographical patterns in different domains of South America (eg. Santos et al., 2012; Oliveira-Filho et al. 2013; Neves et al. 2015; Bueno et al. 2018), conservation (eg. Rezende et al. 2014), delimitation of biodiversity hotspots (Murray-Smith et al. 2008), community phylogenetics (Pennington, Lavin, and Oliveira-Filho 2009) and dispersal syndromes (Almeida-Neto et al. 2008). Using the first version of this dataset, Oliveira-Filho and Ratter (1995) showed how the gallery forests in Central Brazil were the main link between the Amazon Forest and the Atlantic Forest and an important source of tree species diversity in the central region of Brazil. After incorporating more sites to the database, Oliveira-Filho and Fontes (2000) showed that the semi-deciduous and deciduous forests near the coastal Brazilian rainforest should be considered as a part of the Atlantic Forest Domain. In 2006, Oliveira-Filho, Jarenkow, and Rodal pointed out the main phytogeographical patterns related to the Seasonally Dry Tropical Forests in the eastern part of South America.

Since 2017, NeoTropTree has been made fully and freely avaialable at www.neotroptree.info by its creator and curator – Professor Ary Oliveira-Filho.

Figure 1.3: Map of South America, Central America and southern Noth America indicating the location of all data points included in NeoTropTree (NTT), the primary source of biodiversity information used in this thesis. Each point in this map represents a tree species checklist compiled for a circular area with a diameter of 10km. Points are classified into Phytogeographic Domains following the classification proposed by Oliveira-Filho (2015). Check main text for more information on how NeoTropTree was structured and compiled.

1.9 - Thesis' structure

This PhD thesis's main objectives are to delimit the principal biomes of Lowland Tropical South America (LTSA) based on tree species composition, investigate their climatic/edaphic drivers and to quantify their tree species richness. The thesis is structured into five chapters: an introductory chapter followed by three empirical chapters which are then followed by the conclusions chapter. The three empirical data chapters address the three main objectives listed above, and the final chapter is where I discuss all of my findings and comment on possible future research.

In the first data chapter, I have aimed to determine and delimit the main biomes encountered in lowland tropical South America through a hierarchical cluster analysis based on tree species composition. I then investigated if edaphic and climatic data gathered from online databases could distinguish these biomes through a classification tree approach (random forest analysis). My findings lead to the delimitation of five main biomes: Amazon Forest, Atlantic Forest, Savanna, Seasonally Dry Tropical Forest (SDTF) and Chaco. Contrary to how biomes are often portrayed on maps, I show that LTSA biomes overlap substantially in geographic space. With the results obtained from the classification tree analysis, I then show that edaphic and climatic data cannot be used to accurately delimit biomes.

In the second data chapter, I more deeply investigate the environmental drivers behind biome transitions in lowland tropical South America. To this end, I select two target areas/nuclei with high biome heterogeneity: Bolivia and eastern Brazil. Collaborators and I collected soil samples for 182 sites in these nuclei for which I obtained tree species checklists compiled via NTT and other databases. The selection of biomes, sites and nuclei were based on the outcomes of the first chapter. I also obtained data on climatic and fire regime for these sites. All data were analyzed within a Structural Equation Modelling (SEM) framework to explain the transitions among the three main biomes: moist forest, SDTF and savanna. In this chapter, I was able to show that there is intra-continental variation in the environmental drivers of transitions among these three biomes' transitions. I was also able to show that the transition between moist forest and savanna, between SDTF and savanna and between moist forest and SDTF are controlled by different environmental drivers.

In the third data chapter, I quantify, through different means, tree species richness in four core areas of biodiversity in Lowland Tropical South America: the Amazon Forest, the Atlantic Forest, the core SDTF and the core Savanna. For this purpose I have constructed species accumulation and extrapolation curves based on non-parametric diversity estimators. I obtained extrapolated values of tree species richness for each biome. As a proxy for comparing endemic species richness, I have also conducted an indicator species analysis and calculated the number of unique species to each biome and the ones shared among biomes. In this chapter, I show that the wet biomes (the Atlantic and Amazon forests) have a considerably higher number of tree species than the dry biomes (Cerrado and Caatinga) and also a higher proportion of total species that are endemic.

The last chapter is the conclusions chapter in which I synthesize all results from the data chapters and reflect on possible conservation outcomes and future research that needs to be conducted in lowland tropical South America.

During my PhD, I have published three different research articles in collaboration with other researchers. The subjects explored in these contributions are related to the topics I study in my thesis and therefore I have opted to include them in here as appendices (Appendices A, B and C). These three research articles are here included as references only. The first article is Neves et al. (2017) "*Dissection a biodiversity hotspot: The importance of environmentally marginal habitats in the Atlantic Forest Domain of South America*". In this article, we explore the importance of environmentally marginal habitats in the Atlantic Forest to its tree species richness (Appendix A). The second article is Neves et al. (2018) *"Lack of floristic identity in* campos rupestres *– A hyperdiverse mosaic of rocky montane savannas in South America"*. In this contribution, we assess the floristic affiliation of *campos rupestres* a unique environment in South America (Appendix B). The third article is Dexter et al. (2018) *“Inserting tropical dry forests into the discussion on biome transitions in the tropics”.* In this opinion piece, we discuss the Dry Forest biome in depth and define its environmental limits in relation to other biomes (Appendix C).

1.10 - References

Ab'Saber, A. N. 1977. "Os Domínios Morfoclimáticos Da América Do Sul. Primeira Aproximacao." *Geomorfologia* 52: 1–22.

Acebes, Pablo, Juan Traba, Begoña Peco, María L Reus, Stella M Giannoni, and E Juan. 2010. "Abiotic Gradients Drive Floristic Composition and Structure of Plant Communities in the Monte Desert." *Revista Chilena de História Natural* 83: 395–407.

Almeida-Neto, Mário, Flávia Campassi, Mauro Galetti, Pedro Jordano, and Ary Oliveira-Filho. 2008. "Vertebrate Dispersal Syndromes along the Atlantic Forest: Broad-Scale Patterns and Macroecological Correlates." *Global Ecology and Biogeography* 17 (4): 503–13. https://doi.org/10.1111/j.1466-8238.2008.00386.x.

Antonelli, Alexandre, Alexander Zizka, Fernanda Antunes Carvalho, Ruud Scharn, Christine D. Bacon, Daniele Silvestro, and Fabien L. Condamine. 2018. "Amazonia Is the Primary Source of Neotropical Biodiversity." *Proceedings of the National Academy of Sciences*, May, 201713819. https://doi.org/10.1073/pnas.1713819115.

Arens, K. 1963. "As Plantas Lenhosas Dos Campos Cerrados Como Vegetação Adaptada Às Deficiências Minerais Do Solo." In *III Simpósio Sobre o Cerrado*, edited by Edgard Blutcher, 13–115. São Paulo: EDUSP.

Asner, Gregory P., Rainer E. Martin, D. E. Knapp, R. Tupayachi, C. B. Anderson, F. Sinca, N. R. Vaughn, and W. Llactayo. 2017. "Airborne Laser-Guided Imaging Spectroscopy to Map Forest Trait Diversity and Guide Conservation." *Science* 355 (6323): 385–389.

Baker, Timothy R., R. Toby Pennington, Kyle G. Dexter, Paul V.A. Fine, Helen Fortune-Hopkins, Euridice N. Honorio, Isau Huamantupa-Chuquimaco, et al. 2017. "Maximising Synergy among Tropical Plant Systematists, Ecologists, and Evolutionary Biologists." *Trends in Ecology & Evolution* 32 (4): 258–67. https://doi.org/10.1016/j.tree.2017.01.007.

Baldeck, C. A., K. E. Harms, J. B. Yavitt, R. John, B. L. Turner, R. Valencia, H. Navarrete, et al. 2012. "Soil Resources and Topography Shape Local Tree Community Structure in Tropical Forests." *Proceedings of the Royal Society B: Biological Sciences* 280 (1753): 20122532–20122532. https://doi.org/10.1098/rspb.2012.2532.

Beuchle, René, Rosana Cristina Grecchi, Yosio Edemir Shimabukuro, Roman Seliger, Hugh Douglas Eva, Edson Sano, and Frédéric Achard. 2015. "Land Cover Changes in the Brazilian Cerrado and Caatinga Biomes from 1990 to 2010 Based on a Systematic Remote Sensing Sampling Approach." *Applied Geography* 58 (March): 116–27. https://doi.org/10.1016/j.apgeog.2015.01.017.

Boivin, Nicole L., Melinda A. Zeder, Dorian Q. Fuller, Alison Crowther, Greger Larson, Jon M. Erlandson, Tim Denham, and Michael D. Petraglia. "Ecological Consequences of Human Niche Construction: Examining Long-Term Anthropogenic Shaping of Global Species Distributions."

Proceedings of the National Academy of Sciences 113, no. 23 (June 7, 2016): 6388–96. https://doi.org/10.1073/pnas.1525200113.

Borcard, Daniel, Pierre Legendre, and Pierre Drapeau. 1992. “Partialling out the Spatial Component of Ecological Variation.” *Ecology* 73 (3): 1045–55. https://doi.org/10.2307/1940179.

Bridgewater, Samuel, James A. Ratter, and José Felipe Ribeiro. 2004. “Biogeographic Patterns, b-Diversity and Dominance in the Cerrado Biome of Brazil.” *Biodiversity and Conservation* 13: 2295–2318.

Bueno, M. L., D. R. Neves, A. T. Oliveira Filho, C. R. Lehn, and J. A. Ratter. 2013. “A Study in an Area of Transition between Seasonally Dry Tropical Forest and Mesotropic Cerradão, in Mato Grosso Do Sul, Southwestern Brazil.” *Edinburgh Journal of Botany* 70 (3): 469–86. https://doi.org/10.1017/S0960428613000164.

Bueno, Marcelo Leandro, Kyle G. Dexter, R. Toby Pennington, Vanessa Pontara, Danilo Mesquita Neves, James Alexander Ratter, and Ary Teixeira de Oliveira-Filho. 2018. “The Environmental Triangle of the Cerrado Domain: Ecological Factors Driving Shifts in Tree Species Composition between Forests and Savannas.” *Journal of Ecology* 106 (5): 2109–20. https://doi.org/10.1111/1365-2745.12969.

Cabrera, A. L., and A. Willink. 1973. *Biogeografia de América Latina*. Monografía. Washington: Secretaría General de la Oraganización de Estados Americanos.

Cardoso, Domingos, Tiina Särkinen, Sara Alexander, André M. Amorim, Volker Bittrich, Marcela Celis, Douglas C. Daly, et al. 2017. “Amazon Plant Diversity Revealed by a Taxonomically Verified Species List.” *Proceedings of the National Academy of Sciences* 114 (40): 10695–700. https://doi.org/10.1073/pnas.1706756114.

Clapham, W. B. 1973. *Natural Ecosystems*. Vol. 8. New York: Macmillan USA.

Clements, F. E. 1917. “The Development and Structure of Biotic Communities.” In *Meeting at New York City 27-29, 1916*, 5:119–28. 2. New York: British Ecological Society.

———. 1949. *Dynamics of Vegetation*. New York: The H.W. Wilson Co.

Coutinho, Leopoldo Magno. 2000. “O Bioma Do Cerrado.” In *Eugen Warming e o Cerrado Brasileiro - Um Século Depois*, 1st ed., 77–93. Sao Paulo: Editora da UNESP.

———. 2006. “The Biome Concept.” *Acta Botanica Brasilica* 20 (1): 13–23. https://doi.org/10.1590/S0102-33062006000100002.

Cox, C. Barry, Peter D. Moore, and Richard Ladle. 2016. *Biogeography: An Ecological Approach and Evolutionary Approach*. 9th ed. London: Wiley-Blackwell.

CRIA. 2012. “SpeciesLink.” Centro de Referência Em Informação Ambiental. 2012. http://www.splink.org.br/index.

Cruz Ruggiero, P. G., M. a. Batalha, Vânia Regina Pivello, and Sérgio Tadeu Meirelles. 2002. “Soil-Vegetation Relationships in Cerrado (Brazilian Savanna) and Semideciduous Forest, Southeastern Brazil.” *Plant Ecology* 160 (1): 1–16. https://doi.org/10.1023/A:1015819219386.

Dantas, Vinícius de L., Marina Hirota, Rafael S. Oliveira, and Juli G. Pausas. 2016. “Disturbance Maintains Alternative Biome States.” Edited by Marcel Rejmanek. *Ecology Letters* 19 (1): 12–19. https://doi.org/10.1111/ele.12537.

Darwin, Charles. 1859. *On the Origin of Species By Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. 1st ed. London.

Dinerstein, Eric, David Olson, Anup Joshi, Carly Vynne, Neil D. Burgess, Eric Wikramanayake, Nathan Hahn, et al. 2017. “An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm.” *BioScience* 67 (6): 534–45. https://doi.org/10.1093/biosci/bix014.

DRYFLOR, et al. 2016. “Plant Diversity Patterns in Neotropical Dry Forests and Their Conservation Implications.” *Science* 353 (6306): 1383–87. https://doi.org/10.1126/science.aaf5080.

Dornelas, M., N. J. Gotelli, B. McGill, H. Shimadzu, F. Moyes, C. Sievers, and A. E. Magurran. “Assemblage Time Series Reveal Biodiversity Change but Not Systematic Loss.” *Science* 344, no. 6181 (April 18, 2014): 296–99. https://doi.org/10.1126/science.1248484.

Eisenlohr, Pedro Vasconcellos, and Ary Teixeira de Oliveira-Filho. 2015. “Revisiting Patterns of Tree Species Composition and their Driving Forces in the Atlantic Forests of Southeastern Brazil.” *Biotropica* 47 (6): 689–701. https://doi.org/10.1111/btp.12254.

Eisenlohr, Pedro Vasconcellos. 2012. “Padrões Florístico-Estruturais, de Diversidade Alfa e de Distribuição Potencial de Espécies Arbóreas Entre Os Domínios Do Cerrado e Da Amazônia.” Uiversidade Estadual de Campinas.

Fiaschi, Pedro, and José R. Pirani. 2009. “Review of Plant Biogeographic Studies in Brazil.” *Journal of Systematics and Evolution* 47 (5): 477–96. https://doi.org/10.1111/j.1759-6831.2009.00046.x.

Friedl, M.A, D.K McIver, J.C.F Hodges, X.Y Zhang, D Muchoney, A.H Strahler, C.E Woodcock, et al. 2002. “Global Land Cover Mapping from MODIS: Algorithms and Early Results.” *Remote Sensing of Environment* 83 (1–2): 287–302. https://doi.org/10.1016/S0034-4257(02)00078-0.

Friedl, Mark A., Damien Sulla-Menashe, Bin Tan, Annemarie Schneider, Navin Ramankutty, Adam Sibley, and Xiaoman Huang. 2010. “MODIS Collection 5 Global Land Cover: Algorithm

Refinements and Characterization of New Datasets.” *Remote Sensing of Environment* 114 (1): 168–82. https://doi.org/10.1016/j.rse.2009.08.016.

Furley, P. A., J. A. Ratter, and D. R. Gifford. 1988. “Observations on the Vegetation of Eastern Mato Grosso, Brazil. III. The Woody Vegetation and Soils of the Morro de Fumaça, Torixoreu.” *Proc. R. Soc. Lond. B* 235 (1280): 259–80. https://doi.org/10.1098/rspb.1988.0075.

Furley, Peter a. 1999. “The Nature and Diversity of Neotropical Savanna Vegetation with Particular Reference to the Brazilian Cerrados.” *Global Ecology and Biogeography* 8 (3–4): 223–41. https://doi.org/10.1046/j.1365-2699.1999.00142.x.

Furley, Peter A., and James A. Ratter. 1988. “Soil Resources and Plant Communities of the Central Brazilian Cerrado and Their Development.” *Journal of Biogeography* 15 (1): 97–108. https://doi.org/10.2307/2845050.

Gentry, Alwyn H. 1982. “Neotropical Floristic Diversity: Phytogeographical Connections Between Central and South America, Pleistocene Climatic Fluctuations, or an Accident of the Andean Orogeny?” *Annals of the Missouri Botanical Garden* 69 (3): 557–93.

———. 1988. “Changes in Plant Community Diversity and Floristic Composition on Environmental and Geographical Gradients.” *Annals of the Missouri Botanical Garden* 75 (1): 1–34.

———. 1992. “Tropical Forest Biodiversity : Distributional Patterns and Their Conservational Significance Phytogeographic Patterns” 1: 19–28.

Giulietti, A M, R M Harley, L P Queiroz, M R V Barbosa, A L B Neta, and M A Figueiredo. 2002. “Espécies Endêmicas Da Caatinga.” In *Vegetação e Flora Da Caatinga*, edited by E V S B Sampaio, A M Giulietti, J Virginio, and C F L Gamarra-Rojas, 103–18. Recife: Associação Plantas do Nordeste.

Goodland, R, and R Pollard. 1973. “The Brazilian Cerrado Vegetation : A Fertility Gradient.” *Journal of Ecology* 61 (1): 219–24. https://doi.org/10.2307/2258929.

Govaerts, Rafael. 2001. “How Many Species of Seed Plants Are There?” *Taxon* 50 (4): 1085–90.

Hanan, Niall P., Andrew T. Tredennick, Lara Prihodko, Gabriela Bucini, and Justin Dohn. 2014. “Analysis of Stable States in Global Savannas: Is the CART Pulling the Horse?” *Global Ecology and Biogeography* 23 (3): 259–63. https://doi.org/10.1111/geb.12122.

Hasui, Érica, Jean Paul Metzger, Rafael G. Pimentel, Luís Fábio Silveira, Alex A. d A. Bovo, Alexandre C. Martensen, Alexandre Uezu, et al. 2018. “ATLANTIC BIRDS: A Data Set of Bird Species from the Brazilian Atlantic Forest.” *Ecology* 99 (2): 497–497. https://doi.org/10.1002/ecy.2119.

Heywood, Vernon H., and José M. Iriondo. 2003. “Plant Conservation: Old Problems, New Perspectives.” *Biological Conservation* 113 (October): 321–35. https://doi.org/10.1016/S0006-3207(03)00121-6.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. “Very High Resolution Interpolated Climate Surfaces for Global Land Areas.” *International Journal of Climatology* 25: 1965–78.

Hirota, Marina, Milena Holmgren, Egbert H. Van Nes, and Marten Scheffer. 2011. “Global Resilience of Tropical Forest and Savanna to Critical Transitions.” *Science (New York, N.Y.)* 334 (6053): 232–35. https://doi.org/10.1126/science.1210657.

Holdridge, L. R. 1947. “Determination of World Plant Formations From Simple Climatic Data.” *Science* 105 (2727): 367–68. https://doi.org/10.1126/science.105.2727.367.

Hoorn, C, F P Wesselingh, H Steege, M a Bermudez, A Mora, J Sevink, I Sanmartín, et al. 2010. “Amazonia Through Time : Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity.” *Science (New York, N.Y.)* 330 (November): 927–31. https://doi.org/10.1126/science.1194585.

Hopkins, Michael J. G. 2007. “Modelling the Known and Unknown Plant Biodiversity of the Amazon Basin.” *Journal of Biogeography* 34 (8): 1400–1411. https://doi.org/10.1111/j.1365-2699.2007.01737.x.

Humboldt, Count Alexander. 1816. “XCIII. On the Laws Observed in the Distribution of Vegetable Forms.” *The Philosophical Magazine* 47 (218): 446–53. https://doi.org/10.1080/14786441608628504.

IBGE. 1992. *Manual Técnico Da Vegetação Brasileira*. Rio de Janeiro: Estatística Instituto Brasileiro de Geografia e Estatística.

Instituto Brasileiro de Geografia e Estatística, ed. 2012. *Manual Técnico Da Vegetação Brasileira*. 2a edição revista e ampliada. Manuais Técnicos Em Geociências, número 1. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística-IBGE.

John, Robert, James W. Dalling, Kyle E. Harms, Joseph B. Yavitt, Robert F. Stallard, Matthew Mirabello, Stephen P. Hubbell, et al. 2007. “Soil Nutrients Influence Spatial Distributions of Tropical Tree Species.” *Proceedings of the National Academy of Sciences* 104 (3): 864–69. https://doi.org/10.1073/pnas.0604666104.

Junqueira, André Braga, Carolina Levis, Frans Bongers, Marielos Peña-Claros, Charles Roland Clement, Flávia Costa, and Hans ter Steege. “Response to Comment on ‘Persistent Effects of Pre-Columbian Plant Domestication on Amazonian Forest Composition.’” *Science* 358, no. 6361 (October 20, 2017): eaan8837. https://doi.org/10.1126/science.aan8837.

Legendre, Pierre. 1993. “Spatial Autocorrelation: Trouble or New Paradigm?” *Ecology* 74 (6): 1659–73. https://doi.org/10.2307/1939924.

Lehmann, Caroline E. R., Sally A. Archibald, William A. Hoffmann, and William J. Bond. 2011. “Deciphering the Distribution of the Savanna Biome.” *New Phytologist* 191 (1): 197–209. https://doi.org/10.1111/j.1469-8137.2011.03689.x.

Levis, C., F. R. C. Costa, F. Bongers, M. Peña-Claros, C. R. Clement, A. B. Junqueira, E. G. Neves, et al. “Persistent Effects of Pre-Columbian Plant Domestication on Amazonian Forest Composition.” *Science* 355, no. 6328 (March 3, 2017): 925–31. https://doi.org/10.1126/science.aal0157.

Linares-Palomino, Reynaldo, Ary T Oliveira-Filho, and R Toby Pennington. 2011. “Neotropical Seasonally Dry Forests: Diversity, Endemism, and Biogeography of Woody Plants.” In *Seasonally Dry Tropical Forests: Ecology and Conservation*, edited by Rodolfo Dirzo, Hillary S. Young, A. Mooney Mooney, and Gerardo Ceballos, 3–21. Washington: Island Press.

Martini, Adriana Maria Zanforlin, Pedro Fiaschi, André M. Amorim, and José Lima Da Paixão. 2007. “A Hot-Point within a Hot-Spot: A High Diversity Site in Brazil’s Atlantic Forest.” *Biodiversity and Conservation* 16 (11): 3111–28. https://doi.org/10.1007/s10531-007-9166-6.

Milliken, William, Daniela Zappi, Denise Sasaki, Mike Hopkins, and R Toby Pennington. 2010. “Amazon Vegetation : How Much Don’t We Know and How Much Does It Matter ?” *Kew Bulletin* 65: 691–709.

Mittermeier, Russell a., Gustavo a B Da Fonseca, Anthony B. Rylands, and Katrina Brandon. 2005. “A Brief History of Biodiversity Conservation in Brazil.” *Conservation Biology* 19 (3): 601–7. https://doi.org/10.1111/j.1523-1739.2005.00709.x.

Moncrieff, Glenn R., William J. Bond, and Steven I. Higgins. 2016. “Revising the Biome Concept for Understanding and Predicting Global Change Impacts.” *Journal of Biogeography* 43 (5): 863–73. https://doi.org/10.1111/jbi.12701.

Moncrieff, Glenn R., Thomas Hickler, and Steven I. Higgins. 2015. “Intercontinental Divergence in the Climate Envelope of Major Plant Biomes: Intercontinental Biome Divergence.” *Global Ecology and Biogeography* 24 (3): 324–34. https://doi.org/10.1111/geb.12257.

Morrone, Juan J. 2000. “What Is the Chacoan Subregion?” *Neotrópica* 46: 51–68.

Morrone, Juan J. 2006. “Biogeographic Areas and Transition Zones of Latin America and the Caribbean Islands Based on Panbiogeographic and Cladistic Analyses of the Entomofauna.” *Annual Review of Entomology* 51 (125): 467–94. https://doi.org/10.1146/annurev.ento.50.071803.130447.

Myers, Norman, Russell A Mittermeier, Cristina G Mittermeier, Gustavo A B Fonseca, and Jennifer Kent. 2000. “Biodiversity Hotspots for Conservation Priorities” 403 (February): 853–58.

Neves, Danilo M, Kyle G Dexter, R Toby Pennington, Marcelo L Bueno, and Ary T Oliveira-Filho. 2015. “Environmental and Historical Controls on Floristic Composition across the South American Dry Diagonal.”

Neves, Danilo M., Kyle G. Dexter, R. Toby Pennington, Arthur S. M. Valente, Marcelo L. Bueno, Pedro V. Eisenlohr, Marco A. L. Fontes, et al. 2017. “Dissecting a Biodiversity Hotspot: The Importance of Environmentally Marginal Habitats in the Atlantic Forest Domain of South America.” *Diversity and Distributions* 23 (8): 898–909. https://doi.org/10.1111/ddi.12581.

Newbold, Tim, Lawrence N. Hudson, Samantha L. L. Hill, Sara Contu, Igor Lysenko, Rebecca A. Senior, Luca Börger, et al. “Global Effects of Land Use on Local Terrestrial Biodiversity.” *Nature* 520, no. 7545 (April 2015): 45–50. https://doi.org/10.1038/nature14324.

Nigel C. A. Pitman, Hugo Mogollón, Nallarett D´ Avila, Marcos Ríos, Roosevelt García-Villacorta, Juan Guevara, Timothy R. Baker, Abel Monteagudo, Oliver L. Phillips, Rodolfo Vásquez-Martínez, Manuel Ahuite, Milton Aulestia, Carlos E. Cerón, Pierre-André Loizeau, David a. Neill, Percy Nuñez V., Walter a. Palacios, Dairon Cardenas, and Rodolphe Spichiger and Elvis Valderrama. 2008. “Tree Community Change across 700 Km of Lowland Amazonian Forest from the Andean\nFoothills to Brazil.” *Biotropica* 40 (5): 11.

Nihei, Silvio Shigueo, and C. J B De Carvalho. 2007. “Systematics and Biogeography of Polietina Schnabl & Dziedzicki (Diptera, Muscidae): Neotropical Area Relationships and Amazonia as a Composite Area.” *Systematic Entomology* 32 (3): 477–501. https://doi.org/10.1111/j.1365-3113.2006.00376.x.

Odum, E. P. 1971. *Fundamentals of Ecology*. London: W.B. Saunders.

———. 1992. *Ecology*. 2nd ed. London: Holt Rinehart and Winston.

Oliveira, Marcondes A., Alexandre S. Grillo, and Marcelo Tabarelli. “Forest Edge in the Brazilian Atlantic Forest: Drastic Changes in Tree Species Assemblages.” *Oryx* 38, no. 04 (October 2004). https://doi.org/10.1017/S0030605304000754.

Oliveira, Ubirajara, Adriano Pereira Paglia, Antonio D. Brescovit, Claudio J. B. de Carvalho, Daniel Paiva Silva, Daniella T. Rezende, Felipe Sá Fortes Leite, et al. 2016. “The Strong Influence of Collection Bias on Biodiversity Knowledge Shortfalls of Brazilian Terrestrial Biodiversity.” *Diversity and Distributions* 22 (12): 1232–44. https://doi.org/10.1111/ddi.12489.

Oliveira-Filho, A. T. 2014. “NeoTropTree, Flora Arbórea Da Região Neotropical: Um Banco de Dados Envolvendo Biogeografia, Diversidade e Conservação.” 2014. http://www.icb.ufmg.br/treeatlan/.

Oliveira-Filho, A. T., J. C. Budke, J. A. Jarenkow, P. V. Eisenlohr, and D. R. M. Neves. 2013. “Delving into the Variations in Tree Species Composition and Richness across South American Subtropical Atlantic and Pampean Forests.” *Journal of Plant Ecology*, December, 1–23. https://doi.org/10.1093/jpe/rtt058.

Oliveira-Filho, A. T., and J. A. Ratter. 1995. “A Study of the Origin of Central Brazilian Forests by the Analysis of Plant Species Distribution Patterns.” *Edinburgh Jounal of Botany* 52 (2): 141–94.

Oliveira-Filho, Ary T. 2009. “Classificação Das Fitofisionomias Da América Do Sul Cisandina Tropical e Subtropical: Proposta de Um Novo Sistema - Prático e Flexível - Ou Uma Injeção a Mais de Caos?” *Rodriguésia* 60 (2): 237–58.

Oliveira-Filho, Ary T. 2017. “NeoTropTree, Flora Arbórea Da Região Neotropical: Um Banco de Dados Envolvendo Biogeografia, Diversidade e Conservação.” NeoTropTree, Flora Arbórea Da Região Neotropical: Um Banco de Dados Envolvendo Biogeografia, Diversidade e Conservação. 2017. http://www.neotroptree.info/.

Oliveira-filho, Ary T, and Marco Aurelio L Fontes. 2000. “Patterns of Floristic Differentiation among Atlantic Forests in Southeastern Brazil and the Influence of Climate Patterns of Floristic Differentiation among Atlantic Forests in Southeastern Brazil and the Influence of Climate.” *Biotropica* 32 (4): 793–810.

Oliveira-Filho, Ary T, Lojo Andre Jarenkow, and Maria Jesus Nogueira Rodal. 2006. “Floristic Relationships of Seasonally Dry Forests of Eastern South America Based on Tree Species Distribution Patterns.” In *Neotropical Savannas and Seasonally Dry Forests: Plant Diversity, Biogeography, and Conservation*, 159–92.

Oliveira-Filho, Ary Teixeira, and Pedro Vasconcellos Eisenlohr. 2012. “Tree-Atlan 2.1, Flora Arbórea Da América Do Sul Cisandina Tropical e Subtropical: Um Banco de Dados Envolvendo Biogeografia, Diversidade e Conservação.” 2012. http://www.icb.ufmg.br/treeatlan/.

Olson, David M., Eric Dinerstein, Eric D. Wikramanayake, Neil D. Burgess, George V. N. Powell, Emma C. Underwood, Jennifer A. D’amico, et al. 2001. “Terrestrial Ecoregions of the World: A New Map of Life on EarthA New Global Map of Terrestrial Ecoregions Provides an Innovative Tool for Conserving Biodiversity.” *BioScience* 51 (11): 933–38. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2.

Pennington, R T, G P Lewis, and J A Ratter. 2006. “An Overview of the Plant Diversity, Biogeography and Conservation of the Neotropical Savannas and Seasonally Dry Forests.” In *Neotropical Savannas and Seasonally Dry Forests: Plant Diversity, Biogeography and Conservation*, edited by R T Pennington, G P Lewis, and J A Ratter, 1–29. Boca Raton: CRC Press.

Pennington, R. Toby, Matt Lavin, and Ary Oliveira-Filho. 2009. “Woody Plant Diversity, Evolution, and Ecology in the Tropics: Perspectives from Seasonally Dry Tropical Forests.” *Annual Review of Ecology, Evolution, and Systematics* 40 (1): 437–57. https://doi.org/10.1146/annurev.ecolsys.110308.120327.

Pennington, R Toby, Darién E Prado, and Colin A Pendry. 2000. “Neotropical Seasonally Dry Forests and Quaternary Vegetation Changes.” *Journal of Biogeography* 27: 261–73.

Pitman, Nigel C. a., Miles R. Silman, and John W. Terborgh. 2013. “Oligarchies in Amazonian Tree Communities: A Ten-Year Review.” *Ecography* 36 (2): 114–23. https://doi.org/10.1111/j.1600-0587.2012.00083.x.

Pitman, Nigel C. A., John Terborgh, Miles R. Silman, and Percy Nuñez V. 1999. “Tree Species Distribution in an Upper Amazonian Forest.” *Ecology* 80 (8): 2651–61.

Pitman, Nigel C A, John W Terbough, Miles R Silman, Percy Núñez V., David A Neill, Carlos E Cerón, Walter A Palacios, and Milton Aulestia. 2001. “Dominance and Distribution of Tree Species in Upper Amazonian Terra Firme Forests.” *Ecology* 82 (8): 2101–17.

Prado, D. E. 2000. “Seasonally Dry Forests of Tropical South America: From Forgotten Ecosystems to a Ney Phytogeographic Unit.” *Edinburgh Jounal of Botany* 57 (3): 437–61.

Prado, Darien E, and Peter E Gibbs. 1993. “Patterns of Species Distributions in the Dry Seasonal Forests of South America.” *Annals of the Missouri Botanical Garden* 80 (4): 902–27.

Queiroz, L P. 2006. “The Brazilian Caatinga: Phytogeographical Patterns Inferred from Distribution Data of the Leguminosae.” In *Neotropical Savannas and Dry Forests: Plant Diversity, Biogeography and Conservation*, edited by R T Pennington, G P Lewis, and J A Ratter, 113–49. Oxford: Taylor & Francis CRC Press.

Quesada, C. a., J. Lloyd, L. O. Anderson, N. M. Fyllas, M. Schwarz, and C. I. Czimczik. 2011. “Soils of Amazonia with Particular Reference to the RAINFOR Sites.” *Biogeosciences* 8 (6): 1415–40. https://doi.org/10.5194/bg-8-1415-2011.

Quesada, C. a., J. Lloyd, M. Schwarz, T. R. Baker, O. L. Phillips, S. Patiño, C. Czimczik, et al. 2009. “Regional and Large-Scale Patterns in Amazon Forest Structure and Function Are Mediated by Variations in Soil Physical and Chemical Properties.” *Biogeosciences Discussions* 6 (2): 3993–4057. https://doi.org/10.5194/bgd-6-3993-2009.

Quesada, C. a., J. Lloyd, M. Schwarz, S. Patiño, T. R. Baker, C. Czimczik, N. M. Fyllas, et al. 2010. “Variations in Chemical and Physical Properties of Amazon Forest Soils in Relation to Their Genesis.” *Biogeosciences* 7 (5): 1515–41. https://doi.org/10.5194/bg-7-1515-2010.

Quesada, Carlos Alberto, Martin G Hodnett, Lacê M Breyer, Alexandre J B Santos, Sérgio Andrade, Heloisa S Miranda, Antonio Carlos Miranda, and Jon Lloyd. 2008. “Seasonal Variations in Soil Water in Two Woodland Savannas of Central Brazil with Different Fire History.” *Tree Physiology* 28 (3): 405–15.

Rangel - Ch., J. Orlando. 2015. “La Biodiversidad de Colombia: Significado y Distribución Regional.” *Revista de La Academia Colombiana de Ciencias Exactas, Físicas y Naturales* 39 (51): 176. https://doi.org/10.18257/raccefyn.136.

Ratter, J. A., S. Bridgewater, and J. F. Ribeiro. 2003. “Analysis of the Floristic Composition of the Brazilian Cerrado Vegetation III: Comparison of the Woody Vegetation of 376 Areas.” *Edinburgh Jounal of Botany* 60 (1): 57–109.

Rezende, Vanessa Leite, Pedro V. Eisenlohr, André Luís De Gasper, Alexander Christian Vibrans, and Ary Teixeira De Oliveira-Filho. 2014. “Toward a Better Understanding of the Subtropical Atlantic Forest in the State of Santa Catarina, Brazil: Tree Sampling Accuracy, Species Richness and Frequency of Rare Species.” *Acta Botanica Brasilica* 28 (3): 382–91. https://doi.org/10.1590/0102-33062014abb3146.

Rezende, Vanessa Leite, Ary T. de Oliveira-Filho, Pedro V. Eisenlohr, Luciana Hiromi Yoshino Kamino, and Alexander Christian Vibrans. 2014. “Restricted Geographic Distribution of Tree Species Calls for Urgent Conservation Efforts in the Subtropical Atlantic Forest.” *Biodiversity and Conservation*, May. https://doi.org/10.1007/s10531-014-0721-7.

Ribeiro, J. F., and B. M. T. Walter. 2008. “As Principais Fitofisionomias Do Bioma Cerrado.” In *Cerrado: Ecologia e Flora*, edited by S. M. Sano, S. P. Almeida, and J. F Ribeiro, 1st ed., 151–212. Brasília: Embrapa.

Sampaio, E V S B. 1995. “Overview of the Brazilian Caatinga.” In *Seasonally Dry Tropical Forests*, edited by E V S B Sampaio, S H Bullock, H A Mooney, and E Medina, 35–63. Cambridge.

Sanchez, Maryland, Fernando Pedroni, Pedro V. Eisenlohr, and Ary T. Oliveira-Filho. 2013. “Changes in Tree Community Composition and Structure of Atlantic Rain Forest on a Slope of the Serra Do Mar Range, Southeastern Brazil, from near Sea Level to 1000m of Altitude.” *Flora - Morphology, Distribution, Functional Ecology of Plants* 208 (3): 184–96. https://doi.org/10.1016/j.flora.2013.03.002.

Sandel, Brody, Alvaro G. Gutiérrez, Peter B. Reich, Franziska Schrodt, John Dickie, and Jens Kattge. 2015. “Estimating Themissing Species Bias in Plant Trait Measurements.” Edited by Sebastian Schmidtlein. *Journal of Vegetation Science* 26 (5): 828–38. https://doi.org/10.1111/jvs.12292.

Santos, Rubens M, Ary T Oliveira-Filho, Pedro V Eisenlohr, Luciano P Queiroz, Domingos B O S Cardoso, and Maria J N Rodal. 2012. “Identity and Relationships of the Arboreal Caatinga among Other Floristic Units of Seasonally Dry Tropical Forests (SDTFs) of North-Eastern and Central Brazil.” *Ecology and Evolution* 2 (2): 409–28. https://doi.org/10.1002/ece3.91.

Santos, Bráulio A., Víctor Arroyo-Rodríguez, Claudia E. Moreno, and Marcelo Tabarelli. “Edge-Related Loss of Tree Phylogenetic Diversity in the Severely Fragmented Brazilian Atlantic Forest.” Edited by Jerome Chave. *PLoS ONE* 5, no. 9 (September 8, 2010): e12625. https://doi.org/10.1371/journal.pone.0012625.

Särkinen, Tiina, João R V Iganci, Reynaldo Linares-Palomino, Marcelo F Simon, and Darién E Prado. 2011. “Forgotten Forests--Issues and Prospects in Biome Mapping Using Seasonally Dry Tropical Forests as a Case Study.” *BMC Ecology* 11 (1): 27. https://doi.org/10.1186/1472-6785-11-27.

Scarano, F. R. 2002. “Structure, Function and Floristic Relationships of Plant Communities in Stressful Habitats Marginal to the Brazilian Atlantic Rainforest.” *Annals of Botany* 90 (4): 517–24. https://doi.org/10.1093/aob/mcf189.

Scarano, Fabio Rubio. 2009. “Plant Communities at the Periphery of the Atlantic Rain Forest: Rare-Species Bias and Its Risks for Conservation.” *Biological Conservation* 142 (6): 1201–8. https://doi.org/10.1016/j.biocon.2009.02.027.

Schimper, A. F. W. 1903. *Plant-Geography upon a Physiological Basis*. Oxford: Claredon Press.

Silva, José Maria Cardoso da, and Marcelo Tabarelli. “Tree Species Impoverishment and the Future Flora of the Atlantic Forest of Northeast Brazil” 404 (2000): 3.

Schulman, Leif, Tuuli Toivonen, and Kalle Ruokolainen. 2007. “Analysing Botanical Collecting Effort in Amazonia and Correcting for It in Species Range Estimation.” *Journal of Biogeography* 34 (8): 1388–99. https://doi.org/10.1111/j.1365-2699.2007.01716.x.

Sousa-Baena, Mariane S., Letícia Couto Garcia, and a. Townsend Peterson. 2014. “Knowledge behind Conservation Status Decisions: Data Basis for ‘Data Deficient’ Brazilian Plant Species.” *Biological Conservation* 173 (May): 80–89. https://doi.org/10.1016/j.biocon.2013.06.034.

Sousa-Baena, Mariane Silveira, Letícia Couto Garcia, and Andrew Townsend Peterson. 2014. “Completeness of Digital Accessible Knowledge of the Plants of Brazil and Priorities for Survey and Inventory.” Edited by Lluís Brotons. *Diversity and Distributions* 20 (4): 369–81. https://doi.org/10.1111/ddi.12136.

Spichiger, Rodolphe, Clément Calenge, and Bastian Bise. 2004. “Geographical Zonation in the Neotropics of Tree Species Characteristic of the Paraguay-Paraná Basin.” *Journal of Biogeography* 31: 1489–1501.

Staver, A. C., S. Archibald, and S. A. Levin. 2011. “The Global Extent and Determinants of Savanna and Forest as Alternative Biome States.” *Science* 334 (6053): 230–32. https://doi.org/10.1126/science.1210465.

Steege, H. ter, N. C. A. Pitman, D. Sabatier, C. Baraloto, R. P. Salomao, J. E. Guevara, O. L. Phillips, et al. 2013. “Hyperdominance in the Amazonian Tree Flora.” *Science* 342 (6156): 1243092–1243092. https://doi.org/10.1126/science.1243092.

Steege, Hans ter, Nigel C a Pitman, Oliver L Phillips, Jerome Chave, Daniel Sabatier, Alvaro Duque, Jean-François Molino, et al. 2006. “Continental-Scale Patterns of Canopy Tree Composition and Function across Amazonia.” *Nature* 443 (7110): 444–47. https://doi.org/10.1038/nature05134.

Steege, Hans ter, Nigel C a Pitman, Daniel Sabatier, Christopher Baraloto, Rafael P Salomão, Juan Ernesto Guevara, Oliver L Phillips, et al. 2013. “Hyperdominance in the Amazonian Tree Flora.” *Science (New York, N.Y.)* 342 (6156): 1243092. https://doi.org/10.1126/science.1243092.

Steege, Hans ter, Rens W. Vaessen, Dairon Cárdenas-López, Daniel Sabatier, Alexandre Antonelli, Sylvia Mota de Oliveira, Nigel C. A. Pitman, Peter Møller Jørgensen, and Rafael P. Salomão. 2016. "The Discovery of the Amazonian Tree Flora with an Updated Checklist of All Known Tree Taxa." *Scientific Reports* 6 (July): 29549. https://doi.org/10.1038/srep29549.

Stehmann, João Renato, Rafaela Forzza, Alexandre Salino, Sobral, and Luciana Hiromi Yoshino Kamino. 2009. *Plantas Da Floresta Atlântica*. 1st ed. Riode Janeiro: Jardim Botânico do Rio de Janeiro.

Tabarelli, Marcelo, Waldir Mantovani, and Carlos A. Peres. "Effects of Habitat Fragmentation on Plant Guild Structure in the Montane Atlantic Forest of Southeastern Brazil." *Biological Conservation* 91, no. 2 (December 1, 1999): 119–27. https://doi.org/10.1016/S0006-3207(99)00085-3.

Thomas, W, and M R Barbosa. 2008. "Natural Vegetation Types in the Altantic Forest of Northeastern Brazil." In *The Atlantic Coastal Forests of Northeastern Brazil*, edited by W Thomas, 6–20. Memoirs of the New York Botanical Garden.

Tuomisto, Hanna, Gabriela Zuquim, and Glenda Cárdenas. 2014. "Species Richness and Diversity along Edaphic and Climatic Gradients in Amazonia." *Ecography*, no. May (July): n/a-n/a. https://doi.org/10.1111/ecog.00770.

Vanzoline, P. E. 1963. "Problemas Faunísticos Do Cerrado." In *Simpósio Sobre o Cerrado*, edited by M. Ferri. São Paulo: Editora da Universidade de São Paulo.

Vanzolini, Paulo Emílio. 1963. "Problemas Faunísticos Do Cerrado." *Simpósio Sobre o Cerrado* 307 (321): 1.

Velloso, Agnes L, Ana Maria Giulietti, David C Oren, Evaristo Eduardo De Miranda, and Shirley Keel. 2001. "Ecorregiões Propostas Para o Bioma Caatinga." Recife.

Violle, Cyrille, Benjamin Borgy, and Philippe Choler. 2015. "Trait Databases: Misuses and Precautions." *Journal of Vegetation Science* 26 (5): 826–27. https://doi.org/10.1111/jvs.12325.

Wallace, Alfred Russel. 1855. "On the Law Which Has Regulated the Introduction of New Species." *Annals and Magazine of Natural History* 171 (2): 141–53.

Walter, Heinrich, and Elgene Box. 1976. "Global Classification of Natural Terrestrial Ecosystems." *Vegetatio* 32 (2): 75–81. https://doi.org/10.1007/BF02111901.

Werneck, Fernanda P. 2011. "The Diversification of Eastern South American Open Vegetation Biomes: Historical Biogeography and Perspectives." *Quaternary Science Reviews* 30 (13): 1630–48. https://doi.org/10.1016/j.quascirev.2011.03.009.

Werneck, Fernanda P., Gabriel C. Costa, Guarino R. Colli, Darién E. Prado, and Jack W. Sites. 2011. "Revisiting the Historical Distribution of Seasonally Dry Tropical Forests: New Insights

Based on Palaeodistribution Modelling and Palynological Evidencegeb.” *Global Ecology and Biogeography* 20 (2): 272–88. https://doi.org/10.1111/j.1466-8238.2010.00596.x.

Werneck, Fernanda P., Cristiano Nogueira, Guarino R. Colli, Jack W. Sites, and Gabriel C. Costa. 2012. “Climatic Stability in the Brazilian Cerrado: Implications for Biogeographical Connections of South American Savannas, Species Richness and Conservation in a Biodiversity Hotspot.” *Journal of Biogeography* 39 (9): 1695–1706. https://doi.org/10.1111/j.1365-2699.2012.02715.x.

Whittaker, R. H. 1975. *Communities and Ecosystems*. 2nd edition. New York: Macmillan USA.

Wittmann, Florian, Jochen Schöngart, Juan Carlos Montero, Thomas Motzer, Wolfgang J. Junk, Maria T F Piedade, Helder L. Queiroz, and Martin Worbes. 2006. “Tree Species Composition and Diversity Gradients in White-Water Forests across the Amazon Basin.” *Journal of Biogeography* 33 (8): 1334–47. https://doi.org/10.1111/j.1365-2699.2006.01495.x.

Woodward, F I, M R Lomas, and C K Kelly. 2004. “Global Climate and the Distribution of Plant Biomes.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 359 (1450): 1465–76. https://doi.org/10.1098/rstb.2004.1525.

Zomer, R. J., D. A. Bossio, A. Trabucco, L. Yuanjie, D. C. Gupta, and V. P. Singh. 2007. “Trees and Water: Small Holder Agroforestry on Irrigated Lands in Northern India.” Colombo, Sri Lanka.

Zomer, R. J., A. Trabucco, D. A. Bossio, O. van Straaten, and L. V. Verchot. 2008. “Climate Change Mitigation: A Spatial Analysis f Global Land Suitability for Clean Development Mechanism Afforestation and Reforestation.” *Agriculture, Ecosystems & Environment* 126: 67–80.

2

Using tree species inventories to map biomes and assess their climatic overlaps in lowland tropical South America

This chapter has been published in the journal Global Ecology and Biogeography and the co-authors are listed below. I conducted all analyses, with advice from co-authors, and I led the writing of the manuscript, while incorporating feedback from co-authors.

Reference:

Silva de Miranda, Pedro Luiz, Ary T. Oliveira-Filho, R. Toby Pennington, Danilo M. Neves, Timothy R. Baker, and Kyle G. Dexter. 2018. “Using Tree Species Inventories to Map Biomes and Assess Their Climatic Overlaps in Lowland Tropical South America.” *Global Ecology and Biogeography* 27 (8): 899–912. https://doi.org/10.1111/geb.12749.

Chapter 2 - Using tree species inventories to map biomes and assess their climatic overlaps in lowland tropical South America

Pedro Luiz Silva de Miranda, Ary Oliveira-Filho, R. Toby Pennington, Danilo M. Neves, Timothy R. Baker, Kyle G. Dexter

2.1 - Abstract

Aim

To define and map the main biomes of lowland tropical South America (LTSA) using data from tree species inventories and to test the ability of climatic and edaphic variables to distinguish amongst them.

Location

Lowland Tropical South America (LTSA), including Argentina, Bolivia, Brazil, Ecuador, Paraguay, Peru and Uruguay.

Time Period

Present

Major Taxa Studied

Trees

Methods

We compiled a database of 4,103 geo-referenced tree species inventories distributed across LTSA. We used *a priori* vegetation classifications and cluster analyses of floristic composition to assign sites to biome. We mapped these biomes geographically and assessed climatic overlaps amongst them. We implemented classification tree approaches to quantify how well climatic and edaphic data can assign inventories to biome.

Results

Our analyses distinguish savanna and seasonally dry tropical forest (SDTF) as distinct biomes, with the Chaco woodlands potentially representing a third dry biome in LTSA. Amongst the wet forests, we find that the Amazon and Atlantic Forests may represent different biomes as they are distinct in both climate and species composition. Our results show an important environmental overlap amongst biomes, with error rates to classify sites into biomes of 19-21% and 16-18% when only climatic data and with the inclusion of edaphic data, respectively.

Main Conclusions

Tree species composition can be used to determine biome identity at continental scales. We find high biome heterogeneity at small spatial scales, likely due to variation in edaphic conditions and disturbance history. This points to the challenges of using climatic and/or interpolation-based edaphic data or coarse resolution, remotely-sensed imagery to map tropical biomes. From this perspective, we suggest that using floristic information in biome delimitation will allow for greater synergy between conservation efforts centred on species diversity and management efforts centred on ecosystem function.

2.2 - Introduction

The biome concept has existed for over a century with the overarching purpose of delimiting recognisable, ecologically meaningful vegetation units. Humboldt (1816) used the term phytophysigonomy when referring to areas that may be geographically disjunct, but share similar vegetation physiognomy or structure. The link between vegetation structure and climatic conditions was detailed by Schimper (1903), who attributed these similarities to physiological and anatomical adaptations to precipitation and temperature. The relationship between vegetation form and climate permeates the majority of vegetation classification schemes proposed during the 20th century (Clements, 1916; Holdridge, 1947; Walter, 1973; Whittaker, 1975), and climate is still regarded as the main driver of plant and biome distributions (Box, 1995; Prentice et al., 1992; Prentice, 1990). More recently, biomes have been used to categorise the function of ecosystems at large spatial scales, including across continents (Higgins, Buitenwerf, & Moncrieff, 2016; Woodward, Lomas, & Kelly, 2004), and the most prevalent biome concept at present, which we employ here, is that of a widespread vegetation formation with distinct ecosystem function.

The term ‘biome’ itself was first employed by Clements (1916) when referring to the biotic community, or set of species, occupying a certain habitat. However, subsequently, Holdridge (1947), Walter (1973), Whittaker (1975) and Odum (1975) gave more emphasis to the relationship between climate and vegetation structure when proposing classification systems for vegetation formations or biomes, and distanced themselves from the community composition perspective suggested by Clements (1916). These latter authors delimited biomes using standard climatic variables, such as mean annual temperature (MAT) and mean annual precipitation (MAP) (e.g., Whittaker 1975). A motivating factor for these studies was to create practical classification systems that allow researchers to assign sites to biome by simply knowing the MAT and MAP (e.g., as in Qian, Jin, & Ricklefs, 2017; Siepielski et al., 2017). More recently, large-scale remotely sensed data have become available, which has led researchers to map biomes using simple characterisations of vegetation physiognomy or ecosystem function, including average vegetation height, percent tree cover, primary productivity and phenology (Higgins et al., 2016; Hirota, Holmgren, Van Nes, & Scheffer, 2011; Staver, Archibald, & Levin, 2011; Woodward et al., 2004). However, remote sensing approaches can fail when biomes are indinstinguiable from satellite images (Beuchle et al., 2015) or when there is high structural heterogeneity within biomes (Särkinen, Iganci, Linares-Palomino, Simon, & Prado, 2011)

Meanwhile, the different global biome schemes, be they derived from climate or remote sensing, often fail to agree on which are the main biomes (e.g Whitakker, 1975 vs. Friedl et al., 2002 vs. Woodward et al., 2004 vs. Higgins et al., 2016), and can differ dramatically on the mapping of any given biome (Särkinen et al., 2011). Furthermore, the degree to which biome maps actually delimit the spatial distribution of ecosystem function is debated (Moncrieff, Hickler, & Higgins, 2015). The need for more ecologically meaningful definitions of biomes has led some to suggest that functional traits, such as wood density or leaf mass per area of the dominant plant species, should be used to define and delimit biomes (Van Bodegom, Douma, & Verheijen, 2014; Violle, Reich, Pacala, Enquist, & Kattge, 2014). In order to map functional trait distributions at large spatial scales, researchers have used geo-referenced collection localities for species with available trait data (e.g. Engemann et al., 2016; Lamanna et al., 2014). There are challenges with this approach, most importantly, the absence of trait data for many species, especially in tropical vegetation (Baker et al., 2017; Sandel et al., 2015; Violle, Borgy, & Choler, 2015). The premise of this paper is that species occupying distinct biomes have different functional traits and therefore that floristic information can be used to map biomes, avoiding the uncertainties associated with linking species composition to trait databases. Species distribution modelling (a.k.a. ecological niche modelling) of indicator species can be used to map biomes (as in Prieto-Torres & Rojas-Soto, 2016; Särkinen et al., 2011), but such distribution modelling usually uses only climatic variables as predictors and therefore is subject to similar concerns as mapping biomes directly based on climatic data. We argue that, at least for some regions, there are now sufficient species distribution data to map biomes directly using the distribution data themselves.

The mapping of biomes based on floristic information also offers the possibility of synergies with conservation (Whittaker et al., 2005). Bioregionalisation schemes that partition space into geographic units based on species composition and environmental data, such as the global ecoregions proposed by Olson & Dinerstein (1998) and Olson et al. (2001) – recently reviewed and updated by Dinerstein et al. (2017) – have been used by researchers and decision makers in conservation at local and global scales. For example, it was by relying on Olson & Dinerstein's (1998) scheme that Myers et al. (2000) and Mittermeier et al. (1998, 2004) proposed the global biodiversity hotspots, which are biomes or geographic subsets of biomes (i.e. ecoregions), that present high numbers of endemic species and are particularly threatened.

Brazil, which comprises the majority of the land surface of Lowland Tropical South America (LTSA), has proposed its own bioregionalisation scheme, the Domain system, established by Veloso, Rangel Filho, & Lima (1991) and IBGE (2012). The six Domains, which are used to guide conservation and management policy, are the Amazon Forest, Atlantic Forest, Cerrado, Caatinga, Pantanal and Pampa. The first two are wet forests, with the Amazon Forest occupying much of northern LTSA and the Atlantic Forest occurring along the Atlantic coast of South America, principally in Brazil. They are separated by a 'Dry Diagonal' of seasonally dry forests, woodlands and savanna vegetation formations (Neves, Dexter, Pennington, Bueno, & Oliveira Filho, 2015; Vanzolini, 1963). The Cerrado Domain is comprised primarily of savanna and sits in the centre of the Dry Diagonal, occupying much of central Brazil, but there are disjunct patches of savanna found elsewhere in LTSA, particularly within the Atlantic and Amazon Forests (Ratter, Ribeiro, & Bridgewater., 1997). Wet forests intrude into the Cerrado as gallery forests along river courses (Oliveira-Filho & Ratter, 1995). The

Caatinga Domain at the northeast corner of the Dry Diagonal represents the largest extent of seasonally dry tropical forest (SDTF) in LTSA (Prado & Gibbs, 1993). However, SDTF also occurs in disjunct patches throughout the Cerrado on more fertile soils (DRYFLOR, 2016; Pennington, Prado, & Pendry, 2000; Prado & Gibbs, 1993). SDTFs and the Cerrado can be distinguished by physiognomy, function and dissimilarities in phylogenetic composition (Oliveira-Filho, Pennington, Rotella, & Lavin, 2014; Oliveira-Filho et al., 2013). The Chaco woodlands at the southwest of the Dry Diagonal are climatically seasonal and its woodlands do not experience fire. The Chaco woodlands have been considered distinct from SDTF on the basis that they experience regular frost, greater temperature seasonality and often distinct edaphic conditions, e.g. hypersaline soils (DRYFLOR, 2017; Prado & Gibbs, 1993). The Pantanal Domain has heterogeneous vegetation including SDTFs, savanna and swamps, while the Pampa Domain is a largely subtropical grassland that has forest patches along river courses and on certain edaphic conditions.

Lowland Tropical South America, due to its size, diversity and non-continuous geographic distribution of biomes and vegetation types, is an ideal system to study how biomes can be delimited, at a continental scale, through means other than climate and remote sensing. Its complex environmental controls of both climate and soil point to the necessity of developing a new approach for biome delimitation that is better linked to biodiversity. Biome schemes centred on species composition may be more useful for comparative biology, conservation, and enable a better understanding of the possible mechanistic relationships between vegetation and environment.

Here we test the utility and performance of a floristic approach for mapping biomes at a continental scale, with a particular focus on Brazil and neighbouring countries. We use a dataset of 4,103 geo-referenced floristic inventories of tree species that span the major climatic and edaphic gradients of the region. We first test how well climatic data perform in distinguishing among biomes. We hypothesize that climatic data will be able to distinguish wet forests from the dry biomes, but that it will fail to distinguish SDTF from savanna as they are often edaphically differentiated (Ratter et al., 1997). We also test the ability of edaphic data, when considered in conjunction with climate, to increase the accuracy of biome delimitation. Lastly, we assess how our floristic approach to mapping biomes compares with the ecoregion-based classification system of Dinerstein et al. (2017) (a revised version of Olson et al. (2001) system), and then for Brazil only, against the Domain classification of IBGE (2012). Our use of floristics data may allow for the delimitation and mapping of biomes in a manner directly relevant to managing ecosystems and developing conservation strategies, for example by enabling the modelling of future climate change effects on tropical vegetation (Prieto-Torres et al., 2016).

2.3 – Methods

2.3.1 - The NeoTropTree dataset

Floristic inventories of tree communities were obtained from the NeoTropTree (NTT) dataset (Oliveira-Filho, 2017), which contains tree species inventories for more than 6,000 geo-referenced sites across South America. Trees are here defined as free-standing woody plants greater than three metres in height. Every site in the NTT database is based on a tree species list generated via an inventory, phytosociological survey or floristic survey. These data sources are derived from published and unpublished literature (e.g. PhD theses, environmental consultancy reports). Other species are added to the site species list based on surveys of specimens in herbaria in South America, USA and Europe or online (e.g. CRIA, 2012). All entries are carefully checked for doubtful determinations and synonyms by consulting the taxonomic literature, the “Flora do Brasil” (http://floradobrasil.jbrj.gov.br/) and the “Flora del Conosur” (Zuloaga, Belgrano, Zuloaga, & Belgrano, 2015) – http://www.darwin.edu.ar/), with additional direct consultation of taxonomists. Our data excludes checklists with < 10 species, because in lowland tropical regions, this is invariably due to low sampling or collecting efforts, rather than truly low species richness.

The vegetation type for each site, as documented in the original data source, is recorded and standardized to the vegetation types in Oliveira-Filho (2017). When a herbarium voucher of an additional species is noted to come from within a 5 km radius of the original site, the collection label is checked to ensure that the species is found in the same vegetation type. Where two or more sites of different vegetation types co-occur within 10 km (768 sites– 19.13 % of our total), this results in geographically overlapping sites in the NTT database, each for a distinct vegetation type. Further details of NTT history, protocols and data can be found at www.neotroptree.info. We restricted analyses to the tropical and neighbouring subtropical lowlands of South America east of the Andes, and did not include any NTT site above 1,000 m elevation or below 36° S latitude. Montane areas were excluded because biogeographic barriers may be playing significant roles in floristic differentiation. Including subtropical sites allowed us to contextualize our results from the tropics. In total, we included 4,103 individual sites, containing 10,306 tree species from 1,062 genera and 148 families.

2.3.2 - Statistical Analyses

We performed hierarchical clustering based on tree species composition to assign sites to groups in an unsupervised manner (i.e. without reference to any environmental data). For clustering, we used the Simpson floristic distance amongst sites, which is the complement of the number of species shared between two sites divided by the maximum number of species that could be shared between the two sites: $1 - \text{species}_{\text{shared}}/\text{total\_species}_{\text{minimum}}$ (Baselga, 2010). This is identical to the β_{sim} metric (Kreft & Jetz, 2010), but we use the term Simpson distance because of its historical precedence (Baselga, 2010). This metric isolates the effects of species turnover and is not confounded by large differences in species richness amongst sites (Baselga, 2010). We built 1,000 clusters, each after randomising the row order in the matrix (species per site), following the procedure of Dapporto et al. (2013). We removed 24 sites that were unstable in their placement across the 1,000 clusters, which were identified by co-opting an approach used in phylogenetics to identify 'rogue taxa' that reduce resolution in phylogenetic analyses (Aberer, Krompass, & Stamatakis, 2012). In the final consensus cluster, only those groups that were present in at least 50% of the clusters are distinguished (Omland, Cook, & Crisp, 2008). This analysis was performed in R (R Team, 2016) using the "recluster" package (Dapporto et al., 2015).

To determine the biome identity of clusters, we used a reciprocal illumination procedure of assessing the overall structure of the cluster while considering site vegetation types. This process is inherently fractal and one could identify increasingly smaller groups of sites. We focused on defining biomes in the broadest sense in order to increase their generality and utility, and our delimitations were performed in the context of the main biomes that have previously been proposed for LTSA, namely wet or moist tropical forests (hereafter wet forests), SDTF, subtropical forests, savanna and chaco woodlands. In essence, our approach tested if there is floristic integrity to these previously proposed biomes, and we found clear evidence that there was, i.e. higher-level groups were comprised largely of one broad biome type. For heuristic purposes, we constructed a continuous biome map by applying Thiessen's polygons method in ArcGIS 10.4.1 (ESRI, 2017). This approach expands a polygon of a given biome classification for each NTT site until the polygons from neighbouring NTT sites are encountered. If they represent the same biome, then the polygons are fused and this procedure is continued until the entirety of the study area was categorised to biome.

We assessed which sites may be intermediate or transitional between our biomes using a silhouette analysis, via the R package cluster (Maechler, Rousseeuw, Struyf, Hubert, & Hornik, 2016). We also visually assessed where these ambiguously classified sites are located in species compositional space by means of a non-metric multidimensional scaling analysis (NMDS, McCune, Grace, and Urban 2002) of sites in two dimensions based on the Simpson distance amongst sites.

2.3.3 - Using climate and edaphic data to distinguish biomes

To assess if the biomes identified could be distinguished using climatic data, with or without edaphic data, we used a Random Forest classification tree approach (Breiman, 2001), implemented in the randomForest package in the R Statistical Software (Liaw & Wiener, 2002). We used 19 bioclimatic variables developed by Hijmans, Cameron, Parra, Jones, & Jarvis (2005), which quantify various aspects of temperature and precipitation regimes, as well as an estimate of average maximum climatological water deficit (CWD) per year (Chave et al., 2014). As edaphic variables, we included pH (extracted with KCl), cation exchange capacity (cmol/kg) and percentage of sand, silt and clay extracted from SoilGrids v0.5.5 (https://soilgrids.org/, (Hengl et al., 2017) at four different soil depths: 0 cm, 5 cm, 15 cm and 30 cm, which were then averaged. Two different classifications were performed, one considering climatic data alone and another considering both climatic and soil data.

In order to assess the success rate of the classification tree approach in assigning sites to biome and to determine which biomes were incorrectly classified, we generated confusion matrices, which show assignment based on climate alone or climate and soil versus assignment done above based on vegetation type and tree species composition. We also estimated the importance of each variable for distinguishing biomes using Breiman's measure of importance (Breiman, 2001). As we had substantial variation in sample size amongst our biomes that could bias importance measures, we equalized the number of sites across all biomes by rarefying to the number of sites present in the most poorly sampled biome. Rarefactions were performed randomly 100 times and variable importance values were averaged across the 100 replicates. In order to understand climatic overlaps amongst biomes, we additionally plotted sites in a pairwise manner for key climatic variables (MAP, MAT and CWD).

2.3.4 - Comparison to existing biome maps

We compared how two commonly used vegetation maps for South America classify sites to biome compared to our analyses. We focused on the map of Dinerstein et al. (2017), in which ecoregions are grouped into biomes and which is a revised version of Olson et al. (2001), and the Brazilian Domain system (IBGE 2012). We determined which biomes and domains in these systems conceptually correspond to the biomes we established here, and assessed how often these mapping systems gave the same identity to our NTT sites. The ecoregion data layer was obtained from https://ecoregions2017.appspot.com/ and the IBGE Domain data layer from http://www.geoservicos.ibge.gov.br/geoserver/web/ (layer CREN:biomas_5000).

2.4 - Results

2.4.1 - Biomes of Lowland Tropical South America

Hierarchical cluster analysis produced five higher-level groups (Figure 2.1), which we designated as biomes based on *a priori* vegetation type classifications. Wet forests fell into two different groups, which we tentatively treat as separate biomes. One comprises sites in the Amazon and the Guiana Shield, which we refer to as the Amazon Forest biome, and the other is comprised of sites along the Atlantic coast, which we refer to as the Atlantic Forest biome (Figure 2.2). These two biomes are largely concordant with the Amazon and Atlantic Forest Domains, except that they also include semideciduous and gallery forests, found well outside of the geographic areas of the forest Domains (Figure 2.2).

The other three major groups in the cluster are found primarily in the Dry Diagonal, which extends from northeast Brazil to Bolivia, Paraguay and northern Argentina (Figure 2.2). One, which we refer to as Savanna, comprises sites with a grassy understorey found throughout central Brazil and eastern Bolivia, overlapping with the Cerrado Domain, but with disjunct occurrences in the Amazon Forest and Atlantic

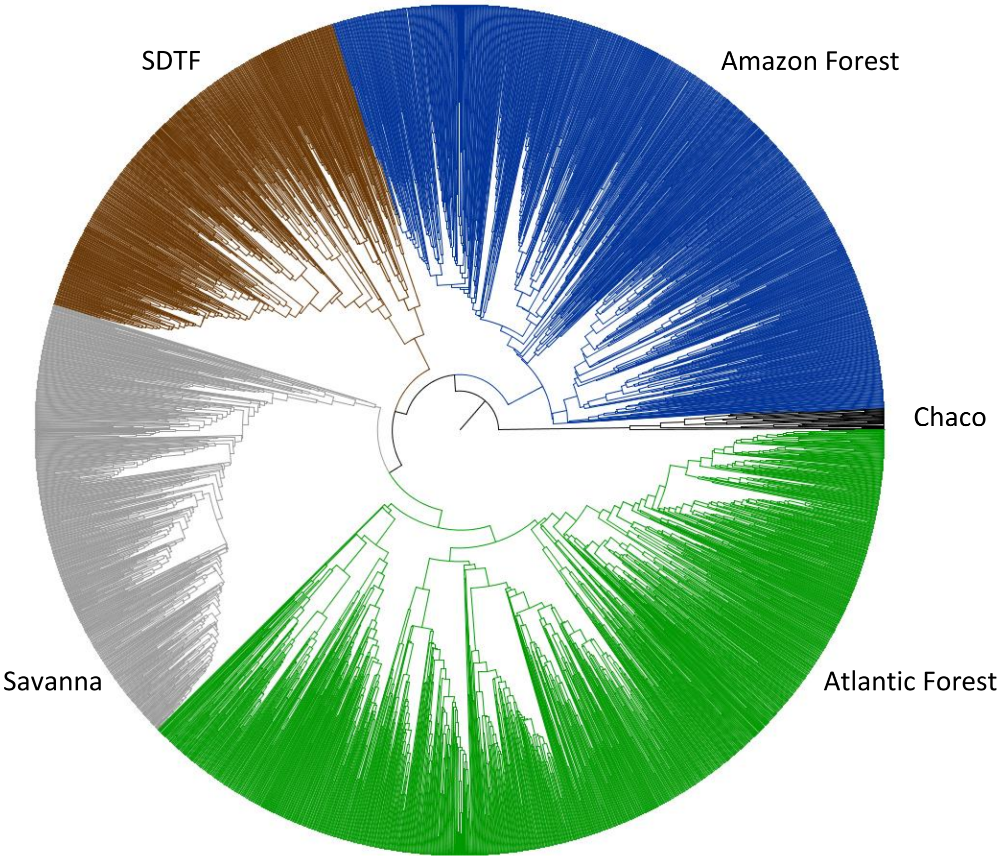

Figure 2.1: Hierarchical cluster of 4,103 sites in lowland (<1,000 m.a.s.l.) tropical South America and neighbouring subtropical areas based on tree species composition. Five principal higher-level groups can be observed, which were refer to as the Amazon Forest (blue), Atlantic Forest (green), Savanna (grey), Seasonally Dry Tropical Forest or SDTF (brown) and Chaco (black) biomes. See main text for details

Forest biomes. The Savanna biome is clearly distinguished floristically from a biome that we term Seasonally Dry Tropical Forest (SDTF), based on the original vegetation classifications of sites (http://www.neotroptree.info/). The SDTF biome has a discontinuous distribution from the Pantanal and Chiquitania in Bolivia and southern Brazil to its largest extension in the Caatinga Domain of northeastern Brazil (Figure 2.2). It is spatially interdigitated with the Savanna biome. The last group, which we distinguish as a separate biome is the Chaco, comprising woodlands in Bolivia, Argentina and Paraguay and extending to the borders of southern Brazil. While most of the sites in the Chaco biome cluster are subtropical and experience frost, there are a significant number of sites found north of 23 degrees latitude that are unlikely to experience freezing and can be considered tropical (Fig. 2.2). See Supplementary Materials (Appendix 1) for further description of the biomes. Our continuous biome map, developed using the Thiessen's polygons method, shows the LTSA biomes' overall spatial distribution and highlights the regions in which they interdigitate (Figure 2.3).

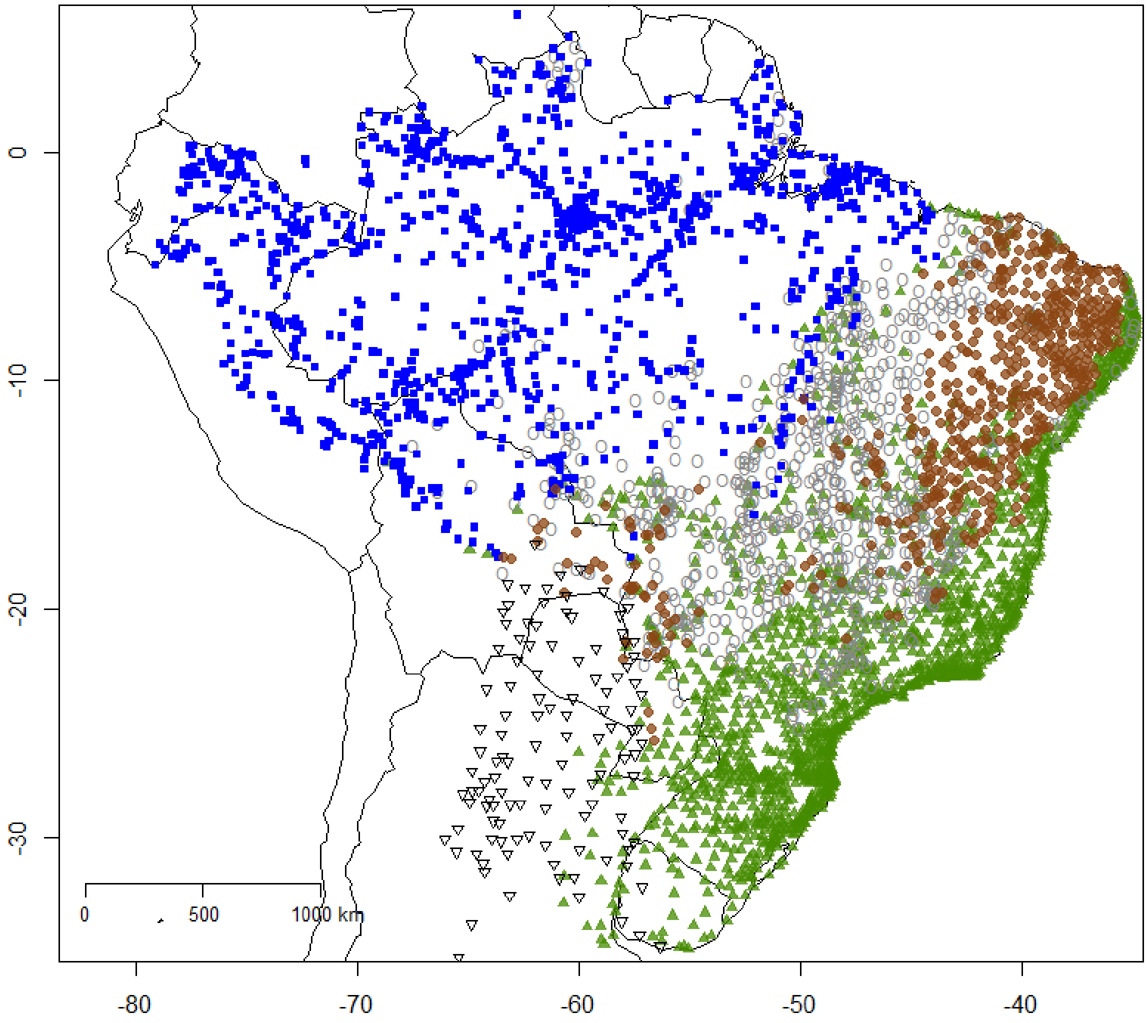

Figure 2.2: Map of Lowland Tropical South America with sites classified into biomes based on hierarchical cluster analysis of tree species composition: Atlantic Forest (green triangles), Seasonally Dry Tropical Forest (brown circles), Savanna (hollow grey circles), Amazon Forest (blue squares), Chaco (inverted hollow black triangles). Sites that were revealed to be more similar floristically to a different biome from the one with which they originally clustered are here given the symbol of the floristically more similar biome

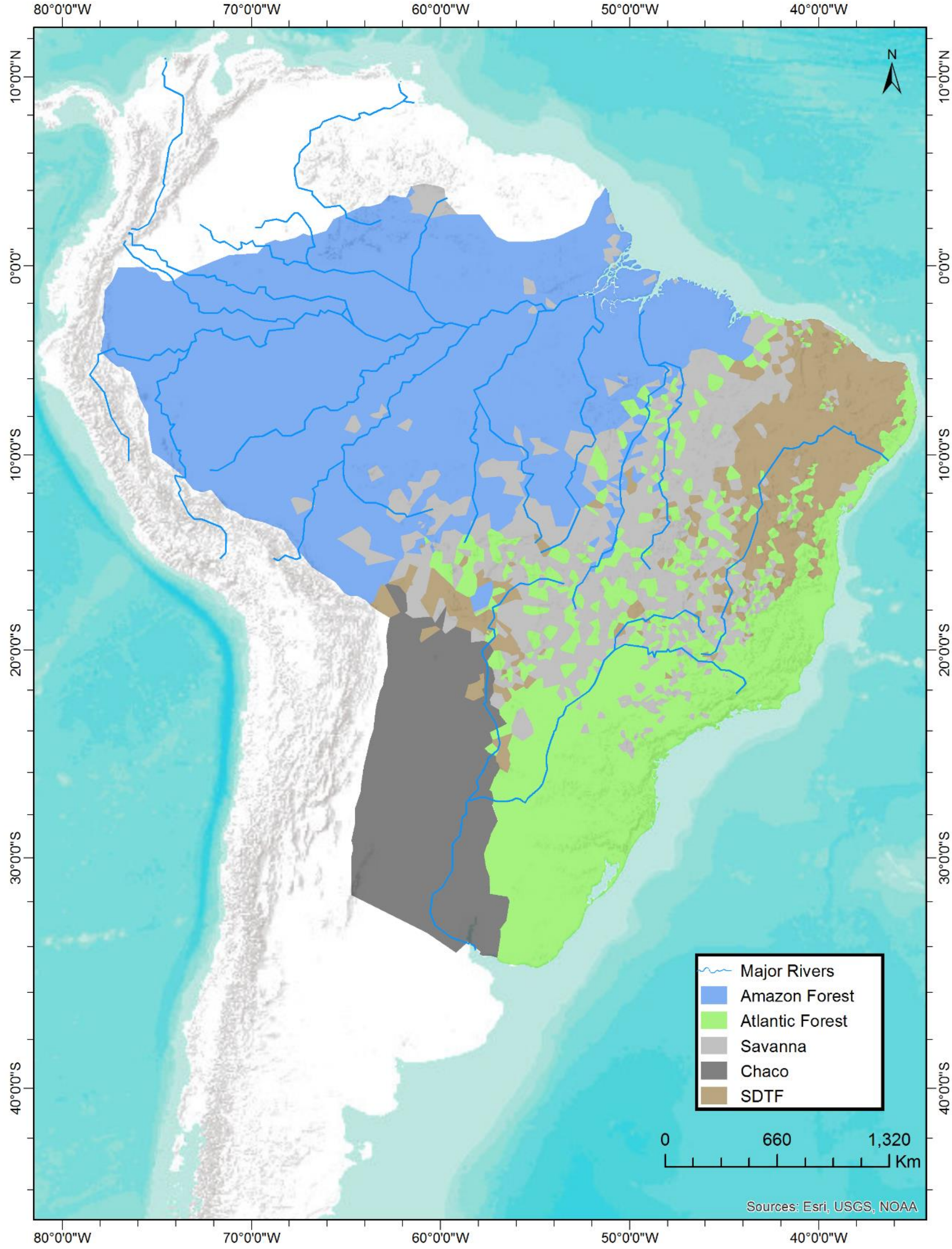

Figure 2.3: Map of South America with a schematic representation of the biomes delimited via hierarchical cluster analysis in the present contribution (Amazon Forest, Atlantic Forest, Savanna, Chaco and Seasonally Dry Tropical Forests – SDTF). The map was created by applying the Thiessen polygons method on the categorised points presented in figure 2. See text for further details.

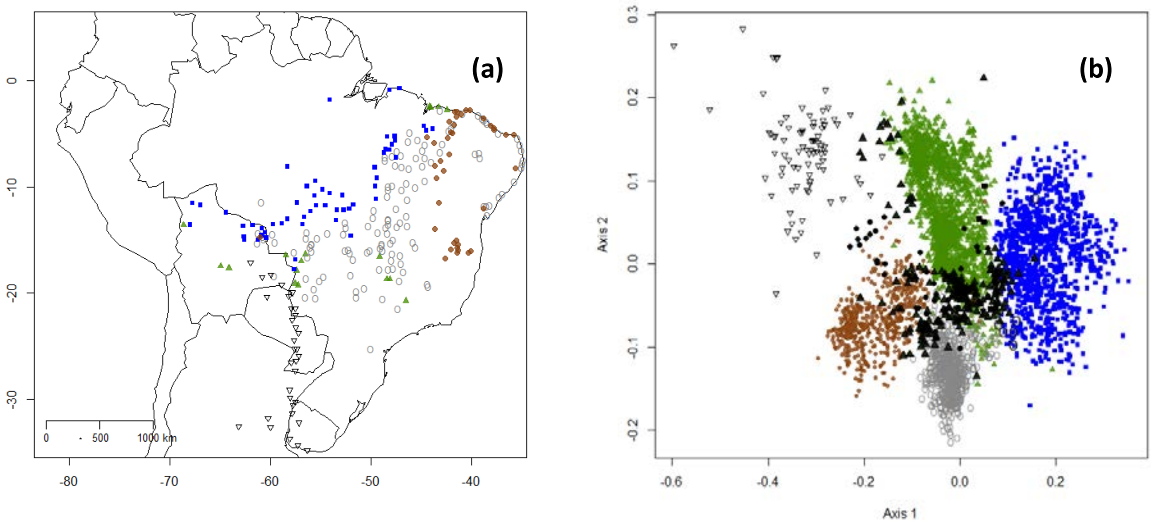

Figure 2.4: NeoTropTree sites which have a transitional/ambiguous floristic identity, as revealed by the silhouette analysis, and how they are distributed in geographic (a) and species compositional (b) spaces. In (a), sites are categorised according the biome to which they are floristically more similar. In (b), correctly classified sites are shown in the same colour scheme as Figure 2.2, whereas misclassified sites are represented in black and in the same shape as the sites of their biome based on the original clustering analysis. Symbols correspond to: Atlantic Forest (green triangles), Seasonally Dry Tropical Forest (brown circles), Savanna (hollow grey circles), Amazon Forest (blue squares), Chaco (inverted hollow black triangles).

Of 4,103 sites, 1,097 were classified as Amazon Forest, 1,566 as Atlantic Forest, 760 as Savanna, 564 as SDTF and 116 as Chaco. Silhouette analysis (Figure S2.1) showed that 271 sites are floristically more similar to a different biome than that with which they were original clustered, which we interpret to indicate that these sites are transitional between two biomes (Figure 2.4a, Table S2.1). An ordination of sites (NMDS with two axes, stress value= 0.1816) also suggests that these sites are compositionally transitional (Figure 2.4b). Floristically transitional sites were common between the Amazon and Atlantic Forest biomes (53 sites), between the Savanna and Atlantic Forest biomes (115 sites), and between the SDTF and Atlantic Forest biomes (49 sites), while they were infrequent between other biomes, including between any pair of dry biomes. Floristically transitional sites are common in the Dry Diagonal (Figure 2.4a), particularly between the Cerrado and the Amazon Forest and between the Chaco and the Atlantic Forest. Many of the gallery forests within the Cerrado Domain also have an ambiguous tree species compositional identity and are therefore difficult to classify.

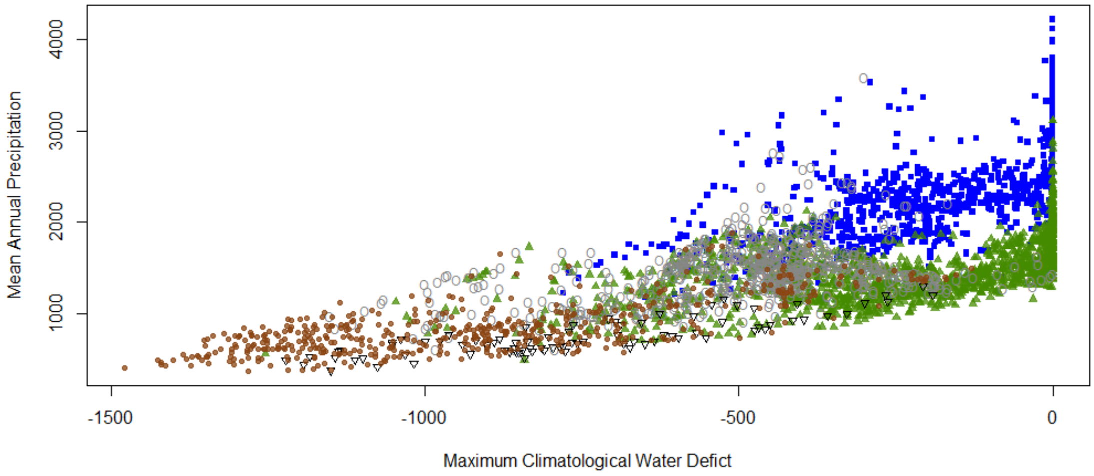

Figure 2.5: Distribution of sites with respect to precipitation regime. Mean annual precipitation values come from worldclim (Hijmans et al. 2005) and maximum climatological water deficit comes from Chave et al. (2014). Symbols correspond to: Atlantic Forest (green triangles), Seasonally Dry Tropical Forest (brown circles), Savanna (hollow grey circles), Amazon Forest (blue squares), Chaco (inverted hollow black triangles). Modelled after Fig. 1 in Malhi *et al.* (2009), which suggested that savannas were drier than seasonal forests, contrary to the pattern here.

2.4.2 - Using climate and edaphic data to distinguish biomes

We find that biomes overlap substantially in climatic space, both in terms of water availability (Figure 2.5) and temperature (Figure 2.6). For example, all five biomes defined here occupy at least two of the climatic biomes proposed by Whittaker (1975) (Figure 2.6). Of the 3,832 sites that are not considered transitional in nature, 712 were misclassified based on climate (18.6% of sites; Table 2.1). Considering all sites together, including transitional ones, we found a slightly higher error rate of 20.7% (Table S2.2). The most common misclassifications involved Amazon or Atlantic Forest sites being classified as belonging to the Savanna biome or vice versa, while climatic misclassifications of SDTF and Savanna were also common (Table 2.1). Sites in the Amazon and Atlantic. Forest wet biomes were distinct climatically. Meanwhile, the Chaco biome was rarely confused climatically with any of the other biomes. These patterns did not change when sites that have centres within 10 km of each other, i.e. overlapping in geographic space, were removed (Table S2.3, error rate: 20.3%).

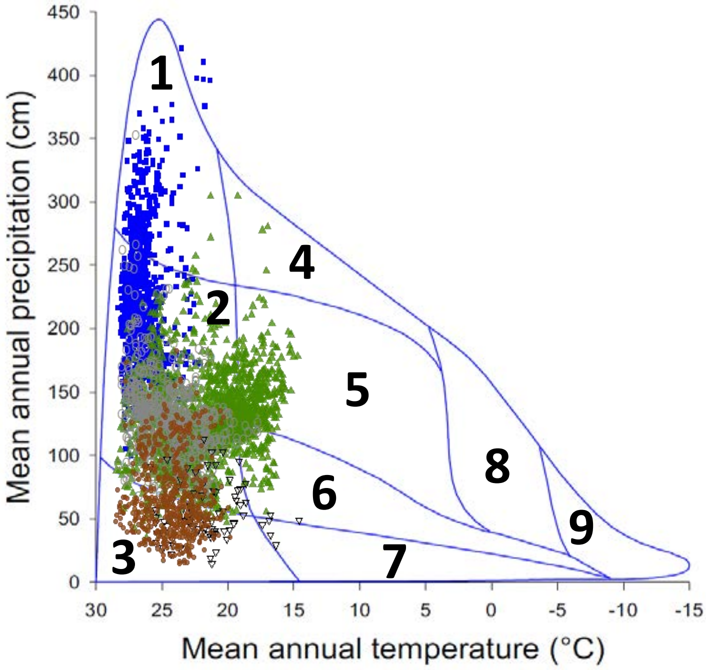

Figure 2.6: Distribution of sites in climatic space across the nine biomes proposed by Wittaker (1975) considering mean annual precipitation (cm) and mean annual temperature (C°). Numbers correspond to: Tropical rain forest (1), Tropical seasonal forest/savanna (2), Tropical and subtropical desert (3), Temperate rainforest (4), Temperate deciduous forest (5), Woodland/scrubland (6), Temperate grassland/dessert (7), Boreal forest (8), and Tundra (9). While symbols and colors correspond to: Atlantic Forest (green triangles), Seasonally Dry Tropical Forest (brown circles), Savanna (hollow gray circles), Amazon Forest (blue squares), and Chaco (inverted hollow black triangles).

The inclusion of edaphic variables slightly increased overall classification success by 3.2% (Table 2.2), and 3% when transitional sites were included (Table S2.4). There were a total of 124 sites that switched from being classified incorrectly (with just climatic data) to being classified correctly (once edaphic data were included; Table 2.2). Most of these were Savanna sites classified as Atlantic Forest and vice-versa.

Whether or not edaphic variables are included, the three main most important variables for classification were Mean Annual Precipitation, Temperature Seasonality and Maximum Climatological Water Deficit (Table 2.3). Overall, climatic variables seem to be more important than edaphic variables for distinguishing biomes, with variables related to precipitation, water availability and temperature seasonality ranking higher than variables related to mean temperature. However, overall we do have fewer edaphic variables and pH and cation exchange capacity (CEC) are among the top 10 variables (Table 2.3).

Table 2.1: Confusion matrix between sites categorised based on floristic composition via hierarchical clustering (rows) and sites categorised using climate and a classification tree approach (columns). The diagonal gives the number of sites that are correctly classified by climate, while the off-diagonal elements give mis-classifications (18.6%). Only non-floristically transitional sites were considered. Accuracy: 81%; Average precision: 81%; Average recall: 80%.

	Amazon Forest	Atlantic Forest	Cerrado	Chaco	SDTF
Amazon Forest	989	6	45	0	0
Atlantic Forest	3	1290	199	5	50
Cerrado	58	167	357	0	50
Chaco	0	7	0	76	1
SDTF	0	51	65	1	408

Table 2.2: Confusion matrix between sites categorised based on floristic composition via hierarchical clustering (rows) and sites categorised using climate + soil and a classification tree approach (columns). The diagonal gives the number of sites that are correctly classified by climate, while the off-diagonal elements give mis-classifications (15.2%). Accuracy: 84%; Average precision: 84%; Average recall: 83%.

	Amazon Forest	Atlantic Forest	Cerrado	Chaco	SDTF
Amazon Forest	1001	4	37	0	0
Atlantic Forest	4	1331	161	4	49
Cerrado	48	121	423	0	40
Chaco	0	7	0	76	1
SDTF	0	55	52	1	417

2.4.3 - Comparison to existing biome maps

The classification systems developed by Olson and Dinerstein et al. (2001, 2017) and IBGE (2012) assigned 74-75% of the NTT sites to the same biomes as they were placed according to our analyses (74.7% Dinerstein et al*.,* 2017, Table S2.5; 74.5% IBGE, 2012, Table S2.6). In Dinerstein's system, the majority of the misclassification results from Atlantic Forest sites being incorrectly classified as Tropical or Subtropical Savannas and Savanna being classified as Tropical Moist Forest (Figure S2.2). In IBGE's system, the error rate stems from SDTF sites being classified as Cerrado and vice-versa (Figure S2.3).

Table 2.3: The mean variable importance value (± one standard error) for all climatic variables included in the Random Forest analysis across 100 runs of the Bremnans' algorithm utilizing rarefactions of the main dataset (116 sites per biome**)**.

Environmental Variables	**Climate**	**Climate + Soil**
	Mean ± SE	**Mean ± SE**
Mean Annual Precipitation (mm)	356.81 ± 1.09	318.8 ± 1.18
Temperature Seasonality (C^{o})	319.73 ± 1.23	287.14 ± 1.13
Maximum Climatological Water Deficit (mm/yr)	273.2 ± 0.69	232.07 ± 0.71
Isothermality (%)	233.29 ± 0.98	211.53 ± 0.87
pH (KCl)	*	188.98 ± 0.84
Mean Temperature of Coldest Quarter (C^{o})	187.06 ± 0.95	163.07 ± 0.97
Precipitation of Wettest Quarter (mm)	155.06 ± 0.56	120.57 ± 0.48
Cation Exchange Capacity (cmol/Kg)	*	119.89 ± 0.23
Precipitation of Driest Quarter (mm)	148.46 ± 0.53	119.37 ± 0.51
Precipitation of Driest Month (mm)	133.16 ± 0.49	109.94 ± 0.44
Mean Annual Temperature (C^{o})	122.75 ± 0.71	96.15 ± 0.66
Precipitation of Wettest Month (mm)	119.83 ± 0.42	91 ± 0.35
Mean Temperature of Driest Quarter (C^{o})	106.46 ± 0.57	81.93 ± 0.49
Amount of Sand (%)	*	81.73 ± 0.17
Maximum Temperature of Warmest Month (C^{o})	103.8 ± 0.33	81.69 ± 0.31
Amount of Silt (%)	*	76.89 ± 0.13
Temperature Annual Range (C^{o})	101.51 ± 0.32	75.32 ± 0.23
Precipitation Seasonality (%)	99.22 ± 0.24	74.3 ± 0.31
Minimum Temperature of Coldest Month (C^{o})	99.21 ± 0.23	73.38 ± 0.37
Precipitation of Warmest Quarter (mm)	98.61 ± 0.3	70.77 ± 0.18
Precipitation of Coldest Quarter (mm)	97.11 ± 0.47	69.21 ± 0.25
Temperature's Diurnal Range (C^{o})	91.45 ± 0.19	68.67 ± 0.16
Amount of Clay (%)	*	65.97 ± 0.13
Mean Temperature of Wettest Quarter (C^{o})	79.01 ± 0.22	61.57 ± 0.24
Mean Temperature of Warmest Quarter (C^{o})	60.71 ± 0.12	46.52 ± 0.16

2.5 - Discussion

Our study demonstrates that using climatic data alone, with or without supplementary edaphic data, to map biomes would result in substantial error, causing misclassification of 15.2 - 20.7% of sites. Such misclassifications are due to pronounced climatic overlap of biomes (Figures 2.5, 2.6) and to edaphic heterogeneity at small spatial scales that is not captured by available data, which are derived via interpolation among relatively sparse soil sampling. Recently, researchers have begun assigning study sites to biomes, generally those of Whittaker (1975) based solely on climatic values, e.g. mean annual precipitation and temperature (e.g. Díaz et al., 2016; Qian & Ricklefs, 2017; Siepielsky et al*.,* 2017). Our results suggest this is potentially problematic (Figure 2.6). For example, the Amazon and Atlantic Forests can both occur in areas that are more seasonal than 'tropical rain forest' (sensu Whittaker), while the Savanna biome can occur in much wetter areas than indicated by Whittaker (1975; see also Lehmann et al., 2014). It is notable that none of our five major biomes are restricted to a single biome in Whittaker's climatic biome classification (Figure 2.6). We were able to employ a floristic approach to mapping biomes at a continental scale. Recent biome maps of LTSA, generally based on remote sensing, either fail to include major biomes (e.g. Seasonally Dry Tropical Forest is absent from Hirota et al., 2011; Staver et al., 2011), or are unable to distinguish amongst the dry tropical biomes of Savanna and SDTF (Beuchle et al., 2015). While floristic approaches to mapping biomes are unlikely to succeed inter-continentally because of the lack of shared species or even genera at this scale (Dexter et al., 2015), the increasing availability of floristic composition and species distribution information (e.g. www.gbif.org, www.forestplots.net, www.neotroptree.info) should allow this approach to be implemented within continents. It is important to note that any complete and continuous (or 'wall-to-wall') map of biome distribution will be inaccurate at small spatial scales due to high edaphic and floristic heterogeneity coupled with incomplete sampling. We have generated a continuous map (Figure 2.3), but its purpose is as a heuristic scheme to understand patterns in the distribution of biomes in LTSA. We do not contend that every point on the map is accurately classified, as that would belie one of the principal outcomes of this study, that of high biome heterogeneity at small spatial scales, as previously noted by Pennington et al. (2006), Werneck (2011) and Collevatti et al. (2013).

2.5.1 - Biomes of Lowland Tropical South America

Our analyses suggest three to five major biomes in LTSA. The Amazon and Atlantic Forests might represent separate biomes, whereas previously they have often been considered as a single tropical wet/moist forest biome. They are floristically distinct and their climatic niches are almost completely non-overlapping. Our floristic circumscription of the Atlantic Forest matches the *sensu-latissimo* definition of Oliveira-Filho, Jarenkow, & Rodal (2006). Our delimitation of the Amazon Forest is similar to previous studies that include the majority of the Amazon Basin drainage and the Guianan Shield (*e.g.*, Prance, 1982; ter Steege et al., 2006), although we note that our sampling of the Guianan Shield is limited.

The Savanna biome is floristically distinct from the other dry biomes, which is expected since it is a uniquely disturbance driven system, strongly influenced by fire (Archibald, Lehmann, Gómez-dans, & Bradstock, 2013; Ratter et al., 1997). Many sites in the SDTF biome are often drier, in terms of MAP and CWD, than the majority of sites in the Savanna biome (Figure 2.5), which runs counter to thinking that tropical wet forest transitions to tropical seasonal forest and then to savanna as water availability declines (e.g. Malhi et al., 2009). Meanwhile, our results from floristic analyses give support to previous studies (DRYFLOR, 2016; Pennington, Lavin, & Oliveira-Filho, 2009; Pennington et al., 2000; Prado & Gibbs, 1993) that have argued that the SDTFs scattered across lowland tropical South America should be regarded as a single biome, with the exclusion of the Chaco. We find that the climatic niches of Chaco and SDTF do not overlap, with the Chaco occurring in a colder climate with much higher temperature seasonality. However, further studies are needed that compare ecosystem function in the Amazon versus Atlantic Forests and in the SDTF versus Chaco to verify their status as distinct biomes. For further discussion of floristic patterns within and across biomes, please refer to the supplementary material (Appendix 2).

2.5.2 - Using climate and edaphic data to distinguish biomes

Mean annual precipitation (MAP), several measures of dry season precipitation and water deficit, temperature variability and soil pH were the most important environmental variables in distinguishing major biomes (Table 2.2). That precipitation-related variables are on average more important than temperature-related variables is to be expected, given that the majority of our sampling and most of the biomes under study are within the tropics, and thus represent a limited range of non-freezing temperature regimes (Augusto, Davies, Delzon, De Schrijver, & Chave, 2014). Nevertheless, it is notable that measures of temperature variability, particularly across seasons, were more important than other temperature measures, including mean annual temperature (MAT) and minimum temperature of the coldest month. This may be because plant species' ranges are often constrained by how much temperature can vary in a given location, and by temperature extremes (O'Sullivan et al., 2017).

2.5.3 - Comparing to existing biome maps

The comparisons between the classification system presented here and those of Dinerstein et al. (2017) and the Domain system (IBGE, 2012) revealed a ~25% misclassification rate for the latter two. These high error rates stem from two sources: the intrusion of SDTF and the Atlantic and Amazon Forests (as gallery forest) into the Savanna biome in the dry diagonal, and the existence of non-equivalent categories among these systems. Dinerstein et al. (2017) and IBGE (2012) recognize tropical and subtropical wetlands (named Pantanal in IBGE's system) as a distinct biome or domain, while the IBGE Domain system also delimits the Pampas (a.k.a. Campos Sulinos - southern Brazilian steppes). These two categories have not been detected and classified by our approach. Rather, the region classified as Pantanal by IBGE (2012) is covered by a mix of different vegetation formations that are floristically similar to SDTFs, Savannas and also the semideciduous portion of the Atlantic Forest. The forests within the area known as the Pampas at South Brazil are floristically similar, in relation to tree species composition, to the rest of the subtropical portion of the Atlantic Forest biome (Oliveira-Filho, Budke, Jarenkow, Eisenlohr, & Neves, 2015).

2.5.4 - Synergies between biodiversity conservation and ecosystem management

Delimiting biomes based on tree species composition offers the possibility of synergy between ecosystem management planning and conservation prioritisation. The biomes we have delimited differ in tree species composition and therefore likely differ in ecosystem function. Ecosystem management plans should therefore be developed separately for each. Similarly, these biomes have almost no species in common, yet have many species unique to them. Our schematic map (Figure 2.3) also indicates how these biomes are distributed at a continental scale, highlighting how discontinuous biome distribution can be in LTSA. These are important observations that must be considered in conservation and management. As an example, it is only recently that the SDTF have been recognised as a biome (Gentry, 1995; Murphy & Lugo, 1986; Prado & Gibbs, 1993), a definition consistent with our analyses, and there is no synthetic conservation plan that addresses the biome as a whole across the Neotropics (though see DRYFLOR 2016 for first steps). Current conservation planning for SDTF in Brazil focuses solely on the Caatinga Domain, but many Brazilian SDTFs are found in disjunct patches outside of this area, especially in the Cerrado, placing them under laws designed to protect savanna diversity. As another example, the Chaco is under great threat due to an increase of habitat destruction and fragmentation during the last 30 years (Hansen et al., 2013, Nori et al. 2016), but if recognised as a separate biome, as our analyses suggest, the urgency of its conservation may be better recognised (Kuemmerle et al., 2017).

2.6 - Conclusions

We have mapped the principal biomes in LTSA by using information on tree species composition of > 4,000 sites. The Savanna, Amazon and Atlantic Forest and SDTF biomes have an interdigitated distribution in central South America and overlap substantially in climatic space. Biome distribution cannot therefore be fully accounted for by climate, suggesting that climate projections alone will be insufficient to predict future biome shifts. Additional, meaningful environmental variables (e.g. available nitrogen, phosphorous, aluminium, etc.) must be measured and accounted for in models. The interdigitiation of biomes, especially in the dry diagonal across Brazil, is not recognised in the current IBGE (2012) system on which Brazilian conservation legislation is based, leading to the neglect of highly threaten SDTF vegetation outside of the Caatinga Domain. Our analyses also show Chaco and SDTF are distinct, which must be considered in land management and conservation. We suggest that species composition can be central to delimiting meaningful biomes for comparative research and conservation.

2.7 - References

Aberer, Andre J, Denis Krompass, and Alexandros Stamatakis. 2012. “Pruning Rogue Taxa Improves Phylogenetic Accuracy :,” 1–16.

Archibald, Sally, Caroline E R Lehmann, Jose L Gómez-dans, and Ross a Bradstock. 2013. “Defining Pyromes and Global Syndromes of Fire Regimes.” *Proceedings of the National Academy of Sciences* 110 (16): 6442–47. https://doi.org/10.1073/pnas.1211466110/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1211466110.

Baker, Timothy R., R. Toby Pennington, Kyle G. Dexter, Paul V.A. Fine, Helen Fortune-Hopkins, Euridice N. Honorio, Isau Huamantupa-Chuquimaco, et al. 2017. “Maximising Synergy among Tropical Plant Systematists, Ecologists, and Evolutionary Biologists.” *Trends in Ecology & Evolution* xx: 1–10. https://doi.org/10.1016/j.tree.2017.01.007.

Baselga, Andrés. 2010. “Partitioning the Turnover and Nestedness Components of Beta Diversity.” *Global Ecology and Biogeography* 19 (1): 134–43. https://doi.org/10.1111/j.1466-8238.2009.00490.x.

Beuchle, René, Rosana Cristina Grecchi, Yosio Edemir Shimabukuro, Roman Seliger, Hugh Douglas Eva, Edson Sano, and Frédéric Achard. 2015. “Land Cover Changes in the Brazilian Cerrado and Caatinga Biomes from 1990 to 2010 Based on a Systematic Remote Sensing Sampling Approach.” *Applied Geography* 58: 116–27. https://doi.org/10.1016/j.apgeog.2015.01.017.

Bodegom, P. M. Van, Jacob C Douma, and Lieneke M Verheijen. 2014. “A Fully Traits-Based Approach to Modeling Global Vegetation Distribution.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (38): 13733–38. https://doi.org/10.1073/pnas.1304551110.

Box, Elgene O. 1995. “Factors Determining Distributions of Tree Species and Plant Functional Types.” *Vegetatio* 121 (1–2): 101–16. https://doi.org/10.1007/BF00044676.

Breiman, Leo. 2001. “Random Forests.” *Machine Learning* 45 (1): 5–32. https://doi.org/10.1023/A:1010933404324.

Chave, Jérôme, Maxime Réjou-Méchain, Alberto Búrquez, Emmanuel Chidumayo, Matthew S. Colgan, Welington B C Delitti, Alvaro Duque, et al. 2014. “Improved Allometric Models to Estimate the Aboveground Biomass of Tropical Trees.” *Global Change Biology* 20 (10): 3177–90. https://doi.org/10.1111/gcb.12629.

Clements, Frederic E. 1916. “The Development and Structure of Biotic Communities.” In *Ecological Society of America*, 5:119–28. New York.

Collevatti, Rosane G., Matheus Souza Lima-Ribeiro, José Alexandre F. Diniz-Filho, Guilherme Oliveira, Ricardo Dobrovolski, and Levi Carina Terribile. 2013. “Stability of Brazilian Seasonally

Dry Forests under Climate Change: Inferences for Long-Term Conservation.” *American Journal of Plant Sciences* 04 (04): 792–805. https://doi.org/10.4236/ajps.2013.44098.

CRIA. 2012. “SpeciesLink.” Centro de Referência Em Informação Ambiental. 2012. http://www.splink.org.br/index.

Dapporto, Author Leonardo, Matteo Ramazzotti, Simone Fattorini, Gerard Talavera, Roger H L Dennis, and Maintainer Leonardo Dapporto. 2014. “Package ‘ Recluster .’”

Dapporto, Leonardo, Matteo Ramazzotti, Simone Fattorini, Gerard Talavera, Roger Vila, and Roger L H Dennis. 2013. “Recluster: An Unbiased Clustering Procedure for Beta-Diversity Turnover.” *Ecography* 36 (10): 1070–75. https://doi.org/10.1111/j.1600-0587.2013.00444.x.

Devisscher, Tahia, Liana O. Anderson, Luiz E. O. C. Aragão, Luis Galván, and Yadvinder Malhi. 2016. “Increased Wildfire Risk Driven by Climate and Development Interactions in the Bolivian Chiquitania, Southern Amazonia.” *PLoS ONE* 11 (9). https://doi.org/10.1371/journal.pone.0161323.

Dexter, K G, B Smart, C Baldauf, T R Baker, ..., and R. T Pennington. 2015. “Vegetation in Seasonally Dry Regions of the Tropics: Floristics and Biogeography.” *International Forestry Review* 17 (S2): 10–32.

Díaz, Sandra, Jens Kattge, Johannes H C Cornelissen, Ian J Wright, Sandra Lavorel, Stéphane Dray, Björn Reu, et al. 2016. “The Global Spectrum of Plant Form and Function.” *Nature* 529 (7585): 167–71. https://doi.org/10.1038/nature16489.

Dinerstein, Eric, David Olson, Anup Joshi, Carly Vynne, Neil D. Burgess, Eric Wikramanayake, Nathan Hahn, et al. 2017. “An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm.” *BioScience* 67 (6): 534–45. https://doi.org/10.1093/biosci/bix014.

DRYFLOR. 2016. “Plant Diversity Patterns in Neotropical Dry Forests and Their Conservation Implications.” *Science* 353 (6306): 1383–87.

DRYFLOR, R. Toby Pennington, Karina Banda-R, Alfonso Delgado-Salinas, Kyle G. Dexter, Luciano Galetti, Reynaldo Linares-Palomino, et al. 2017. “Response - Forest Conservation: Remember Gran Chaco.” *Science* 355 (6324): 465–66.

Engemann, Kristine, Brody Sandel, Brian J. Enquist, Peter MØller JØrgensen, Nathan Kraft, Aaron Marcuse-Kubitza, Brian Mcgill, et al. 2016. “Patterns and Drivers of Plant Functional Group Dominance across the Western Hemisphere: A Macroecological Re-Assessment Based on a Massive Botanical Dataset.” *Botanical Journal of the Linnean Society* 180 (2): 141–60. https://doi.org/10.1111/boj.12362.

ESRI. 2017. “ArcGIS 10.4.1.” Redlands: CA: Environmental Systems Research Institute.

Friedl, M A, D K McIver, J C Hodges, X Y Zhang, D Muchoney, and Strahler. 2002. “Global Land Cover Mapping from MODIS: Algorithms and Early Results.” *Remote Sensing of Environment* 83 (1–2): 287–302.

Gentry, A. H. 1995. “Patterns of Diversity and Floristic Composition in Neotropical Montane Forest.” In *Biodiversity and Conservation of Neotropical Montane Forest*, edited by S. P. Churchill, H. Balslev, E. Forero, and J. L. Luteyn, 103–26. New York: The New York Botanical Garden.

Hansen, M C, P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, et al. 2013. “High-Resolution Global Maps of 21st-Century Forest Cover Change.” *Science* 342 (6160): 850–53. https://doi.org/10.1126/science.1244693.

Hengl, Tomislav, Jorge Mendes de Jesus, Gerard B. M. Heuvelink, Maria Ruiperez Gonzalez, Milan Kilibarda, Aleksandar Blagotić, Wei Shangguan, et al. 2017. “SoilGrids250m: Global Gridded Soil Information Based on Machine Learning.” *PLOS ONE* 12 (2): e0169748. https://doi.org/10.1371/journal.pone.0169748.

Higgins, Steven I., Robert Buitenwerf, and Glenn R. Moncrieff. 2016. “Defining Functional Biomes and Monitoring Their Change Globally.” *Global Change Biology*, 1–11. https://doi.org/10.1111/gcb.13367.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. “Very High Resolution Interpolated Climate Surfaces for Global Land Areas.” *International Journal of Climatology* 25: 1965–78.

Hirota, Marina, Milena Holmgren, Egbert H. Van Nes, and Marten Scheffer. 2011. “Global Resilience of Tropical Forest and Savanna to Critical Transitions.” *Science* 334 (October): 232–35. https://doi.org/10.1126/science.1210657.

Holdridge, L. R. 1947. “Determination of World Plant Formations from Simple Climatic Data.” *Science* 105: 267–68.

Humboldt, Count Alexander. 1816. “XCIII. On the Laws Observed in the Distribution of Vegetable Forms.” *The Philosophical Magazine* 47 (218): 446–53. https://doi.org/10.1080/14786441608628504.

Instituto Brasileiro de Geografia e Estatística, ed. 2012. *Manual Técnico Da Vegetação Brasileira*. 2a edição revista e ampliada. Manuais Técnicos Em Geociências, número 1. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística-IBGE.

Kreft, Holger, and Walter Jetz. 2010. “A Framework for Delineating Biogeographical Regions Based on Species Distributions.” *Journal of Biogeography* 37 (11): 2029–53. https://doi.org/10.1111/j.1365-2699.2010.02375.x.

Kuemmerle, Tobias, Germán Baldi, Micaela Camino, Erika Cuellar, Rosa Leny Cuellar, Julieta Decarre, Sandra Díaz, et al. 2017. “Forest Conservation : Remember Gran Chaco Forest Conservation : Humans ’ Handprints.” *Science* 355 (6324): 465.

Lamanna, Christine, Benjamin Blonder, Cyrille Violle, Nathan J B Kraft, Brody Sandel, Irena Šímová, John C Donoghue, et al. 2014. “Functional Trait Space and the Latitudinal Diversity Gradient.” *Pnas* 111 (38): 13745–50. https://doi.org/10.1073/pnas.1317722111.

Lehmann, Caroline E R, T Michael Anderson, Mahesh Sankaran, Steven I Higgins, Sally Archibald, William A Hoffmann, Niall P Hanan, et al. 2014. “Savanna Vegetation-Fire-Climate Relationships Differ Among Continents.” *Science* 343 (January): 548–53. https://doi.org/10.1126/science.1247355.

Liaw, a, and M Wiener. 2002. “Classification and Regression by RandomForest.” *R News* 2 (December): 18–22. https://doi.org/10.1177/154405910408300516.

Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik. 2016. “Cluster: Cluster Analysis Basics and Extensions.”

Malhi, Yadvinder, L. E. O. C. Aragao, David Galbraith, Chris Huntingford, Rosie Fisher, Przemyslaw Zelazowski, Stephen Sitch, Carol McSweeney, and Patrick Meir. 2009. “Exploring the Likelihood and Mechanism of a Climate-Change-Induced Dieback of the Amazon Rainforest.” *Proceedings of the National Academy of Sciences* 106 (49): 20610–15. https://doi.org/10.1073/pnas.0804619106.

McCune, Bruce, James B. Grace, and Dean L. Urban. 2002. *Analyses of Ecological Communities*. Vol. 28. Gleneden Beach: OR: MjM software design.

Mittermeier, Russel A., Patricio Gil Gil, Michael Hoffmann, John Pilgrim, Thomas Brooks, Cristina Goettsch Mittermeier, John Lamoreux, Gustavo A.B. da Fonseca, LAMOREUX JOHN, and GUSTAVO A.B. DA FONSECA. 2004. *Hotspots Revisited*. 1st ed. Mexico City: CEMEX.

Mittermeier, Russell A., Norman Myers, Jorgen B. Thomsen, Gustavo A. B. da Fonseca, and Silvio Olivieri. 1998. “Biodiversity Hotspots and Major Tropical Wilderness Areas: Approaches to Setting Conservation Priorities.” *Conservation Biology* 12 (3): 516–20. https://doi.org/10.1046/j.1523-1739.1998.012003516.x.

Moncrieff, Glenn R., Thomas Hickler, and Steven I. Higgins. 2015. “Intercontinental Divergence in the Climate Envelope of Major Plant Biomes.” *Global Ecology and Biogeography* 24 (3): 324–34. https://doi.org/10.1111/geb.12257.

Murphy, Peter G, and Ariel E. Lugo. 1986. “Ecology of Tropical Dry Forest.” *Annual Review of Ecology and Systematics* 17: 67–88.

Myers, Norman, Russell A Mittermeier, Cristina G Mittermeier, Gustavo A B Fonseca, and Jennifer Kent. 2000. “Biodiversity Hotspots for Conservation Priorities.” *Nature* 403 (February): 853–58.

Neves, Danilo M., Kyle G. Dexter, R. Toby Pennington, Marcelo L. Bueno, and Ary T. Oliveira Filho. 2015. “Environmental and Historical Controls of Floristic Composition across the South American Dry Diagonal.” *Journal of Biogeography* 42 (8): 1566–76. https://doi.org/10.1111/jbi.12529.

Odum, Eugene P. 1975. *Ecology*. Second. London: Holt Rinehart and Winston.

Oliveira-Filho, A. T., and J. A. Ratter. 1995. “A Study of the Origin of Central Brazilian Forests by the Analysis of Plant Species Distribution Patterns.” *Edinburgh Jounal of Botany* 52 (2): 141–94.

Oliveira-Filho, Ary T. 2017. “NeoTropTree, Flora Arbórea Da Região Neotropical: Um Banco de Dados Envolvendo Biogeografia, Diversidade e Conservação.” NeoTropTree, Flora Arbórea Da Região Neotropical: Um Banco de Dados Envolvendo Biogeografia, Diversidade e Conservação. 2017. http://www.neotroptree.info/.

Oliveira-Filho, Ary T., Jean C. Budke, João A. Jarenkow, Pedro V. Eisenlohr, and Danilo R. M. Neves. 2015. “Delving into the Variations in Tree Species Composition and Richness across South American Subtropical Atlantic and Pampean Forests.” *Journal of Plant Ecology* 8 (3): 242–60. https://doi.org/10.1093/jpe/rtt058.

Oliveira-Filho, Ary T, Lojo Andre Jarenkow, and Maria Jesus Nogueira Rodal. 2006. “Floristic Relationships of Seasonally Dry Forests of Eastern South America Based on Tree Species Distribution Patterns.” In *Neotropical Savannas and Seasonally Dry Forests: Plant Diversity, Biogeography, and Conservation*, 159–92.

Oliveira-Filho, Ary T., R. Toby Pennington, Jay Rotella, and Matt Lavin. 2014. “Exploring Evolutionarily Meaningful Vegetation Definitions in the Tropics: A Community Phylogenetic Approach.” In *Forests and Global Change*, edited by David A. Coomes, David F. R. P. Burslem, and William D. Simonson, 239–60. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107323506.012.

Oliveira-Filho, A.T., D. Cardoso, B.D. Schrire, G.P. Lewis, R.T. Pennington, T.J. Brummer, J. Rotella, and M. Lavin. 2013. “Stability Structures Tropical Woody Plant Diversity More than Seasonality: Insights into the Ecology of High Legume-Succulent-Plant Biodiversity.” *South African Journal of Botany* 89 (November): 42–57. https://doi.org/10.1016/j.sajb.2013.06.010.

Olson, David M., and Eric Dinerstein. 1998. “The Global 200: A Representation Approach to Conserving the Earth’s Most Biologically Valuable Ecoregions.” *Conservation Biology* 12 (3): 502–15. https://doi.org/10.1046/j.1523-1739.1998.012003502.x.

Olson, David M., Eric Dinerstein, Eric D. Wikramanayake, Neil D. Burgess, George V. N. Powell, Emma C. Underwood, Jennifer A. D’amico, et al. 2001. “Terrestrial Ecoregions of the World: A New Map of Life on EarthA New Global Map of Terrestrial Ecoregions Provides an Innovative Tool for Conserving Biodiversity.” *BioScience* 51 (11): 933–38. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2.

Omland, Kevin E, Lyn G Cook, and Michael D Crisp. 2008. “Tree Thinking for All Biology: The Problem with Reading Phylogenies as Ladders of Progress.” *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology* 30 (9): 854–67. https://doi.org/10.1002/bies.20794.

O’sullivan, Odhran S., Mary A. Heskel, Peter B. Reich, Mark G. Tjoelker, Lasantha K. Weerasinghe, Aurore Penillard, Lingling Zhu, et al. 2017. “Thermal Limits of Leaf Metabolism across Biomes.” *Global Change Biology* 23 (1): 209–23. https://doi.org/10.1111/gcb.13477.

Pennington, R. Toby, Matt Lavin, and Ary Oliveira-Filho. 2009. “Woody Plant Diversity, Evolution, and Ecology in the Tropics: Perspectives from Seasonally Dry Tropical Forests.” *Annual Review of Ecology, Evolution, and Systematics* 40 (1): 437–57. https://doi.org/10.1146/annurev.ecolsys.110308.120327.

Pennington, R Toby, Darién E Prado, and Colin A Pendry. 2000. “Neotropical Seasonally Dry Forests and Quaternary Vegetation Changes.” *Journal of Biogeography* 27: 261–73.

Prado, Darien E., and Peter E. Gibbs. 1993a. “Patterns of Species Distributions in the Dry Seasonal Forests of South America.” *Annals of the Missouri Botanical Garden* 80 (4): 902–27.

Prado, Darien E, and Peter E Gibbs. 1993b. “Patterns of Species Distributions in the Dry Seasonal Forests of South America.” *Annals of the Missouri Botanical Garden* 80 (4): 902–27.

Prance, G.T. 1979. *Forest Refuges: Evidence from Woody Angiosperms*.

Prentice, I Colin, Wolfgang Cramer, Sandy P Harrison, Rik Leemans, A Robert, Allen M Solomon, Source Journal, et al. 1992. “A Global Biome Model Based on Plant Physiology and Dominance , Soil Properties and Climate.” *Journal of Biogeography* 19 (2): 117–34.

Prentice, K C. 1990. “Bioclimatic Distribution of Vegetation for General Circulation Models.” *Journal of Geophysical Research* 95: 11811–30. https://doi.org/10.1029/JD095iD08p11811.

Prieto-Torres, David A., and Octavio R. Rojas-Soto. 2016. “Reconstructing the Mexican Tropical Dry Forests via an Autoecological Niche Approach: Reconsidering the Ecosystem Boundaries.” *PLOS ONE* 11 (3): e0150932. https://doi.org/10.1371/journal.pone.0150932.

Qian, Hong, Yi Jin, and Robert E. Ricklefs. 2017. “Patterns of Phylogenetic Relatedness of Angiosperm Woody Plants across Biomes and Life-History Stages.” *Journal of Biogeography*. https://doi.org/10.1111/jbi.12936.

Ratter, J. a., J. F. Ribeiro, and Bridgewater S. 1997. “The Brazilian Cerrado Vegetation and Threats to Its Biodiversity.” *Annals of Botany* 80: 223–30. https://doi.org/10.1006/anbo.1997.0469.

Sandel, Brody, Alvaro G. Gutiérrez, Peter B. Reich, Franziska Schrodt, John Dickie, and Jens Kattge. 2015. “Estimating the Missing Species Bias in Plant Trait Measurements.” *Journal of Vegetation Science* 26 (5): 828–38. https://doi.org/10.1111/jvs.12292.

Särkinen, Tiina, João R V Iganci, Reynaldo Linares-Palomino, Marcelo F Simon, and Darién E Prado. 2011. “Forgotten Forests--Issues and Prospects in Biome Mapping Using Seasonally Dry Tropical Forests as a Case Study.” *BMC Ecology* 11 (1): 27. https://doi.org/10.1186/1472-6785-11-27.

Schimper, A. F. W. 1903. *Plant-Geography upon a Physiological Basis*. Oxford: Clarendon Press.

Siepielski, Adam M, Michael B Morrissey, Mathieu Buoro, Stephanie M Carlson, Christina M Caruso, Sonya M Clegg, Tim Coulson, et al. 2017. “Precipitation Drives Global Variation in Natural Selection Adam.” *Science* 355: 959–62. https://doi.org/10.1126/science.aag2773.

Staver, a Carla, Sally Archibald, and Simon a Levin. 2011. “The Global Extent and Determinants of Savanna and Forest as Alternative Biome States.” *Science (New York, N.Y.)* 334 (6053): 230–32. https://doi.org/10.1126/science.1210465.

Steege, Hans ter, Nigel C. A. Pitman, Oliver L. Phillips, Jerome Chave, Daniel Sabatier, Alvaro Duque, Jean-François Molino, et al. 2006. “Continental-Scale Patterns of Canopy Tree Composition and Function across Amazonia.” *Nature* 443 (7110): 444–47. https://doi.org/10.1038/nature05134.

Team, R Core. 2016. “R: A Language and Environment for Statistical Computing.” Vienna, Austria: R Foundation for Statistical Computing.

Vanzolini, Paulo Emílio. 1963. “Problemas Faunísticos Do Cerrado.” *Simpósio Sobre o Cerrado* 307 (321): 1.

Veloso, Henrique Pimenta, Antonio Lourenço Rosa Rangel Filho, and Jorge Carlos Alves Lima. 1991. *Classificação da vegetação brasileira, adaptada a um sistema universal*. Rio de Janeiro: Ministério da Economia, Fazenda e Planejamento, Fundação Instituto Brasileiro de Geografia e Estatística, Diretoria de Geociências, Departamento de Recursos Naturais e Estudos Ambientais.

Violle, Cyrille, Benjamin Borgy, and Philippe Choler. 2015. “Trait Databases: Misuses and Precautions.” *Journal of Vegetation Science* 26 (5): 826–27. https://doi.org/10.1111/jvs.12325.

Violle, Cyrille, Peter B Reich, Stephen W Pacala, Brian J Enquist, and Jens Kattge. 2014. “The Emergence and Promise of Functional Biogeography.” *Proceedings of the National Academy of Sciences* 111 (38): 13690–96. https://doi.org/10.1073/pnas.1415442111.

Walter, H. 1973. *Vegetation of the Earth in Relation to Climate and the Eco-Physiological Conditions*. New York: Springer.

Werneck, Fernanda P. 2011. “The Diversification of Eastern South American Open Vegetation Biomes: Historical Biogeography and Perspectives.” *Quaternary Science Reviews* 30 (13): 1630–48. https://doi.org/10.1016/j.quascirev.2011.03.009.

Whittaker, R.H. 1975. *Community and Ecosystems*. New York: McMillan.

Whittaker, Robert J, Miguel B Araújo, Paul Jepson, Richard J Ladle, E James, M Watson, Katherine J Willis, et al. 2005. "Conservation Biogeography : Assessment and Prospect." *Diversity and Distributions* 11 (1): 3–23.

Woodward, F I, M R Lomas, and C K Kelly. 2004. "Global Climate and the Distribution of Plant Biomes." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 359 (1450): 1465–76. https://doi.org/10.1098/rstb.2004.1525.

Zuloaga, Fernando O., Manuel J. Belgrano, Fernando O. Zuloaga, and Manuel J. Belgrano. 2015. "The Catalogue of Vascular Plants of the Southern Cone and the Flora of Argentina: Their Contribution to the World Flora." *Rodriguésia* 66 (4): 989–1024. https://doi.org/10.1590/2175-7860201566405.

2.8 - Appendix 1: Main Biomes of Lowland Tropical South America – Brief descriptions

2.8.1 - Wet Forest Biomes (Amazon and Atlantic Forests)

All rain forests, moist forests, evergreen forests and most semideciduous forests fell within two overarching groups in the cluster analysis, which we termed the Atlantic and Amazon Forest biomes. While we have argued that a floristics approach can be used to delimit biomes at continental scales where biogeographic factors are not the main driver of turnover in species composition, it may be that the floristic differentiation between the Atlantic and Amazon Forests is due in part to their biogeographic isolation by the Dry Diagonal. However, to definitely determine whether these forests represent distinct biomes, further comparative research is needed to determine how they compare in terms of ecosystem function.

The Atlantic Forest biome can be further divided into three different floristic groups, a northern group, completely tropical, encompassing all the coastal Atlantic forests ranging from northeast Brazil south to the state of Rio de Janeiro; a second group, largely sub-tropical, beginning at Sao Paulo’s coast and harbouring all of the forests covering the South of Brazil, Uruguay, Southeast Paraguay and portions of Northeast Argentina, especially the Missiones region; and a last group, also tropical, formed by semideciduous forests further inland, scattered mostly across Brazil, but also present as far west as Bolivia. This distribution matches the *sensu-latissimo* definition proposed by (Oliveira-Filho et al., 2006) with the additional inclusion of forest patches amongst the subtropical grasslands in the south of Brazil, Southern Paraguay, most of Uruguay and Northeast Argentina. This region has been distinguished from the Atlantic Forest in the past based on its overall physiognomy of forest patches in a grassy landscape, which contrasts with contiguous forest. However, these forest patches clearly show strong floristic continuity with the Atlantic Forest, as was also observed by Oliveira-Filho et al. (2013), and likely have similar ecosystem function to the now heavily fragmented Atlantic Forest.

The Amazon Forest biome does not show as clear subdivisions as the Atlantic Forest biome. However, there is evident floristic differentiation between "terra firme" and seasonally flooded forests, and these two subgroups can be further divided between sites in the western Amazon (from Peru, Bolivia, Ecuador and the Brazilian state of Acre) and the eastern Amazon (encompassing most of the Brazilian portion of the Amazon Forest, including the states of Amazonas, Pará, Mato Grosso, Maranhão and Roraima). These divisions between eastern and western Amazon and between "terra firme" and seasonally flooded forests have been reported before in the literature (e.g. Prance 1982; ter Steege et al., 2006).

The gallery forests within the Cerrado Domain do not cluster with the prevailing Savanna biome in that Domain, nor do they form their own unique cluster. Instead, they are floristically most similar to the most geographically proximal wet forest biome, either the Atlantic or the Amazon Forest. Similarly, sites found in sandy coastal areas of Brazil, often termed "restingas" or "matas de maré", do not comprise a single group in our hierarchical cluster, but cluster with the closest wet forest biome (Atlantic or Amazon Forest).

2.8.2 - Dry Biomes (Savanna, Seasonally Dry Tropical Forest and Chaco)

Our analyses confirm that the savannas distributed across LTSA form a single floristic unit. There are no clear subdivisions within this Biome. Savanna is a disturbance driven system, which may allow for the ready establishment of dispersing propagules of dominant tree species and a homogenisation of the tree flora over large spatial scales. Indeed, savannas in SA have been shown to possess a consistent set of dominant oligarchic tree species (Bridgewater et al., 2004), which may be why clear subgroups are not evident. In addition, the high disturbance in the system may prevent tree communities from reaching an equilibrium or 'climax' in species composition, which may inhibit sites from converging on similar species composition in similar environments, which could in turn inhibit the formation of clear floristic groups.

Our analyses suggest that the SDTF scattered across lowland tropical South America should be regarded as one single biome, as has been suggested by previous studies (DRYFLOR, 2016; Pennington et al., 2000, 2009; Prado and Gibbs, 1993). As found by Neves et al. (2015) and DRYFLOR (2016), our results suggest two main groups across the Dry Diagonal, one comprising the various forests of the Caatinga Domain and the other comprising SDTF patches scattered throughout the Cerrado Domain and into regions of the Pantanal and Chiquitania. The Misiones floristic group here shows greater floristic affinity with the Atlantic Forest than it does with other SDTF. The Misiones forests receive more rainfall than other STDF (Neves et al. 2015) and are semi-deciduous in nature (DRYFLOR 2016). Meanwhile, the Piedmont forests are found to be floristically more similar to sites in the Chaco than to other SDTF. This is perhaps not surprising given their proximity to the Chaco and that both environments receive significant frost in the winter season (Neves et al. 2015).

The Chaco is floristically different, in terms of tree species composition, from other sites across LTSA. While this difference has been noted in the past, particularly in comparison with SDTF (Pennington et al., 2000; Prado & Gibbs, 1993; Spichiger et al., 2004), it has often been attributed to the Chaco experiencing heavy frost. While many of the sites in our Chaco biome do experience frost, a large number of sites in eastern Bolivia, western Paraguay and south central Brazil (Mato Grosso do Sul state) do not experience frost, and could be considered tropical in nature. We refer to these northern Chaco sites as the ‘tropical Chaco’. It is floristically distinct from other SDTF and may have different ecosystem function, but further research is needed to compare ecosystem function in SDTF versus tropical and subtropical Chaco sites.

2.8.3 - Chiquitania and Pantanal

Two regions that have always been a challenge to place in floristic or biome classification schemes are the Chiquitania and Pantanal regions of eastern Bolivia and southwestern Brazil. The Chiquitania region is the site of contact between savannas (composed mostly of the savanna wetlands from the Pantanal region and the Llanos de Moxos region in Bolivia), Amazon Forest, SDTF and the Chaco (Killeen et al., 2006; Pennington et al., 2009). This region is composed of a mosaic of SDTF mixed with savannas, overlying diverse old geological formations (Navarro, 2011), and its northern portion grades into the Amazon Forest. Chiquitania is notable for its lack of endemic plant species, which is attributed to its recent geological past and to its transitional nature (Killeen et al., 2006). Our analyses show that sites within the Chiquitania's geographic range (Navarro, 2011) alternatively cluster together with the SDTF, Savanna and Amazon Forest biomes, and that perhaps the region should not be considered as a distinct vegetation entity on its own.

The floristic identity of forests and woodlands in the Pantanal also do not stand out as distinct within a continental context, although such was proposed by Veloso et al. (1991) and Navarro (2011). The Brazilian government also classifies it as a unique Domain (IBGE, 2012). However, just like Chiquitania, the Pantanal is composed of sites that belong to the Savanna and SDTF biomes as well as a wet forest biome, but in this case the Atlantic Forest biome. The lack of endemic species in this region is also evidence of its lack of floristic distinctness and recent geological history (Pott et al., 2011).

2.8.4 - References

Bridgewater, S., Ratter, J. A., & Ribeiro, J. F. (2004). Biogeographic patterns , b-diversity and dominance in the cerrado biome of Brazil. *Biodiversity and Conservation*, *13*, 2295–2318.

DRYFLOR. (2016). Plant diversity patterns in neotropical dry forests and their conservation implications. *Science*, *353*, 1383–1387.

IBGE. (2012). *Manual Técnico da Vegetação Brasileira* (segunda ed). Rio de Janeiro: Instituto de Brasileiro de Geografia e Estatística - IBGE.

Killeen, T., Chavez, E., Peña-Claros, M., Toledo, M., Arroyo, L., Caballero, J., … Steininger, M. (2006). The Chiquitano Dry Forest, the Transition between Humid and Dry Forest in Eastern Lowland Bolivia. In R. . T. Pennington, G. P. . Lewis, & J. A. Ratter (Eds.), *Neotropical Savannas and Seasonally Dry Forests Plant Diversity, Biogeography, and Conservation* (pp. 213–233). CRC Press.

Navarro, G. (2011). *Classificación de la Vegetación de Bolivia*. Santa Cruz de la Sierra: Centro de Ecología Difusión Simón I. Patiño.

Neves, D. M., Dexter, K. G., Pennington, R. T., Bueno, M. L., & Oliveira-Filho, A. T. (2015). *Environmental and historical controls on floristic composition across the South American Dry Diagonal*.

Oliveira-Filho, A. T., Budke, J. C., Jarenkow, J. A., Eisenlohr, P. V., & Neves, D. R. M. (2013). Delving into the variations in tree species composition and richness across South American subtropical Atlantic and Pampean forests. *Journal of Plant Ecology, 8*, 242–260.

Oliveira-Filho, A. T., Jarenkow, L. A., & Rodal, M. J. N. (2006). Floristic relationships of seasonally dry forests of Eastern South America based on tree species distribution patterns. In *Neotropical Savannas and Seasonally Dry Forests: Plant Diversity, Biogeography, and Conservation* (pp. 159–192).

Pennington, R. T., Lavin, M., & Oliveira-Filho, A. (2009). Woody Plant Diversity, Evolution, and Ecology in the Tropics: Perspectives from Seasonally Dry Tropical Forests. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 437–457.

Pennington, R. T., Prado, D. E., & Pendry, C. A. (2000). Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography*, *27*, 261–273.

Pott, a, Oliveira, a K. M., Damasceno-Junior, G. a, & Silva, J. S. V. (2011). Plant diversity of the Pantanal wetland. *Brazilian Journal of Biology = Revista Brasleira de Biologia*, *71*(1 Suppl 1), 265–273.

Prado, D. E., & Gibbs, P. E. (1993a). Patterns of Species Distributions in the Dry Seasonal Forests of South America. *Annals of the Missouri Botanical Garden*, *80*, 902–927.

Prado, D. E., & Gibbs, P. E. (1993b). Patterns of species distributions in the dry seasonal forests of South America. *Annals of the Missouri Botanical Garden*, *80*, 902–927.

Prance, G. T. (1982). *Biological Diversity in the Tropics*. (N. York, Ed.). Columbia University Press.

Spichiger, R., Calenge, C., & Bise, B. (2004). Geographical zonation in the Neotropics of tree species characteristic of the Paraguay-Paraná Basin. *Journal of Biogeography*, *31*, 1489–1501.

ter Steege, H., Pitman, N. C. a, Phillips, O. L., Chave, J., Sabatier, D., Duque, A., … Vásquez, R. (2006). Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature*, *443*, 444–447.

Veloso, H. P., Rangel Filho, A. L. R., & Lima, J. C. A. (1991). *Classificação da Vegetação Brasileira Adaptada a um Sistema Universal*. Rio de Janeiro: Fundação Instituto Brasileiro de Geografia e Estatística, IBGE.

2.9 - Appendix 2: Supplementary figures and tables

Table S2.1: Summary of results for silhouette analysis. The rows correspond to totals under the original classification, derived from the hierarchical clustering analysis, while the columns correspond to totals based on looking at the overall similarity of sites to the multidimensional centroid of each major group in the cluster. The diagonal corresponds to sites where the two approaches agree, while the off-diagonal elements correspond to sites where the two approaches disagree, which we consider to indicate sites that are transitional between the two biomes.

	Amazon Forest	Atlantic Forest	Cerrado	Chaco	SDTF	Original Classification
Amazon Forest	1042	7	0	0	0	1049
Atlantic Forest	46	1549	115	19	39	1768
Cerrado	7	0	632	0	0	639
Chaco	0	0	0	84	0	84
SDTF	2	10	13	13	525	563
Corrected Classification	1097	1566	760	116	564	4103

Table S2.2: Confusion matrix between sites categorised based on floristic composition via hierarchical clustering (rows) and sites categorised using climate + soil and a classification tree approach (columns), for all sites including ones identified as transitional via a silhouette analysis. The diagonal gives the number of sites that are correctly classified by climate + soil, while the off-diagonal elements give mis-classifications (20.7%). Accuracy: 79%; average precision: 78%; average **recall** rate: 76%.

	Amazon Forest	Atlantic Forest	Cerrado	Chaco	SDTF
Amazon Forest	1021	8	66	0	2
Atlantic Forest	7	1281	209	11	58
Cerrado	79	179	439	1	62
Chaco	0	22	0	89	5
SDTF	2	60	75	5	422

Table S2.3: Confusion matrix between sites categorised based on floristic composition via hierarchical clustering (rows) and sites categorised using climate and a classification tree approach (columns), for all non-geographically overlapping sites (those with centres >10 km apart). The diagonal gives the number of sites that are correctly classified by climate, while the off-diagonal elements give mis-classifications (20.3%). Accuracy: 79%; average precision: 79%; average recall rate: 77%.

	Amazon Forest	Atlantic Forest	Cerrado	Chaco	SDTF
Amazon Forest	812	6	63	0	1
Atlantic Forest	6	1051	158	7	44
Cerrado	66	137	361	1	56
Chaco	0	14	1	78	4
SDTF	2	42	64	4	353

Table S2.4: Confusion matrix between sites categorised based on floristic composition via hierarchical clustering (rows) and sites categorised using climate + soil and a classification tree approach (columns), for all sites including ones identified as transitional via a silhouette analysis. The diagonal gives the number of sites that are correctly classified by climate + soil, while the off-diagonal elements give mis-classifications (17.7%). Accuracy: 82%; average precision: 81%; average recall rate: 80%.

	Amazon Forest	Atlantic Forest	Cerrado	Chaco	SDTF
Amazon Forest	1038	6	51	0	2
Atlantic Forest	9	1317	174	10	56
Cerrado	69	136	500	1	54
Chaco	0	22	0	87	7
SDTF	3	57	64	4	436

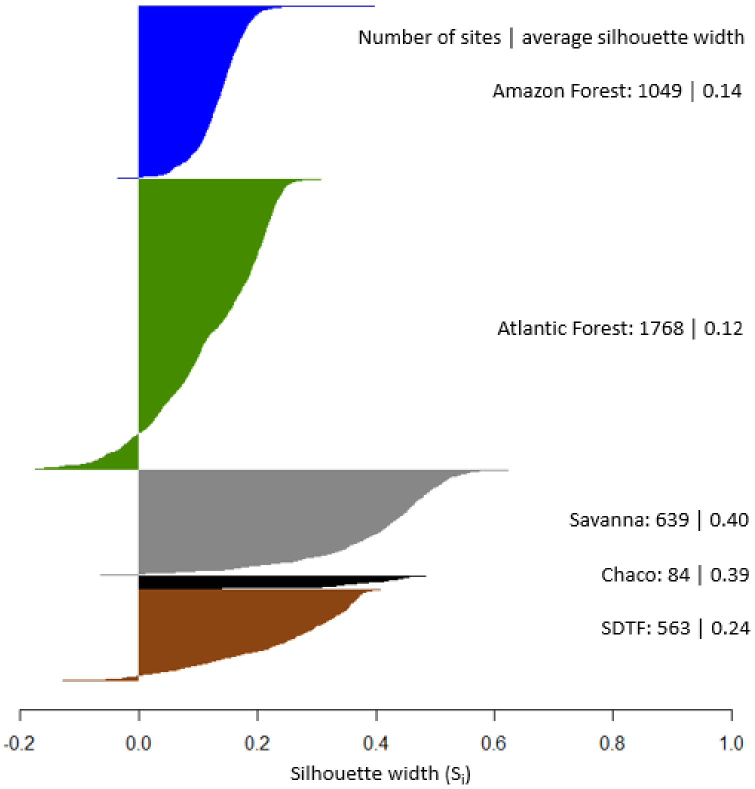

Figure S2.1 – Silhouette plot with all 4103 NeoTropTree sites included in the cluster analysis. Positive Silhouette width values (S_i) indicate that a site is indeed most similar, in terms of tree species composition, to the other sites in the biome it has been assigned to, whereas negative values indicate that a given site is compositionally more similar to one of the other biomes delimited through the cluster analysis than it is to the biome with which it clustered. The plot also presents the number of sites that compose each biome and their average silhouette width value (S_i).

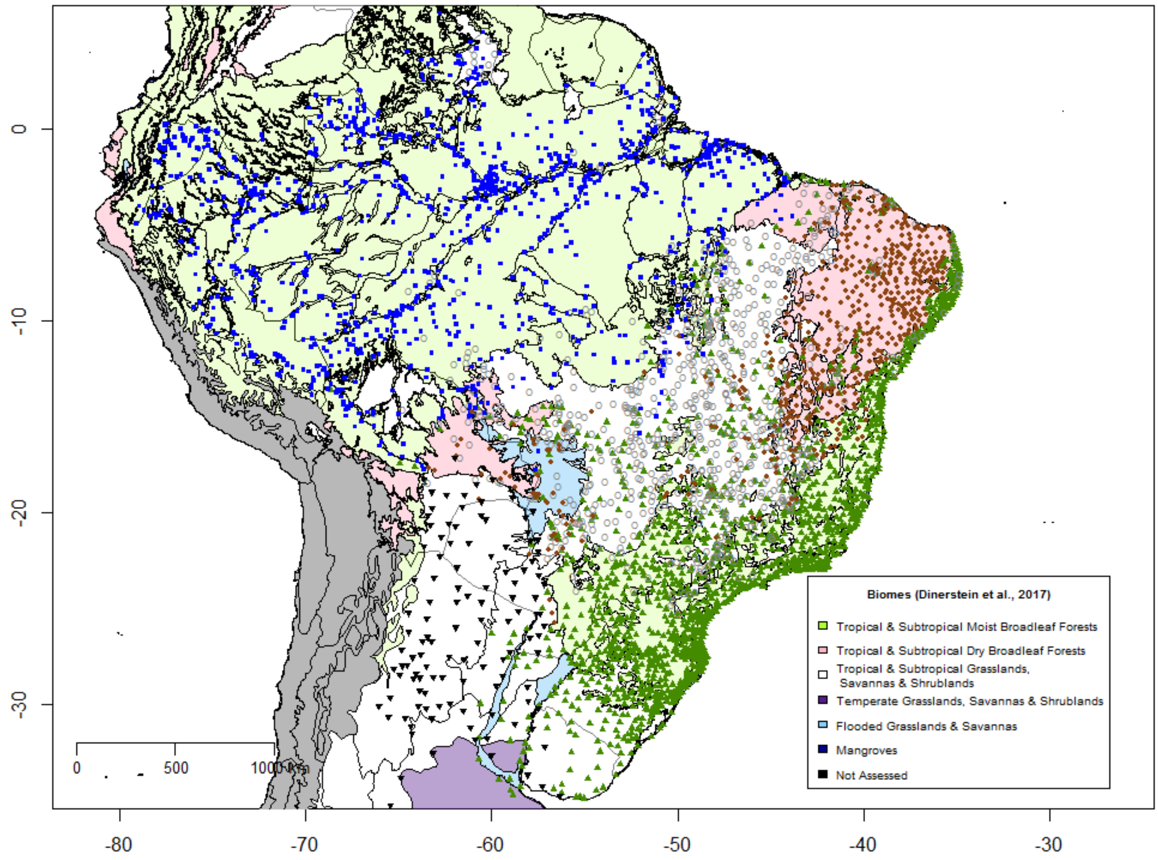

Figure S2.2 – Map of South America with areas coloured according to Dinerstein *et al.,* (2017), which combines ecoregions into biomes and is a reviewed and updated version of Olson *et. al.* (2001). The points on the map are the NeoTropTree tree species inventory sites classified into biomes by this study: Atlantic Forest (green triangles), Seasonally Dry Tropical Forest (brown circles), Savanna (hollow gray circles), Amazon Forest (blue squares), and Chaco (inverted hollow black triangles)

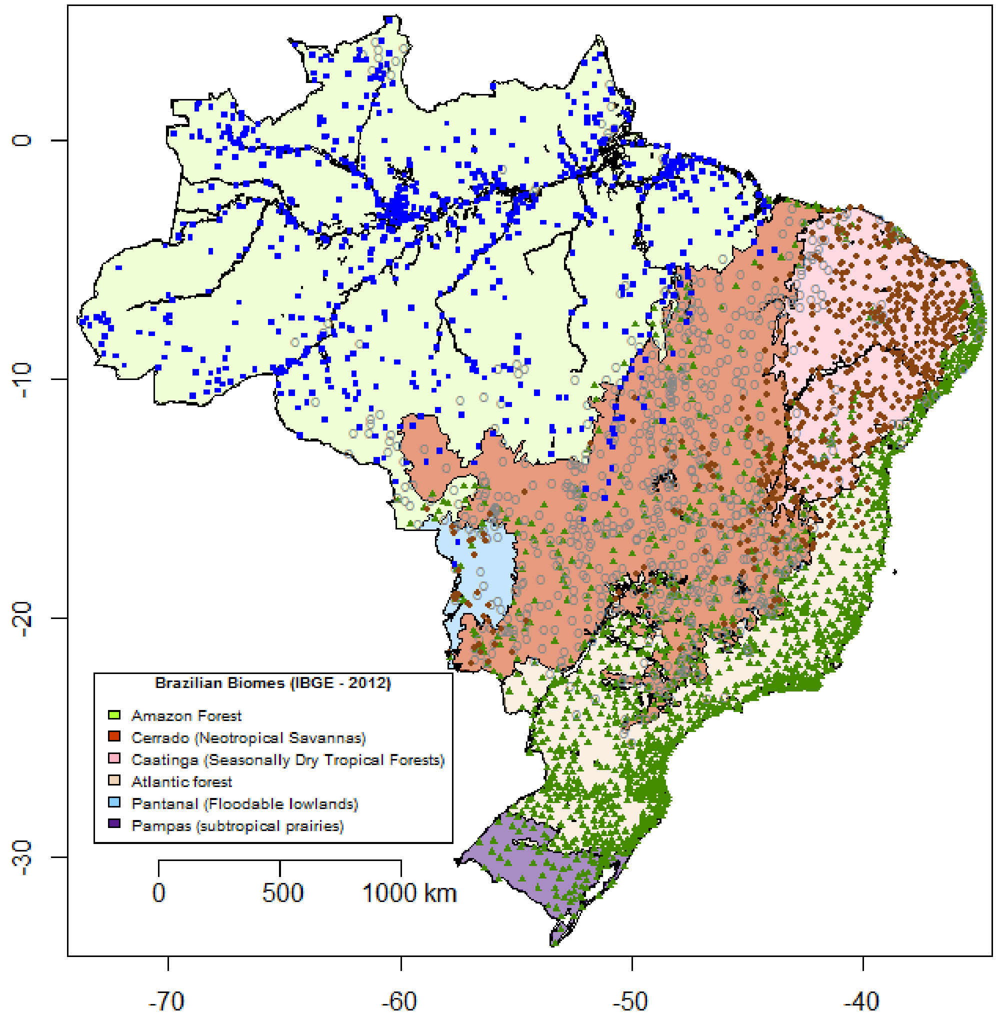

Figure S2.3 – Map of Brazil with areas coloured according to the Domain system of IBGE (2012), which are also sometimes referred to as biomes. The points on the map are the Brazilian NeoTropTree tree species inventory sites classified into biomes by this study: Atlantic Forest (green triangles), Seasonally Dry Tropical Forest (brown circles), Savanna (hollow gray circles), Amazon Forest (blue squares), and Chaco (inverted hollow black triangles).

Table S2.5: Confusion matrix between sites categorised according to Dinerstein *et al.* (2017) biome classification system (rows), which was adapted from Olson *et al.*(2001), and sites as categorised into biomes in this study (columns). The underlined numbers represents the sites that were assigned to matching categories between the two systems. The other elements are treated here as mis-classifications.

	Existing Lowland Tropical South America Biomes according to tree species composition					
Dinerstein *et al.* (2017) - Adapted from Olson *et al.* (2001)	Amazon Forest	Atlantic Forest	Savanna	Chaco	SDTF	Total
Tropical and subtropical moist broadleaf forests (tropical and subtropical humid)	994	994	117	0	41	2146
Tropical and subtropical grasslands, savannas and shrublands (tropical and subtropical semiarid)	57	403	544	99	90	1193
Tropical and subtropical dry broadleaf forests (tropical and subtropical semihumid)	21	88	78	1	420	608
Flooded grasslands and savannas (temperate to tropical fresh or brackish water inundated)	1	18	16	4	15	54
Temperate grasslands, savannas and shrublands (temperate semiarid)	0	10	0	12	0	22
Mangrove (subtropical and tropical salt water inundated)	24	51	5	0	0	80

2.2

Table S2.6: Confusion matrix between sites categorised according to IBGE (2012) biome/phytogeographic domain classification system (rows), which was adapted from Veloso (1992), and sites as categorised into biomes in this study (columns). The underlined numbers represents the sites that were assigned to matching categories between the two systems. The other elements are treated here as mis-classifications.

	Existing Lowland Tropical South America Biomes according to tree species composition					
Brazilian biomes	Amazon Forest	Atlantic Forest	Savanna	SDTF	Chaco	Total
Amazon Forest	635	20	74	1	0	730
Atlantic Forest	0	911	37	31	0	979
Cerrado	35	287	539	114	2	977
Caatinga	0	57	54	369	0	480
Pampa	0	78	0	0	1	79
Pantanal	2	9	21	18	1	51
Continental water mass	85	21	4	10	1	121
Oceanic water mass	6	89	2	2	0	99

3

Environmental controls of biome distribution in Bolivia and Brazil – dissecting the relative importance of water availability, soil fertility and fire

3 - Environmental controls of biome distribution in Bolivia and Brazil – dissecting the relative importance of water availability, soil fertility and fire

3.1 - Abstract

Background:

The distribution of biomes worldwide is often linked to climate. In the tropics, areas with the same climate can support different biomes, demonstrating that factors other than climate, including edaphic conditions and disturbance (such as fire) also shape the limits of biomes. For example, the roles of climate (precipitation) and disturbance in the dynamic transition between savannas and forests have been studied and described. However, the role of edaphic factors in shaping biome distribution has not been properly assessed due to lack of appropriate data and forests are usually treated as a single biome whereas they need to be differentiated into moist forests and seasonally dry tropical forests (SDTFs).

Aims:

I sought to determine the environmental controls distinguishing the limits of moist forest, savanna and SDTF biomes in Bolivia and eastern Brazil – two regions of high biome heterogeneity in lowland tropical South America (LTSA).

Methods:

Soil data were collected for 182 sites across three biomes: moist forest, savanna and SDTF. Edaphic variables collected were soil pH, soil texture and concentrations of acidic cations (H, Al), total phosphorous and nitrogen, and base cations. Water availability was assessed as mean annual precipitation and drought probability, while fire regime was quantified by fire return interval derived from MODIS data. Structural equation modelling (SEM) was then used to statistically assess controls on biome limits.

Results:

I found that environmental controls behind biome distribution differ between the two studied regions in LTSA. In Bolivia, water availability is the main driver behind the distribution of moist forests versus savannas. In eastern Brazil, this transition is controlled by both water availability and soil fertility. In Bolivia, the environmental spaces occupied by savannas and SDTF are statistically indistinguishable. In eastern Brazil, the transition between savannas and SDTFs is linked to soil fertility, water availability and fire regime. The only significant driver involved in the transition between SDTF and moist forests in both Brazil and Bolivia is water availability.

Conclusions:

I found that water availability, soil fertility and fire are all important correlates of biome limits, but that their relative importance varies between regions and among biome transitions. For example, different environmental drivers underlie the transition from savanna to moist forest versus to SDTF. Importantly, environmental controls of a given biome transition can differ between Brazil and Bolivia. Overall, water availability is the most important environmental correlate of biome distributions in our study region. Even so, edaphic conditions in terms of soil fertility are important in determining biome transitions over large geographic scales and should not be disregarded in tropical biogeography and in predicting the impacts of environmental change.

3.2 - Introduction

The distribution of biomes has commonly been linked to climate (e.g. Humboldt 1816; Holdridge 1947; Whittaker 1975). Temperature and rainfall, alongside light availability, determines the length of a growing season for example. Temperature extremes (both high and low), temperature seasonality, and the occurrence of frost can all limit tree species establishment and growth (Sullivan et al. 2017), and in turn are also considered to limit biomes (Lehmann et al. 2011; DRYFLOR 2016). Meanwhile, the frequency, length and severity of drought events will filter species by their tolerance of such extremes (Vicente-Serrano et al. 2013). Previously, variation in mean annual precipitation has been linked, in the tropics, to the division between wetter biomes (namely, moist forests) and drier biomes (savannas and SDTFs) (Holdridge 1947; Whittaker 1975; Odum 1992). However, the transition between savanna and SDTF is not well understood, and has been considered to be a product of rainfall amount, soil fertility, and fire. Indeed, savannas are formed by fire-prone vegetation and SDTFs harbour fire-intolerant plant lineages, such as the Cactaceae and genera within Euphorbiaceae (Lavin et al. 2004; Schrire, Lavin, and Lewis 2005). Because of the latter taxa, SDTFs are also known as the succulent biome (Schrire et al. 2005).

Different studies have been able to show, however, that biome distributions cannot be fully accounted for by climate alone, and this has been exemplified by the studies of Hirota et al. (2011), Staver, Archibald, and Levin (2011), Moncrieff, Hickler, and Higgins (2015), Dantas et al. (2016). These studies found that, under the same climate, various parts of the tropics can be covered by different biomes. For example, in South America's Dry Diagonal (Vanzolini 1963), a belt of dry biomes separating the moist Atlantic and Amazon forests, patches of dry and semideciduous forests can be encountered interspersed within the predominant savanna vegetation (Prado and Gibbs 1993; Pennington, Prado, and Pendry 2000; Neves et al. 2015; Bueno et al. 2018). This example and others lead to the conclusion that other environmental factors, such as edaphic conditions and disturbance are also linked to biome distribution in the tropics, especially in South America (e.g. Silva de Miranda et al., 2018; Dexter et al. 2018).

In this region, numerous soil characteristics, including water-holding capacity, nutrient content, pH and soil structure, have been put forth as possible determinants of vegetation cover and biome limits. In a series of local-scale studies in Neotropical savannas, Furley and Ratter (1988) described savannas as having deep, well-structured, nutrient-poor and acidic soils with high aluminium content. Here, savanna plants have physiological adaptations to cope with aluminium toxicity, which is necessary because of the high concentrations in these Brazilian savanna soils (Arens, 1963; Goodland & Pollard, 1973; Dezzeo et al, 2004). Further, trees in this region tend to have deep root systems capable of exploring deeper soil layers, considered an adaptation to tolerating an extended annual dry season (Oliveira et al. 2005). On the other hand, Neotropical dry forests (SDTFs) are often described as growing on shallow, poorly structured soils, or even on exposed substrate, such as limestone (Velloso, Sampaio, and Pareyn 2001). While such soils are considered nutrient-rich, with pH close to neutral, they have poor water-holding capacity due to the proportion of sand and lack of soil aggregates(Velloso, Sampaio, and Pareyn 2001; Santos et al. 2012; Moro et al. 2016).

Disturbance regime – principally via fire and mammalian herbivory – is known to impact biome distributions in the tropics (Whittaker and Levin 1977; Dublin, Sinclair, and McGlade 1990; Warman and Moles 2009; Staver, Archibald, and Levin 2011). In the Neotropics, fire is the main form of disturbance (due to the loss of the megafauna many tens of thousands of years ago) and its role in driving vegetation change has been well studied (e.g. Durigan and Ratter 2006, 2016). Fire reduces biomass in an ecosystem via consumption and top-kill of plants, especially trees, and favouring the occupation of grasses and herbs that can rapidly regrow lost biomass (Lehmann et al. 2011, 2014). However, the absence or scarcity of fire can facilitate increased tree growth and favour forest establishment. Under such circumstances trees are more likely to escape the “fire-trap” where after reaching a certain height and biomass threshold, most trees will be able to survive fire events (Higgins, Bond, and Trollope 2000; Hoffmann et al. 2012). Because of this, fire return interval (amount of time between fire events) is a better predictor of biome identity than other fire-related variables (e.g fire intensity). This is probably because in longer intervals without any fire events, the higher the chances of saplings and treelets growing sufficiently in height and to escape the fire-trap (Hoffman et al, 2012).

Previous studies exploring biome transitions in the tropics have centred broadly on savanna and forest (e.g. Staver et al. 2011; Dantas et al. 2016), without distinguishing between different forested biomes. Here, I seek to bring a more nuanced and biologically meaningful understanding of biome limits in south America by distinguishing tropical dry forests, also known as Seasonally Dry Tropical Forests (SDTFs) in the Neotropics, and moist forests as distinct entities (Pennington, Prado, and Pendry 2000; Pennington, Lavin, and Oliveira-Filho 2009) as well as integrating detailed soils information. In the Neotropics, SDTFs cover an extensive area where soils are more fertile and climate more seasonal than where savannas occur (Pennington, Prado, and Pendry 2000; Ratter, Bridgewater, and Ribeiro 2003; Dexter et al. 2015; Neves et al. 2015). Importantly, biomass accumulation dynamics in SDTF and relationships with fire events are different to that observed in moist forests, especially during early-successional stages (Rozendaal et al., 2017). These differences in environmental drivers between moist forests and SDTFs will most likely lead to key differences in biome distribution and transition dynamics.

My main objective with this research is to investigate the roles of climatic, edaphic and fire-related conditions on the distribution of savannas, moist forests and SDTFs across lowland Tropical South America (LTSA), through an analytical framework including all three biome states and their environmental correlates. Here, I hypothesize that different environmental drivers underlie the transition between each pair of biomes considered: savannas and SDTFs, savannas and moist forests, and moist forests and SDTFs.

We therefore aim to test quantitatively the following predictions:

a) Water availability is the main driver behind the division between moist forests and the drier biomes (SDTFs and savannas)

b) Edaphic factors will be the main driver behind the division between SDTFs and savannas.

c) Fire will be related to the distribution of savannas.

These predictions suggest that moist forests will occupy areas with higher mean annual precipitation and less precipitation seasonality, whereas savannas and SDTFs will occupy areas that are drier and more seasonal. In relation to soil, SDTFs will tend to grow on areas with more nutrient-rich, alkaline or pH neutral soils and savannas will grow on soils with lower pH and a higher quantity of aluminium. A more frequent fire return interval will be correlated with savanna occurrence.

3.3 - Methods

3.3.1 - Area of Study

Lowland tropical South America encompasses three main biomes: savanna, SDTF and moist forest. Their mains zones of contact are located in Bolivia and Brazil (Dinerstein et al., 2017; Silva de Miranda et al., 2018). The moist forest biome can be further divided into two main blocks: the Amazon and Atlantic Forests (IBGE 2012, Silva de Miranda et al. 2018). The former occupies the majority of LTSA to the east of the Andes and the latter the Brazilian coast line and portions of the interior Brazilian states of Minas Gerais (MG) and Sao Paulo (SP) (Atlantic Forest *sensu-latissimo* of Oliveira-Filho & Fontes, 2000). Savannas is found mainly in central Brazil and southeast Bolivia but also in Venezuela (Eva et al., 2004; Huber, 2006), Colombia (the

Llanos) and in patches in the Amazon Forest (Barbosa et al., 2005). The STDFs are distributed across the diagonal of dry formations (Prado & Gibbs, 1993; Vanzolini, 1963) – a diagonal composed of dry biomes starting in northeast Brazil and ending in northern Argentina. It is in the northeast of this diagonal that South America's SDTFs' largest nucleus, the Brazilian Caatinga, can be found (DRYFLOR et al., 2016; Pennington et al., 2000; Veloso et al., 1991). Outside of the Caatinga, patches of SDTF can also be encountered among the savannas in regions of more fertile soils and also in the Chiquitania region of Bolivia (Navarro 2011), the north coast of Colombia and Venezuela and in inter-Andean dry valleys (DRYFLOR, 2016).

My study focuses on two portions of LTSA with high biome heterogeneity – 1) eastern Brazil: which covers southeast Brazil and parts of the Brazilian states of Bahia, Goiás and Tocantins; and 2) Bolivia: which covers most of tropical lowland Bolivia, but also neighbouring portions of southern Peru and the Brazilian states of Acre and Mato Grosso do Sul (Fig. 3.1). In both regions, a mix of savannas and moist and dry forests can be encountered. In eastern Brazil, the moist forest biome is entirely Atlantic Forest and is located mainly along the Atlantic coast and in the state of Minas Gerais. SDTF can be found in Bahia, the north of Minas Gerais and the northeast of Goiás. Savannas are encountered mainly in Goiás and Tocantins, but patches can also be found in Minas Gerais and Bahia. In the second nucleus (Bolivia), the moist forest biome is largely Amazonian. This biome is mainly located in the Bolivian departments of Santa Cruz and Beni. SDTF are mainly found in the department of Santa Cruz, in the Chiquitania region, and can also be found in the west of Mato Grosso do Sul (Brazil). In this nucleus, savannas can be found mainly in the Santa Cruz department (Fig. 3.1).

3.3.2 - The datasets

3.3.2.1 - The sites

I have selected 182 sites from two different databases: NeoTropTree (available at http://www.neotroptree.info/) and ForestPlots (available at http://www.forestplots.net/) (Fig. 3.1), 98 in eastern Brazil and 84 in Bolivia. All 182 sites were classified into one of three biomes: moist tropical forest, SDTF and savanna. Geographically, NeoTropTree sites are circles with 10km of diameter.

3.3.2.2 - Soil samples:

The Harmonized Soils and SoilGrids databases were assembled through interpolaton procedures in order to ensure a full grid coverage of all landmasses. These interpolation methods uses geographic points where soil properties have been formely assessed in order to estimate values at locations where these properties have not yet been studied. Because of this, most of the values presented by these databases are mathematical products and not laboratory obtained values. Therefore, I have opted to work only with field-collected soil samples.

Collaborators collected soil samples from the eastern Brazil nucleus (98 samples) from June to September 2014 and I collected samples from Bolivia (48 samples) from June to September 2015. Data from Acre (Brazil) and southern Peru (36 samples) come from Quesada et al. (2008; 2009, 2010, 2011). The first 30cm of soil were collected at five different points at a site and then mixed together into a single sample per site. Loose leaf litter were removed before collecting the soil samples. All samples were analysed at the Instituto Nacional de Pesquisas na Amazonia (INPA – Brazil) to ensure protocols employed for sample analysis would be consistent. Samples were analysed for multiple metrics. These were: 1) soil texture (percentage of sand, silt and clay); 2) nutrient content (K, P, Ca and Mg – in cmol/Kg - measured at soil pH using the silver thiourea method – Ag-TU; Pleysier and Juno, 1980); 3) Concentration of acidic cations (H and Al expressed in cmol/Kg, same protocol as used for nutrient content), pH (measured in water); 4) Measurements of Ca, K and Mg (all in cmol/Kg of soil) were used to calculate sum of bases (Na was not included). Details of all protocols for these measurements can be found in Quesada et al (2008, 2009, 2010, 2011).

Within statistical analyses presented here, the soil attributes of pH, sum of base cations, total phosphorous, proportion of sand and concentration of acidic cations (H and Al) were included. The proportion of the three main soil fractions – sand, silt and clay – are correlated, so only one requires inclusion.

3.3.2.3 - Other environmental variables:

I obtained estimates of mean annual precipitation (MAP) and mean annual temperature (MAT) from the Bioclim database (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005). MAP and MAT were used to construct Whittaker plots (Whittaker, 1975) and explore biome distributions in climatic space (Fig. 3.2).

Foley's Drought Index (FDI - Foley, 1957) sums cumulative precipitation anomalies that occurred over a specified timeframe by quantifying monthly precipitation deviations from monthly long-term means. Because of this, it serves as a proxy for interannual precipitation variability and precipitation predictability (Keyantash & Dracup, 2002). FDI values were extracted from the world-wide drought layer produced by Lehmann et al. (2014) and calculated using monthly rainfall values from 1901 to 2002. For each 0.5 by 0.5 degree grid cell, monthly rainfall deficit was calculated via the following equation: (for each month) actual annual rainfall for three years before every month less the expected (long-term average) rainfall for that period, divided by the MAP; this value is than divided by the mean annual precipitation. The data layer registers the largest monthly rainfall deficit identified in the specified timeframe (1901 to 2002).

World-wide fire return interval (FRI) data were estimated by Archibald et al. (2013) by fitting 10 year long fire datasets derived from MODIS data (Roy et al. 2008) to a Weibull distribution (further details are available in Archibald et al. 2013).

Values of MAP, MAT, FDI and FRI were extracted for the central latitude and longitude of each site included in our study. Also, given no differences in mean annual temperature could be identified among biomes and the regions sampled (Fig. 3.2), and with no sites above 1000m elevation (i.e., to exclude areas that may experience frost), temperature was not considered further in these analyses. As can be observed in the early works of Holdridge (1947), Whittaker (1975) and Odum (1992), tropical biomes occur in the same temperature range and considered precipitation regime as the main climatic driver of differentiation among biomes.

3.3.3 - Statistical Analyses:

3.3.3.1 - Wilcoxon signed rank tests

A series of pairwise Wilcoxon signed rank tests were run on all environmental variables to assess differences in environmental means among biomes in each region. Wilcoxon signed rank tests are a non-parametric equivalent of ANOVA. Further, boxplots were associated with the pairwise Wilcoxon signed-rank tests. Alpha significance value was initially set to 0.05 and corrected for multiple testing using Bonferroni's correction (alpha/number of consecutive tests) with the significance level then set at 0.016. Both box-plots and Wilcoxon tests were executed using the R Statistical Software v3.5 using the "ggpubr" package (Kassambara, 2016).

3.3.3.2 - Random Forest Analysis

Random Forest is a non-parametric statistical approach where the predictors are used to classify the response variable via meaningful splits in the distributions of the predictor variables. In this approach, the classes are defined *a-priori* and discontinuities identified taking this into account. Here, environmental variables were predictors and our classification scheme consisted of biomes (described above). Analyses generate importance values for all predictors, which are a measure of their relative importance in classification (the number of times they were employed to classify the data). Importance values were assessed using Breiman's measure of importance (Breiman 2001). The analysis was performed in R (v. 3.5) using the "randomForest" package (Liaw and Wiener 2002).

3.3.3.3 - Structural Equation Models (SEMs)

Environmental controls interact synergistically not individually. Joint interaction of environmental controls determine biome identity and distribution. As such, I used a workflow able to represent and quantify such joint effects and account for the impact of controls on one another. Structural Equation Models (SEMs) are multivariate, hierarchical statistical models based on linear or non-linear regressions (Grace, 2006; Grace, Anderson, Olff, & Scheiner, 2010). SEMs are constructed around the covariance structure of directly measured indicators (observed variables). Through these structures, SEMs derive latent variables – mathematical constructs representing complex or abstract variables that cannot be directly measured, but whose structure can be computed by observing how the indicators making up a given latent variable co-vary (Grace, Anderson, Olff, & Scheiner, 2010).

The construction of these models consists of establishing association and/or causation links among latent variables based on *a-priori* concepts of how the studied systems function (Grace et al., 2006). To that end, I created a base model (Fig. 3.3) around three latent variables representing classes of environmental controls hypothesised to underpin biome distribution: soil fertility, water availability and fire. Soil fertility is indicated by six different indicator variables – pH, quantity of sand (%), phosphorous (cmol/kg), sum of bases (sum of quantities of calcium, magnesium and potassium, cmol/kg), nitrogen (cmol/kg) and aluminium (acidic cations – H, Al. cmol/kg). Water availability combines climatic and edaphic factors and was composed of three indicators – Mean Annual Precipitation (mm), Foley's Drought Index and quantity of sand (%). Fire regime has a single indicator variable, fire return interval (FRI – quantified as expected years between fires). When constructing latent variables, all indicators need to vary in the same direction. To this end, I have inverted the signs for both aluminium and quantity of sand (for the latter, the sign was inverted for both soil fertility and water availability).

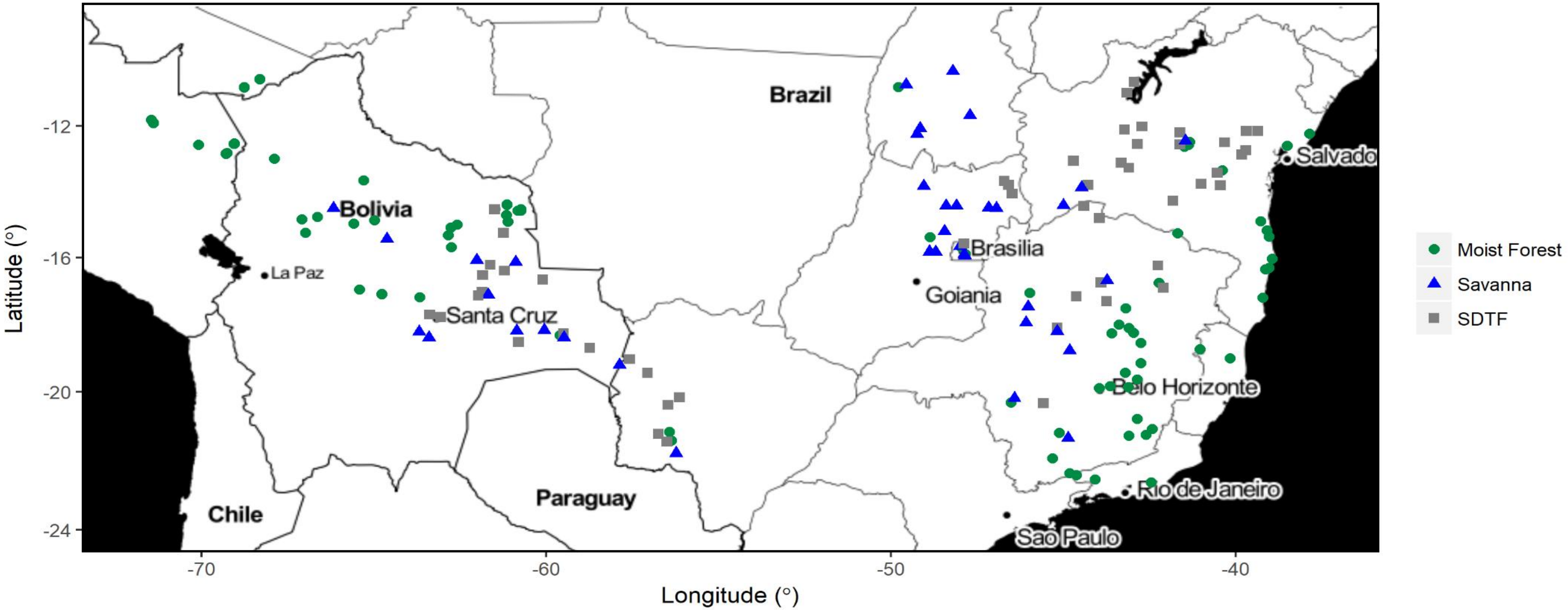

Figure 3.1: Map of South America with both transitional regions studied in the present contribution – 1) Bolivia and 2) eastern Brazil. Moist Forest, Savanna and Seasonally Dry Tropical Forest can be found in both regions.

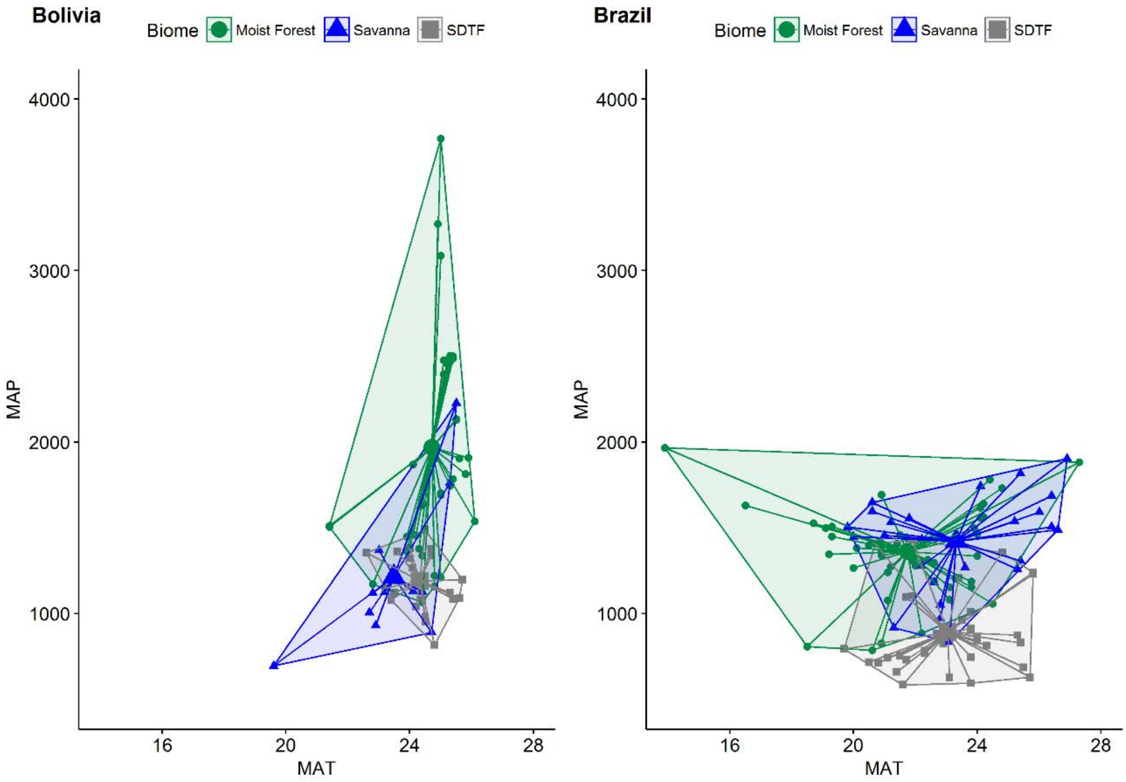

Figure 3.2: Whittaker plots displaying the climatic spaces of Moist Forests, Savannas and SDTFs both in Bolivia and eastern Brazil. (MAT – Mean Annual Temperature; MAP – Mean Annual Precipitation).

I conducted the SEMs in a pairwise fashion for biomes resulting in three different models for both the eastern Brazil and Bolivia nuclei. The models are: savanna versus moist forest; savanna versus SDTF; and SDTF versus moist forest. The three latent exogenous (predictor) variables and the latent endogenous (response) variable – biome identity – form the core of the SEMs. Our model consists of five regressions in total, three of which aim to estimate the relationship between the exogenous variables (predictors) and the endogenous variable. The other two regressions are to account for the effect of water availability over the other two exogenous variables. Water dynamics in these systems will most likely have an impact on the quantity of nutrients available for plants in the soils (since high rainfall will leach nutrients out of the soils) and on biomass accumulation and combustibility (Lehmann et al. 2014).

SEMs are based on covariance structure, therefore all variables were scaled to avoid model/term inflation. The quantity of sand, MAP and FRI were log transformed (natural logarithm), and the standard deviations of all variables were subsequently checked to ensure they were the same order of magnitude. Latent variables are mathematical constructs based on similarities among indicators (observed variables). The correlation structure of indicators comprising each latent variables was investigated to ensure correlations among indicators did not exceed 0.7, which can be problematic for inference (Grace et al. 2010).

Model fitting was performed and assessed in a systematic fashion following the recommendations made by Grace (2006, 2010) and Shipley (2002) and the analytical workflows proposed by Capmourteres & Anand (2016) and Leitão et al. (2018). The model represented in Figure 3.3, the base model, is the starting model for each of the pairwise comparisons. After the first run, I removed non-significant factor loadings (non-significant indicators / observed variables, $p > 0.05$) or those with a factor charge (measure of how important an indicator is to the latent variable) lower than 0.65 or higher than 1. Below 0.65, indicators do not contribute much to the latent variables and, above 1, they highly co-vary with other significant indicators (Grace 2006, 2010; Shipley 2002). I removed one non-significant loading at a time and ran the model again. This procedure was repeated until only significant indicators with a charge between 0.65 and 1 were present. Because our response variable is categorical, I used Robust estimators, which corrects estimations when using non-continuous variables (Savalei 2014).

To test if a proposed model is significantly different from the data, studies usually rely on Chi-square, CFI (confirmatory fit index), RMSEA (root mean square error of approximation) and other more commonly employed measures of model fitness. These metrics require the data to be normally distributed, which is not the case here. Because of this, I have fit our models based on the Bollen-Stine bootstrap values (Bollen and Stine, 1992). Non-significant bootstrap values ($p > 0.05$) indicate that the proposed model does not differ significantly from the structure of the data (Bollen and Stine, 1992). The SEMs were constructed in R using the sem() function in the "lavaan" package (Oberski 2014).

3.4 - Results

3.4.1 - Wilcoxon signed rank tests – describing environmental space:

Eastern Brazil and Bolivia differ in their climatic, edaphic and fire environments. More so, each biome differs in its environmental niche. Bolivia has much wetter moist forests than eastern Brazil (Fig 3.2, Fig S3.1), but savannas and SDTFs in this region have very similar rainfall regimes (Fig 3.3, Fig S3.1). In eastern Brazil, the biome – environment relationships differ; savannas and moist forests have similar precipitation and SDTFs are drier than the other biomes (Fig S3.1). FDI does not show a similar trend. In Bolivia, SDTFs tend to have less inter-annual variation (have bigger negative values of FDI, on average) than SDTFs and savannas (Fig S3.2). On the other hand, eastern Brazil biomes differ in regards to FDI (Fig S3.2). Fire return interval values also differ among regions. In Bolivia, FRI differs between moist forests and SDTFs, but not between savannas and the forested biomes (Fig S3.3). In eastern Brazil, moist forests have higher FRIs than savannas, but not higher than dry forests (Fig S3.3). Finally, in eastern Brazil, there is no difference in FRI between SDTFs and the other two biomes.

Edaphic conditions also differ between regions and among biomes. In Bolivia, concentrations of acidic cations, including aluminium, are only different between moist forests and SDTFs, where they are higher in moist forests (Fig S3.4). In Brazil, acidic cation concentrations are lower in SDTFs than in the other two biomes (Fig S3.4). I detected differences in pH between moist forests and both drier biomes in Bolivia, but not between the latter two (Fig S3.5). Brazil shows the same pattern observed in Bolivia for pH (Fig S3.5). All three biomes in Bolivia do not differ in sum of bases, but in Brazil dry forests tend to have higher values for SB than moist forests and savannas (Fig S3.6). These last two biomes have similar levels of SB (Fig S3.6). Quantity of sand varies across biomes in Bolivia – moist forests tend to have less sandy soils than savannas and SDTFs (Fig S3.7). No difference in quantity of sand can be observed across the eastern Brazilian biomes (Fig S3.7). Apart from a difference between moist forests and savannas in Brazil, no other significant differences in nitrogen concentration can be found in eastern Brazil or Bolivia (Fig S3.8). Concerning phosphorus, Bolivian moist forests tends to have higher concentrations than Bolivian savannas (Fig S3.9). There are no other differences in concentration of phosphorus among biomes in Bolivia or eastern Brazil (Fig S3.9).

3.4.2 - Random forest models

In general, mean annual precipitation, Foley's drought index, quantity of sand, phosphorous and fire return interval were among the three most important variables (high importance values) across all models (Table 3.1 – the three highest importance values for each comparison are underlined). For moist forests and savannas in eastern Brazil, FDI, FRI and nitrogen were the most important variables used to classify our sites into biomes. For this same transition in Bolivia, MAP, sand and phosphorous were the most important. For moist forests and SDTFs in eastern Brazil, MAP, aluminium and FDI were the most important variables. A different pattern can be observed in Bolivia; in this region, MAP, FRI and sand were the three most important environmental drivers. The transition between savannas and dry forests in eastern Brazil is better described by MAP, FDI and sum of bases, whereas in Bolivia, this transition is better linked to pH, sand and phosphorus.

3.4.3 - Structural Equation Models

I produced a total of six pairwise SEMs across Brazil and Bolivia, all of which have non-significant Bollen-Stine bootstrap values (>0.05), indicating that all models are sufficient approximations to the environmental conditions encountered in these transition zones (Figs 3.4, 3.5). Environmental controls change from one biome transition to the other depending on the pair of biomes considered and the region being analyzed (Figs 3.4, 3.5), indicating that the effect the environment has over biome distribution changes across geographic space and is not uniform across LTSA. This also highlights that different environmental factors can underlie the same transition in different parts of LTSA.

3.4.3.1 - Savanna / Moist Forest transition

The transition between savannas and moist forests in Bolivia is driven by water availability alone (savannas are drier than moist forests. Fig 3.4a, r-square: 0.4), whereas in eastern Brazil, this transition is related to soil fertility, water availability and fire together (savannas tend to be drier, less fertile and burn more often than moist forests. Fig 3.5a, r-square: 0.296). There are also differences in how soil fertility was constructed in these two models – in eastern Brazil, it is linked to phosphorus whereas in Bolivia, it is linked to pH and sum of bases. Water availability has been indicated by MAP in both models. In terms of effect sizes, water availability has a bigger impact in Bolivia (0.522) than in eastern Brazil (0.266). Also, in eastern Brazil, soil fertility and water availability have approximately the same effect size on biome identity (soil fertility -> 0.288, water availability -> 0.266).

3.4.3.2 - Savanna / Seasonally Dry Tropical Forest transition

In the transition zone between savannas and SDTFs in eastern Brazil, soil fertility, water availability and fire were important, with savannas being less fertile, wetter and more prone to fire than SDTFs (Fig 3.5b, r-square: 0.751). However, we did not identify any significant environmental controls of this transition in Bolivia (Fig 3.4b, r-square: 0.08). In eastern Brazil, the soil fertility construct was defined by pH, sum of bases and nitrogen. How water availability was constructed also varies from one nucleus to the other. In eastern Brazil, it is defined by FDI and mean annual precipitation, whereas, in Bolivia, it is defined by FDI only. Effect sizes also vary from one nucleus to the other. In eastern Brazil, where all three drivers were considered significant, water availability has the largest effect (0.817), followed by soil fertility (0.416) and fire (0.185).

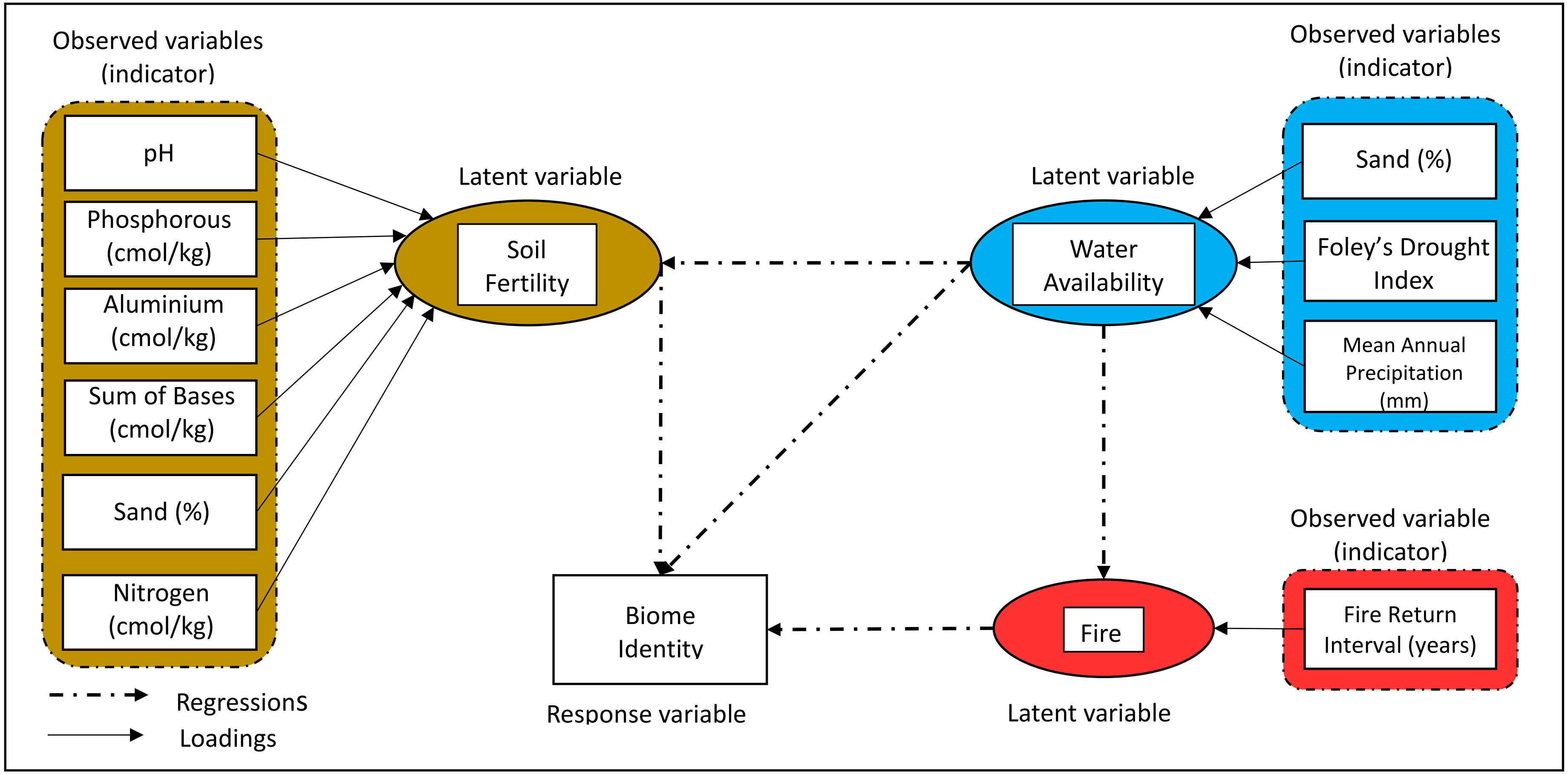

Figure 3.3: Diagram illustrating all hypotheses being tested within a Structure Equation Model framework. Three endogenous latent variables – Soil Fertility, Water Availability and Fire – are created through the interaction of indicator, or observed, variables and utilized as predictors of Biome Identity (exogenous latent, or response, variable).

3.4.3.3 - Moist Forest / Seasonally Dry Tropical Forest transition

The transition zones between moist forests and dry forests in both Bolivia (Fig 3.4c, r-square: 0.485) and eastern Brazil (Fig 3.5c, r-square: 0.629) are driven by water availability only, which has been defined by MAP in both cases. Even though soil fertility is not a significant driver in both scenarios, it has been defined by the interactions of pH and sum of bases in both cases. Water availability has similar effect sizes in both nuclei (eastern Brazil -> 0.669, Bolivia -> 0.623).

3.5 - Discussion

Transitions between SDTF and savanna, SDTF and moist forest, and savanna and moist forest differ in their underlying environmental drivers. Water availability, soil fertility and fire are all significant factors in the transition among these three biomes, but their significance and relative importance differs between eastern Brazil and Bolivia. I demonstrate important regional variation within a continent in the environmental limits of biomes; a factor commonly overlooked in broader scale studies of biome limits (e.g. Hirota et al., 2011; Staver et al., 2011; Dantas et al, 2016).

Among all drivers considered in this study, water availability– indicated by MAP in nearly all cases – was the most important, being a significant driver of biome identity in five of the six models analyzed (Fig 3.4a, b, c; Fig 3.5a, c). This result agrees with the long conceived notions of Humboldt (1816), Holdridge (1947), Whittaker (1975) and Odum (1992). As hypothesized, this environmental driver was the chief difference between moist forests and the drier biomes. However, contrary to our expectations, water availability was also the most important driver in the transition between SDTFs and savannas in eastern Brazil. This indicates that, even though these biomes overlap in geographic space and can be found under similar climatic conditions (Silva de Miranda et al., 2018), SDTFs are, on average, drier and more seasonal than savannas, at least in eastern Brazil. Soil fertility and fire were also considered significant drivers of biome identity and distribution (Fig 3.4a, b), but only in the eastern Brazil region. This again points to the existence of intra-continental variation in biomes' environmental limits.

3.5.1 - Savanna / Moist Forest transition

In both regions, the transition between savannas and moist forests was mainly driven by water availability (Fig. 3.4a, Fig 3.5a), with this driver being more important in Bolivia than in eastern Brazil. This finding is possibly related to how moist forests are circumscribed in both regions. Eastern Brazilian moist forest sites belong to the Brazilian Atlantic Forest, including semi-deciduous forests and gallery forests that occur in drier regions (Oliveira-Filho and Fontes 2000; IBGE 2012), whereas Bolivian moist forests are all Amazonian. This translates into Bolivian moist forests appearing to have higher water availability (Figure 3.1), as water availability was indicated only by MAP in both cases (Figs 3.4a, 3.5a). If measures of local hydrology were included in the study, I may have found equivalently high water availability for moist forests in eastern Brazil (e.g., for gallery forests). Another key difference between the two regions is the significance of soil fertility and fire. In eastern Brazil, soil fertility and fire were also important in distinguishing savanna from moist forest (Fig 3.5a), whereas in Bolivia, only water availability was significant.

Savannas are described as having, in average, less fertile soils that are more acidic and contain higher concentrations of aluminium relative to moist forests and SDTFs (Furley 1994; Ribeiro and Walter 2008; Assis et al. 2011). In fact, moist forests and SDTF tree species tend to be nutrient demanding due to their larger tree size and biomass and do not tend to occur in savanna environments because of this (Pellegrini et al. 2016). Also, due to aluminium toxicity, savanna plant species possess anatomical and ecophysiological adaptations to this cation's presence in the soil solution (Arens, 1963; Goodland & Pollard, 1973; Dezzeo et al, 2004), and such adaptations may not be present in moist forest and SDTF trees. Because of these two factors, moist forests tend to be encountered on more fertile soils (in relation to

Table 3.1: Variable Importance values obtained through a Random Forest. Variables not included in a given model are shown as (*). The pairwise models are the same ones represented in figures 3 and 4.

Variables	Moist Forest vs. SDTF		Savanna vs. SDTF		Moist Forest vs. Savanna	
	Brazil	Bolivia	Brazil	Bolivia	Brazil	Bolivia
Sum of Bases (cmol/Kg)	2.961	1.168	3.545	1.157	2.536	0.713
Sand (%)	2.174	3.250	0.949	2.171	3.080	4.406
Phosphorous (cmol/Kg)	1.747	1.007	1.310	1.967	1.897	1.591
Aluminium (cmol/Kg)	4.927	2.591	3.102	0.834	2.789	0.410
Nitrogen (cmol/Kg)	2.242	1.040	0.780	1.199	4.498	1.165
pH	3.723	2.873	1.834	2.089	3.151	1.085
Fire Return Interval (years)	3.015	4.077	1.737	1.163	4.717	1.259
Foley's Drought Index	4.637	2.341	5.804	1.088	7.611	1.020
Mean Annual Precipitation (mm/year)	12.192	9.796	9.543	1.343	2.283	4.061

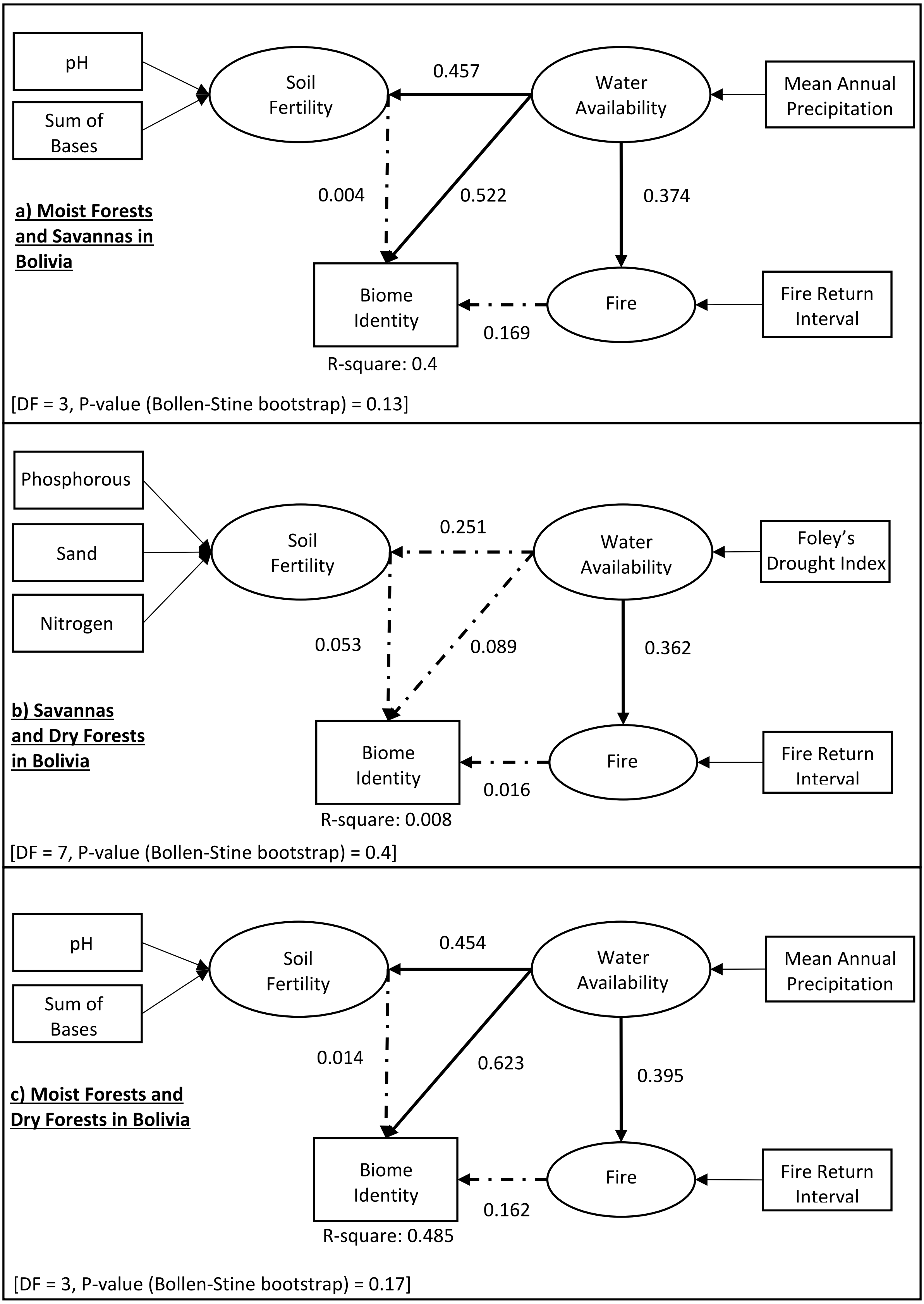

Figure 3.4: Schematic representation of all Bolivian models. Only significant indicator (observed variables) with factor loadings higher than 0.65 were considered. Path strength values are displayed for the regressions. Continuous arrows represent significant paths, dotted arrows represent non-significant paths. R-squared is the sum of all of the effects considered for a given endogenous (response) variable.

Savannas) with lower aluminium concentrations, so the lack of significance of soil fertility in Bolivia is an unexpected result.

There are two possible explanations for this: 1) the understanding of how savannas are related to edaphic conditions came from studies conducted in the Brazilian Cerrado (Brazilian central savannas) and this knowledge may not apply to other savannas in the Neotropics; 2) Most Bolivian savannas are mesotrophic and have similar levels of soil fertility to moist forests (Figure S3.6) and, therefore, soil fertility is not a significant driver of biome identity in this region. Soil fertility has been indicated differently in the models for the two regions. In Bolivia, it is the product of the interaction between pH and Sum of Bases. Bolivian savannas are on average less acidic than moist forests (Fig. S3.5). Also, there is no significant difference in means between these two biomes in relation to sum of bases, probably related to the ample variation in sum of bases in moist forests in Bolivia (Fig. S3.6).

3.5.2 - Savanna / Seasonally Dry Tropical Forest transition

Following Furley and Ratter (1988), Oliveira-Filho and Ratter (2002), Ratter, Bridgewater, and Ribeiro (2003) and Bueno et al. (2018), I have hypothesised that the transition between SDTFs and savannas would be influenced by soil fertility and that dry forests would tend to occupy more fertile soils, whereas savannas would occupy nutrient-poor soils. I was able to confirm this for the eastern Brazil nucleus, but not for the Bolivian nucleus. There are three possible explanations for this, all of which overlap to a certain degree: 1) the number of sites examined in Bolivia for both biomes was not sufficient to detect a significant relationship between biome identity and soil fertility; 2) soil fertility is not linked to savanna distribution in Bolivia, because the majority of the savannas in the Chiquitania are mesotrophic (Villareal et al., 2015), and savanna dynamics are linked to other environmental drivers or; 3) Bolivian savanna history may include greater human influence (Power et al. 2016; Veldman, 2016). The random forest models and the Wilcoxon signed rank tests confirm that indeed there is little to no difference between SDTF and savanna in Bolivia for the

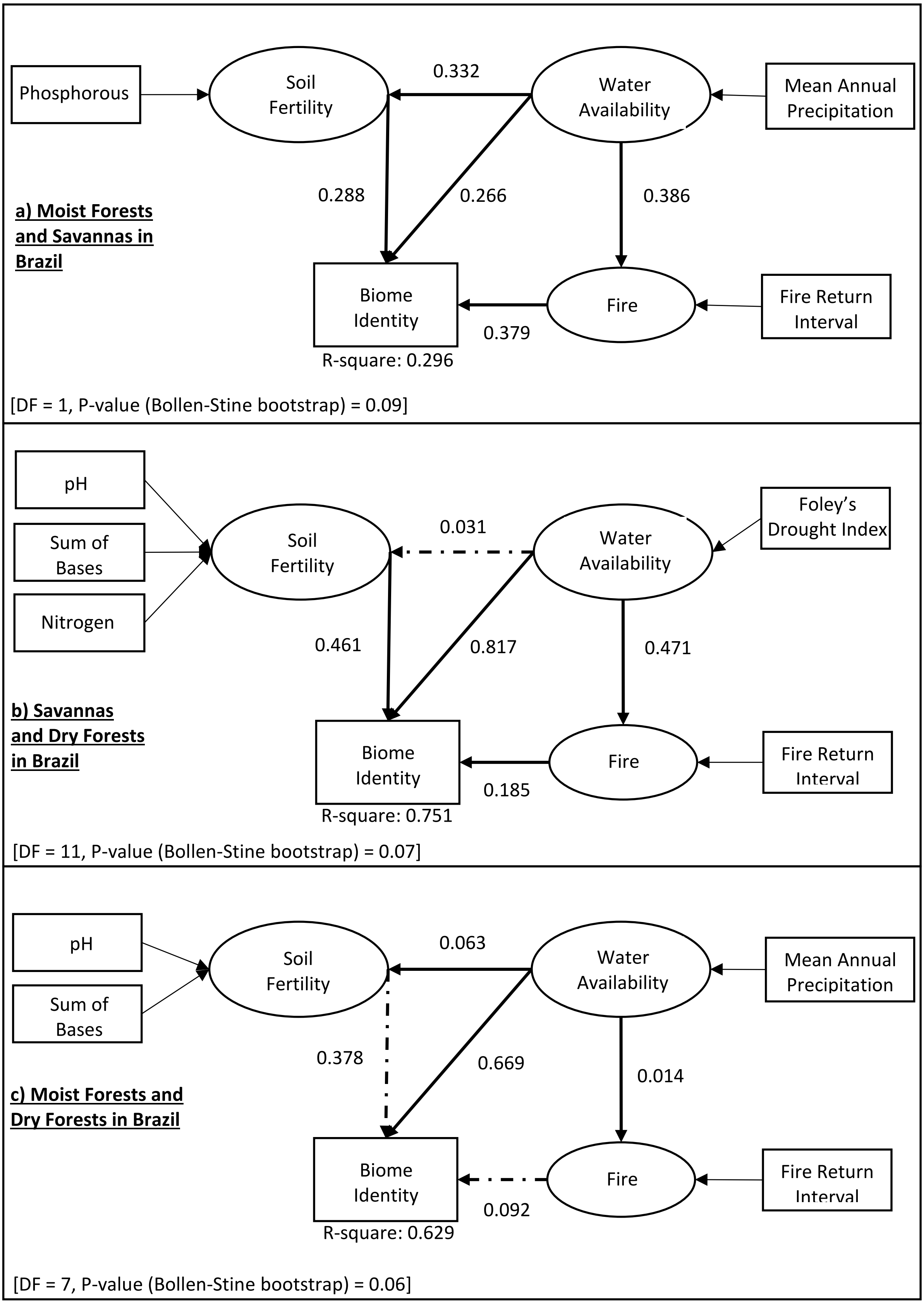

Figure 3.5: Schematic representation of all Brazilian models after fitting. Only significant indicator (observed variables) with factor loadings higher than 0.65 were considered. Path strength values are displayed for the regressions. Continuous arrows represent significant paths, doted arrows represent non-significant paths. R-squared is the sum of all of the effects considered for a given endogenous (response) variable.

variables considered, indicating that the transition between the two must be related to factors not included in our study.

In eastern Brazil, the transition between savannas and SDTFs show water availability as one environmental driver. Water availability was significantly indicated by FDI and MAP, suggesting that SDTFs are drier on average than savannas and experience greater inter-annual precipitation variability, a proxy for drought. SDTFs have been considered to be a transitional biome state between moist forests and savannas (e.g., Whittaker 1975, Malhi et al. 2009), with savannas being drier than SDTFs. However, our work and the research of Lavin et al. (2004) and Schrire et al. (2005) shows that SDTFs are actually drier and experience greater interannual variability in rainfall, on average, than savannas. Pennington et al. (2000, 2006) suggested that these two biomes share the same climatic space and the only driver of differentiation between them would be contrasting edaphic conditions. However, here we show that climatic conditions cannot be completely disregarded when studying this transition, which is aligned with the findings made by Silva de Miranda et al. (2018).

Another key difference between the Bolivian and eastern Brazilian nuclei is the importance of fire – detected as significantly important only in Brazil. SDTF plants lack the adaptations to survive fire. Also, SDTFs are usually slow-growing forests, since these environments will only be able to assimilate carbon during short periods of rain (two months only in some environments – Pennington et al, 2000). In addition, SDTFs' understory flora is not dominated by grass species, which limits the amount of flammable biomass this biome can accumulate and diminishes the possibilities of these forests burning.

3.5.3 - Moist Forest / Seasonally Dry Tropical Forest transition

In eastern Brazil and Bolivia, the transition between moist forests and SDTFs is solely driven by water availability, which has been significantly indicated by MAP in both nuclei. Not surprisingly, soil fertility and fire do not play a role in this transition, since both biomes are sensitive to fire events and have sufficient nutrients cycling in the system to enable tree species to grow and accumulate biomass. The fact that we have sampled mostly semideciduous forests in eastern Brazil and mostly evergreen forests in Bolivia does not seem to have an impact on this result, meaning that even drier portions of the moist forest biome are still wetter than SDTFs (as seen in Fig 3.2). How SDTFs have been defined and delimited in South America/Neotropics has changed during the last 30 years and a contentious area is whether semideciduous forests should be classified as being SDTF or as a part of the moist forest biome (Dexter et al. 2018).

By following Murphy and Lugo's (1995) physiognomic definition of SDTF and by examining Gentry's (1995) list of woody elements commonly found in these forests, Pennington et al. (2000) included the Atlantic's semideciduous forests in central Brazil in their delimitation of South America's SDTFs. In the same year, Oliveira-Filho & Fontes (2000) demonstrated how central Brazil's semideciduous forests are floristically linked to what these authors have called Atlantic Forest *senso-stricto* (moist forests forming a 300km wide band along the Brazilian Atlantic coast). Through extensive tree species checklists compilation, these authors were able to show that semideciduous forests are formed by tree species of the Atlantic Forest that are able to better cope with drier climatic conditions. This view is also aligned with that of (Scarano 2002), and the floristic links between semideciduous forests and the Atlantic Forest domain were also observed by Neves et al. (2017) and Silva de Miranda et al. (2018). Such similarities in tree species composition indicate a possible similarity in ecosystem function. These, in association with our findings in this contribution, suggests that semideciduous forests should be treated as a part of the moist forest biome.

3.6 - Conclusions

Soil fertility, water availability and fire are all important environmental drivers of biome identity in lowland tropical South America. I show the necessity of splitting forests into two biomes, moist forest and SDTF, in studies investigating biome environmental limits. These biomes and their transitions with surrounding biomes are linked to different environment controls. I show the importance of considering multiple biomes and environmental drivers in predictive models of biome cover. In eastern Brazil, all three drivers analysed have an effect on biome identity, whereas in Bolivia only water availability appears to have a significant effect. In eastern Brazil, I showed that the distribution of savannas (Cerrado) and SDTFs is linked to soil fertility and that aluminium concentration does not seem to have a major effect on biome distribution. Also, in eastern Brazil I show that fire has an impact on the transition between savannas and SDTFs. Meanwhile, I was unable to demonstrate that fire and soil fertility are significant drivers of biome identity in Bolivia. Differences in effect and importance of soil fertility between eastern Brazil and Bolivia suggest that generalizations taken from studies conducted in Brazil may not work in other regions of South America or other continents. My results suggest that the soil features at play in these transitions are geography-dependant and generalizations from single regions are difficult. In summary, I was able to show the complexity of LTSA environmental controls of biome distribution and how these controls change in importance across regions.

3.7 - References

Archibald, Sally, Caroline E. R. Lehmann, Jose L. Gómez-Dans, and Ross A. Bradstock. 2013. “Defining Pyromes and Global Syndromes of Fire Regimes.” *Proceedings of the National Academy of Sciences* 110 (16): 6442–47. https://doi.org/10.1073/pnas.1211466110.

Arens, K. 1963. “As Plantas Lenhosas Dos Campos Cerrados Como Vegetação Adaptada Às Deficiências Minerais Do Solo.” In *III Simpósio Sobre o Cerrado*, edited by Edgard Blutcher, 13–115. São Paulo: EDUSP.

Assis, Ana Carolina Cunha, Ricardo Marques Coelho, Eduardo da Silva Pinheiro, and Giselda Durigan. 2011. “Water Availability Determines Physiognomic Gradient in an Area of Low-Fertility Soils under Cerrado Vegetation.” *Plant Ecology* 212 (7): 1135–47. https://doi.org/10.1007/s11258-010-9893-8.

Banda, Karina, Alfonso Delgado-Salinas, Kyle G. Dexter, Reynaldo Linares-Palomino, Ary Oliveira-Filho, Darién Prado, Martin Pullan, Catalina Quintana, Ricarda Riina, and Gina M. Rodríguez. 2016. “Plant Diversity Patterns in Neotropical Dry Forests and Their Conservation Implications.” *Science* 353 (6306): 1383–1387.

Bollen, Kenneth A., and Robert A. Stine. 1992. “Bootstrapping Goodness-of-Fit Measures in Structural Equation Models.” *Sociological Methods & Research* 21 (2): 205–29. https://doi.org/10.1177/0049124192021002004.

Breiman, Leo. 2001. “Random Forests.” *Machine Learning* 45 (1): 5–32.

Bueno, Marcelo Leandro, Kyle G. Dexter, R. Toby Pennington, Vanessa Pontara, Danilo Mesquita Neves, James Alexander Ratter, and Ary Teixeira de Oliveira-Filho. 2018. “The Environmental Triangle of the Cerrado Domain: Ecological Factors Driving Shifts in Tree Species Composition between Forests and Savannas.” *Journal of Ecology* 106 (5): 2109–20. https://doi.org/10.1111/1365-2745.12969.

Capmourteres, Virginia, and Madhur Anand. 2016. “Assessing Ecological Integrity: A Multi-Scale Structural and Functional Approach Using Structural Equation Modeling.” *Ecological Indicators* 71 (December): 258–69. https://doi.org/10.1016/j.ecolind.2016.07.006.

Dantas, Vinícius de L., Marina Hirota, Rafael S. Oliveira, and Juli G. Pausas. 2016. “Disturbance Maintains Alternative Biome States.” Edited by Marcel Rejmanek. *Ecology Letters* 19 (1): 12–19. https://doi.org/10.1111/ele.12537.

Devisscher, Tahia, Liana O. Anderson, Luiz E. O. C. Aragão, Luis Galván, and Yadvinder Malhi. 2016. “Increased Wildfire Risk Driven by Climate and Development Interactions in the Bolivian Chiquitania, Southern Amazonia.” *PLoS ONE* 11 (9). https://doi.org/10.1371/journal.pone.0161323.

Dexter, K.G., B. Smart, C. Baldauf, T.R. Baker, M.P. Bessike Balinga, R.J.W. Brienen, S. Fauset, et al. 2015. “Floristics and Biogeography of Vegetation in Seasonally Dry Tropical Regions.” *International Forestry Review* 17 (2): 10–32. https://doi.org/10.1505/146554815815834859.

Dexter, Kyle G., R. Toby Pennington, Ary T. Oliveira-Filho, Marcelo L. Bueno, Silva de Miranda, Pedro L, and Danilo M. Neves. 2018. “Inserting Tropical Dry Forests Into the Discussion on Biome Transitions in the Tropics.” *Frontiers in Ecology and Evolution* 6. https://doi.org/10.3389/fevo.2018.00104.

Dryflor, Karina Banda-R, Alfonso Delgado-Salinas, Kyle G. Dexter, Reynaldo Linares-Palomino, Ary Oliveira-Filho, Darién Prado, et al. 2016. “Plant Diversity Patterns in Neotropical Dry Forests and Their Conservation Implications.” *Science* 353 (6306): 1383–87. https://doi.org/10.1126/science.aaf5080.

Dublin, Holly T., A.R.E. Sinclair, and J. McGlade. 1990. “Elephants and Fire as Causes of Multiple Stable States in the Serengeti-Mara Woodlands.” *Journal of Animal Ecology* 59 (3): 1147–64. https://doi.org/10.2307/5037.

Durigan, G., and J. A. Ratter. 2006. “Successional Changes in Cerrado and Cerrado/Forest Ecotonal Vegetation in Western Sao Paulo State, Brazil, 1962–2000.” *Edinburgh Journal of Botany* 63 (01): 119. https://doi.org/10.1017/S0960428606000357.

Durigan, Giselda, and James A. Ratter. 2016. “The Need for a Consistent Fire Policy for Cerrado Conservation.” *Journal of Applied Ecology* 53 (1): 11–15. https://doi.org/10.1111/1365-2664.12559.

Eva, Hugh D., Alan S. Belward, Evaristo E. De Miranda, Carlos M. Di Bella, Valery Gond, Otto Huber, Simon Jones, Matteo Sgrenzaroli, and Steffen Fritz. 2004. “A Land Cover Map of South America.” *Global Change Biology* 10 (5): 731–44. https://doi.org/10.1111/j.1529-8817.2003.00774.x.

Foley, J. C. 1957. *Droughts in Australia: Review of Records from Earlierst Yers of Settlement to 1955*. Bull: Australian Bureau of Metereology.

Furley, Peter A. 1994. “Savanna Formations: Ecology and Environment.” *Progress in Physical Geography* 18 (2): 276–94.

Furley, Peter A., and James A. Ratter. 1988. “Soil Resources and Plant Communities of the Central Brazilian Cerrado and Their Development.” *Journal of Biogeography* 15 (1): 97–108. https://doi.org/10.2307/2845050.

Gentry, Alwyn H. 1995. “Diversity and Floristic Composition of Neotropical Dry Forests.” In *Seasonally Dry Tropical Forests*, edited by Stephen H. Bullock, Harold A. Mooney, and Ernesto Medina, 146–94. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511753398.007.

Goodland, R, and R Pollard. 1973. “The Brazilian Cerrado Vegetation : A Fertility Gradient.” *Journal of Ecology* 61 (1): 219–24. https://doi.org/10.2307/2258929.

Grace, J. 2006. *Structural Equation Modeling and Natural Systems*. Cambridge: Cambridge University Press.

Grace James B., Anderson T. Michael, Olff Han, and Scheiner Samuel M. 2010. “On the Specification of Structural Equation Models for Ecological Systems.” *Ecological Monographs* 80 (1): 67–87. https://doi.org/10.1890/09-0464.1.

Higgins, Steven I., William J. Bond, and Winston S. W. Trollope. 2000. “Fire, Resprouting and Variability: A Recipe for Grass–Tree Coexistence in Savanna.” *Journal of Ecology* 88 (2): 213–29. https://doi.org/10.1046/j.1365-2745.2000.00435.x.

Hijmans, Robert J., Susan E. Cameron, Juan L. Parra, Peter G. Jones, and Andy Jarvis. 2005. “Very High Resolution Interpolated Climate Surfaces for Global Land Areas.” *International Journal of Climatology* 25 (15): 1965–78. https://doi.org/10.1002/joc.1276.

Hirota, Marina, Milena Holmgren, Egbert H. Van Nes, and Marten Scheffer. 2011. “Global Resilience of Tropical Forest and Savanna to Critical Transitions.” *Science (New York, N.Y.)* 334 (6053): 232–35. https://doi.org/10.1126/science.1210657.

Hoffmann, William A., Erika L. Geiger, Sybil G. Gotsch, Davi R. Rossatto, Lucas C. R. Silva, On Lee Lau, M. Haridasan, and Augusto C. Franco. 2012. “Ecological Thresholds at the Savanna-Forest Boundary: How Plant Traits, Resources and Fire Govern the Distribution of Tropical Biomes.” Edited by Francisco Lloret. *Ecology Letters* 15 (7): 759–68. https://doi.org/10.1111/j.1461-0248.2012.01789.x.

Holdridge, L. R. 1947. “Determination of World Plant Formations From Simple Climatic Data.” *Science* 105 (2727): 367–68. https://doi.org/10.1126/science.105.2727.367.

Huber, Otto. 2006. “Herbaceous Ecosystems on the Guayana Shield, a Regional Overview.” *Journal of Biogeography* 33 (3): 464–75. https://doi.org/10.1111/j.1365-2699.2005.01454.x.

Humboldt, Count Alexander. 1816. “XCIII. On the Laws Observed in the Distribution of Vegetable Forms.” *The Philosophical Magazine* 47 (218): 446–53. https://doi.org/10.1080/14786441608628504.

Instituto Brasileiro de Geografia e Estatística, ed. 2012. *Manual técnico da vegetação brasileira*. 2a edição revista e ampliada. Manuais técnicos em geociências, número 1. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística-IBGE.

Keyantash, John, and John A. Dracup. 2002. “The Quantification of Drought: An Evaluation of Drought Indices.” *Bulletin of the American Meteorological Society* 83 (8): 1167–1180.

Lavin, Matt, Brian P. Schrire, Gwilym Lewis, R. Toby Pennington, Alfonso Delgado–Salinas, Mats Thulin, Colin E. Hughes, Angela Beyra Matos, and Martin F. Wojciechowski. 2004.

"Metacommunity Process Rather than Continental Tectonic History Better Explains Geographically Structured Phylogenies in Legumes." *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 359 (1450): 1509–22. https://doi.org/10.1098/rstb.2004.1536.

Lehmann, C. E. R., T. M. Anderson, M. Sankaran, S. I. Higgins, S. Archibald, W. A. Hoffmann, N. P. Hanan, et al. 2014. "Savanna Vegetation-Fire-Climate Relationships Differ Among Continents." *Science* 343 (6170): 548–52. https://doi.org/10.1126/science.1247355.

Lehmann, Caroline E. R., Sally A. Archibald, William A. Hoffmann, and William J. Bond. 2011. "Deciphering the Distribution of the Savanna Biome." *New Phytologist* 191 (1): 197–209. https://doi.org/10.1111/j.1469-8137.2011.03689.x.

Leitão, Rafael P., Jansen Zuanon, David Mouillot, Cecília G. Leal, Robert M. Hughes, Philip R. Kaufmann, Sébastien Villéger, et al. 2018. "Disentangling the Pathways of Land Use Impacts on the Functional Structure of Fish Assemblages in Amazon Streams." *Ecography* 41 (1): 219–32. https://doi.org/10.1111/ecog.02845.

Liaw, Andy, and Matthew Wiener. 2002. "Classification and Regression by RandomForest." *R News* 2 (3): 18–22.

Moncrieff, Glenn R., Thomas Hickler, and Steven I. Higgins. 2015. "Intercontinental Divergence in the Climate Envelope of Major Plant Biomes: Intercontinental Biome Divergence." *Global Ecology and Biogeography* 24 (3): 324–34. https://doi.org/10.1111/geb.12257.

Moro, Marcelo Freire, Eimear Nic Lughadha, Francisca Soares de Araújo, and Fernando Roberto Martins. 2016. "A Phytogeographical Metaanalysis of the Semiarid Caatinga Domain in Brazil." *The Botanical Review* 82 (2): 91–148. https://doi.org/10.1007/s12229-016-9164-z.

Murphy, P., and A.E. Lugo. 1995. "Dry Forests of Central America and the Caribbean." In *Seasonally Dry Tropical Forests*, 146–94. Cambridge: Cambridge University Press.

Navarro, Gonzalo. 2011. *Classificación de La Vegetación de Bolivia*. Santa Cruz: Centro de Ecologia Difusión Simón I. Patiño.

Neves, Danilo M., Kyle G. Dexter, R. Toby Pennington, Marcelo L. Bueno, and Ary T. Oliveira Filho. 2015. "Environmental and Historical Controls of Floristic Composition across the South American Dry Diagonal." *Journal of Biogeography* 42 (8): 1566–76. https://doi.org/10.1111/jbi.12529.

Neves, Danilo M., Kyle G. Dexter, R. Toby Pennington, Arthur S. M. Valente, Marcelo L. Bueno, Pedro V. Eisenlohr, Marco A. L. Fontes, et al. 2017. "Dissecting a Biodiversity Hotspot: The Importance of Environmentally Marginal Habitats in the Atlantic Forest Domain of South America." *Diversity and Distributions* 23 (8): 898–909. https://doi.org/10.1111/ddi.12581.

Oberski, Daniel. 2014. “Lavaan.Survey : An R Package for Complex Survey Analysis of Structural Equation Models.” *Journal of Statistical Software* 57 (1). https://doi.org/10.18637/jss.v057.i01.

Odum, E. P. 1992. *Ecology*. 2nd ed. London: Holt Rinehart and Winston.

Oliveira, R. S., L. Bezerra, E. A. Davidson, F. Pinto, C. A. Klink, D. C. Nepstad, and A. Moreira. 2005. “Deep Root Function in Soil Water Dynamics in Cerrado Savannas of Central Brazil.” *Functional Ecology* 19 (4): 574–81. https://doi.org/10.1111/j.1365-2435.2005.01003.x.

Oliveira-Filho, Ary T., and Marco Aurélio L. Fontes. 2000. “Patterns of Floristic Differentiation among Atlantic Forests in Southeastern Brazil and the Influence of Climate1.” *Biotropica* 32 (4b): 793–810. https://doi.org/10.1111/j.1744-7429.2000.tb00619.x.

Oliveira-Filho, Ary T., and James A. Ratter. 2002. “Vegetation Physigonomies and Wood Flora of the Bioma Cerrado.” In *The Cerrados of Brazil: Ecology and Natural History of a Neotropical Savamma*, 91–120. New York: Columbia University Press.

Pellegrini, Adam F.A., Jacob B. Socolar, Paul R. Elsen, and Xingli Giam. 2016. “Trade-Offs between Savanna Woody Plant Diversity and Carbon Storage in the Brazilian Cerrado.” *Global Change Biology* 22 (10): 3373–82. https://doi.org/10.1111/gcb.13259.

Pennington, R. Toby, Matt Lavin, and Ary Oliveira-Filho. 2009. “Woody Plant Diversity, Evolution, and Ecology in the Tropics: Perspectives from Seasonally Dry Tropical Forests.” *Annual Review of Ecology, Evolution, and Systematics* 40 (1): 437–57. https://doi.org/10.1146/annurev.ecolsys.110308.120327.

Pennington, R. Toby, James E. Richardson, and Matt Lavin. 2006. “Insights into the Historical Construction of Species-Rich Biomes from Dated Plant Phylogenies, Neutral Ecological Theory and Phylogenetic Community Structure.” *New Phytologist* 172 (4): 605–16. https://doi.org/10.1111/j.1469-8137.2006.01902.x.

Pennington, R. Toby, Darién E. Prado, and Colin A. Pendry. 2000. “Neotropical Seasonally Dry Forests and Quaternary Vegetation Changes.” *Journal of Biogeography* 27 (2): 261–273.

Prado, Darién E., and Peter E. Gibbs. 1993. “Patterns of Species Distributions in the Dry Seasonal Forests of South America.” *Annals of the Missouri Botanical Garden* 80 (4): 902–27.

Quesada, C. a., J. Lloyd, L. O. Anderson, N. M. Fyllas, M. Schwarz, and C. I. Czimczik. 2011. “Soils of Amazonia with Particular Reference to the RAINFOR Sites.” *Biogeosciences* 8 (6): 1415–40. https://doi.org/10.5194/bg-8-1415-2011.

Quesada, C. a., J. Lloyd, M. Schwarz, T. R. Baker, O. L. Phillips, S. Patiño, C. Czimczik, et al. 2009. “Regional and Large-Scale Patterns in Amazon Forest Structure and Function Are Mediated by Variations in Soil Physical and Chemical Properties.” *Biogeosciences Discussions* 6 (2): 3993–4057. https://doi.org/10.5194/bgd-6-3993-2009.

Quesada, C. a., J. Lloyd, M. Schwarz, S. Patiño, T. R. Baker, C. Czimczik, N. M. Fyllas, et al. 2010. “Variations in Chemical and Physical Properties of Amazon Forest Soils in Relation to Their Genesis.” *Biogeosciences* 7 (5): 1515–41. https://doi.org/10.5194/bg-7-1515-2010.

Quesada, Carlos Alberto, Martin G Hodnett, Lacê M Breyer, Alexandre J B Santos, Sérgio Andrade, Heloisa S Miranda, Antonio Carlos Miranda, and Jon Lloyd. 2008. “Seasonal Variations in Soil Water in Two Woodland Savannas of Central Brazil with Different Fire History.” *Tree Physiology* 28 (3): 405–15.

Ratter, J. A., S. Bridgewater, and J. F. Ribeiro. 2003. “Analysis of the Floristic Composition of the Brazilian Cerrado Vegetation III: Comparison of the Woody Vegetation of 376 Areas.” *Edinburgh Journal of Botany* 60 (01). https://doi.org/10.1017/S0960428603000064.

Ribeiro, J. F., and Bruno M.T. Walter. 2008. “As Principais Fitofisionomias Do Bioma Cerrado.” In *Cerrado: Ecologia e Flora*, 1st ed. Vol. 1. Brasília: Embrapa Cerrados/Embrapa Informação Tecnológica.

Roy, D.P., L. Boschetti, C.O. Justice, and J. Ju. 2008. “The Collection 5 MODIS Burned Area Product — Global Evaluation by Comparison with the MODIS Active Fire Product.” *Remote Sensing of Environment* 112 (9): 3690–3707. https://doi.org/10.1016/j.rse.2008.05.013.

Rozendaal, Danaë M. A., Robin L. Chazdon, Felipe Arreola-Villa, Patricia Balvanera, Tony V. Bentos, Juan M. Dupuy, J. Luis Hernández-Stefanoni, et al. 2017. “Demographic Drivers of Aboveground Biomass Dynamics During Secondary Succession in Neotropical Dry and Wet Forests.” *Ecosystems* 20 (2): 340–53. https://doi.org/10.1007/s10021-016-0029-4.

Santos, Rubens M., Ary T. Oliveira-Filho, Pedro V. Eisenlohr, Luciano P. Queiroz, Domingos B. O. S. Cardoso, and Maria J. N. Rodal. 2012. “Identity and Relationships of the Arboreal Caatinga among Other Floristic Units of Seasonally Dry Tropical Forests (SDTFs) of North-Eastern and Central Brazil: Biogeography of Brazilian Seasonally Dry Tropical Forests.” *Ecology and Evolution* 2 (2): 409–28. https://doi.org/10.1002/ece3.91.

Savalei, Victoria. 2014. “Understanding Robust Corrections in Structural Equation Modeling.” *Structural Equation Modeling: A Multidisciplinary Journal* 21 (1): 149–60. https://doi.org/10.1080/10705511.2013.824793.

Scarano, F. R. 2002. “Structure, Function and Floristic Relationships of Plant Communities in Stressful Habitats Marginal to the Brazilian Atlantic Rainforest.” *Annals of Botany* 90 (4): 517–24. https://doi.org/10.1093/aob/mcf189.

Schrire, B.D., Matt Lavin, and G.P. Lewis. 2005. “Global Distribution Patterns of the Leguminosae: Insights from Recent Phylogenies.” In *Plant Diversity and Complexity Patterns, Local, Regional and Global Dimensions*, 375–422. Copenhagem: The royal Danish Academy of Sciences and Letters.

Shipley, Bill. 2002. *Cause and Correlation in Biology: A User’s Guide to Path Analysis, Structural Equations and Causal Inference*. 1 edition. Cambridge: Cambridge University Press.

Silva de Miranda, Pedro Luiz, Ary T. Oliveira-Filho, R. Toby Pennington, Danilo M. Neves, Timothy R. Baker, and Kyle G. Dexter. 2018. "Using Tree Species Inventories to Map Biomes and Assess Their Climatic Overlaps in Lowland Tropical South America." *Global Ecology and Biogeography* 27 (8): 899–912. https://doi.org/10.1111/geb.12749.

Staver, A. C., S. Archibald, and S. A. Levin. 2011. "The Global Extent and Determinants of Savanna and Forest as Alternative Biome States." *Science* 334 (6053): 230–32. https://doi.org/10.1126/science.1210465.

Sullivan, Martin J. P., Joey Talbot, Simon L. Lewis, Oliver L. Phillips, Lan Qie, Serge K. Begne, Jerôme Chave, et al. 2017. "Diversity and Carbon Storage across the Tropical Forest Biome." *Scientific Reports* 7 (January): 39102. https://doi.org/10.1038/srep39102.

Vanzolini, Paulo Emílio. 1963. "Problemas Faunísticos Do Cerrado." *Simpósio Sobre o Cerrado* 307 (321): 1.

Velloso, Agnes L., Everardo V. S. B. Sampaio, and Frans G. C. Pareyn. 2001. *Ecorregiões propostas para o bioma Caatinga*. Recife: Instituto de Conservação Ambiental The Nature Conservancy.

Veloso, Henrique Pimenta, Antonio Lourenço Rosa Rangel Filho, and Jorge Carlos Alves Lima. 1991. *Classificação da vegetação brasileira, adaptada a um sistema universal*. Rio de Janeiro: Ministério da Economia, Fazenda e Planejamento, Fundação Instituto Brasileiro de Geografia e Estatística, Diretoria de Geociências, Departamento de Recursos Naturais e Estudos Ambientais.

Vicente-Serrano, S. M., C. Gouveia, J. J. Camarero, S. Begueria, R. Trigo, J. I. Lopez-Moreno, C. Azorin-Molina, et al. 2013. "Response of Vegetation to Drought Time-Scales across Global Land Biomes." *Proceedings of the National Academy of Sciences* 110 (1): 52–57. https://doi.org/10.1073/pnas.1207068110.

Villarroel, Daniel, Cassia B R Munhoz, and Carolyn E B Proença. 2016. "Campos y Sabanas del Cerrado en Bolivia: Delimitación, síntesis terminológica y sus caracteristicas fisionómicas," 34.

Warman, Laura, and Angela T. Moles. 2009. "Alternative Stable States in Australia's Wet Tropics: A Theoretical Framework for the Field Data and a Field-Case for the Theory." *Landscape Ecology* 24 (1): 1–13. https://doi.org/10.1007/s10980-008-9285-9.

Whittaker, R. H. 1975. *Communities and Ecosystems*. 2nd edition. New York: Macmillan USA.

Whittaker, R. H., and S. A. Levin. 1977. "The Role of Mosaic Phenomena in Natural Communities." *Theoretical Population Biology* 12 (2): 117–39. https://doi.org/10.1016/0040-5809(77)90039-9.

3.8 - Appendix 3: supplementary figures to chapter 3

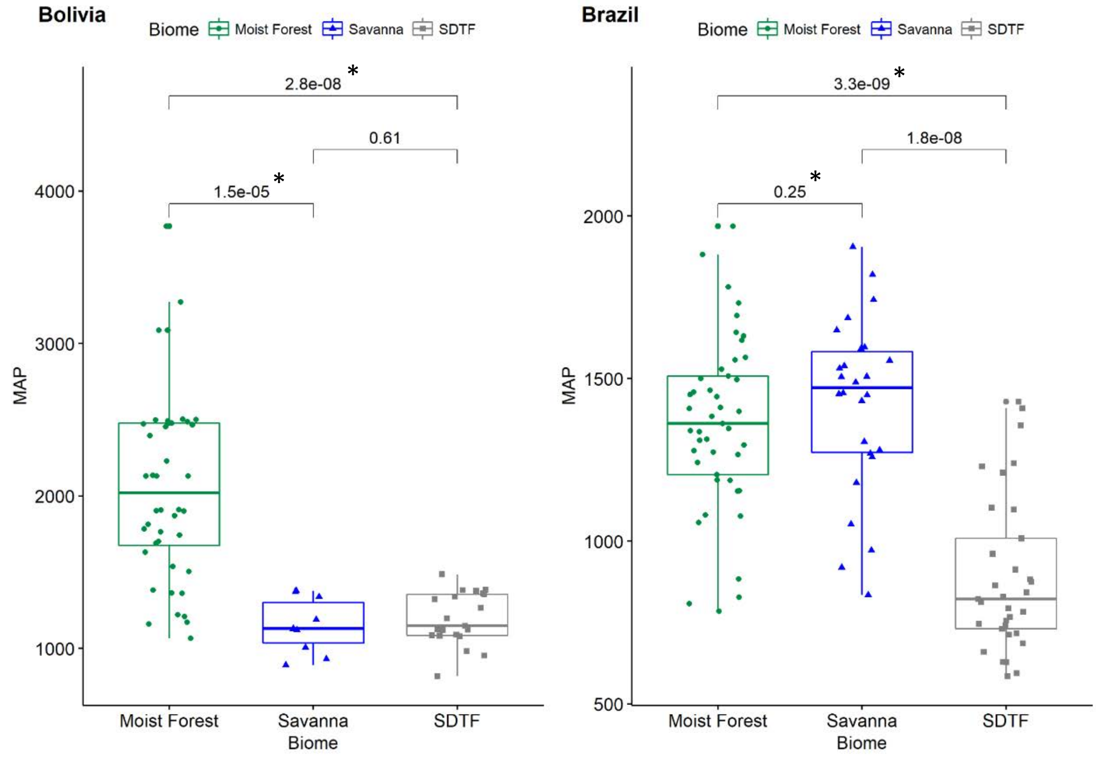

Fig S3.1: Box-plots of mean annual precipitation associated with pairwise Wilcoxon's signed rank test for Bolivia and eastern Brazil. Significant comparisons are signed with (*). P<0.016. Note the different scales on the y-axes for the left and right-hand panels.

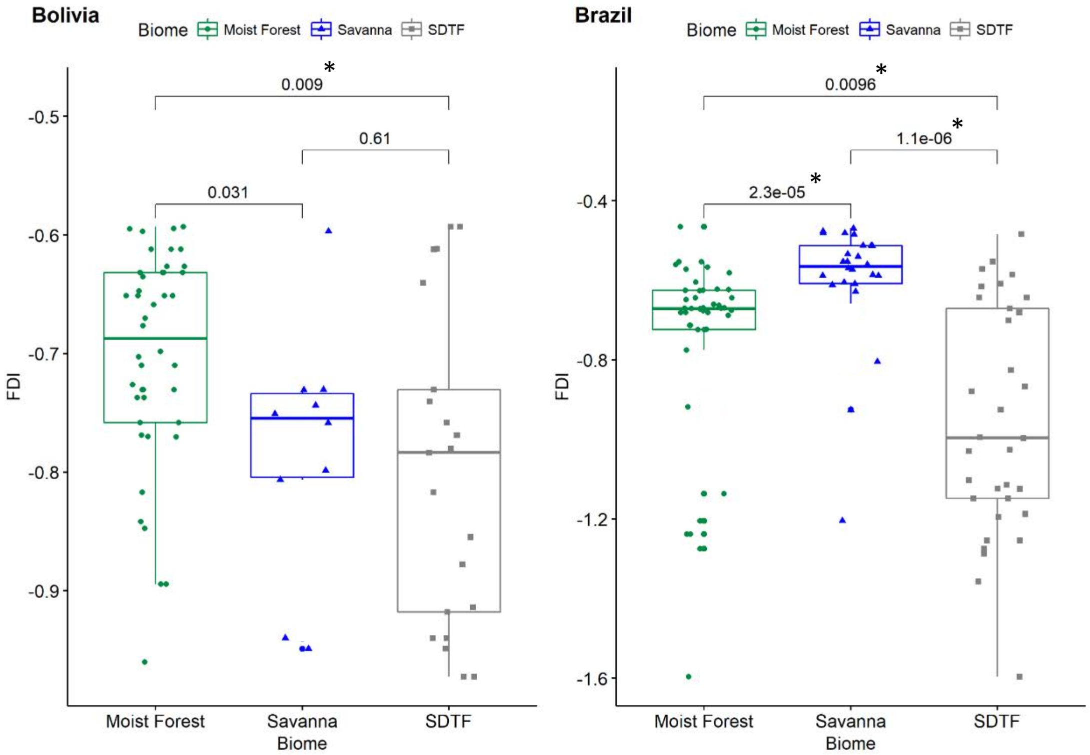

Fig S3.2: Box-plots of Foley's drought index (FDI) associated with pairwise Wilcoxon's signed rank test for Bolivia and eastern Brazil. Significant comparisons are signed with (*). P<0.016. Note the different scales on the y-axes for the left and right-hand panels.

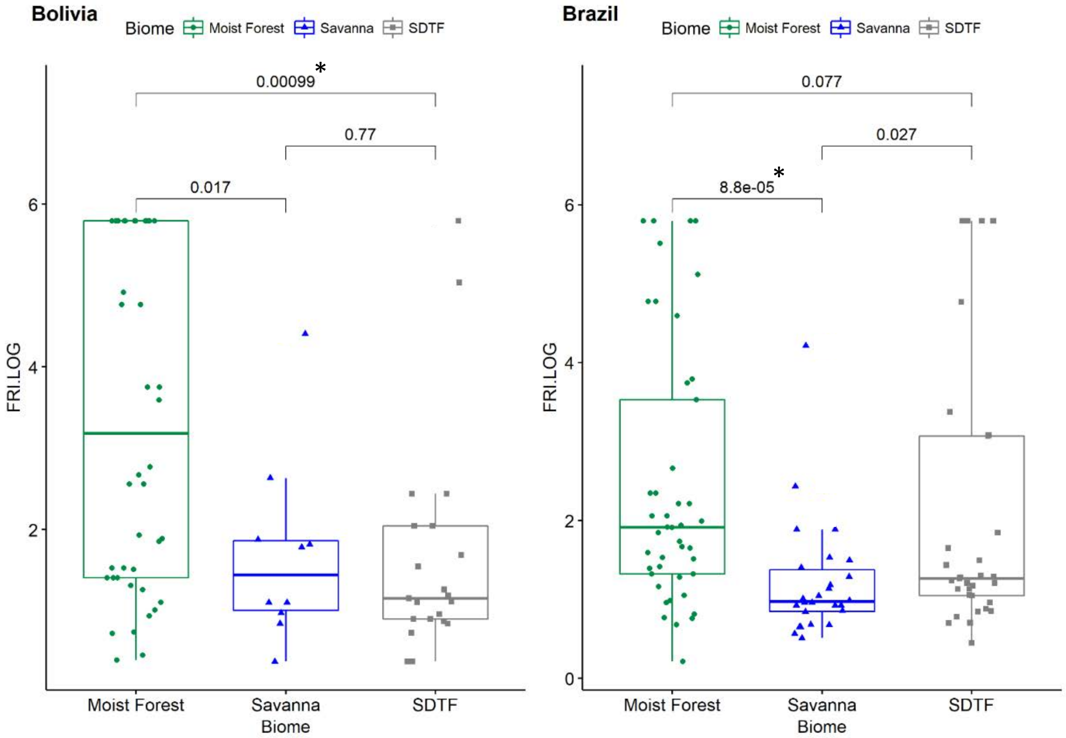

Fig S3.3: Box-plots of fire return interval (log-transformed, base e) associated with pairwise Wilcoxon's signed rank test for Bolivia and eastern Brazil. Significant comparisons are signed with (*). P<0.016.

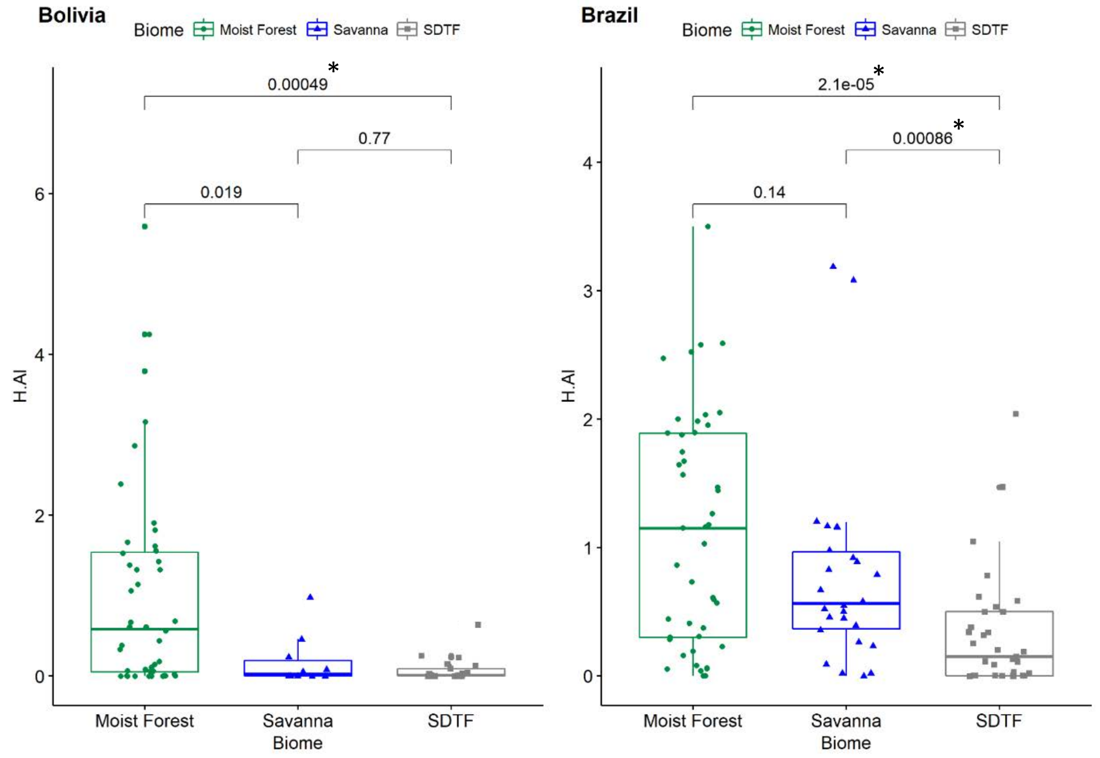

Fig S3.4: Box-plots of acidic cations (H,Al) associated with pairwise Wilcoxon's signed rank test for Bolivia and eastern Brazil. Significant comparisons are signed with (*). P<0.016. Note the different scales on the y-axes for the left and right-hand panels.

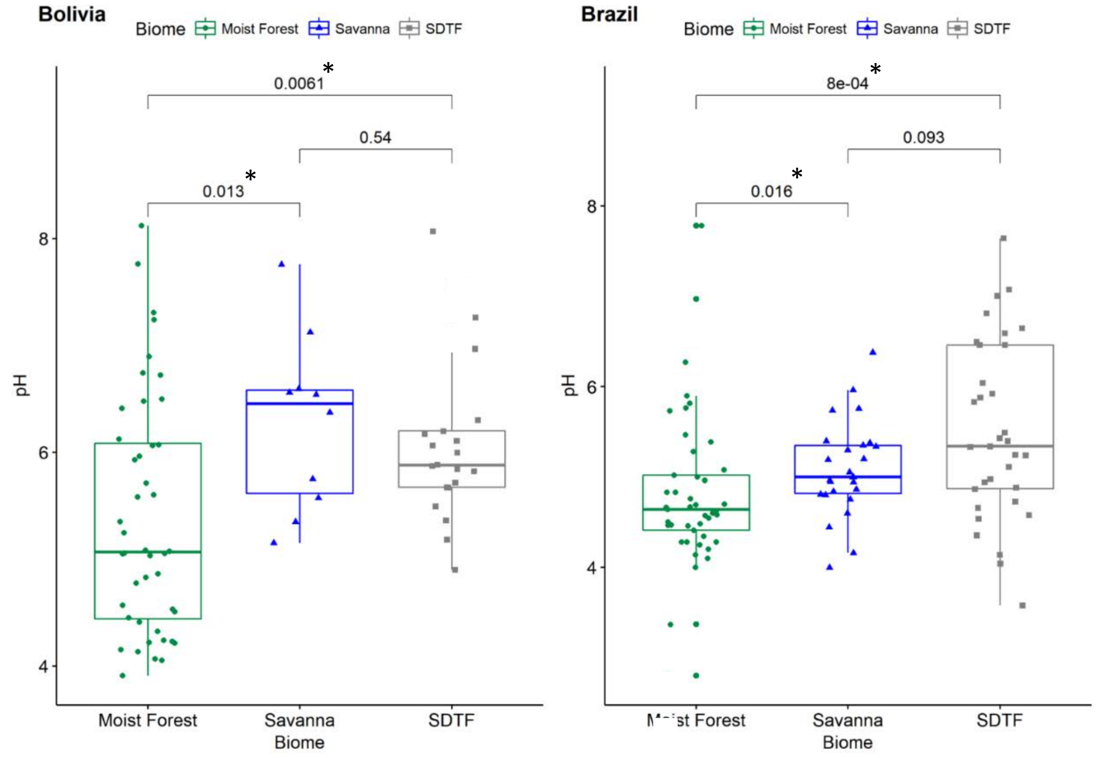

Fig S3.5: Box-plots of pH associated with pairwise Wilcoxon's signed rank test for Bolivia and eastern Brazil. Significant comparisons are signed with (*). P<0.016. Note the different scales on the y-axes for the left and right-hand panels.

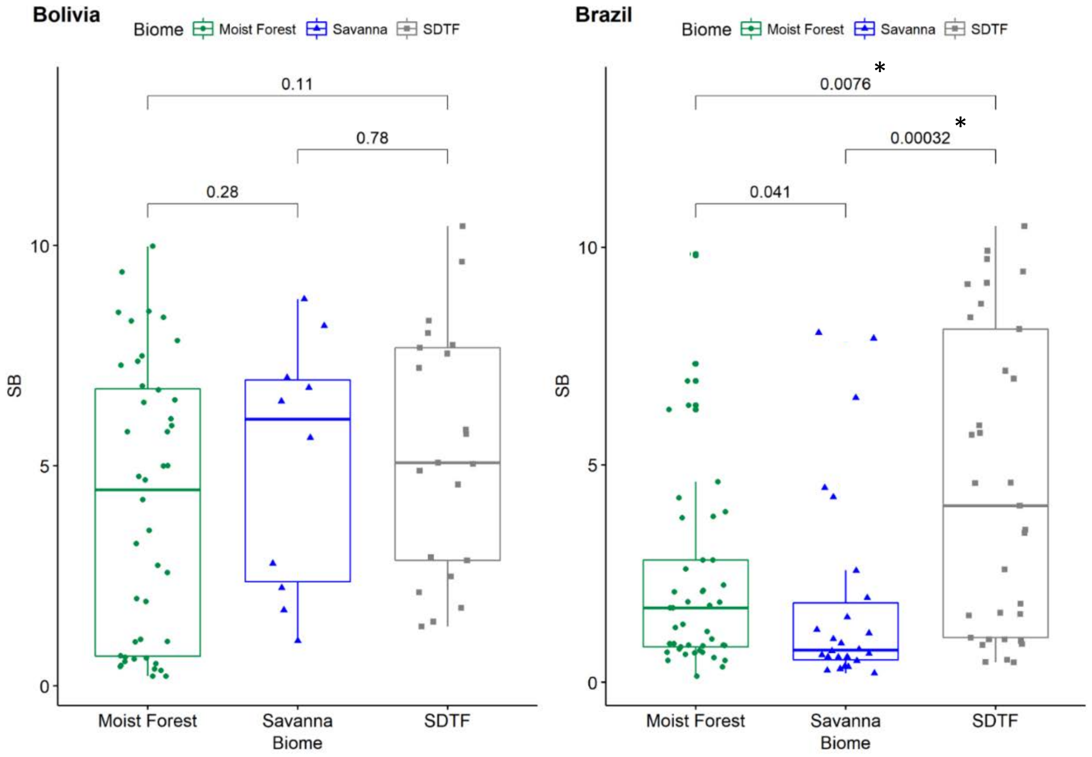

Fig S3.6: Box-plots of sum of bases (SB) associated with pairwise Wilcoxon's signed rank test for Bolivia and eastern Brazil. Significant comparisons are signed with (*). P<0.016.

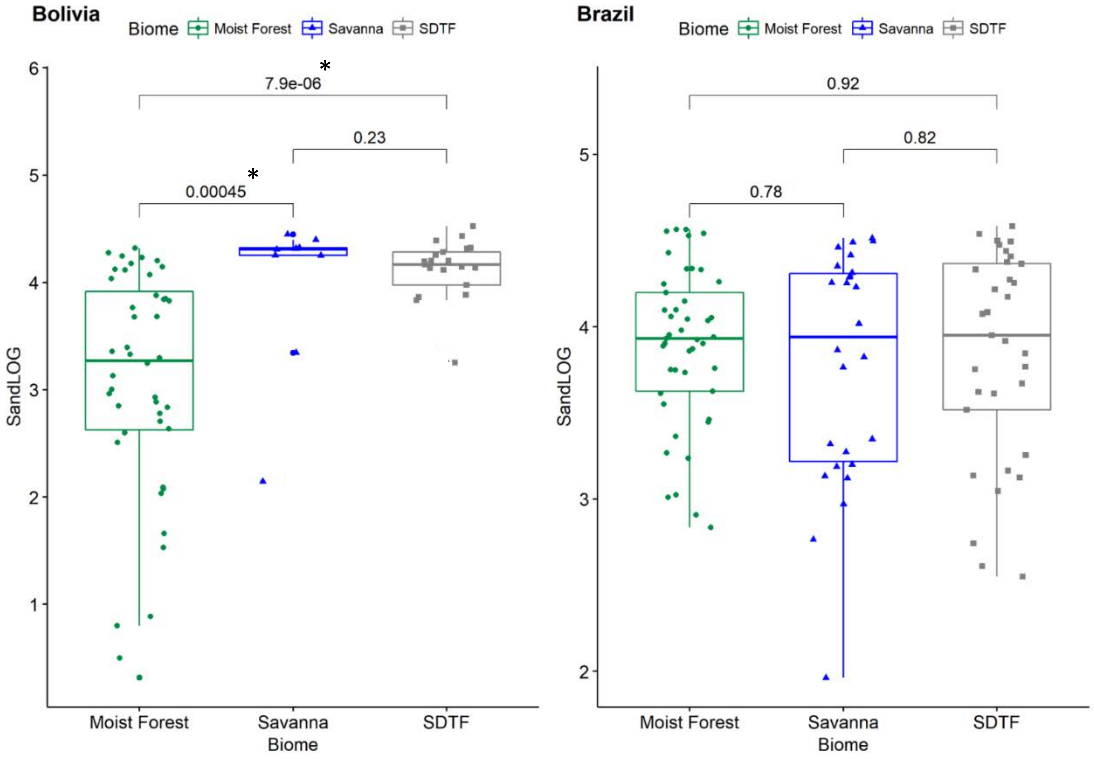

Fig S3.7: Box-plots of quantity of sand (log-transformed, base e) associated with pairwise Wilcoxon's signed rank test for Bolivia and eastern Brazil. Significant comparisons are signed with (*). P<0.016. . Note the different scales on the y-axes for the left and right-hand panels.

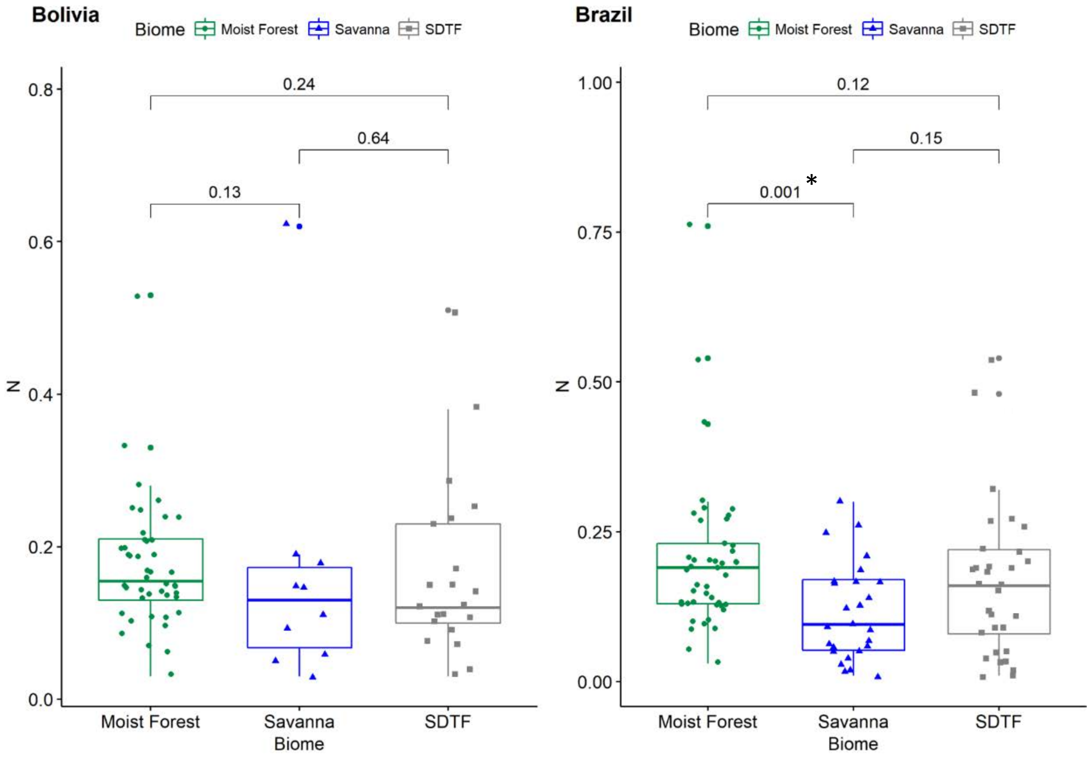

Fig S3.8: Box-plots of Nitrogen associated with pairwise Wilcoxon's signed rank test for Bolivia and eastern Brazil. Significant comparisons are signed with (*). P<0.016. Significant comparisons are signed with (*). P<0.016. . Note the different scales on the y-axes for the left and right-hand panels.

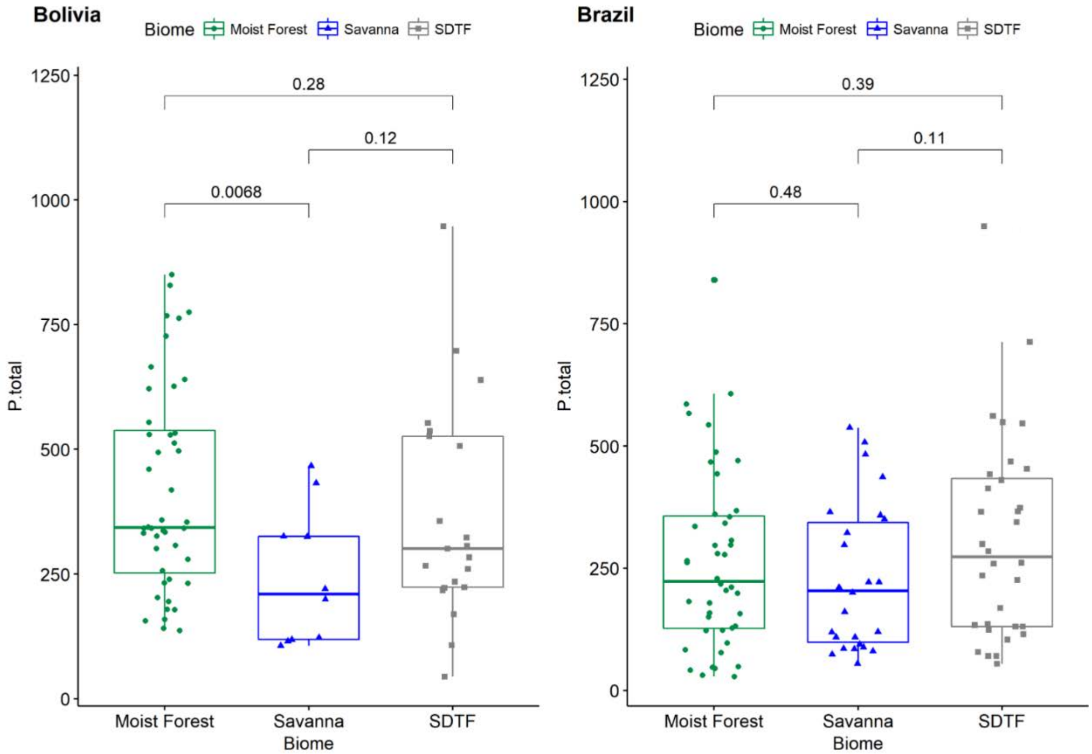

Fig S3.9: Box-plots of total phosphorous associated with pairwise Wilcoxon's signed rank test for Bolivia and eastern Brazil. Significant comparisons are signed with (*). P<0.016.

4

A multi-biome comparison of tree species richness, endemism and sampling sufficiency in lowland tropical South America

4 - A multi-biome comparison of tree species richness, endemism and sampling sufficiency in lowland tropical South America

4.1 - Abstract

Background

Biomes of Lowland tropical South America (LTSA) are known for their biodiversity and complex patterns of species distribution. However, controversy still surrounds these biomes' tree species richness and number of endemics. Previous investigations have suggested that sensitivity to drought and biogeographic history are the main factors influencing tree species richness and sendemism levels in LTSA, but this has only been tested at specific geographic locations or through simulation studies. My aims with this study are to quantify and compare the total and endemic tree species richness of the main biomes of lowland tropical South America (LTSA) and relate these patterns to possible environmental and historic drivers. I also estimate how well sampled is the tree species richness of different biomes.

Methods

I extracted 4,540 geo-referenced tree species inventories from the NeoTropTree dataset (Oliveira-Filho, 2017), which are distributed across the four main biomes of LTSA: Amazon Forest, Atlantic Forest, Savanna and Seasonally Dry Tropical Forest (SDTF) and defined their core area of occurrence based on the literature. Following that, I performed three different analyses: 1) documentation of how total species richness accumulated with increased sampling of biomes and their respective core areas; 2) estimation of the upper and lower bounds on the number of endemic species per core area through indicator species analysis and counts of species unique to them, respectively; 3) estimation of unsampled tree species richness of the core areas through species extrapolation curves. Analyses were conducted for core areas of the biomes as well as extended geographic areas that additionally include patches of a given biome found within the core area of other biomes.

Results

The analyses show that the Amazon Forest has 8033 tree species (core area 7845), the Atlantic Forest 4955 (core area 3936), Savannas 3550 (core area 2008) and SDTFs 1845 (core area 1031). Endemic tree species for core area counts followed a similar trend. The Amazon Forest has between 4844 and 5476 endemic tree species, the Atlantic Forest between 2214 and 2484, the Savanna between 163 and 404 and the SDTF, between 141 and 417. Total tree species richness estimations for all core areas showed that the Amazon Forest and the Savanna have the largest gap between observed and estimated total species richness (at least 508 and 584 potential new species, respectively), whereas this difference for the SDTF and the Atlantic Forest is not as substantial (235 and 79 potentially unrecorded species, respectively). However, proportionally, savanna stood out as having the lowest sampling completeness with only 78% of its tree species having been sampled compared to 93-94% for the other three biomes.

Main Conclusions

I show that the Amazon Forest is the most tree species-rich biome in LTSA, followed by the Atlantic Forest, Savannas and SDTFs. The tree flora of the Amazon and Atlantic forests is mainly composed of endemic species whereas that is not the case for the savanna and the SDTF. These results indicate that drought sensitivity and biogeographic history are possibly linked to patterns of tree species distribution in LTSA. The especies extrapolation curves indicate that the Amazon and Atlantic forests and the SDTF have had ~94% of their tree flora documented, which is likely an underestimate due to methodological constraints. Also according to this analysis, only ~77% of the Savanna's tree flora has been documented so far. However, this may be related to the incursion of other biomes into core areas of Savannas, which may amplify the apparent proportion of species that are rare and in turn leads to statistical results of fairly incomplete sampling.

4.2 - Introduction

The Neotropics is the most species rich biogeographic region in the world (Antonelli and Sanmartín 2011) and all South American tropical countries are known for their high species richness. Brazil alone hosts more than 33.000 flowering plant species (“Flora Do Brasil 2020”) and, in South America, Brazil is followed by Colombia (26.500), Peru (17.200), Ecuador (16.000) and Venezuela (15820) in total plant species richness (Rangel - Ch. 2015). However, much of the Neotropics has not been completely explored scientifically and therefore its biodiversity remains only partially documented. For example, in Brazil, almost 3000 new species of plants have been described between 1990 and 2006 (Sobral and Stehmann 2009). Between 2010 and 2015, 1467 more angiosperms were described in this same country (Forzza et al. 2012; Zappi et al. 2015). The description of new species is a pressing matter, since at least half of the plant species threatened with extinction in this area are still unknown to science, according to one estimate (Joppa, Roberts, and Pimm 2010).

These high levels of diversity have been the focus of research attempting to understand them (e.g. Rangel et al. 2018 and references therein). To this end, various theories relating species richness to environmental factors, productivity and evolutionary time have been put forward. Studies that link species richness and biodiversity distribution at the continental scale to geological and climatic events and environmental stability are quite common (Riddle 2016). For example, there are approximately 30 hypotheses aiming to explain the latitudinal gradient in species richness that is observed globally and in the Neotropics (Vázquez and Stevens 2004). However, a full understanding of such patterns is still elusive given their complex nature (Arita and Vázquez-Domínguez 2008 and references therein).

One key mechanism underpinning tree species distribution in the Neotropics may be related to drought-tolerance. The works of Gentry (1988), Ter Steege et al. (2003) and Esquivel-Muelbert et al. (2017) show that differences in tree species richness across and within biomes in the Neotropics are related to drought-sensitivity.

According to Esquivel-Mulbert et al. (2017), tree diversity is negatively correlated to water-stress in the western Neotropics, suggesting that low tolerance to drier conditions is common for tree species in that region. Another possible factor connected to species distribution in the Neotropics' is its biogeographic history. Antonelli et al. (2018) showed that the Amazon Forest served as a "lineage cradle" for the other Neotropical biomes, revealing this biome's importance for Neotropical diversity. In conjunction, the studies of Esquivel-Mulbert et al (2017) and Antonelli et al (2018) points to the possibility that higher tree species richness can be found in wet biomes, especially the Amazon.

Sensitivity to drought and biogeographic history are possibly linked to tree species levels of endemism in the Neotropics as well. High endemism levels are usually attributed to environmental and habitat stability over time in association with species' low capacity to disperse (Graham, Moritz, and Williams 2006). Therefore, endemic species usually have narrow distribution ranges and small populations (Sandel et al. 2011). In a simulation study, Rangel et al. (2018), has shown that persistence and diversification of lineages in specific regions of the South American continent (diversity cradles) seem to be the product of the joint effects of climatic and topographic complexities during the last 800.000 years (Rangel et al 2018). In LTSA, these authors highlight the western Amazon and the coastal Atlantic Forest – the two main moist biomes in this region – as biodiversity cradles and potentially harbouring high levels of diversity and endemism. This again points to the potential role of drought-sensitivity as a driver of tree species endemism levels. However, research on endemism at broad scales usually defines endemic species as having narrow distribution ranges only (e.g. Sandel et al 2011) and does not focus on species that are unique to a given geographic locality or region. Because of this, potential environmental drivers of endemism remain largely unknown and, the number of endemic tree species to LTSA biomes are yet to be accurately calculated.

Knowledge of diversity is unevenly geographically distributed due to biases in plant species collection and sampling effort across different biogeographical (e.g. biomes, formations) and political units (e.g. countries and states). As shown by Sousa-Baena, Garcia, and Peterson (2014) and Oliveira et al. (2016) for Brazil and by Jimenez et al. (2009) for the tropical Andean countries, plant species collection effort is small overall and spatially biased, tending to be centred around roads, rivers, universities and research centres (a.k.a the museum effect, Ponder et al. 2001). Gaps in sampling are a product of how remote and inaccessible much of tropical South America remains and of how expensive collection expeditions are (Jiménez, Distler, and Jørgensen 2009; Oliveira et al., 2017).

At the biome level, these gaps prevent a comparison of how many species are shared or are unique to biomes', their core areas (largest continuous area of occurrence) and their geographically disjunct or marginal areas (areas of the same biome not connected to the biome's core area). For example, little is still known about the differences and similarities between the Brazilian Cerrado and the other savannic ecosystems in LTSA (e.g. the Llanos of Venezuela and Colombia), as studies tend to focus on one or the other (e.g. Ratter, Bridgewater, and Ribeiro 2003; Bueno et al. 2018). This was also the case for the Neotropical SDTFs. However, DRYFLOR et al. (2016) recently united SDTF from across the tropics, and showed that there is a high tree species turnover across SDFTs' nuclei in the Neotropics. DRYFLOR (2016) indicates that biome's marginal areas, which are not commonly assessed in species richness studies, can potentially harbour species that cannot be found in the biome's core area of occurrence. This might as well be true for the Savanna and the Atlantic Forest, since studies in these biomes often do not encompass their entire distribution range.

Importantly, such gaps can lead to inaccurate assessments of species richness for each biome and prevent standardized cross-biome comparisons of species richness and endemism levels, especially of biomes extending across national political boundaries. To that end, there is a need for standardized estimations of total species richness and endemism spanning over biologically meaningful biomes. The most efficient way to perform such analyses would be through species accumulation/extrapolation curves based on reliable biodiversity estimators and statistical models (e.g. Magurran 2003; Chao et al. 2004, 2014; ter Steege et al. 2016; Chen and Shen 2017).

Here, I investigate tree species richness of lowland tropical South America's four main biomes (Amazon Forest, Atlantic Forest, Savanna and SDTF) and their core areas to address the following hypotheses:

a) Previous studies suggest that tree species richness is linked to water stress and biogeographic history. I therefore predict that:

a.1) The most tree species rich biomes in LTSA are the Amazon and Atlantic Forests as these are moist environments with lower average drought stress).

a.2) The Amazon will have more tree species than other biomes, since the Amazon is the main source of lineages for the other Neotropical biomes.

b) Biomes' tree species endemism levels are also most likely linked to tolerance to water stress and biogeographic history. Therefore, the wet biomes – the Amazon and Atlantic Forests – harbour more endemic species than the drier biomes – Savanna and SDTF.

c) Biomes' geographically marginal areas are an important reservoir of tree species richness and the inclusion of such areas may significantly increases tree species counts. Therefore, the marginal areas of the Atlantic Forest, Amazon Forest, SDTF and Savanna will increase these biomes' species counts.

4.3 - Methods

4.3.1 - The dataset – NeoTropTree

I retrieved tree species composition information from the NeoTropTree (NTT) dataset – a compilation of taxonomically updated/standardized floristic checklists for South, Central and southern North America (Oliveira-Filho, 2017). NTT currently contains tree species inventories for more than 8000 geo-referenced sites. In this database, trees are defined as free-standing woody plants greater than three metres in height. In this dataset, every site is created using tree species inventories and phytosociological/floristic surveys. These are recovered from published (e.g. scientific papers, official documents) and unpublished literature (e.g. PhD theses, environmental consultancy reports). Herbaria records of trees from within the area of a site are obtained through the consultation of plant vouchers deposited across the Americas, Europe and online resources (e.g. CRIA, 2012) and then added to these checklists. Taxonomic standardization and actualization is made by consulting the taxonomic literature, the "Flora do Brasil" (http://floradobrasil.jbrj.gov.br/) and the "Flora del Conosur" (Zuloaga & Belgrano, 2015) – http://www.darwin.edu.ar/), with additional direct consultation of taxonomists.

I obtained two sets of data from NTT. Firstly, I obtained 4539 sites in lowland tropical South America which were then classified into four biomes according to their tree species composition – Atlantic Forest, Amazon Forest, Savanna and Seasonally Dry Tropical Forest. This was achieved through a cluster analysis based on pairwise floristic distance among NTT sites (see Chapter 2 for details). Secondly, I removed all sites that were not a part of the four biomes' largest continuous area in South America. The sites obtained in this extraction are a subset of the pool of sites obtained in the first extraction and is composed of 3840 sites. Geographically, the first group of NTT sites correspond to the biomes' core areas (largest continuous area of occurrence, more details below) and also encompasses these biome's disjunct occurrences within the study area. The second group of NTT sites correspond to these biomes' core areas only.

The first group of sites were used to determine the total tree species richness of the four biomes considered in this contribution – these sites compose the whole area of occurrence of each biome in South America, east of the Andes and south of northern Colombia and Venezuela. The second group of sites were used to determine the tree species richness of each one of these biomes' core areas – the largest continuous area of occurrence of each biome analysed here. These were defined considering the works of Oliveira-Filho (2017), IBGE (2012) and Villarroel, Munhoz, and Proença (2016) and corresponds to the phytogeographic domains defined by these studies (continuous plant formations with a prevailing vegetation type). In analysing core areas, I excluded patches of other biomes found within them (e.g. gallery forests and SDTF patches in the Savannas). As the means to ensure all core areas were compositionally coherent, I compared how such sites were classified by Oliveira-Filho (2017) and Silva de Miranda et al. (2018) and removed the ones that were not assigned to the same biome/area by both studies.

4.3.2 - Area of Study

Our study area, lowland tropical South America, comprises for main biomes (Fig. 4.1), described below:

4.3.2.1 - Amazon Forest

The Amazon Forest is the largest and the most diverse biome in South America, and has seen an increase in publications concerning its diversity during recent years (e.g. ter Steege et al. 2013; Cardoso et al. 2017). This biome is known mainly for its lowland forests which encompass two main ecosystems: terra firme (non-flooded) and seasonally flooded forests, both of which are known for their great plant diversity (Tuomisto, Zuquim, and Cárdenas 2014). They tend to form high, closed canopies. Near the northern and southern borders of this biome, some trees lose their leaves during a short dry season. This biome also contains white-sand forests and vegetation found on Tepuís and other geological formations.

This biome’s core area was delimited in accordance to Oliveira-Filho (2017) and maps produced by Cardoso et al. (2017) and ter Steege et al. (2013). This biome extended area of occurrence encompasses gallery forests and other forest enclaves within the Cerrado (central South American savannas), as documented by Silva de Miranda et al. (2018).

4.3.2.2 - Atlantic Forest

The core area of the Atlantic Forest ranges from Rio Grande do Norte state (northeast Brazil) southwards to Rio Grande do Sul state (south Brazil) and reaching eastern Paraguay and northeastern Argentina. It consists of a belt of forests from the coast inwards that is wider in the south than the north (Fiaschi and Pirani 2009). This ecosystem encompasses the moist forests closer to the Atlantic coast, the semi-deciduous forests in southeast Brazil and the sub-tropical Atlantic forests (with the gymnosperm tree *Araucaria araucana* (Molina) K. Koch) in south of Brazil, NE Argentina and Paraguay (Oliveira-Filho and Fontes 2000; Oliveira-Filho et al. 2015). This biome is a biodiversity hotspot (Myers et al. 2000) and is known for its high plant species richness. At present, the Atlantic forest has only approximately 8,5% of its original vegetation left (Fundação SOS MAta Atlântica 2016) and much of it is not well protected (Mittermeier et al. 2005).

The Atlantic Forest's extended area of occurrence encompasses gallery forests and semideciduous forest enclaves in the Cerrado and in the Pampas (sub-tropical grasslands) that are floristically similar to this biome's core area. This was done in accordance to what was observed by Silva de Miranda et al. (2018).

4.3.2.3 - Savannas

The savannas' core area correspond to the delimitation of the South American Cerrado region, which encompasses the continuous extent of savannas spanning central Brazil and eastern Bolivia (Villarroel, Munhoz, and Proença 2016; Oliveira-Filho, 2017). This region covers ~20% of the Brazilian territory and ~10% of Bolivia (Villarroel, Munhoz, and Proença 2016). The SDTF patches and gallery forests (belonging to moist forest biomes), which can be found within Cerrado's geographic limits, were not classified as Savanna in this study. The savanna vegetation itself varies widely from very sparse woodland (termed campo sujo) to much denser woodlands (termed cerradão) (Ribeiro and Walter 2008). The number of vegetation types varies from author to author according to what they believe can be considered as Cerrado or not, but there is consensus on the limits of its overall geographic distribution (IBGE 2012; Villarroel, Munhoz, and Proença, 2016).

The extended area of the Savanna biome (as defined by Silva de Miranda, 2018) also considers the savanna patches occurring within the other biomes (Brazilian Caatinga, Amazon and Atlantic Forests). There are other savannas in South America, such as the Gran Savanna in southeastern Venezuela and the Llanos in Colombia and Venezuela, but these were not considered in the present study.

4.3.2.4 - Seasonally Dry Tropical Forest (SDTF)

The SDTF biome has a patchy distribution across the Neotropics (Prado and Gibbs 1993; Pennington, Prado, and Pendry 2000; DRYFLOR, 2016). In the present contribution, the Brazilian Caatinga is treated as the SDTF biome’s core area due to this region being the largest continuous area of SDTF in the Neotropics (Pennington et al., 2000; DRYFLOR, 2016). This region is located in northeast Brazil and borders the Amazon forest to the northwest, Cerrado to the west and Atlantic Forest to the east and south (IBGE, 2012). Other important regions of SDTF in South America are enclaves within the Cerrado and in inter-Andean valleys and also along the Caribbean coast of Colombia and Venezuela (Prado & Gibbs 1993; DRYFLOR 2016). Out of the biomes being studied in this contribution, SDTF’s core area is the only one whose distribution is restricted entirely to Brazil.

The SDTF’s extended area of occurrence considered here encompasses the Caatinga and the SDTF enclaves throughout the Cerrado and the Bolivian Chiquitania (as shown by Silva de Miranda et al., 2018). Vegetation in this area can vary from shrub dominated landscapes to dry forests with low or occasionally high canopies and patchy grasses and bromeliads in the understory (Velloso et al. 2001).

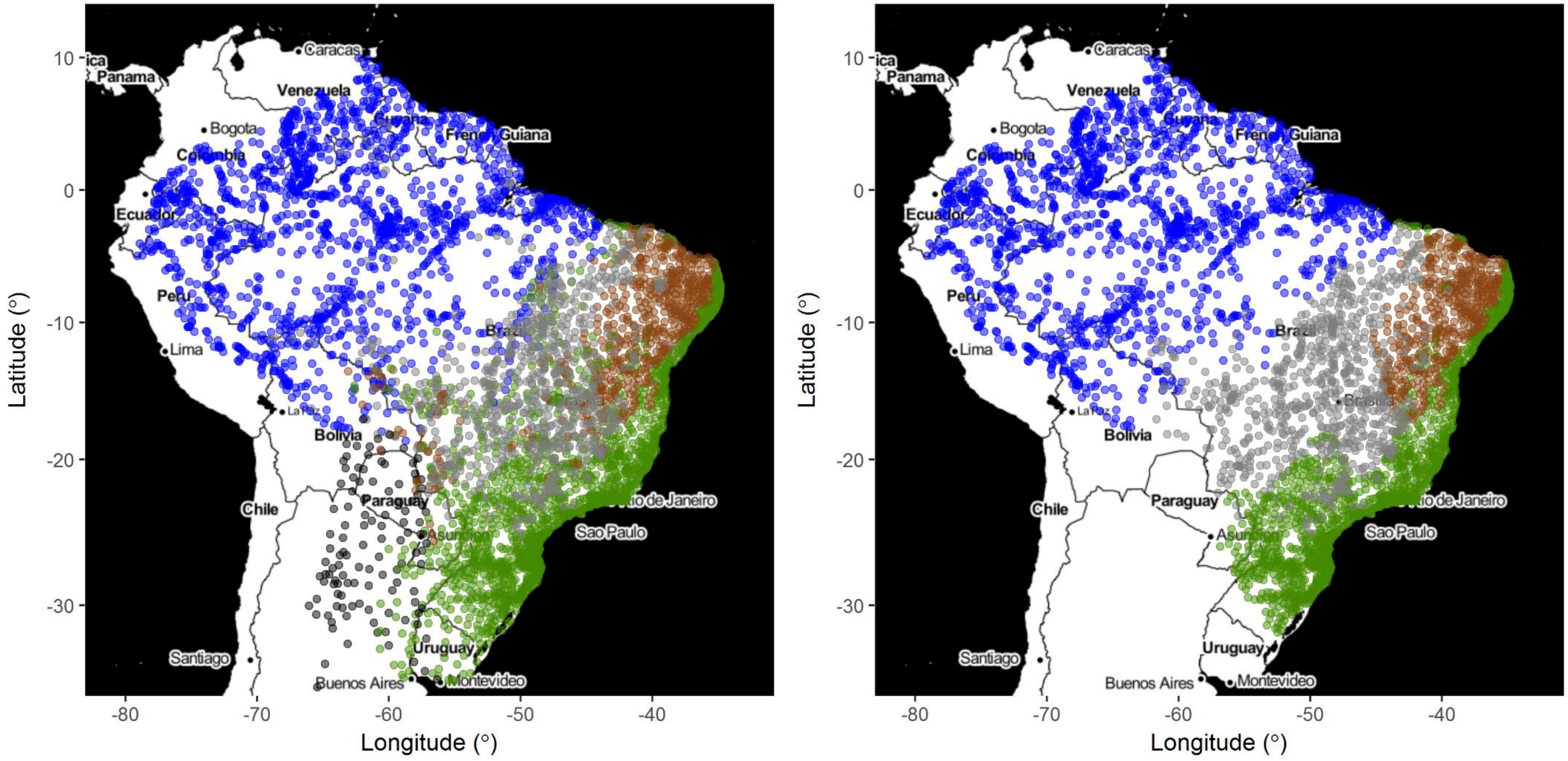

Figure 4.1: Lowland tropical South America’s main biomes (a) and their core areas of occurrence (b) according to their tree species composition. Biomes in (a) were extracted from Silva de Miranda et al. (2018) and core areas in (b) were delimited according to Oliveira-Filho (2017) and Silva de Miranda (2018). [blue – Amazon Forest; Green – Atlantic Forest *senso-latissimo*/*senso-lato*; Brown – dry-diagonal seasonally dry tropical forest/Caatinga; Gray – dry-diagonal and Amazonic savannas/Cerrado; Black – Chaco]

4.3.3 - Species Accumulation/Extrapolation Curves

I constructed species accumulation curves for both core and extended areas of the Amazon Forest, Atlantic Forest, Savanna and Seasonally Dry Tropical Forest. The accumulation curve's 95% confidence intervals were constructed through non-parametric boostrapping, which consisted of resampling of sites for a given biome (1000 times). The accumulation curves were created in R (R Core Team 2018) by using the "Vegan" package (specaccum (), Okasem, 2018).

The tree species extrapolation curves and endemism analyses were constructed for the core area of each biome only. I have excluded geographically marginal areas/sites to enable this analysis to be centred on geographically coherent units. Areas outside the biomes' core areas are usually vegetation patches of various sizes located within another biome (e.g. as vegetation enclaves and gallery forests). Because of this, such areas are prone to harbouring tree species from other biomes via mass effects (Shmida and Wilson 1985). Therefore, the inclusion of marginal areas/sites would lead to the inclusion of low frequency species that would influence the species extrapolation curves and analyses of endemism.

The extrapolation curves were based on the Chao2 estimator derived for incidence (frequency) data (Chao 1987; Chao et al. 2009, 2014). This estimator is based on the number of species/taxa registering only one (singletons) or two occurrences (doubletons) in a group of samples (Chao et al., 2009). Other sample-based estimators such as Chao1, Jackknife1, Jackknife2 and ICE (incidence coverage estimator) also generate diversity estimations based on numbers of singletons and doubletons. The Chao2 estimator also considers the total number of samples and the ratio between total species richness and rare species (singletons and doubletons). Studies do not agree on which are the best diversity estimators to use as their behaviour can vary according to factors such as sample size and sampling effort. In here, I use the Chao2 estimator as this is the estimator recommended by Colwell et al. (2012) when working with incidence data with a high number of rare species. Also, for Chao2, Cowell et al. (2012) and Chao et al. (2014) have developed the means to estimate confidence intervals based on the Bernoulli probability distributions. The tree species richness estimates obtained via the Chao2 approach should be understood as being minimum estimates of diversity.

The extrapolation curves were created in the R software with the package “iNEXT”, by using the iNEXT () function (Hsieh, Ma, and Chao 2016). Following recommendations made by Chao et al. (2014) and Hsieh, Ma, and Chao. (2016), all core areas had their tree species richness extrapolated to twice the sample size of sites within each core area.

4.3.4 - Tree species endemism

Even though endemism is a simple concept in ecology, being defined as a taxon or lineage which can only be encountered within a geographic delimited area/region. Its application can be rather difficult. Biome’s geographic delimitations can be a subject of contention among studies, and this will have a direct impact on estimates of endemism. In this contribution, I estimate endemism levels for core areas only, as the limits of these are better defined in the literature and represent continuous areas of biome occurrence.

I estimate endemism levels based on incidence/frequency data. With these types of data, it is not possible to distinguish whether species are definitely established in a biome or are present due to mass effects or episodic dispersal events from other biomes. Because of this, I performed two analyses related to endemism: an indicator species analyses (described in the following section) and a count of shared and unique tree species for each biome. This was done in order to produce an upper bound (indicator species) and a lower bound (unique/shared species counts) of endemism estimates, with the true number of endemic species likely lying between these two bounds. The indicator species analyses will identify all species that are significantly, positively associated with a given biome, but these species could be statistically identified as such and still occur at low frequencies in other biomes.

Therefore, this analyses provides an upper bound for estimates of the number of endemic species. Meanwhile, using simple occurrence in one biome with zero occurrences in any other biome represents a much stricter interpretation of endemism. A species may be essentially present in only one biome, with a rare occurrence in another biome due to mass effects (i.e., it occurs in another biome, but could not successfully reproduce or establish a population there) or due to taxonomic misidentification (i.e. not even a real occurrence!). Thus, strict counts of species entirely restricted to one biome would underestimate endemism and represent a lower bound on estimates of the number of endemic species.

4.3.5 - Indicator Species Analyses

To determine species biome affiliation, we used a modification of the phi coefficient of (Tichy and Chytrý 2006) that is based on presence/absence data and accounts for variation in sampling effort amongst groups. Specifically, we used the the r^g correlation index of de Cáceres and Legendre (2009) that varies from -1 to 1, with positive values indicating a non-random association of a species with a specified set of sites, in our case those belonging to a single biome. To determine diagnostic/endemic species, we focused our analytical workflow on the positive associations (values from 0 to 1) between species and groups only. To test if the associations between a given species and biomes were significant, we randomized occurrences across sites a thousand times and assessed if a species was found more or less frequently in a biome than expected by chance (using a 0.05 alpha significance threshold). The species with a significant and positive association with any biome are the indicator/diagnostic species for that biome. When more than two biomes are being considered (and here we consider four), a species can be positively associated with more than one biome. I therefore tallied species that indicated only a single biome as well as species that indicated the focal biome and another biome (shared indicator species, see below). Indicator species analyses were conducted using functions in the “indicspecies” package (strassoc() and signassoc() functions) for the R Statistical Software (De Cáceres & Legendre, 2009). Indicator/diagnostic species calculations were based on tree species checklists for the core areas of each biome only.

4.4 - Results

I have registered a total of 11,129 tree species across the four core areas encompassed in this study and a total of 11,354 tree species when the extended areas are considered. The core areas of the Amazon and the Atlantic Forests are the most rich in tree species in Lowland tropical South America, with 7845 and 3963 tree species respectively (Table 4.1). Among the dry biomes' core areas, Savanna is the most rich in tree species, with a total of 2008 tree species, while SDTF's core area registers 1031 tree species (Table 4.1). Most of the diversity of these biomes is found in the core areas. Apart from the Amazon Forest, the marginal areas of all biomes significantly contribute to their tree species richness (Table 4.1; Figure 4.2). The marginal areas of the Atlantic Forest, Savanna and SDTF increase these biomes species totals by 992, 1542 and 814 species respectively. On the other hand, Amazon Forest's marginal areas only add an extra 188 tree species to this biome's tree species richness.

4.4.1 - Species accumulation and extrapolation curves

At twice the current level of sampling, visual exploration of the tree species extrapolation curves constructed by using the Chao2 shows that the Amazon Forest and the Savanna core areas do not reach an asymptote, while those of the Atlantic Forest and SDTF appear to (Figure. 4.3). Also at twice the current level of sampling, the Amazon Forest's core area is estimated to have 8353 tree species (+- 48), the Atlantic Forest's core area has 4171 tree species (+-33), the Savanna's core area has 2592 tree species (+-70) and the SDTF's core area has 1110 tree species (+-18). These minimum estimates indicate that NeoTropTree has registered around 93 to 94% of the Atlantic, Amazon and SDTF's tree floras and, at least, 77.5% of the Savanna's tree flora. The confidence interval for the Savanna's estimate indicates a higher uncertainty surrounding this estimation in relation to the other biomes' core areas.

4.4.2 - Endemic Species – core areas

In general, the wetter biomes’ core areas register more endemic species than the drier biomes’ core areas for both upper (indicator species analyses) and lower (species counts) bound estimates. The Amazon Forest’s number of endemic tree species lies between 4844 and 5476 species, which is approximately 70% of the core area’s tree species richness. The Amazon Forest registers the highest level of endemism out of all biomes analyzed here. The Atlantic Forest’s number of endemic tree species lies between 2214 and 2484, which is less than the Amazon, but higher than the drier biomes’ core areas. The endemic tree species in the Atlantic Forest account for approximately 63% of its total species richness. Core Savanna and core SDTF register similar total estimates of tree species endemism, although the proportional endemism is twice as high in SDTF. The former has around 163 to 404 endemic tree species (approximately 20% of its total tree species richness) and the latter has around 141 to 417 endemic tree species (approximately 40% of its total species richness). Considered together, the moist core areas (Atlantic + Amazon) register a total of 10750 tree species, out of which 6204-8726 are unique to moist forests. The drier core areas (Savanna + SDTF) together register 2403 tree species in total out of which 379-809 species can only be found in these environments (Figure 4.4).

Table 4.1: Tree species richness, number of indicator/diagnostic species and estimated number of tree species for the four main biomes in lowland tropical South America.

Biomes	Amazon Forest	Atlantic Forest	Savanna	Seasonally Dry Tropical Forest (SDTF)
Tree species richness (core + marginal sites)	8033	4955	3550	1845
Tree species richness (core sites)	7845	3936	2008	1031
Tree species richness (marginal sites)	188	1019	1542	814
Estimated tree species richness (core areas)	8353	4171	2592	1110
Minimum estimate of unsampled richness (Chao2 - core area)	508	235	584	79
Proportion of total species that have been sampled in core area of biome	94%	94%	77.50%	93%

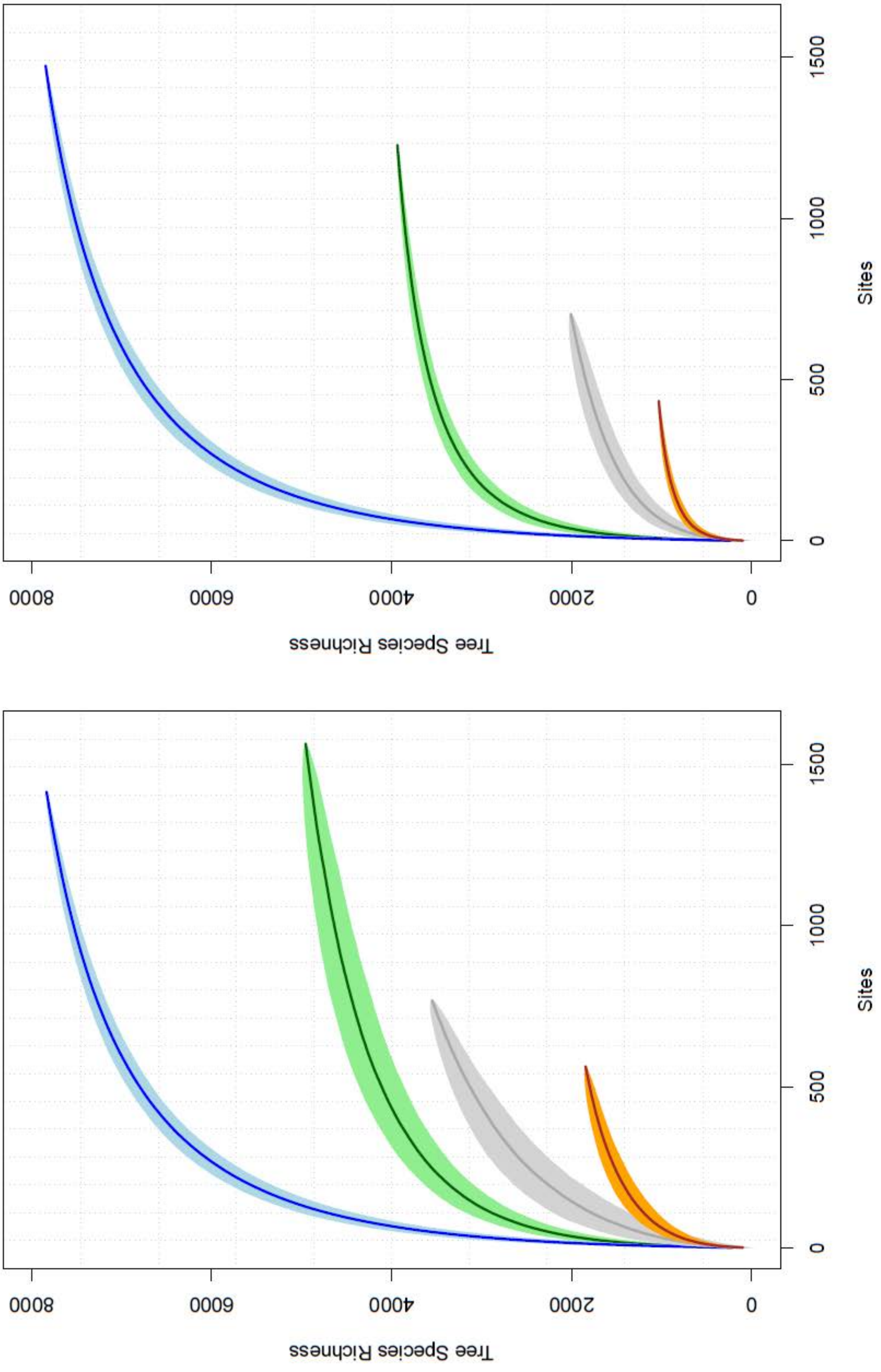

Figure 4.2: Tree species accumulation curves for the main lowland tropical South American biomes. (a) shows tree species accumulation curves for the biome's extended area of occurrence and (b) shows tree species accumulation curves for these same biomes' core areas. [blue – Amazon Forest; Green – Atlantic Forest *senso-latissimo/senso-lato*; Brown – dry-diagonal seasonally dry tropical forest/Caatinga; Gray – dry-diagonal and Amazonic savannas/Cerrado]

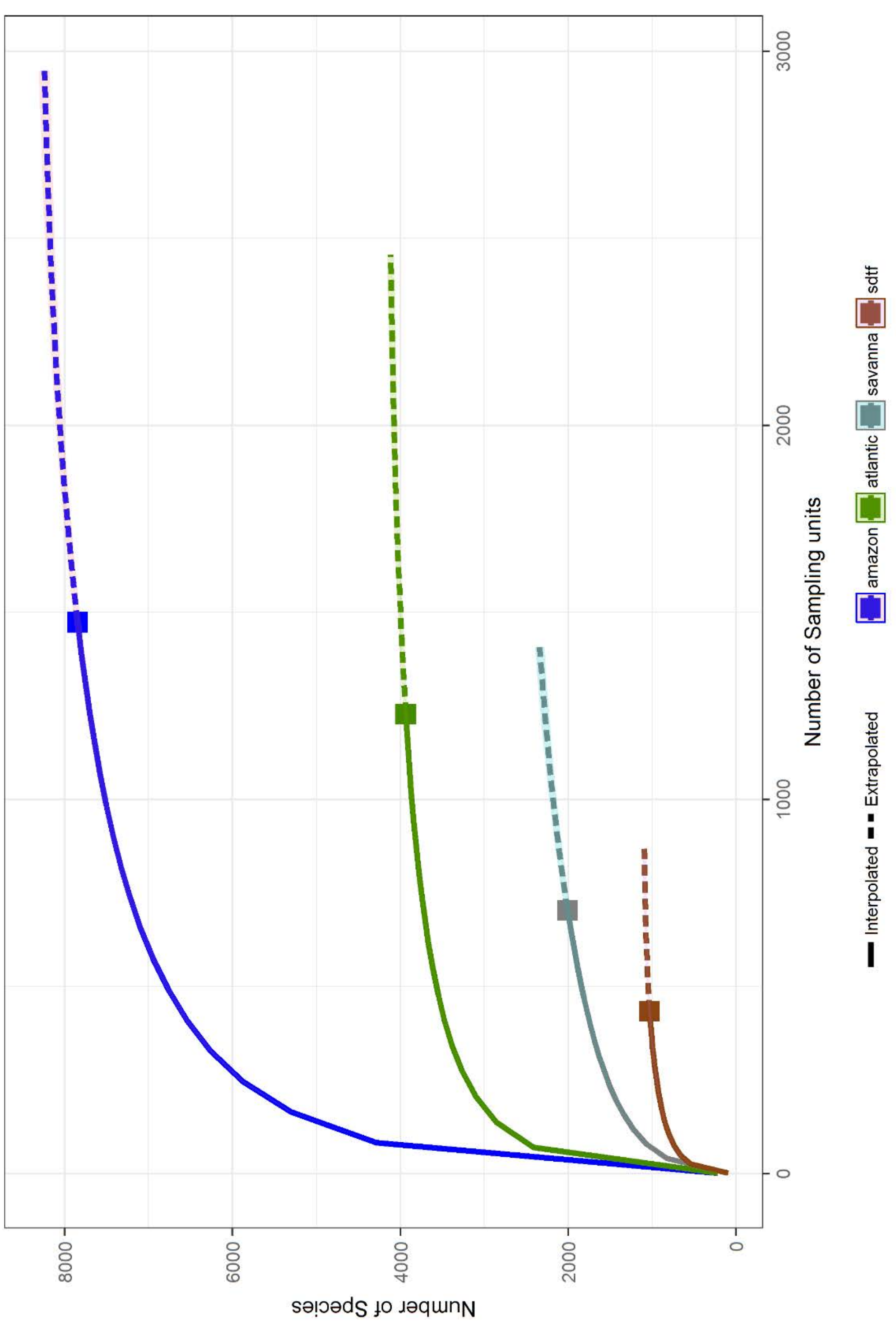

Figure 4.3: Tree species accumulation and extrapolation curves for the four main biomes of lowland tropical South America. Curves were constructed using the Chao2 diversity estimator. See text for more details

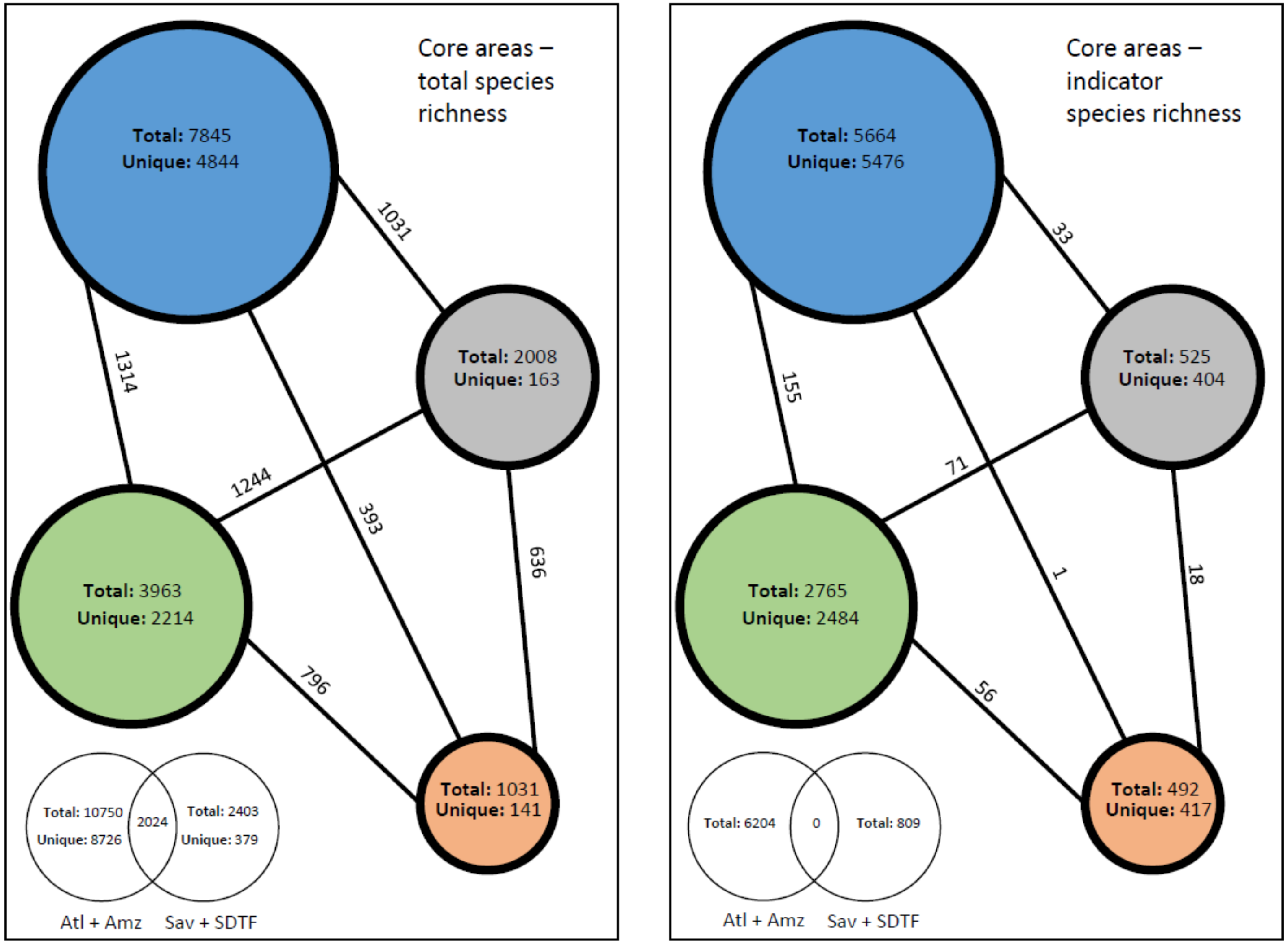

Figure 4.4: Diagram charts representing the number of total/unique species (numbers in balloons) in each biome's core area and shared species (numbers over lines). (a) gives values related to all tree species in each core area and (b) shows values referring to the number of indicator species for each biome determined through a statistical, indicator species analyses (please refer to text for details. [blue – Amazon Forest; Green – Atlantic Forest *senso-latissimo*/*senso-lato*; Brown – dry-diagonal seasonally dry tropical forest/Caatinga; Gray – dry-diagonal and Amazonic savannas/Cerrado].

4.5 - Discussion

The high tree species richness and endemism levels of the Amazon and Atlantic Forests indicate that patterns of tree species richness and distribution are indeed related to drought sensitivity. The fact that more than 60% of the Amazon Forest's tree flora is composed of endemic species highlights the importance of biogeographic history in shaping patterns of tree species distribution and supports the idea of the Amazon being a "diversity cradle" within LTSA. Importantly, all biomes apart from the Savannas have high and similar sampling sufficiency levels, indicating the reliability of these results and of the comparisons stemming from them. As hypothesized, biomes' geographically marginal areas are an important reservoir of tree species diversity and their inclusion in the analyses increased all biomes' tree species richness levels. However, only 225 tree species are unique to these geographically marginal areas (not present in the core areas), indicating that the richness surplus observed in the individual biomes is made up of species that have already been registered in the other biomes analyzed.

4.5.1 - Total and estimated tree species richness

The Amazon Forest's core area has 7845 tree species with between 4844 and 5476 being endemic to this region. The closest value to the one found here was obtained by Cardoso et al. (2017), who registered 6627 tree species for the entire Amazon. These authors compiled a taxonomically thoroughly checked tree species checklist for the same area of the Amazon I have delimited, meaning that this difference is likely related to their rather strict definition of trees in comparison to NTT's definition. Cardoso et al. (2017) defined trees as woody plants that have a diameter at breast height greater than 10cm, whereas NTT defines trees as free-standing woody plants that can achieve 3 m in height and does not establish a DBH threshold.

Importantly, when studying floristic patterns in the Amazon Forest, Oliveira-Filho et al. (*in prep.*) cross-checked NTT's tree species checklist for this area against the checklist produced by Cardoso et al. (2017) and standardized NTT's checklist for this biome. This ensures that differences observed in this contribution are not related to differences in taxonomic standardization. The numbers produced by Lista de espécies da flora do Brasil 2020 (2018) for this biome are also aligned with our tree species count. For Brazil only, they registered a total of 4775 tree species, which shows that many Amazonian tree species can only be found outside Brazil (e.g. Colombia, Ecuador and Peru). This fact points to the importance of transcending national boundaries in order to obtain accurate information for research and conservation purposes.

The estimated tree species richness for the Amazon Forest's core area in twice the number of samples that I analysed is 8353, which shows that there are at least 508 species that have not yet been recorded in this region. Even though this number is higher than the one provided by Cardoso et al. (2017) and lower than the estimates produced by Hubbell et al. (2008) and ter Steege et al. (2013). This is most likely related to differences in how these estimates were obtained. I used the Chao2 estimator (adapted to incidence data), which gives a minimum estimate of species richness, and Hubbel et al. (2008) and ter Steege et al. (2013) fitted a Fisher's log-series model to abundance data. The former is a non-parametric estimator and the latter is a parametric one. Hubbel et al. (2008) estimated approximately 12,500 tree species for the whole Amazon and that the Brazilian Amazon alone would have ~11,210 tree species (Table 4.2). Also by fitting Fisher's log-series models to species abundance data gathered from the ATDN plot network, ter Steege et al. (2013) encountered similar tree species richness for the Amazon. In their 2013 study, they estimated a total of ~ 16,000 tree species for the entire biome (Table 4.2).

Fisher's log-series is considered a reasonable option to estimate total species richness of an ecosystem (Hubbell et al., 2015; ter Steege et al., 2017), because of its effectiveness at different scales and in dealing with species rarity. However, this model is based on neutral theory (Hubbell 2001), showing that there are unrealistic assumptions underlying this method. Importantly, this model can only be applied to abundance data, which makes it difficult to apply to regions and taxonomic groups with poor past sampling history. Such high quality abundance data exists for the Amazon Forest (e.g ATDN - http://atdn.myspecies.info/), but not for the other biomes of LTSA. Therefore, in a context such as this one – high quality incidence (frequency) data spanning over large biogeographic limits – non-parametric estimators such as Chao2 are more useful and viable (Chao et al., 2009; Cowell et al., 2012; Chao et al., 2014) and enable comparison of estimations among biomes even when there is a difference in coverage.

As for the other biomes' core areas (Atlantic Forest, Savanna and SDTF), only a handful of studies have aimed to estimate their overall tree species richness. Françoso, Haidar, and Machado (2016) estimated the tree species richness of the Brazilian Cerrado (part of Savanna core area delimited here) by relying on these same well-known non-parametric estimators: Jacknife1 and 2, Chao 1 and 2 and Bootstraping (Magurran 2003; Chao et al. 2009). By relying on a dataset encompassing 909 observed tree species, these authors determined that the probable number of tree species ranges from 1019 to 1233 (Table 4.2). These numbers are much lower than what I estimated. Françoso et al. (2016) based their estimates on an updated version of the dataset used by Ratter et al. (2003), which was created as the means to study biogeographic patterns within the Brazilian Cerrado. To that end, the authors rightfully eliminated rare occurrences of species known to not be typical of the Brazilian Cerrado according to their previous experience. The elimination of such records and the fact that Françoso et al. (2016) did not include the Bolivian Cerrado may underlie the divergence between tree species estimations made here and by the aforementioned authors.

Contrary to expectations, the Savanna's core area presented the highest proportion of unsampled tree species richness alongside with the Amazon Forest's core area. This is most likely related to two factors: 1) savanna's core area borders all three other biomes encompassed in this study plus the Chaco; 2) patches of Amazon and Atlantic forests and SDTFs are present within savanna's core area, as gallery forests or otherwise. Both factors extend the possibilities of the savanna harboring tree species coming from surrounding biomes through mass effects and dispersal. Since the estimator I used – Chao2 – is affected strongly by rare occurrences (species present in one or two samples only), species maintained by mass effects have the potential to increase the gap between observed and total species richness, and thus inflate the number of unobserved species.

By also using Jacknife1 and 2 and Chao 1 and 2 and bootstrapping, Moro et al. (2014) estimated the probable number of woody species present in the Caatinga (SDTF core area delimited here). These authors utilized a dataset registering 1714 woody species for this biome and estimated its probable woody species richness to range from 2475 to 2921 (Table 4.2), which is much higher than my estimate for the same biome. This is mostly related to Moro (2014) focusing on all woody species and my analysis focusing on trees only (excluding shrubs and lianas). Also for this biome, the lista de espécies da flora do Brasil 2020 (2018) registers 901 tree species for the Caatinga, while DRYFLOR (2016) counts 1112 tree species in their dataset (Table 2). Both numbers are similar to my estimates.

Surprisingly, only Neves et al. (2017) brings information on tree species richness counts and estimates for the whole of the Atlantic Forest (core + marginal areas) that also encompass the patches of this biome spanning over Paraguay and northern Argentina. However, their species count is entirely based on NTT data as well and thus not independent from my estimate in this study. On the other hand, there are very good estimates of flowering plant species richness for the Brazilian Atlantic Forest (Table 4.2). However, only lista de espécies da flora do Brasil 2020 (2018) provides a tree species count for this biome. They register 3372 tree species in the Brazilian Atlantic Forest, which is, as expected, a lower number in comparison to what I have encountered because of the differences in geographic circumscription.

Another point that must be emphasized when discussing species richness estimates is the total area covered by the species assemblages being assessed. In our case, even though the Amazon forest harbors more tree species than the Atlantic forest, the former has a species density of 0.001 species/km2, whereas the later has a species density of 0.003 species/km2. This shows that, at the present moment, the Atlantic forest (core + marginal) harbors more tree species per unit of area than the Amazon forest. The same can be observed for the core SDTF and core Savanna. Even though the core SDTF has around half the number of tree species encountered in the core Savanna, core SDTF has 0.001 species/km2 while core Savanna's tree species density is 0.0009 species/km2.

Regardless of species density, some general trends can be drawn, especially for Brazil. According to lista de espécies da flora do Brasil 2020 (2018), Brazil presents a similar pattern to what I have observed for LTSA – the Brazilian Amazon has more species than the Brazilian Atlantic Forest and the Brazilian core Savanna has more species than the core SDTF. This trend persists when shrub species richness is also taken into account. However, this changes when Angiosperm diversity is taken into account. Zappi et al. (2015) shows that Angiosperm diversity in the Atlantic Forest and in the Cerrado is higher than in the Amazon Forest, a different pattern than what I observed here. This highlights the importance of producing species counts and estimations considering the biome's full range of occurrence and not limit analyses to political boarders.

4.5.2 - Marginal areas and species richness

The marginal areas of Savanna, Atlantic Forest and SDTF harbor a significant number of tree species that are not encountered in the core areas. This finding highlights these marginal area's capacity of maintaining high levels of species diversity, including species that are not encountered in their respective biomes' core areas. However, there are only 225 tree species that are unique to these marginal areas and are not found in the core areas analyzed. Therefore, tree species richness of biomes' marginal areas seem to be mainly composed of species that can also be found in other biomes. Only a small proportion of these areas' richness is composed of endemic species. Even though these marginal areas' endemism levels are low at the continental scale, they contribute to high beta-diversity within biomes and exemplify LTSA's complex biogeography.

4.5.3 - Endemic species

Through an indicator species analyses and by quantifying the number of unique and shared species among biomes’ core areas, I was able to estimate upper and lower bounds on the number of endemic tree species in the core areas’ of Amazon Forest, Atlantic Forest, Savanna and SDTF. As predicted, the Amazon and the Atlantic forests have more endemic tree species than the Savanna and the SDTF, which is also aligned to total tree species richness observed for these areas. Approximately 60 to 70% of the moist core areas’ tree richness is composed of endemics, whereas drier core areas’ tree endemism levels range from 20% to 40%. This indicates that endemism levels must also be related to differences in drought tolerance levels in association with climatic and topographic complexities as proposed by Esquivel-Mulbert et al. (2016), Rangel et al. (2018) and Antonelli et al. (2018).

However, factors such as geographic positioning must also be considered when interpreting these results. In LTSA, the Amazon and the Atlantic Forests are separated from each other by the Caatinga (core SDTF), the Cerrado (core Savanna) and the Chaco – the tree regions forming the diagonal of dry formations (Vanzolini 1963). The two moist environments have been separated from each other and the only routes for plant species to migrate between them is through the “brejos de altitude” in northeast Brazil and the gallery forests going across the Cerrado domain (Oliveira-Filho and Ratter 1995). Therefore, isolation and speciation could be important factors behind these environments high rate of endemism.

The SDTF and Savanna’s core areas’ low levels of tree endemism are also probably related to geography. Unlike the moist biomes, these two core areas border each other and share a large proportion of their tree species richness (Figure 4.4). Mass effects and species migration between biomes could then be the two main reasons for the low tree species endemism observed in these two biomes, although even together, their estimates of tree species endemism are markedly lower than the combined estimate for the moist biomes.

Table 4.2: Tree species totals and estimations gathered from recent literature. Differences in geographic limits, taxonomic/habit groups are advised when necessary.

	Amazon Forest	**Atlantic Forest**	**Savanna**	**Seasonally Dry Tropical Forest (SDTF)**
This study (2018)	8033 (core + marginal), 7845 (core)	4955 (core + marginal), 3936(core)	3550 (core + marginal), 2008 (core)	1845 (core + marginal), 1031 (core)
Zappi et al. (2015)	11896 species of Angiosperms in the Brazilian Amazon	15001 species of Angiosperms in the Brazilian Atlantic Forest	12097 species of Angiosperms (Brazilian Cerrado)	4657 species of Angiosperms (Caatinga)
Lista de espécies da flora do Brasil 2020 (2018)	4775 tree species in the Brazilian Amazon	3372 tree species in the Brazilian Atlantic Forest	1833 tree species in the Brazilian Cerrado	901 species of trees in the Caatinga
Stehmann et al. (2009)	***	13708 species of Angiosperms (*senso-lato* - Brazilian)	***	***
Domingos et al. (2017)	6627 species of trees and 14003 species of Angiosperms in the whole Amazon	***	***	***
Hubbel et al. (2008)	12500 tree species in the whole Amazon Forest and 11210 tree species in the Brazilian Amazon	***	***	***

Table 4.2: Tree species totals and estimations gathered from recent literature. Differences in geographic limits, taxonomic/habit groups are advised when necessary.

	Amazon Forest	**Atlantic Forest**	**Savanna**	**Seasonally Dry Tropical Forest (SDTF)**
Moro et al. (2014)	***	***	***	1714 woody plant species recorded. Number of woody plant species was also estimated - ICE index: 2639, Chao 2: 2681, Jackknife 1: 2475, Jackknife 2: 2921.
Françoso et al. (2016)	***	***	909 tree species in the Brazilian Cerrado. Number of tree species was also estimated - Chao 1: 1223, Jackknife 1: 1159, Jackknife: 1308, Bootstrap: 1019.	***
Neves et al. (2017)		4431 tree species registered in their dataset		
Klink & Machado (2005)	***	***	~7000 species of plants registered in the Brazilian Cerrado	***

Table 4.2: Tree species totals and estimations gathered from recent literature. Differences in geographic limits, taxonomic/habit groups are advised when necessary.

	Amazon Forest	**Atlantic Forest**	**Savanna**	**Seasonally Dry Tropical Forest (SDTF)**
ter Steege et al. (2016)	11187 valid tree species names retrieved from collections made during 1707 till 2015.	***	***	***
DRYFLOR (2016)	***	***	***	1112 tree species in the Caatinga and 1334 species in SDTF enclaves in the Cerrado
Ratter et al. (2003)	***	***	951 tree species in the Brazilian Cerrado	***
ter Steege et al. (2013)	4962 tree species in dataset and ~16000 estimated tree species in the whole Amazon through Fisher's alpha based in a log-series	***	***	***

Biogeographic history must also be taken into account. The SDTFs are known for their species endemism and beta diversity, but I have only counted 141-417 endemic tree species in its core area – the Brazilian Caatinga. According to Queiroz (2006), this region can be further divided into two main formations in relation to geology – crystalline Caatinga and quartzitic/sandy Caatinga. The former occupies 73% of Caatinga's area of occurrence and the latter the remaining 27%. Crystalline Caatinga was formed during the pediplanation (forming of plains due to sediment loss of pediments) that occurred in this area during the Tertiary period, which allowed this area to be occupied by SDTF elements from the surrounding patches. Quartzitic/sandy Caatinga, however, shelters a much older vegetation and plant lineages and, therefore, has a higher rate of endemism. According to Queiroz (2006), this replacement of old lineages by younger ones is the reason why Caatinga does not harbour many endemic species. However, his study focused only on legumes and does not provide an accurate estimate of endemism. On the other hand, savanna's core area is known for its rather low tree species turnovers and homogeneous tree species distribution across its extent (Oliveira-Filho and Ratter 2002; Bridgewater, Ratter, and Ribeiro 2004). This, in association with the fact that this core area is the one region in LTSA that is surrounded by all other biomes, may explain this biome's low level of tree endemism.

4.6 - Conclusions

In this contribution, I provided evidence supporting that tree species richness and endemism levels of LTSA biomes are most likely related to drought-sensitivity – the Amazon and Atlantic Forests are more tree species rich and host a higher number of endemics than the Savanna and the SDTF. This reveals that differences in tree species richness among LTSA biomes are not only related to turnover, they are also related to nestedness differences in total precipitation and precipitation seasonality. Tree species richness estimates have revealed that an increase in sampling/collection effort in the Amazon Forest can potentially lead to new species discoveries and/or new occurrence records for species from other biomes. Overall, the cross biome/core area comparison points to the necessity of proposing conservation initiatives that are able to function at an almost continental-scale and encompass geographically marginal areas, which are often forgotten from such plans. For example, even though Brazil has the largest Amazon Forest cover in LTSA, it seems to hold only half of this biome's tree species richness. Even though Savannas and SDTFs have a smaller number of endemic species in comparison to the moist biomes, their levels of endemism are also meaningful and justify specific conservation efforts and policy tailored to these environments. I also emphasize that the present contribution does not encompass the entirety of the LTSA SDTFs and savannas, and further studies are needed in order to include more of these biomes' patches and marginal are.

4.7 - References

Antonelli, Alexandre, and Isabel Sanmartín. 2011. “Why Are There so Many Plant Species in the Neotropics?” *Taxon* 60 (2): 403–14.

Antonelli, Alexandre, Alexander Zizka, Fernanda Antunes Carvalho, Ruud Scharn, Christine D. Bacon, Daniele Silvestro, and Fabien L. Condamine. 2018. “Amazonia Is the Primary Source of Neotropical Biodiversity.” *Proceedings of the National Academy of Sciences*, May, 201713819. https://doi.org/10.1073/pnas.1713819115.

Arita, Héctor T., and Ella Vázquez-Domínguez. 2008. “The Tropics: Cradle, Museum or Casino? A Dynamic Null Model for Latitudinal Gradients of Species Diversity.” *Ecology Letters* 11 (7): 653–63. https://doi.org/10.1111/j.1461-0248.2008.01197.x.

Bridgewater, Samuel, James A Ratter, and Jose Felipe Ribeiro. 2004. “Biogeographic Patterns, b-Diversity and Dominance in the Cerrado Biome of Brazil.” *Biodiversity and Conservation* 13: 2295–2318.

Bueno, Marcelo Leandro, Kyle G. Dexter, R. Toby Pennington, Vanessa Pontara, Danilo Mesquita Neves, James Alexander Ratter, and Ary Teixeira de Oliveira-Filho. 2018. “The Environmental Triangle of the Cerrado Domain: Ecological Factors Driving Shifts in Tree Species Composition between Forests and Savannas.” *Journal of Ecology* 106 (5): 2109–20. https://doi.org/10.1111/1365-2745.12969.

Cáceres, Miquel De, and a N D Pierre Legendre. 2009. “Associations between Species and Groups of Sites: Indices and Statistical Inference.” *Ecology* 90 (12): 3566–3574. https://doi.org/10.1890/08-1823.1.

Cardoso, Domingos, Tiina Särkinen, Sara Alexander, André M. Amorim, Volker Bittrich, Marcela Celis, Douglas C. Daly, et al. 2017. “Amazon Plant Diversity Revealed by a Taxonomically Verified Species List.” *Proceedings of the National Academy of Sciences* 114 (40): 10695–700. https://doi.org/10.1073/pnas.1706756114.

Chao, Anne. 1987. “Estimating the Population Size for Capture-Recapture Data with Unequal Catchability.” *Biometrics* 43 (4): 783–91. https://doi.org/10.2307/2531532.

Chao, Anne, Robin L. Chazdon, Robert K. Colwell, and Tsung-Jen Shen. 2004. “A New Statistical Approach for Assessing Similarity of Species Composition with Incidence and Abundance Data: A New Statistical Approach for Assessing Similarity.” *Ecology Letters* 8 (2): 148–59. https://doi.org/10.1111/j.1461-0248.2004.00707.x.

Chao, Anne, Robert K. Colwell, Chih-Wei Lin, and Nicholas J. Gotelli. 2009. “Sufficient Sampling for Asymptotic Minimum Species Richness Estimators.” *Ecology* 90 (4): 1125–33. https://doi.org/10.1890/07-2147.1.

Chao, Anne, Nicholas J. Gotelli, T. C. Hsieh, Elizabeth L. Sander, K. H. Ma, Robert K. Colwell, and Aaron M. Ellison. 2014. “Rarefaction and Extrapolation with Hill Numbers: A Framework for Sampling and Estimation in Species Diversity Studies.” *Ecological Monographs* 84 (1): 45–67. https://doi.org/10.1890/13-0133.1.

Chen, Youhua, and Tsung-Jen Shen. 2017. "Rarefaction and Extrapolation of Species Richness Using an Area-Based Fisher's Logseries." *Ecology and Evolution* 7 (23): 10066–78. https://doi.org/10.1002/ece3.3509.

Colwell, Robert K., Anne Chao, Nicholas J. Gotelli, Shang-Yi Lin, Chang Xuan Mao, Robin L. Chazdon, and John T. Longino. 2012. "Models and Estimators Linking Individual-Based and Sample-Based Rarefaction, Extrapolation and Comparison of Assemblages." *Journal of Plant Ecology* 5 (1): 3–21. https://doi.org/10.1093/jpe/rtr044.

Dryflor, Karina Banda-R, Alfonso Delgado-Salinas, Kyle G. Dexter, Reynaldo Linares-Palomino, Ary Oliveira-Filho, Darién Prado, et al. 2016. "Plant Diversity Patterns in Neotropical Dry Forests and Their Conservation Implications." *Science* 353 (6306): 1383–87. https://doi.org/10.1126/science.aaf5080.

Esquivel-Muelbert, Adriane, Timothy R. Baker, Kyle G. Dexter, Simon L. Lewis, Hans ter Steege, Gabriela Lopez-Gonzalez, Abel Monteagudo Mendoza, et al. 2017. "Seasonal Drought Limits Tree Species across the Neotropics." *Ecography* 40 (5): 618–29. https://doi.org/10.1111/ecog.01904.

Fiaschi, Pedro, and José R. Pirani. 2009. "Review of Plant Biogeographic Studies in Brazil." *Journal of Systematics and Evolution* 47 (5): 477–96. https://doi.org/10.1111/j.1759-6831.2009.00046.x.

"Flora Do Brasil 2020." n.d. http://floradobrasil.jbrj.gov.br/reflora.

Forzza, Rafaela C., José Fernando A. Baumgratz, Carlos Eduardo M. Bicudo, Dora A. L. Canhos, Anibal A. Carvalho, Marcus A. Nadruz Coelho, Andrea F. Costa, et al. 2012. "New Brazilian Floristic List Highlights Conservation Challenges." *BioScience* 62 (1): 39–45. https://doi.org/10.1525/bio.2012.62.1.8.

Françoso, Renata D., Ricardo F. Haidar, and Ricardo B. Machado. 2016. "Tree Species of South America Central Savanna: Endemism, Marginal Areas and the Relationship with Other Biomes." *Acta Botanica Brasilica* 30 (1): 78–86. https://doi.org/10.1590/0102-33062015abb0244.

Fundação SOS MAta Atlântica. 2016. "SOS Mata Atlântica: Relatório Anual 2016." São Paulo.

Gentry, Alwyn H. 1988. "Changes in Plant Community Diversity and Floristic Composition on Environmental and Geographical Gradients." *Annals of the Missouri Botanical Garden* 75 (1): 1. https://doi.org/10.2307/2399464.

Graham, Catherine H., Craig Moritz, and Stephen E. Williams. 2006. "Habitat History Improves Prediction of Biodiversity in Rainforest Fauna." *Proceedings of the National Academy of Sciences* 103 (3): 632–36. https://doi.org/10.1073/pnas.0505754103.

Hsieh, T. C., K. H. Ma, and Anne Chao. 2016. "INEXT: An R Package for Rarefaction and Extrapolation of Species Diversity (Hill Numbers)." *Methods in Ecology and Evolution* 7 (12): 1451–56. https://doi.org/10.1111/2041-210X.12613.

Hubbell, S. P., F. He, R. Condit, L. Borda-de-Agua, J. Kellner, and H. ter Steege. 2008. "How Many Tree Species Are There in the Amazon and How Many of Them Will Go Extinct?" *Proceedings of the National Academy of Sciences* 105 (Supplement 1): 11498–504. https://doi.org/10.1073/pnas.0801915105.

Hubbell, Stephen P. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton: Princeton University Press.

Instituto Brasileiro de Geografia e Estatística, ed. 2012. *Manual técnico da vegetação brasileira*. 2a edição revista e ampliada. Manuais técnicos em geociências, número 1. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística-IBGE.

Jiménez, Iván, Trisha Distler, and Peter M. Jørgensen. 2009. "Estimated Plant Richness Pattern across Northwest South America Provides Similar Support for the Species-Energy and Spatial Heterogeneity Hypotheses." *Ecography* 32 (3): 433–48. https://doi.org/10.1111/j.1600-0587.2008.05165.x.

Joppa, Lucas N., David L. Roberts, and Stuart L. Pimm. 2010. "How Many Species of Flowering Plants Are There?" *Proceedings of the Royal Society of London B: Biological Sciences*, July, rspb20101004. https://doi.org/10.1098/rspb.2010.1004.

Magurran, Anne E. 2003. *Measuring Biological Diversity*. 1 edition. Malden, Ma: Wiley-Blackwell.

Mittermeier, Russell a., Gustavo a B Da Fonseca, Anthony B. Rylands, and Katrina Brandon. 2005. "A Brief History of Biodiversity Conservation in Brazil." *Conservation Biology* 19 (3): 601–7. https://doi.org/10.1111/j.1523-1739.2005.00709.x.

Moro, Marcelo Freire, Eimear Nic Lughadha, Denis L. Filer, Francisca Soares de Araújo, and Fernando Roberto Martins. 2014. "A Catalogue of the Vascular Plants of the Caatinga Phytogeographical Domain: A Synthesis of Floristic and Phytosociological Surveys." *Phytotaxa* 160 (1): 1–118. https://doi.org/10.11646/phytotaxa.160.1.1.

Myers, Norman, Russell A. Mittermeier, Cristina G. Mittermeier, Gustavo A. B. da Fonseca, and Jennifer Kent. 2000. "Biodiversity Hotspots for Conservation Priorities." *Nature* 403 (6772): 853–58. https://doi.org/10.1038/35002501.

Neves, Danilo M., Kyle G. Dexter, R. Toby Pennington, Arthur S. M. Valente, Marcelo L. Bueno, Pedro V. Eisenlohr, Marco A. L. Fontes, et al. 2017. "Dissecting a Biodiversity Hotspot: The Importance of Environmentally Marginal Habitats in the Atlantic Forest Domain of South America." *Diversity and Distributions* 23 (8): 898–909. https://doi.org/10.1111/ddi.12581.

Oliveira, Ubirajara, Adriano Pereira Paglia, Antonio D. Brescovit, Claudio J. B. de Carvalho, Daniel Paiva Silva, Daniella T. Rezende, Felipe Sá Fortes Leite, et al. 2016. "The Strong Influence of Collection Bias on Biodiversity Knowledge Shortfalls of Brazilian Terrestrial Biodiversity." *Diversity and Distributions* 22 (12): 1232–44. https://doi.org/10.1111/ddi.12489.

Oliveira-Filho, A. T., and J. A. Ratter. 1995. “A Study of the Origin of Central Brazilian Forests by the Analysis of Plant Species Distribution Patterns.” *Edinburgh Journal of Botany* 52 (2): 141–94. https://doi.org/10.1017/S0960428600000949.

Oliveira-Filho, Ary T. 2017. “NeoTropTree, Flora Arbórea Da Região Neotropical: Um Banco de Dados Envolvendo Biogeografia, Diversidade e Conservação.” NeoTropTree, Flora Arbórea Da Região Neotropical: Um Banco de Dados Envolvendo Biogeografia, Diversidade e Conservação. 2017. http://www.neotroptree.info/.

Oliveira-Filho, Ary T., Jean C. Budke, João A. Jarenkow, Pedro V. Eisenlohr, and Danilo R. M. Neves. 2015. “Delving into the Variations in Tree Species Composition and Richness across South American Subtropical Atlantic and Pampean Forests.” *Journal of Plant Ecology* 8 (3): 242–60. https://doi.org/10.1093/jpe/rtt058.

Oliveira-Filho, Ary T., and Marco Aurélio L. Fontes. 2000. “Patterns of Floristic Differentiation among Atlantic Forests in Southeastern Brazil and the Influence of Climate1.” *Biotropica* 32 (4b): 793–810. https://doi.org/10.1111/j.1744-7429.2000.tb00619.x.

Oliveira-Filho, Ary T., and James A. Ratter. 2002. “Vegetation Physigonomies and Wood Flora of the Bioma Cerrado.” In *The Cerrados of Brazil: Ecology and Natural History of a Neotropical Savamma*, 91–120. New York: Columbia University Press.

Pennington, R. Toby, Darién E. Prado, and Colin A. Pendry. 2000. “Neotropical Seasonally Dry Forests and Quaternary Vegetation Changes.” *Journal of Biogeography* 27 (2): 261–73. https://doi.org/10.1046/j.1365-2699.2000.00397.x.

Ponder, W. F., G. A. Carter, P. Flemons, and R. R. Chapman. 2001. “Evaluation of Museum Collection Data for Use in Biodiversity Assessment.” *Conservation Biology* 15 (3): 648–57. https://doi.org/10.1046/j.1523-1739.2001.015003648.x.

Prado, Darién E., and Peter E. Gibbs. 1993. “Patterns of Species Distributions in the Dry Seasonal Forests of South America.” *Annals of the Missouri Botanical Garden* 80 (4): 902–27.

Queiroz, L P. 2006. “The Brazilian Caatinga: Phytogeographical Patterns Inferred from Distribution Data of the Leguminosae.” In *Neotropical Savannas and Dry Forests: Plant Diversity, Biogeography and Conservation*, edited by R T Pennington, G P Lewis, and J A Ratter, 113–49. Oxford: Taylor & Francis CRC Press.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Rangel - Ch., J. Orlando. 2015. “La Biodiversidad de Colombia: Significado y Distribución Regional.” *Revista de La Academia Colombiana de Ciencias Exactas, Físicas y Naturales* 39 (51): 176. https://doi.org/10.18257/raccefyn.136.

Rangel, Thiago F., Neil R. Edwards, Philip B. Holden, José Alexandre F. Diniz-Filho, William D. Gosling, Marco Túlio P. Coelho, Fernanda A. S. Cassemiro, Carsten Rahbek, and Robert K. Colwell. 2018. “Modeling the Ecology and Evolution of Biodiversity: Biogeographical Cradles, Museums, and Graves.” *Science (New York, N.Y.)* 361 (6399). https://doi.org/10.1126/science.aar5452.

Ratter, J. A., S. Bridgewater, and J. F. Ribeiro. 2003. “Analysis of the Floristic Composition of the Brazilian Cerrado Vegetation III: Comparison of the Woody Vegetation of 376 Areas.” *Edinburgh Journal of Botany* 60 (01). https://doi.org/10.1017/S0960428603000064.

Ribeiro, J. F., and Bruno M.T. Walter. 2008. “As Principais Fitofisionomias Do Bioma Cerrado.” In *Cerrado: Ecologia e Flora*, 1st ed. Vol. 1. Brasília: Embrapa Cerrados/Embrapa Informação Tecnológica.

Riddle, Brett R. 2016. “Comparative Phylogeography Clarifies the Complexity and Problems of Continental Distribution That Drove A. R. Wallace to Favor Islands.” *Proceedings of the National Academy of Sciences* 113 (29): 7970–77. https://doi.org/10.1073/pnas.1601072113.

Sandel, B., L. Arge, B. Dalsgaard, R. G. Davies, K. J. Gaston, W. J. Sutherland, and J.- C. Svenning. 2011. “The Influence of Late Quaternary Climate-Change Velocity on Species Endemism.” *Science* 334 (6056): 660–64. https://doi.org/10.1126/science.1210173.

Shmida, Avi, and Mark V. Wilson. 1985. “Biological Determinants of Species Diversity.” *Journal of Biogeography* 12 (1): 1–20. https://doi.org/10.2307/2845026.

Silva de Miranda, Pedro Luiz, Ary T. Oliveira-Filho, R. Toby Pennington, Danilo M. Neves, Timothy R. Baker, and Kyle G. Dexter. 2018. “Using Tree Species Inventories to Map Biomes and Assess Their Climatic Overlaps in Lowland Tropical South America.” *Global Ecology and Biogeography* 27 (8): 899–912. https://doi.org/10.1111/geb.12749.

Sobral, Marcos, and João Renato Stehmann. 2009. “An Analysis of New Angiosperm Species Discoveries in Brazil (1990-2006).” *Taxon* 58 (1): 227–32.

Sousa-Baena, Mariane Silveira, Letícia Couto Garcia, and Andrew Townsend Peterson. 2014. “Completeness of Digital Accessible Knowledge of the Plants of Brazil and Priorities for Survey and Inventory.” *Diversity and Distributions* 20 (4): 369–81. https://doi.org/10.1111/ddi.12136.

Steege, H. ter, N. C. A. Pitman, D. Sabatier, C. Baraloto, R. P. Salomao, J. E. Guevara, O. L. Phillips, et al. 2013. “Hyperdominance in the Amazonian Tree Flora.” *Science* 342 (6156): 1243092–1243092. https://doi.org/10.1126/science.1243092.

Steege, Hans ter, Rens W. Vaessen, Dairon Cárdenas-López, Daniel Sabatier, Alexandre Antonelli, Sylvia Mota de Oliveira, Nigel C. A. Pitman, Peter Møller Jørgensen, and Rafael P. Salomão. 2016. “The Discovery of the Amazonian Tree Flora with an Updated Checklist of All Known Tree Taxa.” *Scientific Reports* 6 (1). https://doi.org/10.1038/srep29549.

Ter Steege, Hans, Nigel Pitman, Daniel Sabatier, Hernan Castellanos, Peter Van Der Hout, Douglas C. Daly, Marcos Silveira, et al. 2003. “A Spatial Model of Tree α-Diversity and Tree Density for the Amazon.” *Biodiversity & Conservation* 12 (11): 2255–77. https://doi.org/10.1023/A:1024593414624.

Tichy, Lubomír, and Milan Chytrý. 2006. “Statistical Determination of Diagnostic Species for Site Groups of Unequal Size.” *Journal of Vege* 17: 809–18.

Tuomisto, Hanna, Gabriela Zuquim, and Glenda Cárdenas. 2014. "Species Richness and Diversity along Edaphic and Climatic Gradients in Amazonia." *Ecography*, July, n/a-n/a. https://doi.org/10.1111/ecog.00770.

Vanzolini, Paulo Emílio. 1963. "Problemas Faunísticos Do Cerrado." *Simpósio Sobre o Cerrado* 307 (321): 1.

Vázquez, Diego P., and Richard D. Stevens. 2004. "The Latitudinal Gradient in Niche Breadth: Concepts and Evidence." *The American Naturalist* 164 (1): E1-19. https://doi.org/10.1086/421445.

Velloso, Agnes L., Everardo V. S. B. Sampaio, and Frans G. C. Pareyn. 2001. *Ecorregiões propostas para o bioma Caatinga*. Recife: Instituto de Conservação Ambiental The Nature Conservancy.

Villarroel, Daniel, Cassia B R Munhoz, and Carolyn E B Proença. 2016. "Campos y Sabanas del Cerrado en Bolivia: Delimitación, síntesis terminológica y sus caracteristicas fisionómicas," 34.

Zappi, Daniela C., Fabiana L. Ranzato Filardi, Paula Leitman, Vinícius C. Souza, Bruno M.T. Walter, José R. Pirani, Marli P. Morim, et al. 2015. "Growing Knowledge: An Overview of Seed Plant Diversity in Brazil." *Rodriguésia* 66 (4): 1085–1113. https://doi.org/10.1590/2175-7860201566411.

Zuloaga, Fernando O., Manuel J. Belgrano, Fernando O. Zuloaga, and Manuel J. Belgrano. 2015. "The Catalogue of Vascular Plants of the Southern Cone and the Flora of Argentina: Their Contribution to the World Flora." *Rodriguésia* 66 (4): 989–1024. https://doi.org/10.1590/2175-7860201566405.

5

Conclusions

5 - Conclusions

The main objectives of this thesis were to demonstrate how mapping biomes using species composition data provides clear advantages over conventional approaches; to use this compositional approach, in concert with spatially extensive soil sampling, to provide a thorough and novel assessment of the relative roles of edaphic, climatic and pyrogenic factors in driving biome distributions; and to compare and contrast the tree species richness, endemism and adequacy of species sampling in the principal biomes of lowland tropical South America. To achieve these objectives, we have employed a dataset of thoroughly checked tree species checklists, named NeoTropTree (NTT).

5.1 - Delimiting biomes in Lowland Tropical South America through tree species composition

In the chapter 2, I delimited and mapped five biomes in Lowland Tropical South America (LTSA) – Amazon Forest, Atlantic Forest, Savanna, Seasonally Dry Tropical Forest (SDTF) and Chaco. We showed that each one of these biomes is composed of unique tree species assemblages and that biomes are not continuously distributed across geographic space as they are often displayed in the literature. Rather, they interdigitate and are distributed in patches, which is particularly true for Savanna, SDTF and the semideciduous portion of the Atlantic Forest biome. Also, there is an important decoupling between biome distribution and climatic and edaphic conditions, at least given publicly available environmental data. This study shows the challenges of mapping biomes at the continental scale, especially when relying on publicly available environmental data alone. Importantly, it also shows that predicting future biome distribution through changes in climate will be likely to lead to inaccurate predictions.

Biome delimitation and mapping has been a goal of ecologists and biogeographers as long as the concept of the biome has existed. At first, this was conducted using climatic conditions as a proxy for biome distribution, as can be seen in the works of

Whittaker (1975) and Odum (1992). The biome/climate paradigm shifted to an ecosystem function/biome paradigm with the development of better remote sensing technologies and the necessity of delimiting biomes that can be monitored at the global scale and are able to interface with Dynamic Global Vegetation Models (DGVMs). Since then, studies such as Friedl et al. (2002) Woodward, Lomas, and Kelly (2004) and Higgins, Buitenwerf, and Moncrieff (2016), have linked broad-scale vegetation functional traits, such as canopy phenology, height and openness, to biome distribution by reasoning that biomes are defined by similarities in vegetation physiognomy that relate to ecosystem dynamics (Moncrieff, Hickler, and Higgins 2015). This trait mapping strategy offered the possibility of monitoring vegetation cover change, allowing the prediction of vegetation cover under different environmental change scenarios.

Mapping biomes by monitoring ecosystem traits, however, does not account for variation in species distributions across geographic space. Such knowledge is vital for the development of conservation strategies and to place biome distribution within a biodiversity context, rather than only within the climate/environmental change perspective. In the first chapter, delimiting biomes by relying on tree species composition revealed differences among biomes that have gone unnoticed in previous publications which utilized only remote-sensing tools (e.g. Higgins et al., 2016).

Firstly, I was able to show that the Atlantic and Amazon Forests are two different entities in regards to tree species composition and the climatic space they occupy. Even though these two biomes are usually placed together within a moist forest category (e.g Olson et al. 2001; Higgins, Buitenwerf, and Moncrieff 2016; Dinerstein et al. 2017), the Atlantic Forest is actually drier and cooler than the Amazon. Differences in tree species composition and environmental space they occupy suggests the possibility of these two forests functioning differently (e.g. differences in canopy height, leaf-flush and biomass accumulation). Secondly, SDTFs and Chaco are composed of different tree species assemblages. This distinction was already suggested based upon flora and environment in the works of Prado and Gibbs (1993), Pennington, Prado, and Pendry (2000) and Pennington, Richardson, and Lavin (2006), but it has been controversial (e.g., Kuemmerle et al. 2017). Here, I was able to demonstrate this difference using quantitative floristic analyses at the scale of all LTSA and show that the Chaco has a peculiar tree species composition that likely reflects its distinct environmental conditions.

More importantly, my results and maps show how interdigitated biomes are in LTSA, especially in the central part of South America. In this area, Atlantic Forest, particularly its semi-deciduous portion, Savanna and SDTF co-occur. This high level of interdigitation has been ignored in previously published biome, biogeographic and ecoregion delimitations made for South America (e.g. Olson et al. 2001, Morrone 2014, Higgins et al. 2016, Dinerstein et al. 2017). The important implication of this interdigitation of biomes in a small geographic space is that it shows that climatic and interpolated edaphic data cannot be used to map biomes with precision and that maps presenting biomes as continuous entities are potentially incorrect. It also shows that assigning species to biomes by linking geographic occurrences to continuous biome units can potentially cause species' biome affiliations to be incorrect.

Such tight spatial interdigitation of biomes is likely to occur in other continents as well. This indicates that biomes simply do not respect any sort of geographic boundaries that are not major barriers to dispersal. Therefore, the species comprising these biomes seem to be able to disperse across lowland continental areas (c.f. Dexter et al. 2017), with biomes able to appear anywhere within a continent where appropriate environmental conditions exist. Because of this, methods capable of tracking differences in species composition enables tracking and mapping of biomes with more precision.

Acknowledging the fact that biomes can occur in scattered areas allows research on their biogeographic history and environmental limits to address their true complexity better. For example, it may shed light on areas that possibly acted as refugia or as routes of species migration, such as the gallery forests passing across the central Brazilian savannas. These are shown to be floristically similar to the Amazon and Atlantic Forests and are potential routes of species exchange among these two biomes (Oliveira-Filho and Ratter 1995).

Considering the results obtained in this chapter, I propose a new perspective on how to delimit and map biomes at the continental level. Now that great effort is being put into compiling and constructing robust and broad biodiversity databases, such as NeoTropTree and GBIF, species/taxa distribution can be aligned with ecosystem function data gathered from remote-sensing and species trait databases towards more accurate and informative biome maps. Such maps will be of great value for further studies in macro-ecology, environmental change and biogeography.

5.2 - Environmental controls of biome identity

In chapter 3, I showed that biome transitions are environmentally complex in lowland Tropical South America, with edaphic, climatic and pyrogenic factors playing important roles. Their relative importance, however, changes in accordance to the biome transition involved, and geographic location. I tested soil fertility, water availability and fire as environmental drivers of biome identity in lowland tropical South America in two different regions of high biome heterogeneity – eastern Brazil and Bolivia. The biomes included were Savanna, SDTF and Moist Forest. My results suggest that all three drivers are important, but that their importance varies by region and by which pair of biomes are being considered. This points to the complexity of biome transitions in LTSA and to how their environmental limits change at the intra-continental scale. As I have indicated in the second chapter, edaphic factors are important drivers of biome identity and distribution at large scales in LTSA, but their importance can only be perceived with field collected data (i.e. not with publicly available, interpolated data). In chapter 3, I also show how fire is a meaningful driver of biome identity, especially for Savannas.

One of the central questions in biogeography is to understand how the distribution of individual species, species assemblages (communities) and biomes relates to their surrounding environments. The link between vegetation formations/biomes and climatic conditions has been acknowledged and explored since Humboldt (1816) seminal work, culminating with Whittaker's (1975) and Odum's (1992) biome classification systems. However, the relevance of other environmental conditions only began to be better explored in more recent years. In South America, specifically in Brazil, systematic quantification of edaphic conditions began with the creation of EMBRAPA, in 1973, and the development of research such as that by Dobereiner's group (e.g. Dobereiner 1966; Day, Neves, and Döbereiner 1975; Von Bulow and Dobereiner 1975) on how to cultivate crops in the Cerrado and other Brazilian biomes. However, it was through Ratter et al. (1973, 1978), Furley and Ratter (1988), Oliveira-Filho and Fontes (2000) Ribeiro and Walter (2008); Assis et al. (2011), Quesada et al. (2010, 2011, 2012), and others that the role of soil as a driver of biome identity began to be understood in LTSA. However, all of these studies were conducted at the local scale or within a single biome. Little is still known about the importance of edaphic factors at driving biome identity and distribution at large scales.

Research on environmental limits to biomes' geographic distributions has shown that such limits can vary from one continent to the other, while also showing that biomes can overlap in environmental space (e.g. overlaps in mean annual precipitation shown by Lehmann et al. (2014), Moncrieff, Hickler, and Higgins (2015), Silva de Miranda et al. (2018), Dexter et al. (2018). Because of such overlaps, substantial controversy has existed in the literature around transitions from forests to savannas world-wide (Lehmann et al., 2014 and references therein). Much has been discussed about the roles played by soil and disturbance (fire and herbivory) when climatic conditions can lead to more than one biome state. Much of this work was conducted under the Alternative State Theory (AST) framework, which postulated that disturbance can result in more than one biome within the same climatic envelope (e.g. Hirota et al. (2011), Staver, Archibald, and Levin (2011), Dantas et al. (2016). However, these studies only focused on biomes defined based on tree and C4 grass coverage and often relied on poor edaphic data derived from interpolation. Because of this, the role played by edaphic variables at large/continental scales is largely unknown and the relationship among different environmental drivers is poorly understood.

Therefore, in this chapter, I proposed to explore the environmental limits surrounding LTSA's moist forests, seasonally dry tropical forests and savannas – which were delimited through tree species composition. By using a statistical approach able to encompass the complexity of these ecosystems – structural equation modelling (SEM) – I investigated the environmental correlates to these biome's distribution based on field-collected edaphic data from nearly 200 sites and GIS-based climatic and fire data.

I was able to show that water availability, which ended up being solely indicated by climate in our analytical workflow, is the most important driver of biome identity, being more significant than soil fertility and fire in most models (although fire was the most important driver in the transition between savannas and moist forest in Brazil). However, this finding does not exclude a role for fire and soil fertility. In this chapter, I was able to demonstrate that edaphic factors are indeed important at broad scales. SDTFs grow on more fertile soils than Savannas and Moist Forests in eastern Brazil, showing that soil fertility is also associated with this ecosystem's distribution. Also, my results show that fire is invariably linked to savanna distribution regardless of what forested biome it boarders.

One of the main values of this contribution is to show that the relationship between environmental controls and biome identity varies at an intra-continental scale and according to the biomes being examined. Importantly, I conveyed how fundamentally different the transition from savanna to moist forest and savanna to SDTF are. This finding has an important impact on research focusing on biome transitions, especially those made within the AST framework. These studies usually tends to draw generalizations by gathering data at the global scale. However, my results show that biomes' environmental limits and drivers vary significantly at an intra-continental scale and also are strongly linked to the biomes being studied. Therefore, future work must be mindful of these particularities and propose analytical frameworks and geographic delimitations that account for these features in order to avoid biased results. For example, in the particular case of LTSA, forested biomes have to be further divided into moist forests and SDTFs (Dexter et al., 2018). Otherwise, fundamental differences in their climatic and edaphic conditions in relation to savannas will go unnoticed.

Also, the fact that there is intra-continental variation in environmental controls affects research aiming to model biome distribution and investigate its environmental drivers. As we have shown, biome environmental limits change within LTSA and, henceforth, such regional differences should be addressed in studies attempting to predict biome cover change under different environmental change scenarios, such as the ones conducted by Higgins et al. (2016) and Langan, Higgins, and Scheiter (2017). It seems that the biogeographic contingencies within South America are as complex as those discussed at an intercontinental scale by Lehmann et al. (2014).

5.3 - Lowland Tropical South America tree species richness – an overview and cross-biome comparisons

In chapter 4, I show that the Amazon Forest contains the largest overall number of tree species, and of endemic tree species, followed by the Atlantic Forest, Savanna and SDTF. The same can be observed in relation to tree species endemism levels. Both Atlantic and Amazon Forest's core areas have more than 60% of their tree species richness composed of endemics whereas 40% of the core SDTF and 20% of the core Savanna tree species richness are formed by endemic tree species. Through the species extrapolation curve approach, I show that approximately 94% of the

Amazon and Atlantic forests and SDTFs have already been sampled and recorded. Only ~77% of the Savanna's tree flora, in the other hand, have been sampled so far, which is debatable. This is likely to be related to this region bordering four different biomes that can also be encountered in patches within the core savanna area (intrusions). Overall, this study establishes a baseline for comparative diversity studies in lowland tropical South America.

Knowledge of tree species richness across geographic units, whether these are ecological or political, is of importance for conservation and the investigation of diversity patterns, such as the latitudinal gradient of species richness and species-area relationships. LTSA is known for its tree species diversity and high levels of endemism. However, knowledge of tree species richness in this part of the world is often restricted to the country level and sampling effort varies considerably from one country to another. There has been little effort to compile tree species checklists for biomes extending beyond individual countries. Exceptions are efforts to compile complete tree species checklists for Amazonia (e.g.ter Steege et al. 2013; Cardoso et al. 2017), but other LTSA biomes still lack the same attention. Therefore, this chapter's main objective was to compile a comparative overview of tree species richness and endemism across the four main biomes and their respective areas of occurrence in LTSA.

The results enabled me to highlight patterns and answer long-standing questions in regard to LTSA tree species diversity and its levels of endemism in different biomes. I confirmed that the Amazon Forest is indeed the most tree species rich biome in LTSA and that more collection and sampling expeditions needs to be undertaken in order for scientists to have a better understanding of its full tree species richness. Even though a high-quality and accurate tree species checklist was produced for it by Cardoso et al. (2017), my results points to a minimum of 500 tree species that have not yet been recorded for this biome. However, differences between the numbers I obtained for the Amazon the ones obtained by Cadoso et al. (2017) are likely mostly related to differences in how trees were defined (Cardoso et al., 2017 have a larger size-threshold for trees, which is DBH=10cm), as was observed by Oliveira-Filho et al. (*in prep.)*.

The core areas of both Amazon and Atlantic forests have much higher tree species richness and endemism levels than the drier biomes. This shows that factors such as climatic stability, tolerance to drought and biogeographic history are possibly involved in establishing these patterns. As shown by Esquivel-Muelbert et al. (2017), drought-tolerance is a relatively rare adaptation and, since only a few tree species can cope with drier conditions, drier areas tend to be occupied by fewer species. This hypothesis, however, is yet to be tested at a continental scale and my study is the first to show that this potentially holds true for LTSA. Another factor that might be playing a role is biogeographic history. As pointed out by Antonelli et al. (2018), the Amazon Forest served as a “lineage cradle” at various moments after the rise of the Andes. This, in association with the fact that the Amazon and Atlantic Forests are connected through gallery forests through which tree species can migrate (e.g. Oliveira-Filho and Ratter 1995) can explain why these two biomes are more tree species rich than the Savanna and SDTF.

Another factor to be considered is the areal extent that I covered in this study. The Amazon and Atlantic forests were fully encompassed whereas parts of the Neotropical savannas and SDTFs were not considered. DRYFLOR et al. (2016) reported that the SDTF biome across the entirety of the Neotropics encompass between 4660 up to 6958 tree species based on just 1602 inventories, showing that perhaps the whole of the SDTF biome can harbour just as many species as the Amazon Forest. My study revealed that central South America savannas seem to be far more species rich in trees than previously thought. Both Françoso, Haidar, and Machado (2016) and Ratter, Bridgewater, and Ribeiro (2003) have reported that there are approximately a thousand species in the Brazilian Cerrado. The higher number I report in this study could be related to differences in geographic circumscription. Such difference is most likely related to how the datesets were assembled, because Françoso, Haidar, and Machado (2016) and Ratter, Bridgewater, and Ribeiro (2003) focused only on investigating floristic patterns within the Brazilian Cerrado and, to that end, did not consider any tree species records that could be considered as coming from a surrounding biome and woody elements that have not been recorded to grow more than 1.5m in height and 3cm in diameter.

By providing a clear outline of tree species richness and endemism in the four main biomes of LTSA, my analyses set the groundwork for further investigation of patterns of tree species richness and sampling effort. It also informs on which biomes botanists should focus their collection expeditions and provides vital background for conservation efforts aiming to work beyond political boundaries. Future work, however, should focus on covering the full area of these biomes in South America, and in the Neotropics, as the addition of new areas is likely to have an impact on how many species are found in these biomes, as suggested by Pennington, Lehmann, and Rowland (2018).

5.4 - Transition/Ecotonal Zones in LTSA:

An ecotone is a transition area between two biomes, but this term is more commonly used to indicate areas where two distinct ecological communities meet and integrate. They can be sharp or gradual. Even though these zones are frequently mentioned in the literature, they are rarely delimited or mapped in the most important biome schemes (e.g. Olson et al. 2001, Dinerstein et al., 2017). This is also true for South America, for which most biome classification systems will acknowledge the existence of such areas, but will not indicate where they are located or discuss their main features (e.g Rivás-Martinéz et al., 2011, Morrone et al., 2014). In fact, even though ecotonal zones are often pointed to as being locations of high diversity, there is very little evidence to support this statement (Senft, 2009). Overall, these zones remain largely understudied, especially at the biome level.

The evidence I present in this thesis adds another layer of complexity to this topic. In chapter 2, I showed how biomes in LTSA are not large and continuous units. Instead, they present a patchy distribution, especially at the centre of LTSA. Since a patchy distribution will increase the perimeter of each biome, the area occupied by ecotonal zones is potentially being underestimated in current biome schemes. In chapter 3, I estimated the importance of different environmental drivers in transitions between biomes in Brazil and Bolivia. I demonstrated how the importance of water availability, soil fertility and fire changes across geographic space and also according to the biomes involved in each transition. In chapter 4, I estimated the tree species richness of biome patches that are not connected to their core area of distribution – the geographically marginal areas. Even though these marginal patches harbour tree species that cannot be found in the core area of occurrence, only ~200 tree species are endemic to these sites, meaning that the include many species from other biomes and could be considered ‘transitional’ to some degree.

In summary, my findings show that ecotonal zones should not be treated as uniform entities in LTSA and that the environmental factors behind these areas’ distributions vary at the regional scale, within continents. More importantly, my results show how difficult future endeavours to map all ecotonal zones in LTSA would be and how little we still know about these sites.

5.5 - A bit on conservation:

The evidence presented in Chapters 2, 3 and 4 are also of relevance for the planning and implementation of conservation-related policies and strategies in LTSA. As explained in Chapter 2, an accurate biome/ecoregion delimitation, and its associated classification system, allows the targeting of specific biomes and covering their full area of occurrence. In South American countries where the environmental law systems were built considering the particularities of each biome type (e.g. Brazil), a detailed biome delimitation allows for the appropriate legal system to be applied, making conservation efforts more effective.

Results presented in Chapter 3 are also important for conservation. In this data chapter, I was able to show that fire is linked to the distribution of savannas in Brazil. Therefore, as a means to avoid forest encroachment into savanna, savanna conservation policy needs to include fire management and promote controlled fires when and where necessary. This will ensure conservation of these ecosystems' endemic species and their overall biodiversity.

The results obtained in chapter 4 are the most directly linked to conservation efforts in LTSA. In this chapter, I showed how wet biomes are more tree species rich and have more endemic tree species than drier biomes. However, I also demonstrated that the drier biomes – the SDTFs and the Savannas – are also impressively diverse and hold a high number of endemic tree species and, therefore, must not be neglected.

Targeting tree species richness is one of the most achievable ways to highlight how diverse LTSA biomes are and draw attention from the public as to why they need to be conserved. This is due to the high quality of the datasets on tree species distribution available now (such as NeoTropTree) and on how easily trees can be counted and sampled in comparison to other groups. However, conservation also needs to encompass other biodiversity groups, and their distribution, richness and endemism levels. LTSA is not only known for its high tree diversity, but also for its herb, liana and epiphyte species richness (Antonelli and Sanmartín, 2011). In addition, most animal groups, such as mammals, birds, reptiles, amphibians and fishes (Loyola et al., 2009; Reis et al., 2016), are incredibly diverse in South America and have a significant number of species threatened with extinction due to pressing habitat fragmentation and hunting. Importantly, conservation strategies and policies must also consider the ecosystem services the biomes presented here provide, such as water supply to human settlements and recreational purposes. For example, both the Amazon and Atlantic Forests and even SDTFs, can store large quantities of carbon, demonstrating the urgency of conserving these ecosystems as a possible way to mitigate the rise of CO2 in the atmosphere.

5.6 - References

Antonelli, Alexandre, and Isabel Sanmartín. “Why Are There so Many Plant Species in the Neotropics?” *Taxon* 60, no. 2 (2011): 403–14.

Antonelli, Alexandre, Alexander Zizka, Fernanda Antunes Carvalho, Ruud Scharn, Christine D. Bacon, Daniele Silvestro, and Fabien L. Condamine. 2018. “Amazonia Is the Primary Source of Neotropical Biodiversity.” *Proceedings of the National Academy of Sciences*, May, 201713819. https://doi.org/10.1073/pnas.1713819115.

Assis, Ana Carolina Cunha, Ricardo Marques Coelho, Eduardo da Silva Pinheiro, and Giselda Durigan. 2011. “Water Availability Determines Physiognomic Gradient in an Area of Low-Fertility Soils under Cerrado Vegetation.” *Plant Ecology* 212 (7): 1135–47. https://doi.org/10.1007/s11258-010-9893-8.

Cardoso, Domingos, Tiina Särkinen, Sara Alexander, André M. Amorim, Volker Bittrich, Marcela Celis, Douglas C. Daly, et al. 2017. “Amazon Plant Diversity Revealed by a Taxonomically Verified Species List.” *Proceedings of the National Academy of Sciences* 114 (40): 10695–700. https://doi.org/10.1073/pnas.1706756114.

Dantas, Vinícius de L., Marina Hirota, Rafael S. Oliveira, and Juli G. Pausas. 2016. “Disturbance Maintains Alternative Biome States.” Edited by Marcel Rejmanek. *Ecology Letters* 19 (1): 12–19. https://doi.org/10.1111/ele.12537.

Day, J.M., M.C.P. Neves, and J. Döbereiner. 1975. “Nitrogenase Activity on the Roots of Tropical Forage Grasses.” *Soil Biology and Biochemistry* 7 (2): 107–12. https://doi.org/10.1016/0038-0717(75)90007-3.

Dexter, Kyle G., Mathew Lavin, Benjamin M. Torke, Alex D. Twyford, Thomas A. Kursar, Phyllis D. Coley, Camila Drake, Ruth Hollands, and R. Toby Pennington. 2017. “Dispersal Assembly of Rain Forest Tree Communities across the Amazon Basin.” *Proceedings of the National Academy of Sciences*, 201613655.

Dexter, Kyle G., R. Toby Pennington, Ary T. Oliveira-Filho, Marcelo L. Bueno, Silva de Miranda, Pedro L, and Danilo M. Neves. 2018. “Inserting Tropical Dry Forests Into the Discussion on Biome Transitions in the Tropics.” *Frontiers in Ecology and Evolution* 6. https://doi.org/10.3389/fevo.2018.00104.

Dinerstein, Eric, David Olson, Anup Joshi, Carly Vynne, Neil D. Burgess, Eric Wikramanayake, Nathan Hahn, et al. 2017. “An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm.” *BioScience* 67 (6): 534–45. https://doi.org/10.1093/biosci/bix014.

Dryflor, Karina Banda-R, Alfonso Delgado-Salinas, Kyle G. Dexter, Reynaldo Linares-Palomino, Ary Oliveira-Filho, Darién Prado, et al. 2016. “Plant Diversity Patterns in Neotropical Dry Forests and Their Conservation Implications.” *Science* 353 (6306): 1383–87. https://doi.org/10.1126/science.aaf5080.

Dobereiner, Johanna. 1966. “Manganese Toxicity Effects on Nodulation and Nitrogen Fixation of Beans (Phaseolus Vulgaris L.), in Acid Soils.” *Plant and Soil* 24 (1): 153–66. https://doi.org/10.1007/BF01373081.

Esquivel-Muelbert, Adriane, Timothy R. Baker, Kyle G. Dexter, Simon L. Lewis, Hans ter Steege, Gabriela Lopez-Gonzalez, Abel Monteagudo Mendoza, et al. 2017. “Seasonal Drought Limits Tree Species across the Neotropics.” *Ecography* 40 (5): 618–29. https://doi.org/10.1111/ecog.01904.

Françoso, Renata D., Ricardo F. Haidar, and Ricardo B. Machado. 2016. “Tree Species of South America Central Savanna: Endemism, Marginal Areas and the Relationship with Other Biomes.” *Acta Botanica Brasilica* 30 (1): 78–86. https://doi.org/10.1590/0102-33062015abb0244.

Friedl, M.A, D.K McIver, J.C.F Hodges, X.Y Zhang, D Muchoney, A.H Strahler, C.E Woodcock, et al. 2002. “Global Land Cover Mapping from MODIS: Algorithms and Early Results.” *Remote Sensing of Environment* 83 (1–2): 287–302. https://doi.org/10.1016/S0034-4257(02)00078-0.

Furley, Peter A., and James A. Ratter. 1988. “Soil Resources and Plant Communities of the Central Brazilian Cerrado and Their Development.” *Journal of Biogeography* 15 (1): 97–108. https://doi.org/10.2307/2845050.

Higgins, Steven I., Robert Buitenwerf, and Glenn R. Moncrieff. 2016. “Defining Functional Biomes and Monitoring Their Change Globally.” *Global Change Biology* 22 (11): 3583–93. https://doi.org/10.1111/gcb.13367.

Hirota, Marina, Milena Holmgren, Egbert H. Van Nes, and Marten Scheffer. 2011. “Global Resilience of Tropical Forest and Savanna to Critical Transitions.” *Science (New York, N.Y.)* 334 (6053): 232–35. https://doi.org/10.1126/science.1210657.

Humboldt, Count Alexander. 1816. “XCIII. On the Laws Observed in the Distribution of Vegetable Forms.” *The Philosophical Magazine* 47 (218): 446–53. https://doi.org/10.1080/14786441608628504.

Kuemmerle, Tobias, Mariana Altrichter, Germán Baldi, Marcel Cabido, Micaela Camino, Erika Cuellar, Rosa Leny Cuellar, et al. 2017. “Forest Conservation: Remember Gran Chaco.” *Science* 355 (6324): 465.1-465. https://doi.org/10.1126/science.aal3020.

Langan, Liam, Steven I. Higgins, and Simon Scheiter. 2017. “Climate-Biomes, Pedo-Biomes or Pyro-Biomes: Which World View Explains the Tropical Forest-Savanna Boundary in South America?” *Journal of Biogeography* 44 (10): 2319–30. https://doi.org/10.1111/jbi.13018.

Lehmann, C. E. R., T. M. Anderson, M. Sankaran, S. I. Higgins, S. Archibald, W. A. Hoffmann, N. P. Hanan, et al. 2014. “Savanna Vegetation-Fire-Climate Relationships Differ Among Continents.” *Science* 343 (6170): 548–52. https://doi.org/10.1126/science.1247355.

Loyola, Rafael D., Umberto Kubota, Gustavo A. B. da Fonseca, and Thomas M. Lewinsohn. “Key Neotropical Ecoregions for Conservation of Terrestrial Vertebrates.” *Biodiversity and Conservation* 18, no. 8 (July 2009): 2017–31. https://doi.org/10.1007/s10531-008-9570-6.

Moncrieff, Glenn R., Thomas Hickler, and Steven I. Higgins. 2015. “Intercontinental Divergence in the Climate Envelope of Major Plant Biomes: Intercontinental Biome Divergence.” *Global Ecology and Biogeography* 24 (3): 324–34. https://doi.org/10.1111/geb.12257.

Morrone, Juan J. 2014. “Biogeographical Regionalisation of the Neotropical Region.” *Zootaxa* 3782 (1): 1. https://doi.org/10.11646/zootaxa.3782.1.1.

Odum, E. P. 1992. *Ecology*. 2nd ed. London: Holt Rinehart and Winston.

Oliveira-Filho, A. T., and J. A. Ratter. 1995. “A Study of the Origin of Central Brazilian Forests by the Analysis of Plant Species Distribution Patterns.” *Edinburgh Journal of Botany* 52 (2): 141–94. https://doi.org/10.1017/S0960428600000949.

Oliveira-Filho, Ary T., and Marco Aurélio L. Fontes. 2000. “Patterns of Floristic Differentiation among Atlantic Forests in Southeastern Brazil and the Influence of Climate1.” *Biotropica* 32 (4b): 793–810. https://doi.org/10.1111/j.1744-7429.2000.tb00619.x.

Olson, David M., Eric Dinerstein, Eric D. Wikramanayake, Neil D. Burgess, George V. N. Powell, Emma C. Underwood, Jennifer A. D’amico, et al. 2001. “Terrestrial Ecoregions of the World: A New Map of Life on EarthA New Global Map of Terrestrial Ecoregions Provides an Innovative Tool for Conserving Biodiversity.” *BioScience* 51 (11): 933–38. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2.

Pennington, R. Toby, Caroline E.R. Lehmann, and Lucy M. Rowland. 2018. “Tropical Savannas and Dry Forests.” *Current Biology* 28 (9): R541–45. https://doi.org/10.1016/j.cub.2018.03.014.

Pennington, R. Toby, Darién E. Prado, and Colin A. Pendry. 2000. “Neotropical Seasonally Dry Forests and Quaternary Vegetation Changes.” *Journal of Biogeography* 27 (2): 261–73. https://doi.org/10.1046/j.1365-2699.2000.00397.x.

Pennington, R. Toby, James E. Richardson, and Matt Lavin. 2006. “Insights into the Historical Construction of Species-Rich Biomes from Dated Plant Phylogenies, Neutral Ecological Theory and Phylogenetic Community Structure.” *New Phytologist* 172 (4): 605–16. https://doi.org/10.1111/j.1469-8137.2006.01902.x.

Prado, Darién E., and Peter E. Gibbs. 1993. “Patterns of Species Distributions in the Dry Seasonal Forests of South America.” *Annals of the Missouri Botanical Garden* 80 (4): 902–27.

Quesada, C. A., J. Lloyd, L. O. Anderson, N. M. Fyllas, M. Schwarz, and C. I. Czimczik. 2011. “Soils of Amazonia with Particular Reference to the RAINFOR Sites.” *Biogeosciences* 8 (6): 1415–40. https://doi.org/10.5194/bg-8-1415-2011.

Quesada, C. A., J. Lloyd, M. Schwarz, S. Patiño, T. R. Baker, C. Czimczik, N. M. Fyllas, et al. 2010. “Variations in Chemical and Physical Properties of Amazon Forest Soils in Relation to Their Genesis.” *Biogeosciences* 7 (5): 1515–41. https://doi.org/10.5194/bg-7-1515-2010.

Quesada, C. A., O. L. Phillips, M. Schwarz, C. I. Czimczik, T. R. Baker, S. Patiño, N. M. Fyllas, et al. 2012. “Basin-Wide Variations in Amazon Forest Structure and Function Are Mediated by Both Soils and Climate.” *Biogeosciences* 9 (6): 2203–46. https://doi.org/10.5194/bg-9-2203-2012.

Ratter, J. A., G. P. Askew, R. F. Montgomery, and D. R. Gifford. 1978. “Observations on the Vegetation of Northeastern Mato Grosso. II. Forests and Soils of the Rio Suia-Missu Area.” *Proceedings of the Royal Society of London. Series B, Biological Sciences* 203 (1151): 191–208.

Ratter, J. A., S. Bridgewater, and J. F. Ribeiro. 2003. “Analysis of the Floristic Composition of the Brazilian Cerrado Vegetation III: Comparison of the Woody Vegetation of 376 Areas.” *Edinburgh Journal of Botany* 60 (01). https://doi.org/10.1017/S0960428603000064.

Ratter, J. A., P. W. Richards, G. Argent, and D. R. Gifford. 1973. “Observations on the Vegetation of Northeastern Mato Grosso: I. The Woody Vegetation Types of the Xavantina-Cachimbo Expedition Area.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 266 (880): 449–92.

Reis, R. E., J. S. Albert, F. Di Dario, M. M. Mincarone, P. Petry, and L. A. Rocha. “Fish Biodiversity and Conservation in South America.” *Journal of Fish Biology* 89, no. 1 (July 1, 2016): 12–47. https://doi.org/10.1111/jfb.13016.

Ribeiro, J. F., and Bruno M.T. Walter. 2008. “As Principais Fitofisionomias Do Bioma Cerrado.” In *Cerrado: Ecologia e Flora*, 1st ed. Vol. 1. Brasília: Embrapa Cerrados/Embrapa Informação Tecnológica.

Rivas Martínez, Salvador, Gonzalo Navarro Sánchez, Ángel Penas Merino, and Manuel Costa Talens. “Biogeographic Map of South America. A Preliminary Survey.” *International Journal of Geobotanical Research* 1, no. 1 (November 2011): 21–40. https://doi.org/10.5616/ijgr110002.

Silva de Miranda, Pedro Luiz, Ary T. Oliveira-Filho, R. Toby Pennington, Danilo M. Neves, Timothy R. Baker, and Kyle G. Dexter. 2018. “Using Tree Species Inventories to Map Biomes and Assess Their Climatic Overlaps in Lowland Tropical South America.” *Global Ecology and Biogeography* 27 (8): 899–912. https://doi.org/10.1111/geb.12749.

Staver, A. C., S. Archibald, and S. A. Levin. 2011. “The Global Extent and Determinants of Savanna and Forest as Alternative Biome States.” *Science* 334 (6053): 230–32. https://doi.org/10.1126/science.1210465.

Senft, Amanda Ruth. “Species Diversity Patterns at Ecotones.” University of North Carolina at Chapel Hill, 2009.

Steege, H. ter, N. C. A. Pitman, D. Sabatier, C. Baraloto, R. P. Salomao, J. E. Guevara, O. L. Phillips, et al. 2013. “Hyperdominance in the Amazonian Tree Flora.” *Science* 342 (6156): 1243092–1243092. https://doi.org/10.1126/science.1243092.

Von Bulow, J. F. W., and J. Dobereiner. 1975. “Potential for Nitrogen Fixation in Maize Genotypes in Brazil.” *Proceedings of the National Academy of Sciences* 72 (6): 2389–93. https://doi.org/10.1073/pnas.72.6.2389.

Whittaker, R. H. 1975. *Communities and Ecosystems*. 2nd edition. New York: Macmillan USA.

Woodward, F I, M R Lomas, and C K Kelly. 2004. “Global Climate and the Distribution of Plant Biomes.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 359 (1450): 1465–76. https://doi.org/10.1098/rstb.2004.1525.

Appendix A

Dissecting a Biodiversity Hotspot: The Importance of Environmentally Marginal Habitats in the Atlantic Forest Domain of South America

---

Reference:

Neves, Danilo M., Kyle G. Dexter, R. Toby Pennington, Arthur S. M. Valente, Marcelo L. Bueno, Pedro V. Eisenlohr, Marco A. L. Fontes, et al. 2017. "Dissecting a Biodiversity Hotspot: The Importance of Environmentally Marginal Habitats in the Atlantic Forest Domain of South America." *Diversity and Distributions* 23 (8): 898–909. https://doi.org/10.1111/ddi.12581.

DOI: 10.1111/ddi.12581

BIODIVERSITY RESEARCH

WILEY Diversity and Distributions A Journal of Conservation Biogeography

Dissecting a biodiversity hotspot: The importance of environmentally marginal habitats in the Atlantic Forest Domain of South America

Danilo M. Neves[1] | Kyle G. Dexter[2,3] | R. Toby Pennington[3] | Arthur S. M. Valente[4] | Marcelo L. Bueno[5] | Pedro V. Eisenlohr[6] | Marco A.L. Fontes[7] | Pedro L. S. Miranda[2] | Suzana N. Moreira[8] | Vanessa L. Rezende[8] | Felipe Z. Saiter[9] | Ary T. Oliveira-Filho[8]

[1]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA

[2]School of GeoSciences, The University of Edinburgh, Edinburgh, UK

[3]Royal Botanic Garden Edinburgh, Edinburgh, UK

[4]Instituto Estadual de Florestas, Minas Gerais, Ubá, Brazil

[5]Laboratório de Ecologia e Evolução de Plantas, Departamento de Biologia Vegetal, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil

[6]Laboratório de Ecologia, Universidade do Estado de Mato Grosso, Alta Floresta, Brazil

[7]Departamento de Ciências Florestais, Universidade Federal de Lavras, Lavras, Brazil

[8]Programa de Pós-Graduação em Biologia Vegetal, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

[9]Instituto Federal do Espírito Santo, Santa Teresa, Espírito Santo, Brazil

Correspondence
Danilo M. Neves, Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA.
Emails: dneves@email.arizona.edu and danilormn@gmail.com

Funding information
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Grant/Award Number: BEX 13197-13-4; Fundação de Amparo à Pesquisa do Estado de Minas Gerais; Natural Environment Research Council, Grant/Award Number: NE/I028122/1; Conselho Nacional de Desenvolvimento Científico e Tecnológico, Grant/Award Number: 151002/2014-2 and 301644/88-8

Editor: Kenneth Feeley

Abstract

Aim: We aimed to assess the contribution of marginal habitats to the tree species richness of the *Mata Atlântica* (Atlantic Forest) biodiversity hotspot. In addition, we aimed to determine which environmental factors drive the occurrence and distribution of these marginal habitats.

Location: The whole extension of the South American Atlantic Forest Domain plus forest intrusions into the neighbouring Cerrado and Pampa Domains, which comprises rain forests ("core" habitat) and five marginal habitats, namely high elevation forests, rock outcrop dwarf-forests, riverine forests, semideciduous forests and *restinga* (coastal white-sand woodlands).

Methods: We compiled a dataset containing 366,875 occurrence records of 4,431 tree species from 1,753 site-checklists, which were a priori classified into 10 main vegetation types. We then performed ordination analyses of the species-by-site matrix to assess the floristic consistency of this classification. In order to assess the relative contribution of environmental predictors to the community turnover, we produced models using 26 climate and substrate-related variables as environmental predictors.

Results: Ordination diagrams supported the floristic segregation of vegetation types, with those considered as marginal habitats placed at the extremes of ordination axes. These marginal habitats are associated with the harshest extremes of five limiting factors: temperature seasonality (high elevation and subtropical riverine forests), flammability (rock outcrop dwarf-forests), high salinity (*restinga*), water deficit severity (semideciduous forests) and waterlogged soils (tropical riverine forests). Importantly, 45% of all species endemic to the Atlantic Domain only occur in marginal habitats.

Main conclusions: Our results showed the key role of the poorly protected marginal habitats in contributing to the high species richness of the Atlantic Domain. Various types of environmental harshness operate as environmental filters determining the distribution of the Atlantic Domain habitats. Our findings also stressed the importance of fire, a previously neglected environmental factor.

KEYWORDS
campo rupestre, climate, conservation assessment, flammability, rain forests, restinga, stress gradients, variation partitioning

1 | INTRODUCTION

The Atlantic Forest of South America, or the *Mata Atlântica* as it is known in Brazil where it largely occurs, stretches for over 3,500 km across equatorial, tropical and subtropical latitudes, and is renowned world-wide for being one of the 35 biodiversity hotspots for conservation prioritization (Myers, Mittermeier, Mittermeier, Fonseca, & Kent, 2000). Its importance is also demonstrated by its designation as one of the five primary vegetation "Domains" of Brazil (Ab'Sáber, 2003; IBGE, 1993), the others being the Caatinga, Cerrado, Pampa and Amazon Domains. The Atlantic Forest Domain (hereafter Atlantic Domain) borders all the other Domains except for the Amazon. The prevailing land cover of these bordering Domains are semi-arid thorn woodlands in the Caatinga, woody savannas in the Cerrado and prairies in the Pampa. Species from rain forests, the habitat that originally prevailed in the Atlantic Domain, become a minor component of the landscape in these neighbouring Domains, and they are only found in riverine or high elevation forest enclaves.

Environmental restriction to the establishment of the rain forest habitat is certainly operating at the boundaries of the Atlantic Domain. In a seminal paper, Scarano (2009) proposed a list of five key factors limiting the occurrence and distribution of rain forest species in the Atlantic Domain, which at its harshest extremes give rise to distinct habitats (one for each factor), referred to as marginal habitats. Therefore, the rain forest is placed by Scarano (2009) as the "core" expression of the Atlantic Domain, where deep shade plays the chief role as a limiting factor for competing plants. The five marginal habitats are high elevation forests, rock outcrop dwarf-forests, riverine forests, seasonally dry forests and *restinga* (coastal white-sand woodlands). Most of these marginal habitats have a relatively high density of trees and can be considered forests, albeit not as well developed structurally as rain forests. High elevation forests are primarily associated with frost, with secondary limitation imposed by drought (leeward rain-shadow) and high-light intensity. Cloud forests and *Araucaria*-dominated forests are the main vegetation types of highlands in the Atlantic Domain. Rock outcrop dwarf-forests, found at lower elevations (and even at the seashore), are primarily limited by the paucity, or even lack, of soil and related poor water retention. Meanwhile, riverine forests are associated with waterlogging on lowland plains and riverbeds. Seasonally dry forests (either deciduous or semideciduous) replace rain forests where seasonal rainfall regimes bring regular periods of drought. Finally, environmental harshness for *restinga* is primarily associated with salinity, with secondary limitations imposed by drought and low fertility in mineral nutrients (Scarano, 2009) (Figure 1).

Within limited areas, some studies have confirmed the leading role of Scarano's limiting factors as distribution filters for plants. These studies addressed tree species composition for particular sectors of the Atlantic Domain, such as the south-east (Oliveira-Filho & Fontes, 2000; Eisenlohr & Oliveira-Filho, 2015), the subtropical South (Oliveira-Filho, Budke, Jarenkow, Eisenlohr, & Neves, 2015) and the highly biodiverse central region in eastern Bahia state, northeastern Brazil (Saiter, Eisenlohr, Barbosa, Thomas, & Oliveira-Filho, 2016). However, the whole of the Atlantic Domain has only been investigated for epiphytic angiosperms (Menini-Neto, Furtado, Zappi, Oliveira-Filho, & Forzza, 2016). Also, the Atlantic Domain is affected by fire in much of its distribution (Archibald, Lehmann, Goméz-Dans, & Bradstock, 2013), though to a lesser extent than in surrounding Domains, such as in central (Cerrado woody savannas) and southern Brazil (Pampa prairies). Nevertheless, the potential effect of fire in limiting plant species distribution across the Atlantic Domain is yet to be investigated. Here we bring together a novel and comprehensive dataset assembled on the composition of tree communities across the whole Domain (c. 2,000 community surveys across core and marginal habitats, with >1,000 sites representing surveys not used in the aforementioned studies), combined with environmental data, focusing on testing Scarano's proposed limiting variables as well as factors that were neglected in previous studies (e.g., fire).

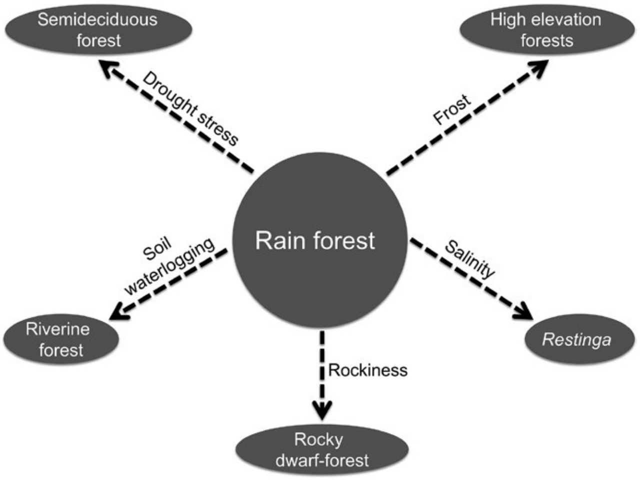

FIGURE 1 Environmental variables (arrows) hypothesized in Scarano (2009) as key factors limiting plant species distribution across the Atlantic Domain of South America. The harshest extremes give rise to distinct vegetation types, referred to as marginal habitats. Coastal white-sand woodlands are called *restinga* in Brazil

Besides the importance for community ecology, understanding the degree to which limiting factors drive community differentiation is inherently relevant for conservation. The Atlantic Domain houses c.18,000 plant species (REFLORA, 2017), but the current high levels of fragmentation and the continuous habitat loss throughout the Domain have raised several concerns in the scientific community (Galindo-Leal, Jacobsen, Langhammer, & Olivieri, 2003; Joly, Metzger, & Tabarelli, 2014; Tabarelli, Pinto, Silva, Hirota, & Bedê, 2005; Tabarelli, Silva, & Gascon, 2004). Therefore, we believe the time is ripe for studies aiming to test the overall importance of environmental conditions in controlling the occurrence and distribution of plant species across the whole extent of the Atlantic Domain and, more importantly, across both its core and marginal habitats.

We addressed the following questions: (1) Are the patterns of tree species distribution across the Atlantic Domain, and its intrusions into neighbouring Domains, limited by factors associated with water deficit (via both soil depth and dry season), water excess (via waterlogging), frosts (via low temperature) and soil salinity? If previously unrecognized environmental conditions are the main factors explaining the patterns of tree species distribution, Scarano's (2009) limiting factors should account for a small proportion of the variation in community composition explained by environmental factors; (2) are these limiting factors leading to floristically distinct marginal habitats? If the community composition of the marginal habitats is simply a nested subset of the more diverse Atlantic Domain rain forest, species turnover should account for a small fraction of the dissimilarity between rain forest and marginal habitats; and (3) what is the contribution of these marginal habitats to the overall high species richness of the Atlantic Domain?

2 | METHODS

2.1 | Study area

The Atlantic Forest, designated as one of the five phytogeographical "Domains" of Brazil (Ab'Sáber, 2003; IBGE, 1993), occurs primarily along the Atlantic coast and is bordered by the Pampa Domain (woody prairies) of southern Brazil and by the "dry diagonal," a corridor that includes three other phytogeographical Domains: Caatinga (largely semi-arid thorn woodlands) of north-eastern Brazil, Cerrado (largely woody savannas) of central Brazil, and Chaco (largely semi-arid thorn woodlands) of Paraguay–Argentina–Bolivia (IBGE, 1993, Prado & Gibbs 1993, Neves, Dexter, Pennington, Bueno, & Oliveira-Filho, 2015). The South American Atlantic Forest Domain (hereafter Atlantic Domain) has a history of controversies over its geographical circumscription and associated terminology. The controversy may be summarized by three main concepts of Atlantic Domain habitats: the sensu stricto, sensu lato and sensu latissimo concepts (Oliveira-Filho, Jarenkow, & Rodal, 2006). The first, and most restrictive concept, includes only the tracts of rain forests that occur as a narrow band along the coast (<100 km wide and up to 2500 m elevation) and stretches all through the Domain, though with two main interruptions, the São Francisco Gap and Campos dos Goytacazes Gap. The former is a semi-arid nucleus at the mouth of the São Francisco River (~10°30′S), and the latter is a seasonally dry region extending from southern Espírito Santo to northern Rio de Janeiro (RJ) States, with its driest extreme at Cabo Frio/RJ (~22°50′S).

The sensu lato concept of Atlantic Domain habitats, which is currently prevalent, includes other habitats adjacent to rain forests, such as the much more extensive semideciduous forests that cover increasingly larger areas towards the south and become wide enough to reach eastern Paraguay and north-eastern Argentina. *Araucaria*-dominated forests are also a very important component of the sensu lato concept, followed by coastal woodlands on white-sand substrates (termed *restingas*) and three highland dwarf-forests: rocky cloud dwarf-forests, rocky semideciduous dwarf-forests and rocky highland savannas (termed *campos rupestres*).

The sensu latissimo concept of Atlantic Domain habitats proposed by Oliveira-Filho et al. (2006) surpasses the geographical limits of the Atlantic Domain to include riverine and deciduous forest tracts occurring in the neighbouring Domains as a secondary component of the landscape, though with a typically Atlantic Domain flora. In the present contribution, we adopt this concept because it allows a more complete inclusion of marginal habitats. However, deciduous forests found in the Cerrado and Pampa Domains, one of the forest types in the sensu lato concept (IBGE, 1993), were not included in this contribution because previous studies (e.g., Eisenlohr & Oliveira-Filho, 2015; Oliveira-Filho et al., 2006) have demonstrated that their flora is distinct and more closely related to that of semi-arid woodlands (e.g., in the Caatinga Domain).

2.2 | Dataset

We extracted the dataset from the NeoTropTree (NTT) database (http://prof.icb.ufmg.br/treeatlan), which consists of tree species checklists (trees defined here as freely standing woody plants >3 m in height) compiled for geo-referenced sites, extending from southern Florida (U.S.A.) and Mexico to Patagonia. NTT currently holds 5,126 sites/checklists, 14,878 woody plant species and 920,129 occurrence records. A site/checklist in NTT is defined by a single habitat, following the classification system proposed by Oliveira-Filho (2015), contained in a circular area with a 10-km diameter. Therefore, where two or more habitats co-occur in one 10-km area, there may be two geographically overlapping sites in the NTT database, each for a distinct habitat.

The data were originally compiled from an extensive survey of published and unpublished (e.g., PhD theses) literature, particularly those on woody plant community surveys and floristic inventories. Moreover, new species occurrence records obtained from both major herbaria and taxonomic monographs have been added to the checklists when they were collected within the 10-km diameter of the original NTT site and within the same habitat. All species and their occurrence records were checked regarding current taxonomic and geographical circumscriptions, as defined by the team of specialists responsible for the online projects *Flora do Brasil* and *Flora del Conosur* (available at http://floradobrasil.jbrj.gov.br/ and http://www.darwin.edu.ar/, respectively). NTT does not, therefore, include occurrence records with doubtful identification, location or habitat, nor sites with an indication of high anthropogenic disturbance. The latter is assessed by taking into account the information available in the studies that comprise the checklists, and by direct observation of site surface on Google Earth©. It also excludes checklists with low species richness (<20 species), because this is often due to low sampling/collecting efforts, which results in poor descriptive power.

This study used a subset of tree inventories from the NTT database, consisting of 328 rain forest sites and 1,425 sites representing the limiting environmental factors and marginal habitats proposed by Scarano (2009), namely seasonally dry (663 semideciduous forests), high elevation (193 *Araucaria*-dominated forests and 61 cloud forests), rock outcrops (49 rocky cloud dwarf-forests, 31 rocky semideciduous dwarf-forests and 41 *campos rupestres*), high salinity (181 restingas—with only forests and dwarf-forests of the mosaic included) and waterlogged soils (133 tropical riverine forests and 73 subtropical riverine forests). Note that marginal habitats associated with seasonal drought and high salinity are represented by one vegetation type, whereas high elevation, rock outcrops and waterlogged soils are represented by more than one vegetation type. The final species matrix contained presence/absence data for 4,431 tree species across 1,753 sites, with a total of 366,875 presences (see Figure 2a and b).

The NTT database also included 26 environmental variables for all its sites, derived from multiple sources (at a 30 arc-second resolution; detailed below). The resolution used in this study was particularly appropriate (1 km^2) because all sites are more than 1 km distant from each other (only 124 of 1,753 sites are less than 5 km distant from another site, and the mean distance between all sites is >1,000 km). Elevation at the NTT site centre was used as an integrative environmental variable. Mean annual temperature, mean daily temperature range, isothermality, temperature seasonality, maximum temperature of the warmest month, minimum temperature of the coldest month, temperature annual range, mean annual precipitation, precipitation of the wettest month, precipitation of the driest month and precipitation seasonality were obtained from WorldClim 1.4 data layers (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005). WorldClim monthly temperatures and precipitation were also interpolated to obtain values for 5-day intervals by applying sinusoidal functions centred at day 15 of each month. These functions yielded values for days 1, 5, 10, 20, 25 and 30, which were used to generate Walter's Climate Diagrams (Walter, 1985) and, thus, four additional variables: duration (days) and severity (days) of both the water deficit and water excess periods. Frost frequency (days) and cloud interception (mm) were obtained from interpolating known values as response variables (data obtained from 135 and 57 Brazilian Meteorological Stations measuring frost frequency and cloud interception, respectively) with elevation, latitude and the WorldClim layers as predicting variables. Potential evapotranspiration (mm) and the aridity index (annual precipitation/potential evapotranspiration) were obtained from Zomer et al. (2007), Zomer, Trabucco, Bossio, van Straaten, and Verchot (2008).

Surface rockiness (% exposed rock), soil coarseness (% sand), soil fertility (% base saturation) and soil salinity (ds/m) were obtained from the Harmonized World Soil Database v 1.2 (available at http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/) and ranked afterwards by mid-class percentage. The use of classes was adopted to add robustness to the data because of the high local soil heterogeneity that makes raw figures unrealistic. Soil drainage classes were obtained following EMBRAPA's protocol (Santos et al., 2013), which combines soil type, texture and depth with landforms. Soil drainage classes, mean annual precipitation (Hijmans et al., 2005) and the aforementioned indices of water deficit and excess were also combined to produce a hyperseasonality index. Grass coverage (%) was used as a proxy of fire return interval (i.e., frequency). Previous studies give support to grass coverage as a good proxy of fire frequency (Archibald et al., 2013; Hoffmann et al., 2012; Lehmann et al., 2014), although further quantification of fire regime is clearly needed (c.f. Archibald et al., 2013). Grass coverage was obtained by direct observation of site surface on Google Earth© images in five 100 × 100 m areas, one at the central coordinates of the NTT site and four at 2.5 km away from it and towards the NE, SW, NW and SE.

Further details of NTT history, products and protocols can be found at http://prof.icb.ufmg.br/treeatlan.

2.3 | Analyses of community turnover

We first explored the patterns of floristic differentiation between rain forest and marginal habitats by performing non-metric multidimensional scaling (NMDS) (McCune & Grace, 2002). We then assessed

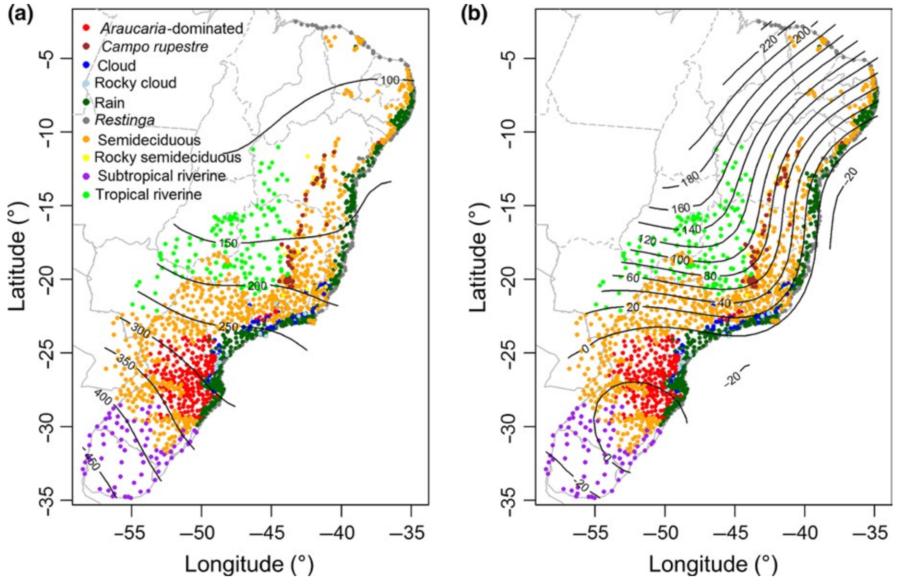

FIGURE 2 Distribution of 1,753 Atlantic Domain sites with their a priori classification into vegetation types (symbols). Variations in (a) temperature seasonality (standard deviation × 100) and (b) water deficit severity (mm) were fitted across geographical space by generalized additive models. Dashed lines represent Brazilian state borders

the relative importance of turnover and nestedness to floristic differentiation between rain forest and each of the marginal habitats. This analysis was performed by first calculating Jaccard pairwise distances, which range from 0 (identical in community composition) to 1 (completely different in community composition). These pairwise distances are then decomposed into dissimilarity due to species turnover (i.e., only compositional changes) and dissimilarity due to differences in species richness. The latter is the difference between Jaccard distance and the dissimilarity due to species turnover (Baselga, 2010). The ordination and the dissimilarity partitioning analyses were conducted in the statistical packages vegan (Oksanen et al., 2016) and betapart (Baselga & Orme, 2012), respectively, both in the R Statistical Environment (R Development Core Team, 2015).

We assessed whether Scarano's (2009) limiting factors are the key environmental factors driving variation in community composition, and then explored the results visually by plotting the habitats in geographical or ordination (NMDS) space and then fitting the values of the most important environmental variables via generalized additive models (GAM) and generalized linear models (GLM), respectively. This routine follows methods similar to those proposed by Blanchet, Legendre, and Borcard (2008) and Legendre, Borcard, and Roberts (2012), which comprise (1) the exclusion of 300 singletons (species found at a single site), as they commonly increase the noise in most analyses without contributing information (Lepš & Šmilauer, 2003); (2) the Hellinger transformation of the binary presence/absence data (Legendre & Gallagher, 2001), which reduces the effect of widespread species; (3) the independent compilation of significant spatial and environmental variables through a forward selection method for redundancy analysis (RDA), after first checking that the respective global models were significant (Blanchet et al., 2008); (4) an additional and progressive elimination of collinear variables based on their variance inflation factor (VIF) and ecological relevance, until maintaining only those with VIF <4 (Quinn & Keough, 2002); and (5) an RDA-based partitioning of variation in the community composition matrix due to environmental variables, spatial autocorrelation and their combined, statistically indistinguishable effects. As spatial variables, we used principal coordinates of neighbour matrices (PCNMs; Borcard, Legendre, Avois-Jacquet, & Tuomisto, 2004), which represent the spatial structure of the sampling units at multiple spatial scales without considering any environmental variation (Borcard, Legendre, & Drapeau, 1992; Borcard et al., 2004; Legendre et al., 2002). We tested the overall significance of the environmental fraction (controlled for spatial autocorrelation) by applying ANOVA permutation tests (999 permutations) for RDA (Peres-Neto, Legendre, Dray, & Borcard, 2006). The variable selection, variation partitioning, NMDS, GLM and GAM analyses were conducted using the fields (Nychka, Furrer, Paige, & Sain, 2015), spacemakeR (Dray, 2010) and vegan (Oksanen et al., 2016) packages in the R Statistical Environment (the variation partitioning script is available as supporting information). The maps were designed using the package maptools (Lewin-Koh & Bivand, 2012) in the R Statistical Environment.

We also calculated patch statistics to test whether floristic differentiation can be modulated by habitat quality (a proxy for anthropogenic effect). We used the PatchStat function—available in the SDMTools package (VanDerWal, Falconi, Januchowski, Shoo, & Storlie, 2014) in the R Statistical Environment—and identified configuration metrics of landscapes (e.g., patch area, edge perimeter) for 95% of our sites using the vegetation map of the Brazilian Atlantic Domain (http://mapas.sosma.org.br/). We found that the effect of habitat quality was negligible in explaining variation in tree community composition across rain forests and marginal habitats (see Table S1 for further details).

2.4 | Conservation assessment

We assessed how well the floristic diversity is captured in our dataset by calculating the expected species accumulation curves for rain forest and marginal habitats, using sample-based rarefaction (Colwell et al., 2012) with the "specaccum" function in the statistical package vegan (Oksanen et al., 2016). We also explored levels of endemism for Atlantic Domain habitats. We obtained the lists of endemic species (woody + non-woody) from Reflora (http://floradobrasil.jbrj.gov.br), which is the most comprehensive study of the patterns of plant species richness and endemism for phytogeographical Domains in eastern South America. Afterwards, we conducted an assessment of the conservation status of the Atlantic Domain habitats by overlaying the distribution of our 1,753 sites on to the coverage of protected areas across South America. We used conservation units from the World Database on Protected Areas (IUCN & UNEP—WCMC, www.protectedplanet.net) and *Cadastro Nacional de Unidades de Conservação* (Ministério do Meio Ambiente—Brazil, www.mapas.mma.gov.br). Species accumulation curves are provided for rain forest and marginal habitats as SI (Fig. S1).

Lastly, we used the main environmental variables emerging from the community turnover models to create site groups discriminating the marginal habitats and then processed the species matrix following the procedure proposed by Tichý and Chytrý (2006) to produce sets of diagnostic species, which are provided as supporting information (Table S2). This procedure is particularly suitable to quantify the fidelity of species to groups that have unequal sizes, that is, different numbers of sampling units, as is the case with our study. After the groups are equalized, a coefficient of fidelity is calculated and the significance of each diagnostic species is obtained with 999 Monte Carlo permutations.

3 | RESULTS

3.1 | Floristic patterns

The distribution of the sites in the ordination space yielded by NMDS (Figure 3a and b) largely segregated rain forests and marginal habitats. The ordination placed "marginal" vegetation types at the extremes of the first three ordination axes. Axis 1 segregated, at negative scores, the shoreline-associated *restinga* and, at positive scores, the vegetation types associated with low-temperature extremes of higher elevations and latitudes further from the equator (*Araucaria*-dominated forests and subtropical riverine forests). Axis 2 segregated, at positive scores, vegetation types associated with rock outcrops (rocky cloud dwarf-forests,

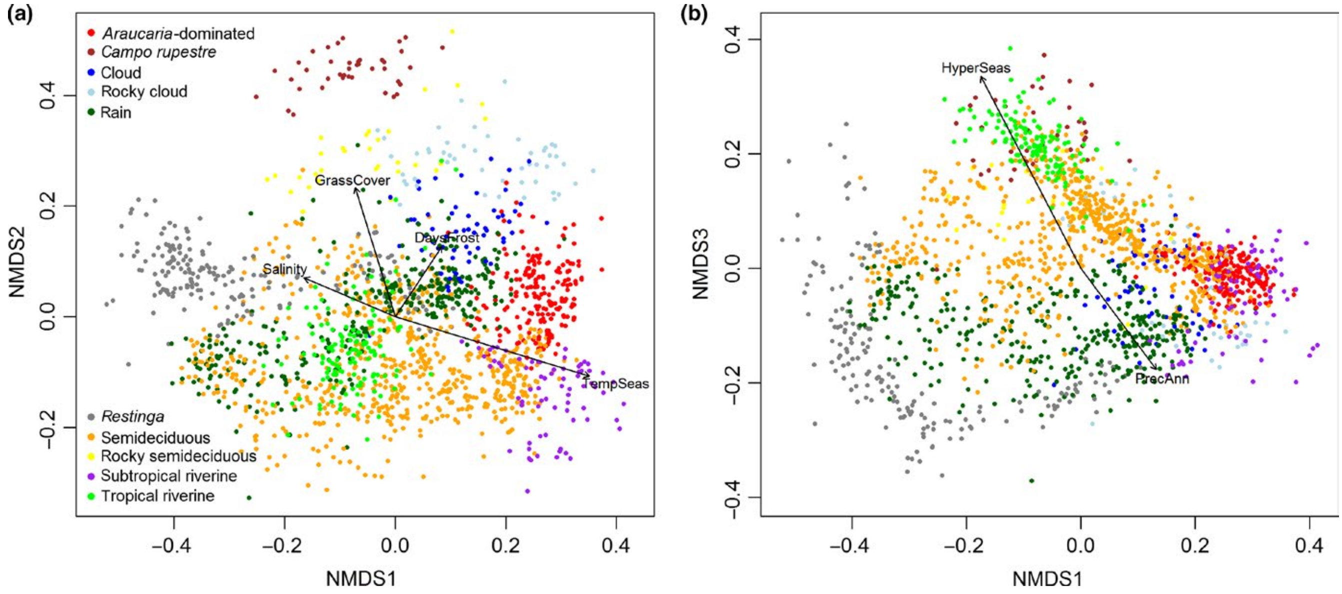

FIGURE 3 Ordination of 1,753 Atlantic Domain sites yielded by non-metric multidimensional scaling (NMDS) of their tree species composition with their a priori classification into vegetation types (symbols). Diagrams are provided for axes 1 × 2 (a) and 1 × 3 (b). Arrows in each diagram represent the correlations between the most explanatory environmental variables and ordination scores. TempSeas, temperature seasonality; DaysFrost, days of frost; Salinity, soil salinity; GrassCover, grass coverage; HyperSeas, water hyperseasonality; PrecAnn, mean annual precipitation

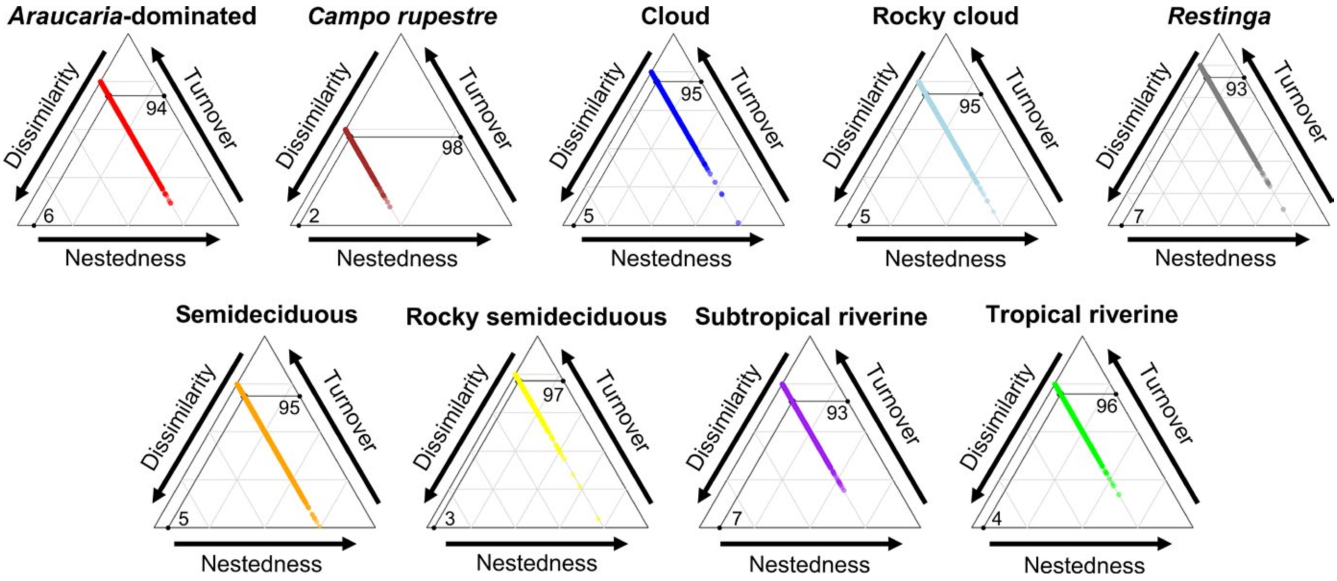

FIGURE 4 Decomposition of the pairwise floristic dissimilarity of rain forest and marginal habitat sites of the South American Atlantic Domain (e.g., bullets in the *Araucaria*-dominated triangle represent pairwise dissimilarities between each of the 193 *Araucaria*-dominated sites and all the 328 rain forest sites, i.e., 63,304 pairwise dissimilarity values). Numbers represent the mean turnover (%) and nestedness (%) components of the Jaccard dissimilarity for each marginal habitat

rocky semideciduous dwarf-forests and *campos rupestres*). Axis 1 further segregated rock outcrop vegetation types into warmer sites (rocky semideciduous dwarf-forests and *campos rupestres*), at positive scores, and colder sites (rocky cloud dwarf-forests), at negative scores. Axis 3 placed the habitat associated with seasonal drought (semideciduous forests) at intermediate scores and the habitat associated with waterlogged soils at positive scores (tropical riverine forests).

The floristic composition of marginal habitats is not simply a nested subset of the more species rich rain forest. The turnover component accounts for most of the floristic dissimilarity of each marginal habitat in relation to rain forests (Figure 4). Nestedness is higher than the turnover component in very few cases (i.e., few marginal habitat sites are simply a subset of another rain forest site; see semideciduous forest triangle in Figure 4). More specifically, vegetation types associated

TABLE 1 Variables selected for the analysis of environmental controls of tree community composition in the Atlantic Domain of South America.

	adj. R^2 cum.	ΔAIC	F	VIF
Temperature seasonality	0.068	−508.02	128.96	3.51
Grass coverage	0.174	−716.16	34.28	1.28
Salinity	0.199	−767.24	27	2.04
Water deficit severity	0.209	−787.86	22.65	3.13
Hyperseasonality	0.222	−816.58	15.42	3.82
Mean annual precipitation	0.234	−840.26	13.41	2.57
Days of frost	0.242	−856.91	8.87	1.76
Elevation	0.251	−863.48	8.52	3.83
Temperature daily range	0.251	−875.73	7.8	2.64
Cloud interception	0.257	−887	4.89	3.27
Soil fertility	0.26	−892.36	4.6	1.46
Water excess duration	0.263	−896.43	3.73	3.11
Sandiness	0.264	−897.48	3	1.74

The variables shown were selected through a forward selection method for redundancy analysis and are ordered by the amount of explained variation in species composition across rain forest and marginal habitats. Goodness-of-fit of the predictor variables was assessed through adjusted coefficients of determination, Akaike information criterion (AIC), F-values and significance tests ($p < .01$ in all cases). VIF, variance inflation factor, obtained using the r-squared value of the regression of one variable against all other explanatory variables. adj. R^2 cum. = cumulative adjusted coefficient of correlation.

with rock outcrops (including *campos rupestre*) have the highest fraction of dissimilarity attributed to turnover while *restinga* and subtropical riverine forest have the lowest fraction attributed to turnover.

3.2 | Variation partitioning analyses

The forward selection procedure retained 13 environmental variables in the model to explain the variation in tree species composition (Table 1). In partitioning the variation explained by the retained environmental and spatial predictors, we found that the environmental fraction explained 27% of the variation, 5% of which was independent of spatial autocorrelation ($p < .01$). The environmental predictors could not account for a spatially structured variation of 12% ($p < .01$), and 61% of the variation remained unexplained (see discussion for more details).

The harshest extremes of the retained environmental variables (Table 1) do lead to distinct habitats, treated here in the context of "marginal" vegetation types. A north to south increase in temperature seasonality was congruent with a latitudinal gradient in community turnover, which represents the floristic differentiation of *Araucaria*-dominated forests and subtropical riverine forests (Figures 2a and 3a) from all other vegetation types. Grass coverage, a proxy for fire frequency (see Methods), was congruent with the floristic differentiation of the vegetation types associated with rock outcrops (including *campos rupestres*) from all other vegetation types (Figure 3a). Within the rock outcrop habitat, the frequency of frost was associated with the floristic differentiation of rocky cloud dwarf-forests from the other rocky vegetation types. Soil salinity was congruent with a coast to inland gradient in community turnover, which represents the floristic differentiation of *restinga* from all other vegetation types (Figure 3a). Another coast to inland gradient is evident in the tropical section of the Atlantic Domain, where water deficit severity and mean annual precipitation, proxies for drought-stress, explained the floristic differentiation of everwet vegetation types, namely rain forest, cloud forests and rocky cloud dwarf-forests, from *campos rupestres*, semideciduous forests, rocky semideciduous dwarf-forests and tropical riverine forests (Figures 2b and 3b). At the harshest extreme of the drought-stress gradient (Figure 3b), water-related hyperseasonality (i.e., ranging from water shortage to soil waterlogging) segregates *campo rupestres* and tropical riverine forests from semideciduous forests. These factors represent the seven most explanatory environmental variables (Table 1) and they accounted for a large fraction of the variation in community composition attributed to environmental predictors (adjusted $R^2 = .242$; Table 1), which is nearly the same as the value for all 13 variables retained in the variation partitioning model (adjusted $R^2 = .264$; Table 1).

3.3 | Conservation assessment

The species accumulation curves showed a levelling off at larger sample sizes for all vegetation types, although no curve actually reached an asymptote. Species accumulation curves levelled off less in vegetation types associated with rock outcrops (including *campos rupestres*) and in *Araucaria*-dominated forest (see Fig. S1). Because the overall floristic dissimilarity between cloud forests and rain forests was relatively low (Figure 3), we assessed the rates of endemism considering these two vegetation types as "core" habitats (wet forests in Table 2 and Figure 5). Despite the fact that wet forests have twice as much protection as marginal habitats (45% and 26%, respectively; Table 2 and Figures 5, 6 and 7), almost half of all species endemic to the Atlantic Domain are only found in marginal habitats (Table 2).

TABLE 2 Wet forests (rain forest + cloud forest) and marginal habitats of the South American Atlantic Domain ranked by their level of endemism in plant species (total endemics/total species richness).

	All				Endemics					
	Angiosperms	Pteridophyta	Gymnosperms	Total	Angiosperms	Pteridophyta	Gymnosperms	Total endemics	%	PA (%)
wet forests	8,938	755	2	9,695	3,740	199	–	3,939	41	45
campos rupestres	4,936	57	–	4,993	1,953	15	–	1,968	39	54
rocky cloud dwarf-forest	2,037	97	2	2,136	429	19	–	448	21	73
restinga	2,490	38	2	2,530	297	1	–	298	12	51
semideciduous forest	3,362	165	1	3,528	243	4	–	247	7	19
rocky semideciduous dwarf-forest	878	21	1	900	8	–	–	8	1	52
Araucaria-dominated forest	1,348	155	4	1,507	81	6	–	87	6	17
tropical riverine forest	2,495	61	5	2,561	101	2	1	104	4	21
subtropical riverine forest	231	2	1	234	–	–	–	–	–	1

PA = percentage of NeoTropTree sites in protected areas (see Figures 5, 6 and 7). Lists of plant species (woody + non-woody) were obtained from the Reflora project (http://floradobrasil.jbrj.gov.br).

4 | DISCUSSION

Both the variation partitioning and the ordination support the importance of the set of limiting conditions proposed by Scarano (2009) as the factors controlling tree community composition of rain forests and marginal habitats, which are treated here in the context of "marginal" vegetation types (*question 1*). We also showed that these limiting factors lead to floristically distinct tree communities, thus indicating that the marginal habitats are not simply a nested subset of the more diverse Atlantic Domain rain forest (*question 2*). In fact, marginal habitats shelter nearly half the endemic plant species in the Atlantic Domain (*question 3*).

4.1 | Limiting factors

A north to south increase in temperature seasonality is the major factor associated with a wide-scale floristic differentiation between tropical habitats and those that are mainly comprised of cold-tolerant species (see Figure 2a and Table 1). Interestingly, this is consistent even within the subtropical section of the Atlantic Domain (Oliveira-Filho et al., 2015), where variation in community composition along the temperature seasonality gradient is congruent with increasing foliage deciduousness, a trait associated with frost-tolerance (Oliveira-Filho et al., 2015). A similar trend in species turnover and foliage deciduousness is also found in the tropical and equatorial sections of the Atlantic Domain, but the main driving force there is rainfall seasonality and the associated dry season (Eisenlohr & Oliveira-Filho, 2015; Saiter et al., 2016). Contrary to our expectations, temperature seasonality showed stronger explanatory power than the frequency of frosts, believed to be a chief factor limiting species distribution across temperature gradients (see Oliveira-Filho et al., 2015; Rundel, Smith, & Meinzer, 1994; Scarano, 2009; Zanne et al., 2014). Nevertheless, within rock outcrop habitats (Figure 3b), the occurrence of frost in rocky cloud dwarf-forests seems to be limiting the establishment of species from *campos rupestres* and rocky semideciduous dwarf-forests, suggesting that the frequency of frosts is an important factor underpinning the distribution of marginal habitats in the Atlantic Domain, though at smaller spatial scales.

Periods of water shortage represented by seasonal droughts are indeed the chief factor driving species turnover in the tropical and equatorial sections of the Atlantic Domain (see Figure 2b), while other local factors may also affect water availability to plants (Pontara et al., 2016). The substrate often either favours or restricts water drainage via landforms and soil depth and texture, while strong winds may add to the water deficit stress, particular nearer to the coast, where *restingas* occur. In this coastal marginal habitat, which was identified as one of the most floristically differentiated (see Figure 3a), the stress due to water deficit is increased by a sandy substrate with high salinity, and by salt spray coming directly from the ocean (Cerqueira, 2000). In addition, although nutrient-poor soils prevail all over the Domain, the edaphic conditions in *restingas* represent an extreme of particularly low soil fertility (most NTT sites of the dataset were classified as "dystrophic" while most *restingas* were "hyperdystrophic").

When assessing whether soil waterlogging leads to a floristically distinct marginal habitat, we found that the intrusions of riverine

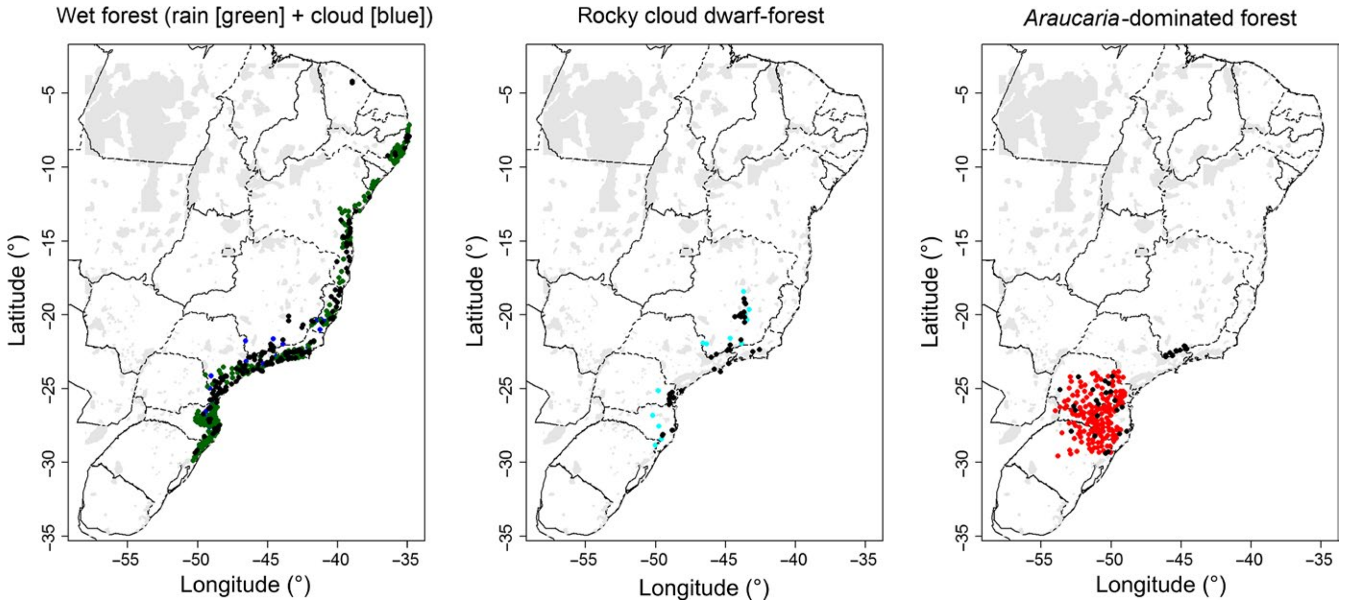

FIGURE 5 Conservation assessment of wet forests (rain + cloud), rocky cloud dwarf-forest and *Araucaria*-dominated forests of the South American Atlantic Domain. Black bullets represent woody plant communities occurring within protected areas. Grey areas represent the current network of protected areas across South America. Dashed lines represent Brazilian state borders

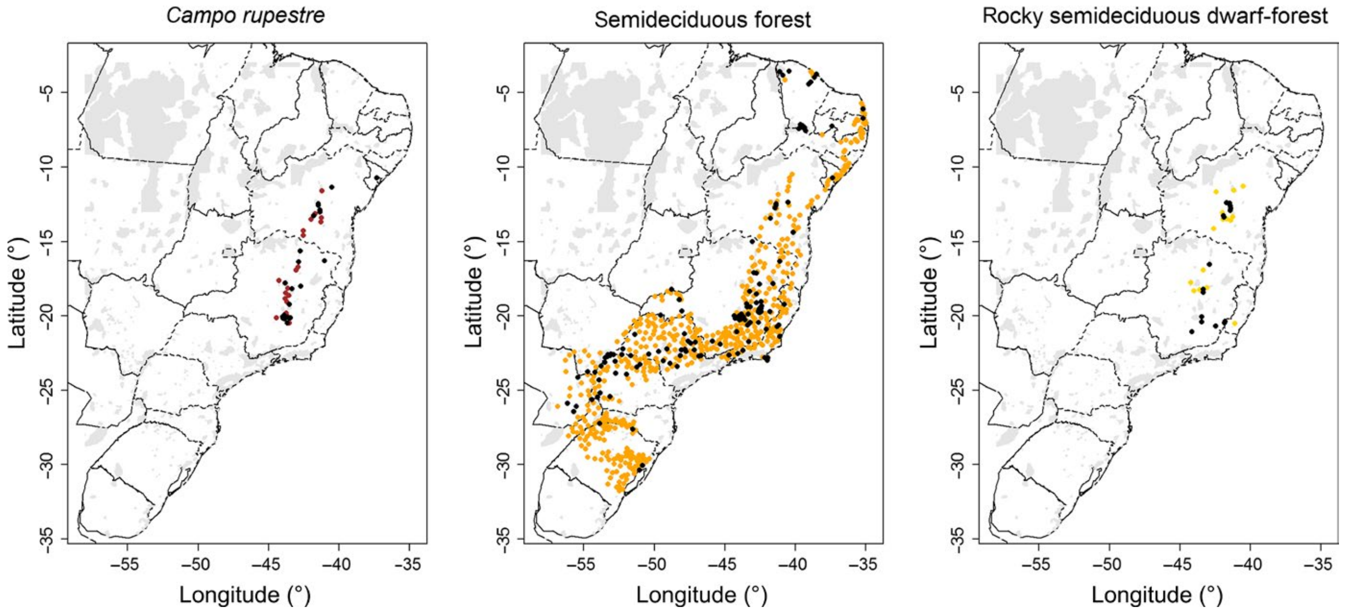

FIGURE 6 Conservation assessment of *campo rupestre*, semideciduous forests and rocky semideciduous dwarf-forests of the South American Atlantic Domain. Black bullets represent woody plant communities occurring within protected areas. Grey areas represent the current network of protected areas across South America. Dashed lines represent Brazilian state borders

forests into poorly drained soils of the Cerrado Domain showed only a weak differentiation from their neighbouring semideciduous forests (see Figure 3). Kurtz, Valentin, and Scarano (2015) also found that riverine habitats of the Atlantic Domain are indistinguishable as a floristic unit from non-flooded habitats, and that their flora is essentially an extract of the regional species pool. These trends may result from a particular feature of the Atlantic Domain. Unlike the Amazon Domain, where a wide net of rivers lead to large areas of seasonally flooded habitats, rivers in the Atlantic Domain represent a minor component of the landscape. In the Amazon, seasonal flooding over wide alluvial beds is known as one of the main sources of floristic differentiation among habitat types and an important driver of tree species distribution patterns (Wittmann et al., 2013), whereas in the Atlantic Domain, the tiny areas of riverine forest are swamped with immigration from the non-flooded habitats. On the other hand, the intrusions of subtropical riverine forests into poorly drained soils of the Pampa Domain seems to have a comparatively stronger floristic differentiation (see Figure 3a), but primarily associated with high temperature seasonality.

For *campos rupestres* we were able to document fire as an important factor limiting tree species distribution across the Atlantic

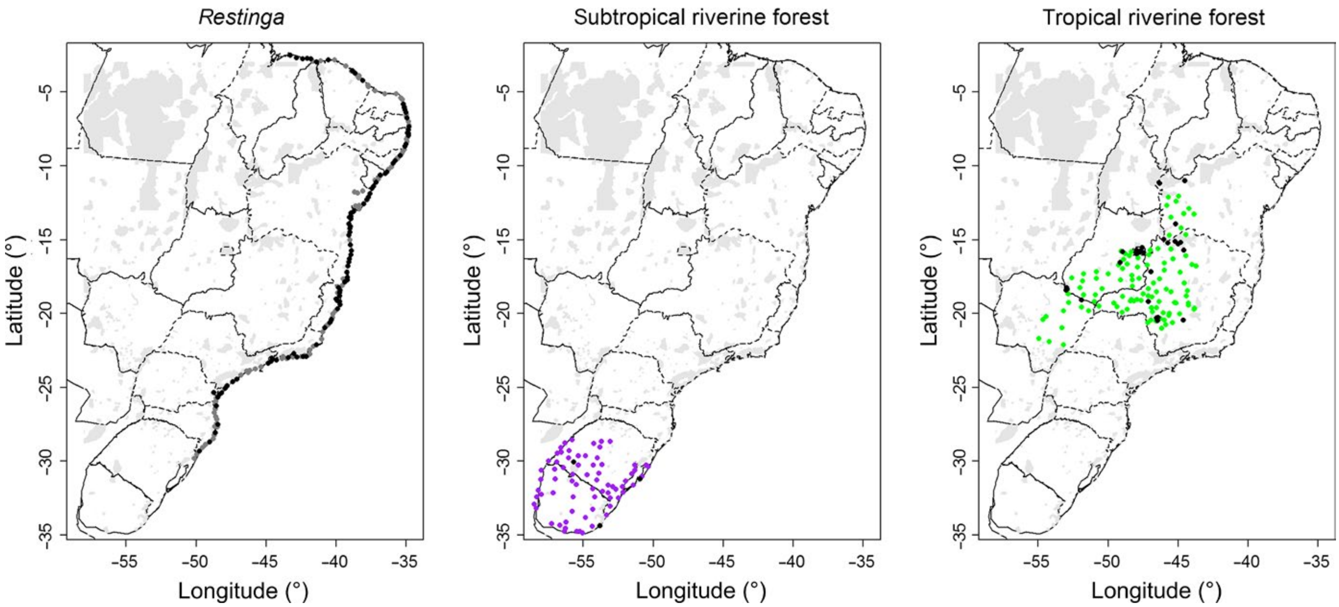

FIGURE 7 Conservation assessment of *restinga*, subtropical riverine forests and tropical riverine forests of the South American Atlantic Domain. Black bullets represent woody plant communities occurring within protected areas. Grey areas represent the current network of protected areas across South America. Dashed lines represent Brazilian state borders. Coastal white-sand woodlands are called *restinga* in Brazil

Domain (see Figure 3a). This is consistent with previous studies showing that forest-savanna boundaries in tropical savannas are driven by fire, though generally in interaction with other factors (Archibald et al., 2013; Dantas, Batalha, & Pausas, 2013; Hoffmann et al., 2012). Within the Atlantic Domain, however, fire frequency is low relative to the surrounding savanna formations (see detailed maps in Archibald et al., 2013) and has therefore been neglected in previous studies. Nevertheless, here we show that fire is actually an important component shaping macroscale patterns of floristic variation across the Atlantic Domain and thus deserves further attention. The congruence between floristic turnover and grass coverage, a proxy for fire frequency, across rocky semideciduous dwarf-forests and *campos rupestres* (Figure 3a) indicates that fire plays a key role in determining the mosaic of rock outcrop habitats in the Atlantic Domain. Rocky semideciduous dwarf-forests seem to represent a transition between rain forests and *campos rupestres* (see Figure 3a), which is likely to be mediated by fire history and local factors contributing to either increase or decrease flammability, particularly topography and soil depth.

4.2 | Spatial structure and unexplained variation

While the relevance of the environmental fraction in controlling community turnover was straightforward to interpret, the variation that either remained unexplained or was attributed to spatial structure independent of the measured environmental factors (61% and 12%, respectively) deserves further attention. Rain forests and marginal habitats are often geographically segregated (Figure 2), suggesting that there may be a role for spatially structured dispersal limitation and historical biogeography in driving some of the observed floristic differentiation. However, given the clear floristic segregation of rock outcrop dwarf-forests from semideciduous and rain forests, despite their spatial interdigitation (e.g., in south-eastern Brazil; Figure 2), we believe it is more parsimonious to attribute the positive spatial autocorrelation, a proxy of distance decay in community similarity (Nekola & White, 1999), to niche-based controls (e.g., unmeasured spatially structured variables describing environmental conditions, natural enemies and competition). Regarding the large fraction of unexplained variation, it may suggest that ecological drift (cf. Hubbell, 2001) is driving stochastic rearrangements of species distribution ranges through time. However, a high proportion of unexplained variation, ranging from 40% to 80% (e.g., Legendre et al., 2009; Neves et al., 2015; reviewed by Soininen, 2014), is a common outcome in studies of floristic composition over similar spatial scales, and could also be attributed to statistical noise (ter Braak, 1986; Guisan, Weiss, & Weiss, 1999) or unmeasured non-spatially structured environmental conditions.

4.3 | Conservation implications

Here we showed the uneven distribution of protected areas across the Atlantic Domain with wet forests having twice as much protection. Marginal habitats receive considerably lower protection, despite harbouring almost half of the 7,099 species endemic to the Atlantic Domain. These 3,160 endemic species are not found anywhere else in the world, including in the rain forests of the Atlantic Domain. This demonstrates that different marginal habitats, characterized by environmental harshness, underpin the patterns of high species richness across the Atlantic Domain as a whole. Therefore, we emphasize that these marginal habitats need better consideration by conservationists and biodiversity scientists, based on their (1) high level of endemism; (2) lower level of protection; and (3) less data (see species accumulation curves of vegetation types associated with rock outcrops in Fig. S1).

4.4 | Concluding remarks

The distribution of the Atlantic Forest marginal habitats is associated with low-temperature extremes (i.e., ranging from winter frosts to summer maxima higher than 40°C), soil salinity, drought-stress and soil waterlogging. Additionally, grass coverage, a proxy for flammability and a previously unappreciated environmental factor in the Atlantic Domain, is amongst the principal factors explaining the patterns of tree species distribution. For conservation purposes, the *restinga* is strikingly distinct both floristically and environmentally (see Figures 3a and b), suggesting the need for further investigation. If *restingas* are indeed a distinct phytogeographical region, instead of an extension of rain forests into saline white-sand environments, they may be much more threatened than assumed based upon classifications that places these two habitats together. *Restinga* has suffered massive fragmentation due to high human occupation in coastal areas and a rapidly developing tourism industry.

ACKNOWLEDGEMENTS

D.M.N., K.G.D. and R.T.P. were supported by the National Environmental Research Council (grant NE/I028122/1). AOF and MLB were supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico—Brazil (CNPq) (grants 301644/88-8 and 151002/2014-2, respectively). ASMV and SNM were supported by the Fundação de Amparo à Pesquisa de Minas Gerais – Brazil (FAPEMIG). PLSM thanks the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brazil (CAPES) for supporting a full PhD at the University of Edinburgh under the Science Without Borders Programme (grant BEX 13197-13-4).

AUTHOR CONTRIBUTIONS

A.O.F compiled the database and conceived the idea; D.M.N. and K.G.D. designed the manuscript; D.M.N. analysed the data; D.M.N. and A.O.F. led the writing with substantial input from K.G.D. and R.T.P. All authors commented on the manuscript and approved the final version.

DATA ACCESSIBILITY

Additional accessibility data are provided as supporting information.

REFERENCES

Ab'Sáber, A. N. (2003). *Os domínios de natureza no Brasil: Potencialidades paisagísticas*. São Paulo: Ateliê Editorial.

Archibald, S., Lehmann, C. E. R., Goméz-Dans, J. L., & Bradstock, R. A. (2013). Defining pyromes and global syndromes of fire regimes. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 6442–6447.

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, *19*, 134–143.

Baselga, A., & Orme, D.I. (2012). Betapart: An R package for the study of beta diversity. *Methods in Ecology and Evolution*, *3*, 808–812.

Blanchet, F. G., Legendre, P., & Borcard, D. (2008). Forward selection of explanatory variables. *Ecology*, *89*, 2623–2632.

Borcard, D., Legendre, P., Avois-Jacquet, C., & Tuomisto, H. (2004). Dissecting the spatial structure of ecological data at multiple scales. *Ecology*, *85*, 1826–1832.

Borcard, D., Legendre, P., & Drapeau, P. (1992). Partialling out the spatial component of ecological variation. *Ecology*, *73*, 1045–1055.

ter Braak, C. J. F. (1986). Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis. *Ecology*, *67*, 1167–1179.

Cerqueira, R. (2000). Biogeografia das restingas. In F. A. Esteves & L. D. Lacerda (Eds.), *Ecologia de restingas e lagoas costeiras* (pp. 65–75). Macaé, Brasil: NUPEM/UFRJ.

Colwell, R. K., Chao, A., Gotelli, N. J., Lin, S.-Y., Mao, C. X., Chazdon, R. L., & Longino, J. T. (2012). Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology*, *5*, 3–21.

Dantas, V. L., Batalha, M. A., & Pausas, J. G. (2013). Fire drives functional thresholds on the savanna–forest transition. *Ecology*, *94*, 2454–2463.

Dray, S. (2010). *SpacemakeR: Spatial modelling*. R package version 473 0.0–5/r101. Retrieved from: http://R-Forge.R-project.org/projects/sedar/

Eisenlohr, P. V., & Oliveira-Filho, A. T. (2015). Revisiting patterns of tree species composition and their driving forces in the Atlantic forests of southeastern Brazil. *Biotropica*, *47*, 689–701.

Galindo-Leal, C., Jacobsen, T. R., Langhammer, P. F., & Olivieri, S. (2003). State of the hotspots: The dynamics of biodiversity loss. In C. Galindo-Leal C & I. G. de Câmara (Eds.), *The Atlantic Forest of South America: Biodiversity status, threats, and outlook* (pp. 12–23). Washington: Center for Applied Biodiversity Science and Island Press.

Guisan, A., Weiss, S. B., & Weiss, A. D. (1999). GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology*, *143*, 107–122.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965–1978.

Hoffmann, W. A., Geiger, E. L., Gotsch, S. G., Rossatto, D. R., Silva, L. C. R., Lau, O. L., … Franco, A. C. (2012). Ecological thresholds at the savanna-forest boundary: How plant traits, resources and fire govern the distribution of tropical biomes. *Ecology Letters*, *15*, 759–768.

Hubbell, S. P. (2001). *The unified neutral theory of biodiversity and biogeography*. Princeton, NJ: Princeton University Press.

IBGE (1993). *Mapa de Vegetação do Brasil*. Rio de Janeiro: Fundação Instituto Brasileiro de Geografia e Estatística, Ministério da Agricultura.

Joly, C. A., Metzger, J. P., & Tabarelli, M. (2014). Experiences from the Brazilian Atlantic forest: Ecological findings and conservation initiatives. *New Phytologist*, *204*, 459–473.

Kurtz, B. C., Valentin, J. L., & Scarano, F. R. (2015). Are the Neotropical swamp forests a distinguishable forest type? Patterns from Southeast and Southern Brazil. *Edinburgh Journal of Botany*, *72*, 191–208.

Legendre, P., Borcard, D., & Roberts, D. W. (2012). Variation partitioning involving orthogonal spatial eigenfunction submodels. *Ecology*, *93*, 1234–1240.

Legendre, P., Dale, M. R. T., Fortin, M. J., Gurevitch, J., Hohn, M., & Myers, D. (2002). The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography*, *25*, 601–615.

Legendre, P., & Gallagher, E. D. (2001). Ecological meaningful transformations for ordination of species data. *Oecologia*, *129*, 271–280.

Legendre, P., Mi, X., Ren, H., Ma, K., Yu, M., Sun, I., & He, F. (2009). Partitioning beta diversity in a subtropical broad-leaved forest of China. *Ecology*, *90*, 663–674.

Lehmann, C. E. R., Anderson, T. M., Sankaran, M., Higgins, S. I., Archibald, S., Hoffmann, W. A., … Bond, W. J. (2014). Savanna vegetation-fire-climate relationships differ among continents. *Science*, *343*, 548–552.

Lepš, J., & Šmilauer, J. P. (2003). *Multivariate analysis of ecological data using CANOCO*. Cambridge, UK: Cambridge University Press.

Lewin-Koh, N. J., & Bivand, R. (2012). *Maptools: Tools for reading and handling spatial objects*. R package version 0.8–17/r238. Retrieved from: http://R-Forge.R project.org/projects/maptools/

McCune, B., & Grace, J. B. (2002). *Analysis of ecological communities*. Gleneden Beach, Oregon: MjM Software Design.

Menini-Neto, L., Furtado, S. G., Zappi, D., Oliveira-Filho, A. T., & Forzza, R. C. (2016). Biogeography of epiphytic Angiosperms in the Brazilian Atlantic forest, a world biodiversity hotspot. *Brazilian Journal of Botany*, *39*, 261–273.

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

Nekola, J. C., & White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, *26*, 867–878.

Neves, D. M., Dexter, K. G., Pennington, R. T., Bueno, M. L., & Oliveira-Filho, A. T. (2015). Environmental and historical controls of floristic composition across the South American Dry Diagonal. *Journal of Biogeography*, *42*, 1566–1576.

Nychka, D., Furrer, R., Paige, J., & Sain, S. (2015). *Fields: Tools for spatial data*. R package version 8.3-5. Retrieved from: http://CRAN.R-project.org/package=fields

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., … Wagner, H. (2016). *Vegan: Community ecology package*. R package version 2.0–3. Retrieved from: http://CRAN.R-project.org/package=vegan

Oliveira-Filho, A. T. (2015). Um Sistema de classificação fisionômico-ecológica da vegetação Neotropical. In P. V. Eisenlohr, J. M. Felfili, M. M. R. F. Melo, L. A. Andrade & J. A. A. Meira-Neto (Eds.), *Fitossociologia no Brasil: Métodos e estudos de casos*, Vol. 2 (pp. 452–473). Viçosa, Brazil: Editora UFV.

Oliveira-Filho, A. T., Budke, J. C., Jarenkow, J. A., Eisenlohr, P. V., & Neves, D. R. M. (2015). Delving into the variations in tree species composition and richness across South American subtropical Atlantic and Pampean forests. *Journal of Plant Ecology*, *8*, 242–260.

Oliveira-Filho, A. T., Jarenkow, J. A., & Rodal, M. J. N. (2006). Floristic relationships of seasonally dry forests of eastern South America based on tree species distribution patterns. In R. T. Pennington, J. A. Ratter, & G. P. Lewis (Eds.), *Neotropical savannas and dry forests: Plant diversity, biogeography and conservation* (pp. 151–184). Boca Raton, USA: CRC Press.

Oliveira-Filho, A. T., & Fontes, M. A. L. (2000). Patterns of floristic differentiation among Atlantic Forests in Southeastern Brazil and the influence of climate. *Biotropica*, *32*, 793–810.

Peres-Neto, P. R., Legendre, P., Dray, S., & Borcard, D. (2006). Variation partitioning of species data matrices: Estimation and comparisons of fractions. *Ecology*, *87*, 2614–2625.

Pontara, V., Bueno, M. L., Garcia, L. E., Oliveira-Filho, A. T., Pennington, R. T., Burslem, D. F. R. P., & Lemos-Filho, J. P. (2016). Fine-scale variation in topography and seasonality determine radial growth of an endangered tree in Brazilian Atlantic forest. *Plant and Soil*, *403*, 115–128.

Prado, D. E., & Gibbs, P. E. (1993). Patterns of species distribution in the dry seasonal forests of South America. *Annals of the Missouri Botanical Garden*, *80*, 902–927.

Quinn, G. P., & Keough, M. J. (2002). *Experimental design and data analysis for biologists*. Cambridge, UK: Cambridge University Press.

R Core Team (2015). *R: A language and environment for statistical computing. Version 3.1.0*. Vienna: R Foundation for Statistical Computing. Retrieved from: http://www.Rproject.org/

REFLORA (2017). *Lista de espécies da flora do Brasil*. Rio de Janeiro, Brazil: Jardim Botânico do Rio de Janeiro. Retrieved from: http://floradobrasil.jbrj.gov.br/ (accessed 19.01.17).

Rundel, P. W., Smith, A. P., & Meinzer, F. C. (1994). *Tropical alpine environments: Plant form and function*. Cambridge, UK: Cambridge University Press.

Saiter, F. Z., Eisenlohr, P. V., Barbosa, M. R., Thomas, W. W., & Oliveira-Filho, A. T. (2016). From evergreen to deciduous tropical forests: How energy-water balance, temperature, and space influence the tree species composition in a high diversity region. *Plant Ecology & Diversity*, *9*, 45–54.

Santos, H. G., Jacomine, P. K. T., Anjos, L. H. C., Oliveira, V. A., Lumbreras, J. F., Coelho, M. R., … Oliveira, J. B. (2013). *Sistema brasileiro de classificação de solos*, 3rd ed. Brasília: Embrapa.

Scarano, F. R. (2009). Plant communities at the periphery of the Atlantic rain forest: Rare-species bias and its risks for conservation. *Biological Conservation*, *142*, 1201–1208.

Soininen, J. (2014). A quantitative analysis of species sorting across organisms and ecosystems. *Ecology*, *95*, 3284–3292.

Tabarelli, M., Pinto, L. P., Silva, J. M. C., Hirota, M., & Bedê, L. (2005). Challenges and opportunities for biodiversity conservation in the Brazilian Atlantic Forest. *Conservation Biology*, *19*, 695–700.

Tabarelli, M., Silva, J. M. C., & Gascon, C. (2004). Forest fragmentation, synergisms and the impoverishment of neotropical forests. *Biodiversity and Conservation*, *13*, 1419–1425.

Tichý, L., & Chytrý, M. (2006). Statistical determination of diagnostic species for site groups of unequal size. *Journal of Vegetation Science*, *17*, 809–818.

VanDerWal, J., Falconi, L., Januchowski, S., Shoo, L., & Storlie, C. (2014). *SDMTools: Species distribution modelling tools: Tools for processing data associated with species distribution modelling exercises*. R package version 1.1-221. Retrieved from: https://CRAN.R-project.org/package=SDMTools

Walter, H. (1985). *Vegetation of the earth and ecological systems of the geo-biosphere*, 3rd ed. Berlin: Springer-Verlag.

Wittmann, F., Householder, E., Piedade, M. T. F., Assis, R. L., Schöngart, J., Parolin, P., & Junk, W. J. (2013). Habitat specificity, endemism and the neotropical distribution of Amazonian white-water floodplain trees. *Ecography*, *36*, 690–707.

Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., FitzJohn, R. G., … Beaulieu, J. M. (2014). Three keys to the radiation of angiosperms into freezing environments. *Nature*, *506*, 89–92.

Zomer, R. J., Bossio, D. A., Trabucco, A., Yuanjie, L., Gupta, D. C., & Singh, V. P. (2007). *Trees and water: Smallholder agroforestry on irrigated lands in northern India. IWMI Research Report 122*. Colombo, Sri Lanka: International Water Management Institute.

Zomer, R. J., Trabucco, A., Bossio, D. A., van Straaten, O., & Verchot, L. V. (2008). Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agriculture Ecosystems and Environment*, *126*, 67–80.

BIOSKETCH

Danilo M. Neves is a postdoctoral research fellow at the University of Arizona. He is interested in the evolutionary dimension of community ecology, with an emphasis on historical biogeography of terrestrial biomes.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Neves DM, Dexter KG, Pennington RT, et al. Dissecting a biodiversity hotspot: The importance of environmentally marginal habitats in the Atlantic Forest Domain of South America. *Diversity Distrib*. 2017;23:898–909. https://doi.org/10.1111/ddi.12581

Appendix B

Lack of floristic identity in campos rupestres – A hyperdiverse mosaic of rocky montane savannas in South America

---

Reference:

Neves, Danilo M., Kyle G. Dexter, R. Toby Pennington, Marcelo L. Bueno, Pedro L.S. de Miranda, and Ary Teixeira Oliveira-Filho. 2018. "Lack of Floristic Identity in Campos Rupestres — A Hyperdiverse Mosaic of Rocky Montane Savannas in South America." *Flora* 238 (January): 24–31. https://doi.org/10.1016/j.flora.2017.03.011.

Flora 238 (2018) 24–31

Contents lists available at ScienceDirect

Flora

journal homepage: www.elsevier.com/locate/flora

Original article

Lack of floristic identity in *campos rupestres*—A hyperdiverse mosaic of rocky montane savannas in South America☆

Danilo M. Neves[a,*], Kyle G. Dexter[b,c], R. Toby Pennington[c], Marcelo L. Bueno[d], Pedro L.S. de Miranda[b], Ary Teixeira Oliveira-Filho[e]

[a] *Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA*
[b] *School of Geosciences, The University of Edinburgh, Edinburgh EH9 3JN, UK*
[c] *Royal Botanic Garden Edinburgh, Edinburgh EH6 5LR, UK*
[d] *Laboratório de Ecologia e Evolução de Plantas, Departamento de Biologia Vegetal, Universidade Federal de Viçosa, Viçosa, 36570-000 Minas Gerais, Brazil*
[e] *Programa de Pós-Graduação em Biologia Vegetal, Universidade Federal de Minas Gerais, Belo Horizonte 31270-090, Brazil*

ARTICLE INFO

Article history:
Received 29 December 2016
Received in revised form 18 March 2017
Accepted 20 March 2017
Edited by Fernando A.O. Silveira
Available online 23 March 2017

Keywords:
rupestrian grasslands
Espinhaço range
cangas
Floristic composition
Environmental heterogeneity
Cluster analysis

ABSTRACT

The rocky montane savannas of South America, known as *campos rupestres* in Brazil, where they largely occur, represent a hyperdiverse habitat housing c.15% of the Brazilian vascular flora in less than 1% of the Brazilian territory. Amongst other factors, the remarkable plant diversity in *campos rupestres* has been attributed to its occurrence as many isolated patches and to floristic influences from surrounding habitats, including lowland woody savannas (*cerrado*), Atlantic rain forests, seasonally dry woodlands and Amazonian rain forests. However, no study has assessed the degree to which the putative floristic influence from surrounding habitats drives compositional variation in *campos rupestres*. Here, we used a dataset on the composition of South American woody plant communities (4,637 community surveys, with 115 representing *campos rupestres*), combined with environmental data, with the aim of characterising and explaining compositional variation of the *campos rupestres* woody flora. Our results showed that all *campos rupestres*, including the sites occurring in Amazonian ironstone formations, are more similar to *cerrado* woody savannas than to any other South American vegetation formations covered in our dataset. Also, multiple *campo rupestre* floristic groups may be recognized based on distinct species composition and environmental conditions, primarily related to substrate and climate. We stress the importance of considering this floristic heterogeneity in conservation, management and research planning.

1. Introduction

The rocky montane savannas of South America, known as *campos rupestres* in Brazil, where they largely occur, are found on quartzite, sandstone and ironstone formations, mostly above 900 m (a.s.l.) and up to 2033 m (Giulietti et al., 1997; Fernandes et al., 2014; Silveira et al., 2015). Its core area is spread along the highlands of eastern Brazil (Giulietti et al., 1997; Hughes et al., 2013; Silveira et al., 2015). Disjunct areas also occur along mountain ranges in central-western Brazil (Frisby and Hind, 2014; Mews et al., 2014; Silveira et al., 2015), eastern Bolivia (Saravia, 2008) and in the Amazon Forest (Silveira et al., 2015). The *campos rupestres* are a growing focus of attention because they have been recently proposed as one of the world's old climatically-buffered infertile landscapes (OCBILs, e.g., the *fynbos* of the Cape Floristic Region; Silveira et al., 2015).

In Brazil, the *campos rupestres* contain c.5000 vascular plant species (REFLORA, 2016), corresponding to a remarkable c.15% of the Brazilian vascular flora in less than 1% of the Brazilian territory (Fernandes et al., 2014; Silveira et al., 2015). Amongst other factors, this outstanding floristic diversity in *campos rupestres* – a pattern common to OCBILs (Hopper, 2009) – has been attributed to its high levels of local endemism (Hensold, 1988; Echternacht et al., 2011a) as well as to the geographically disjunct distribution of *campo rupestre* sites and, hence, the associated floristic influence from distinct habitats (Giulietti et al., 1997), namely *cerrado* woody savannas, Atlantic rain forests, seasonally dry woodlands and Amazonian rain forests. However, to our knowledge, no study has assessed the degree to which this alleged floristic influence from surrounding habitats drives compositional differentiation of *campos rupestres*.

☆ This article is part of a special issue entitled Plant life in campo rupestre: new lessons from an ancient biodiversity hotspot published at the journal FLORA 238C.

* Corresponding author.
E-mail addresses: dneves@email.arizona.edu, danilormn@gmail.com (D.M. Neves).

https://doi.org/10.1016/j.flora.2017.03.011
0367-2530/

Attempts to address this knowledge gap could be of importance for effective conservation strategies. If the floristic variation of *campos rupestres* is high, with multiple distinct floristic groups, future conservation assessments could highlight, for instance, that a large number of separate conservation areas are needed to fully protect *campos rupestres* diversity. Here we go a step further in data refinement and analysis by using a large dataset on the composition of South American woody plant communities (>4,000 community surveys, with >100 representing *campos rupestres*), combined with environmental data, in order to elucidate the spatial floristic patterns of *campos rupestres*. We address the following hypotheses stemming from the literature (Giulietti et al., 1997; Echternacht et al., 2011b): (h_1) multiple *campo rupestre* floristic groups may be recognized based on distinct species composition; (h_2) community composition differentiation amongst *campo rupestre* floristic groups can be predicted by variation in environmental conditions.

2. Material and methods

2.1. Study area

The South American rocky montane savannas (henceforth *campo rupestre* for a single site, and *campos rupestres* for multiple sites) cover c.65,000 km^2 (Fernandes et al., 2014) and occur primarily on quartzite and sandstone formations in the highlands (mostly >900 m a.s.l) of eastern Brazil as well as in scattered mountain ranges in central-western Brazil (Frisby and Hind, 2014; Mews et al., 2014; Silveira et al., 2015). These highlands border three primary vegetation 'Domains' (IBGE, 1993; Ab'sáber, 2003): the Atlantic Domain to the east and south (known as *Mata Atlântica* in Brazil), the Caatinga Domain to the north and the Cerrado Domain to the west (see Giulietti et al., 1997; Hughes et al., 2013; Silveira et al., 2015). The prevailing land cover of these bordering Domains are rain forest in the *Mata* Atlântica, semi-arid thorn woodlands in the Caatinga and woody savannas in the Cerrado. *Campos rupestres* are also found in ironstone formations of south-eastern and central-western Brazil, eastern Bolivia and the south-eastern Amazon Forest (known as *cangas* in Brazil; Jacobi and Carmo, 2011; Silveira et al., 2015). *Campo rupestre* landscapes also house patches of transitional vegetation (e.g., parkland savannas, riverine forests), but here we adopt the *sensu stricto* definition of *campos rupestres*, which comprises only the grassy-shrubby savannas on quartzite, sandstone or ironstone rock outcrops (Alves et al., 2014). Many *campo rupestre* sites in this contribution were not included in previous studies (Fernandes et al., 2014; Silveira et al., 2015), especially those found on quartzite and sandstone outcrops across Goiás state (central-western Brazil; Mews et al., 2014) and the ironstone-associated *campos rupestres* found in Mato Grosso do Sul state, near the Brazil-Bolivia border (Neves and Damasceno-Junior, 2011). Mountaintop grasslands (*campos de altitude*), which are found nearer to the Atlantic coast (Ribeiro et al., 2007), were not included in this contribution because their flora is distinct and more closely related to that of the *páramos* in the Andes (Safford, 2007).

2.2. Dataset

We extracted the dataset from the NeoTropTree (NTT) database (http://prof.icb.ufmg.br/treeatlan), which consists of checklists of woody, freestanding (i.e., lianas excluded) plant species, compiled for geo-referenced sites, extending from southern Florida (U.S.A.) and Mexico to Patagonia. NTT currently holds 5,126 sites/checklists, 14,878 woody plant species and 920,129 occurrence records. A site/checklist in NTT is defined by a single vegetation type, following the classification system proposed by Oliveira-Filho (2015), contained in a circular area with a 10-km diameter. Where two or more vegetation types co-occur in one 10-km area, there may be two geographically overlapping sites in the NTT database, each for a distinct vegetation type.

The data were originally compiled from an extensive survey of published and unpublished literature (e.g., PhD theses), particularly those comprising floristic surveys and forest inventories. New species occurrence records obtained from major herbaria and taxonomic monographs have been added to the checklists when they come from within the 10-km diameter of the original NTT site, and within the same vegetation type. All species and their occurrence records were checked regarding current taxonomic and geographical circumscriptions, as defined by the team of specialists responsible for the online project *Flora do Brasil* (available at http://floradobrasil.jbrj.gov.br/). The compilation of NTT avoided, therefore, the inclusion of occurrence records with doubtful identification, location or vegetation type. It also excluded checklists with very low species richness (<20 species), because this is often due to low sampling/collecting efforts, which results in poor descriptive power.

The dataset extracted from NTT consisted of 4,637 South American woody plant community surveys, of which 115 were *campos rupestres* from eastern and central western Brazil, south-eastern Brazilian Amazon and eastern Bolivia. The full species matrix contained presence/absence data for 11,954 woody plant species, with a total of 883,258 presences, and the *campos rupestres* species matrix contained presence/absence data for 1,055 woody plant species, with a total of 12,801 presences.

The NTT database also includes 24 environmental variables for all sites, obtained from multiple sources. Elevation (m above sea level) at the site centre, obtained from WorldClim 1.4 (Hijmans et al., 2005), was used as an integrative environmental variable. Isothermality, maximum temperature of warmest month, mean annual temperature, mean annual precipitation, mean daily temperature range, minimum temperature of coldest month, precipitation of driest month, precipitation of wettest month, precipitation seasonality, temperature annual range and temperature seasonality were obtained from WorldClim 1.4 data layers (Hijmans et al., 2005). Cloud interception (mm) and frost frequency (days) were obtained from modelling known values as response variables (data obtained from 135 and 57 Brazilian Meteorological Stations measuring frost frequency and cloud interception, respectively), and elevation, latitude and the aforementioned WorldClim layers as predicting variables. Duration (days) and severity (mm) of water deficit were extracted from Walter's Climate Diagrams (Walter, 1985) generated by interpolating 5-day intervals of monthly temperatures and precipitation (WorldClim 1.4; Hijmans et al., 2005). Potential evapotranspiration (mm) and an aridity index (annual precipitation/potential evapotranspiration) were derived by Zomer et al. (2007, 2008) from WorldClim data.

Six variables were substrate-related: grass coverage (%), obtained by direct observation of the site surface via Google Earth© images in five 100 × 100 m areas, one at the central coordinates of the NTT site and four at 2.5 km away from it and towards the NE, SW, NW and SE; soil coarseness (% sand), soil fertility (% base saturation) soil salinity (ds/m) and surface rockiness (% exposed rock), obtained from the Harmonized World Soil Database v 1.2 (available at http://www.fao.org/soils-portal/soil-survey) and then ranked by mid-class percentage (raw figures were unrealistic due to local soil heterogeneity); and soil drainage classes, obtained following EMBRAPA's protocol (Santos et al., 2013), which combines soil type, texture and depth with land forms.

2.3. Analyses of community composition

We used Simpson distance as the dissimilarity metric and unweighted paired groups as the linkage method in a hierarchical clustering analysis (McCune and Grace, 2002). We built 1000 clusters, with each cluster being built after randomising the row order in the species composition matrix (species per site), following the procedure proposed by Dapporto et al. (2013). The final cluster is assembled following the majority consensus rule: if a given group is represented in at least 50% of the trees built using a given set of samples, that group is represented in the final consensus tree (Omland et al., 2008). This analysis was conducted using the recluster package (Dapporto et al., 2015) in the R Statistical Environment (R Core Team, 2016).

We assessed the overall patterns of floristic identity in *campos rupestres* by (i) analyzing species occupancy (i.e., with species incidences rather than abundances), and (ii) performing an ordination of *campo rupestre* woody plant communities (115 sites) by non-metric multidimensional scaling (NMDS) of their species composition (McCune and Grace, 2002) using Simpson distance as the dissimilarity metric. Following methods similar to those proposed by Kreft and Jetz (2010), the colours blue, green, yellow and red were assigned to the four corners of the two-dimensional ordination plot in clockwise order from the origin. NMDS scores were then mapped in geographic space by assigning a colour to each site according to its position in the two-dimensional ordination space. Beforehand, the ordination was rescaled to axes ranging from 0 to 1. Rescaling is possible with NMDS results since ordination axes as such have no meaning and only the relative position of points in ordination space matters. The NMDS and the colour assignment were conducted in the statistical packages vegan (Oksanen et al., 2016) and recluster (Dapporto et al., 2015), respectively, both in the R Statistical Environment (R Core Team, 2016).

We tested whether variation in environmental conditions can predict differentiation in *campos rupestres* community composition, and then explored the results visually by (i) plotting the NMDS scores in ordination and geographic space, and (ii) fitting the values of the most important environmental variables by generalized linear models (GLM) and generalized additive models (GAM), respectively. This routine follows methods similar to those proposed by Blanchet et al. (2008) and Legendre et al. (2012), which comprise (i) the exclusion of 261 unicates (species found at a single site), as they commonly increase the noise in most analyses without contributing information (Lepš and Šmilauer, 2003); (ii) the Hellinger transformation of the binary presence/absence data (Legendre and Gallagher, 2001), which reduces the weight of widespread species and their inherent effect in ordination analyses; (iii) the independent compilation of significant spatial and environmental variables through a forward selection method for redundancy analysis (RDA), after first checking that the respective global models were significant (Blanchet et al., 2008); (iv) an additional and progressive elimination of collinear variables based on their variance inflation factor (VIF) and ecological relevance, until maintaining only those with VIF <10 (Quinn and Keough, 2002); and (v) variation partitioning of the community composition matrix with respect to the significant spatial and environmental variables. As spatial variables, we used principal coordinates of neighbour matrices (PCNMs; Borcard et al., 2004), which represent the spatial structure of the sampling units at multiple spatial scales without considering any environmental variation (Borcard et al., 1992; Legendre et al., 2002; Borcard et al., 2004). We tested the overall significance of the environmental fraction (controlled for spatial autocorrelation) by applying ANOVA permutation tests (999 permutations) for RDA (Peres-Neto et al., 2006). The variable selection, variation partitioning and GLM/GAM analyses were conducted using the fields (Nychka et al., 2015), spacemakeR (Dray, 2010) and vegan (Oksanen et al., 2016) packages in the R Statistical Environment.

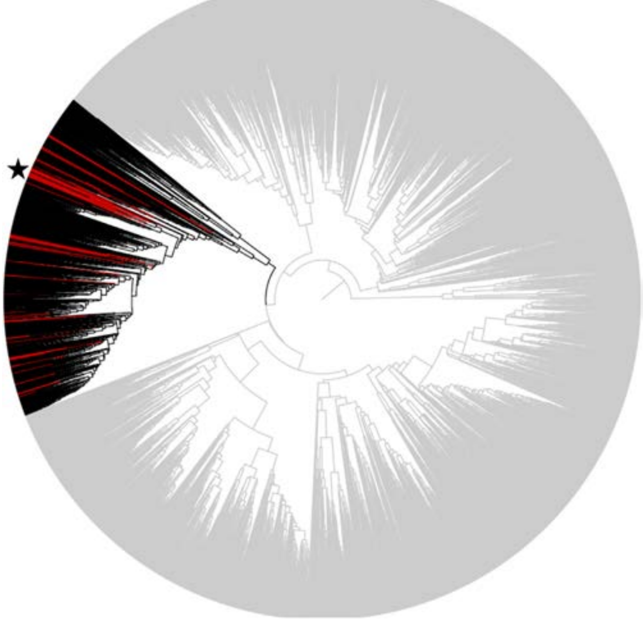

Fig. 1. Hierarchical clustering for 4,637 South American woody plant communities based on their species composition. The dissimilarity measure and linkage methods used were Simpson and unweighted group average, respectively. The woody plant communities are discriminated by different colours: black, 621 *cerrado* woody savannas; red, 115 *campos rupestres*; grey, 3,901 other South American woody plant communities. The star highlights a cluster comprising many *campo rupestre* sites in southeastern and central-western Brazil. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Finally, we conducted an assessment of the conservation status of *campos rupestres* by overlaying the distribution of our 115 sites on to the coverage of protected areas across South America. We used conservation units from the *Cadastro Nacional de Unidades de Conservação* (Ministério do Meio Ambiente – Brazil, www.mapas.mma.gov.br) and World Database on Protected Areas (IUCN & UNEP – WCMC, www.protectedplanet.net). All maps were designed using the package maptools (Lewin-Koh and Bivand, 2012) in the R Statistical Environment.

3. Results

3.1. Floristic patterns

The hierarchical clustering showed that *campos rupestres*, including the sites occuring in Amazonian ironstone formations, are more similar to *cerrado* woody savannas than to any of the other South American vegetation formations (Fig. 1). These *campo rupestre* sites share 95% of their woody plant species with other habitats in our community dataset, ranging from 74% of species shared with Cerrado woody plant formations to less than 40% of species shared with woody plant formations from the Caatinga Domain. *Campo rupestre* sites do not form a single cluster but are scattered within a broad *cerrado* cluster (Fig. 1). The assessment of species incidences revealed that 80% of species have relatively low occupancy across *campos rupestres* (dashed line in Fig. 2). The distribution of *campos rupestres* in the ordination space yielded by NMDS suggests a compositional segregation into four relatively distinct floristic groups (Fig. 3), namely northern (blue spectrum), south-eastern (brown-yellow spectrum), north-eastearn (green spectrum) and central-western (red-purple spectrum) *campos rupestres*. Bolivian *campos rupestres* are floristically related to

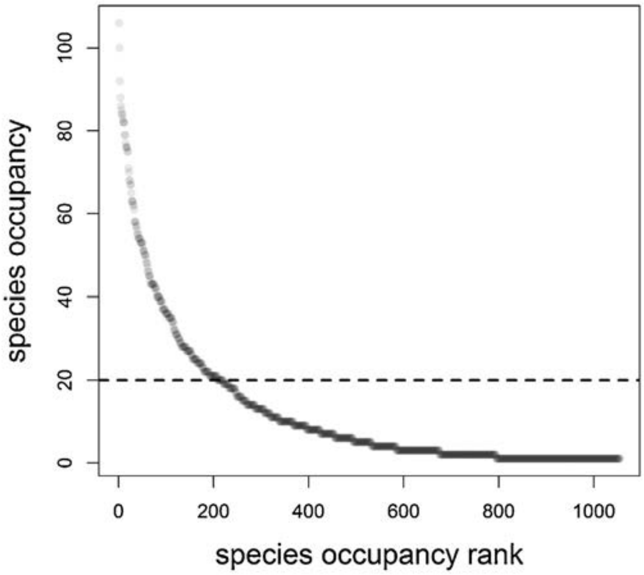

Fig. 2. Rank occupancy of *campo rupestre* woody plant species. Each grey circle represents a *campo rupestre* species in our dataset. Darker shades of grey indicate overlapping circles (i.e., two or more species have similar occupancies). Circles below the dashed line occur in 20 or fewer *campo rupestre* sites (80% of the 1,055 species).

the central-western group (see similarity in Fig. 3b), and the floristic differentiation between eastern groups is comparatively more subtle.

The conservation status of *campo rupestre* sites is uneven across these four floristic groups (Fig. 4a). Many central-western *campos rupestres* are unprotected, while northern *campos rupestres* are better protected (Fig. 4a). The conservation assessment revealed that the current network of protected areas does not cover the entire floristic space of *campo rupestre* woody plant communities (Fig. 4b).

3.2. Environmental drivers of community turnover

The forward selection procedure retained six PCNMs for modelling variation in *campo rupestre* community composition (adjusted $R^2 = 0.158$, which is fairly close to the value for all 38 PCNMs without any selection, adjusted $R^2 = 0.159$). These selected spatial vectors are amongst the first PCNMs, which represent broad-scale, positive spatial autocorrelation. Regarding environmental variables, the forward selection retained 13 environmental variables (adjusted $R^2 = 0.271$, which is near the value for all 24 environmental variables without any forward selection, adjusted $R^2 = 0.304$) for modelling variation in *campo rupestre* community composition (Table 1). Altitude, aridity index, isothermality, maximum temperature of hottest month, minimum temperature of coldest month, potential evapotranspiration, precipitation of wettest month, salinity, temperature annual range, and water deficit duration and severity were the excluded environmental variables.

When partitioning the variation explained by the retained environmental and spatial predictors, we found that the environmental fraction explained 27% of the variation, 15% of which was independent of spatial autocorrelation ($P < 0.01$). The environmental predictors could not account for 3% of the spatially structured variation ($P < 0.01$), and 70% of the variation remained unexplained. By fitting the values of the most important environmental variables in ordination and geographic space (Fig. 3a and b, respectively), we observed a strong east to west gradient related to decreasing surface rockiness (Fig. 3a and b), a proxy for soil water deficit, thus segregating eastern *campos rupestres* from northern and central-western *campos rupestres*. A south-east to north gradient was related to increasing mean annual temperature (MAT) and decreasing temperature seasonality (TempSeas), with northeastern and central-western *campos rupestres* occurring in intermediate MAT and TempSeas (Fig. 3a and b). Mean annual precipitation (MAP) was the third most important variable and was associated with the floristic differentiation of north-eastern from northern *campos rupestres*, with south-eastern and central-western occurring in intermediate MAP (Fig. 3b).

4. Discussion

4.1. Floristic identity of campos rupestres

The first hypothesis was clearly supported by our results. Multiple *campo rupestre* floristic groups may be recognized, based on

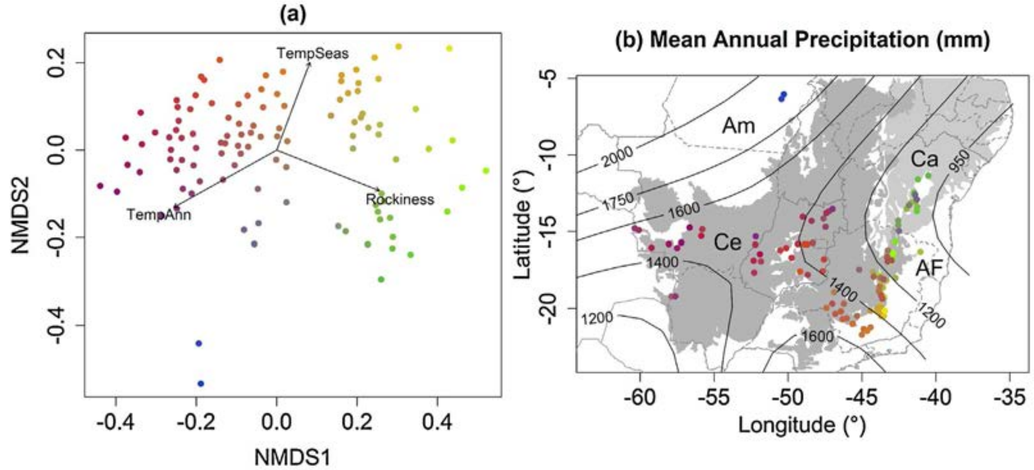

Fig. 3. Ordination of 115 sites of *campo rupestre* inferred from non-metric multidimensional scaling (NMDS) of their woody plant species composition (a), and geographical variation of species composition and mean annual precipitation (b). NMDS scores were plotted in the ordination diagram after assigning a colour to each site according to its position in the two-dimensional ordination space (axes 1 × 2). Variation in surface rockiness, mean annual temperature and temperature seasonality were fitted in ordination space by generalized linear models. Colours of circles plotted across geographic space are identical to the colours of circles in the NMDS scatter plot. Variation in mean annual precipitation was fitted across geographic space by a generalized additive model. Dashed lines in (b) represent Brazilian state borders. AF = Atlantic Forest Domain (white along the Atlantic coast); Am = Amazon Domain (white in northwestern South America); Ca = Caatinga Domain (light grey); Ce = Cerrado Domain (dark grey).

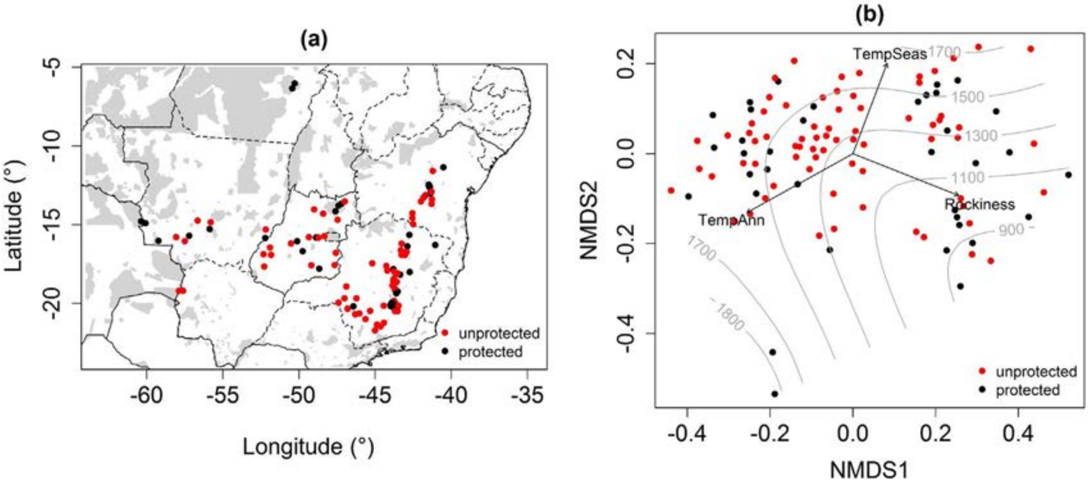

Fig. 4. Conservation assessment of *campo rupestre* woody plant communities. (a) Distribution of protected and unprotected *campos rupestres* in South America. Grey areas represent the current network of protected areas across South America. Dashed lines represent Brazilian state borders. (b) Conservation status of the two-dimensional floristic space of *campo rupestre* woody plant communities. Circles represent the position of *campo rupestre* sites in ordination space inferred from non-metric multidimensional scaling (NMDS axes 1 × 2) and are identical to the position of *campos rupestres* in Fig. 3a. Variation in surface rockiness, mean annual temperature and temperature seasonality were fitted in ordination space by generalized linear models. Contours representing mean annual precipitation were fitted in ordination space by a generalized additive model. Dashed lines in (a) represent Brazilian state borders.

Table 1
Variables selected for the study of environmental drivers of community turnover across *campos rupestres*. The variables shown were selected through a forward selection method for redundancy analysis and are ordered by the amount of explained variation in species composition across *campo rupestre* sites. Goodness-of-fit of the predictor variables was assessed through adjusted coefficients of determination, Akaike Information Criterion (AIC), F-values and significance tests (p-value). VIF, variance inflation factor, was obtained using the r-squared value of the regression of one variable against all other explanatory variables.

	cumulative adjusted R^2	ΔAIC	F-value	p-value	VIF
Rockiness	0.12	−55.55	15.86	0.002	2.64
Mean annual temperature	0.15	−59.75	6.2	0.002	7.20
Mean annual precipitation	0.18	−63.03	5.22	0.002	1.69
Temperature seasonality	0.21	−65.07	3.93	0.002	3.77
Precipitation seasonality	0.22	−66.34	3.14	0.002	5.59
Grass coverage	0.23	−67.44	2.96	0.002	1.78
Soil drainage	0.24	−67.61	2.03	0.002	1.29
Soil fertility	0.25	−67.79	2.03	0.002	4.17
Days of frost	0.25	−67.71	1.77	0.002	1.29
Mean daily temperature range	0.26	−67.62	1.74	0.002	1.50
Precipitation of driest month	0.26	−67.45	1.65	0.002	6.21
Cloud interception	0.27	−67.2	1.57	0.002	4.42
Sandiness	0.27	−66.66	1.29	0.028	4.75

distinct woody plant species composition. Instead of representing a single floristic group across South America, the *campos rupestres* form several separate groups within a wider *cerrado* savannas group. This is the first attempt to show the degree to which the geographically disjunct distribution of *campo rupestre* sites, and its associated environmental heterogeneity, is underpinning the outstanding floristic diversity in *campos rupestres*. Despite the fact that our dataset only comprises woody plants, we predict that subsequent studies focusing on herbs (a large component of *campos rupestres* floristic diversity) and/or animals will reinforce this claim; i.e., influence from surrouding habitats is an important factor shaping overall species composition in *campos rupestres*. Also, we hypothesize that the high level of local endemism found in the non-woody component of *campos rupestres* (Hensold, 1988; Mello-Silva, 1989; Echternacht et al., 2011a) leads to even greater floristic heterogeneity amongst *campos rupestres*; i.e., floristic dissimilarity amongst *campos rupestres* is even higher if considering the non-woody component.

Our results show that individual *campo rupestre* groups share more species with surrounding lowland *cerrados* than they do with other *campo rupestre* groups. On the other hand, the *campo rupestre* sites in southeastern and central-western Brazil represent a large and relatively cohesive floristic group of *campos rupestres* (see star in Fig. 1), in agreement with a considerable degree of floristic similarity between *campos rupestres* from the *Espinhaço* range (southeastern Brazil) and the disjunct mountain ranges from central-western Brazil (Feres et al., 2009).

4.2. *Environmental drivers of community turnover*

The second hypothesis was also supported by our results. Community composition differentiation amongst *campo rupestre* floristic groups can be predicted by variation in environmental conditions. Our results show that northern *campos rupestres* occur in wet and warm environments with lower surface rockiness (i.e., low soil water deficit). The northeastern group occurs in the driest extreme of the precipitation space occupied by *campos rupestres*, whereas southeastern and central-western *campos rupestres* are found in intermediate, moist enviroments. The later two groups diverge over two other important gradients: the southeastern group occurs in environments with lower mean annual temperature and higher surface rockiness, while the central-western group has lower rockiness and intermediate mean annual temperature.

Variation in environmental conditions across the geographically disjunct distribution of *campos rupestres* seems to be the main factor leading to floristic divergence of *campo rupestre* woody plant communities. However, understanding species-environment relationships in *campo rupestre* woody plant communities is complex, as it partly depends on understanding the floristic relationships between *campos rupestres* and their surrounding lowland *cerrados*. On the one hand, the environmental gradients found across *campos rupestres* have given rise to a pattern of low species occupancy such that many *campo rupestre* woody plant species inhabiting this gradient can be said to belong to only one of the four floristic groups. On the other hand, most of these species also occur in other habitats of our woody plant community database, suggesting that the environmental similarity between *campos rupestres* and the surrounding lowland *cerrados* has allowed a regular exchange of woody plant species between these two habitats. This is in agreement with a previous study showing that *campos rupestres* and lowland *cerrados* in Goiás state, a portion of our central-western group, differ in population structure of their woody plant species but not in composition (Mews et al., 2014). From an ecological perspective, *campos rupestres* and their surrounding lowland *cerrados* are likely to form a continuous metacommunity with spatial variation in woody plant population sizes being mainly driven by source-sink dynamics (Pulliam and Danielson 1991); i.e., species that are better adapted to lowland *cerrados* (source habitat for this species) are also found in *campos rupestres* (sink habitat for this species), though in smaller populations, since species better adapted to rocky substrate and shallower soils will prevail in population size.

4.3. *Spatial structure*

The *campo rupestre* floristic groups are largely geographic, thus suggesting that there may be a role for spatially structured dispersal limitation and historical biogeography in driving floristic differentiation. Nevertheless, our results indicate that environmental conditions are better predictors of community turnover (a proxy for niche-based dispersal limitation) than are geographical factors (i.e., community composition/differentiation of unsampled *campos rupestres* would be better predicted based on environmental similarity than by geographic proximity). This is supported by the negligible unique variation attributed to positive spatial autocorrelation in *campos rupestres*, a proxy for a distance decay in community similarity (Nekola and White, 1999), and by the fact that it is more parsimonious to attribute the spatially structured environmental variation to niche-based controls (cf. Legendre et al., 2009; Neves et al., 2015). The niche-based dispersal limitation in *campos rupestres* is further supported by two other results: (i) the comparatively high compositional variation in southeastern Brazil is most likely to be associated with the role of environmental heterogeneity in underpinning the occurrence of three floristic groups, regardless of geographic proximity; and (ii) 95% of woody plant species in *campos rupestres* are also found in other habitats, but have restricted distributions across *campos rupestres*, likely because environmental conditions are more similar between *campos rupestres* and surrounding lowland *cerrados* than between geographically distant *campo rupestre* groups (spatially structured environmental variation).

4.4. *Conservation implications*

Threats to *campo rupestre* biodiversity are many, and include mining, unplanned urbanisation, high frequency of anthropogenic fire, uncontrolled harvesting of ornamental plants, eucalyptus plantations, selective logging and unplanned tourism (Giulietti et al., 1997; Jacobi et al., 2007, 2011; Fernandes et al., 2014; Silveira et al., 2015). Considering the pervasive nature of most of these threats, conservation strategies for *campos rupestres* need to be urgent and well-informed scientifically. We believe our findings fit the 'well-informing' criteria and are therefore of relevance for conservation planning. Here we show that *campos rupestres* are in fact segregated into three or four compositionally distinct floristic groups, which dictates that each group deserves separate conservation planning. In doing so, future assessments may call attention to the distribution of protected areas within each of these *campo rupestre* groups. Recent studies have shown that at smaller geographic scales (e.g., *Espinhaço* Range in eastern Brazil), several areas could be distinguished based on taxonomic and evolutionary uniquiness of plants (Echternacht et al., 2011b; Bitencourt and Rapini, 2013; Souza et al., 2013; Echternacht et al., 2014).

We also showed that *campo rupestre* floristic groups are unevenly protected and that geographical gaps in the distribution of conservation units result in a failure to protect important parts of the *campo rupestre* floristic space. More specifically, *campos rupestres* found at intermediate values of the floristic space summarized by the first NMDS axis are largely unprotected. These are *campos rupestres* occurring under intermediate mean annual precipitation (c.1500 mm) in western Goiás state (central-western group) and southern Minas Gerais state (southeastern group). We also call attention to the *campos rupestres* found in Mato Grosso do Sul state, near the Bolivia border. These *campos rupestres*, occurring in the ironstone formations of the Urucum plateau, are largely unprotected, poorly studied and highly threatened by opencast mining (Neves and Damasceno-Junior, 2011).

An alternative, and important, route forward in conservation planning lies in addressing the evolutionary history of these *campo rupestre* groups. Previous studies indicate that plant lineages from multiple biogeographical origins have colonized *campos rupestres* many times over evolutionary history. For instance, some bromeliad (Versieux et al., 2012) and orchid species (Gustafsson et al., 2010) found in *campos rupestres* seem to have an Atlantic rain forest origin (i.e., sister taxa are mainly found in Atlantic rain forests), while some legume species seem to have an origin in seasonally dry woodlands (Souza et al., 2013). In addition, others have stressed the idea that *campos rupestres* have acted as 'species pump' for the surrounding lowland habitats (Simon et al., 2009; Silveira et al., 2015). Either way, future studies intending to quantify these evolutionary shifts could shed light into the historical assembly of the *campo rupestre* flora and, potentially, emphasize the necessity of conservation strategies aiming to protect distinct *campo rupestre* groups along with associated surrounding habitats.

5. Conclusion

We found an overall lack of compositional identity across the *campos rupestres* woody flora, which is driven by their geographically disjunct distribution with its associated environmental heteregoneity and floristic influence from surrounding habitats. Therefore, we stress the necessity of considering such floristic and environmental heterogeneity in conservation, management and research planning and emphasize the need for multiple protected areas across the separate floristic groups of *campos rupestres*. Our findings also indicate that *campos rupestres* and their surrounding lowland *cerrados* exchange woody plant species regularly and, therefore, merit simultaneous conservation attention. Conservation units aiming to protect *campo rupestre* biodiversity should not be limited to *campo rupestre* areas. Rather, effective protected areas should function as ecological corridors connecting multiple *campos rupestres* through lowland *cerrados*. We predict that future studies will confirm that lowland *cerrados* are linking geographically distant woody plant populations, thus improving ecological functionality of *campos rupestres*, such as gene flow between *campo rupestre* sites.

Acknowledgements

AOF and MLB were supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brazil (CNPq) (grants 301644/88-8 and 151002/2014-2, respectively). D.M.N. and R.T.P. were supported by the National Environmental Research Council (grant NE/I028122/1). PLSM thanks the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) for supporting a full PhD at the University of Edinburgh under the Science Without Borders Programme (grant BEX 13197-13-4). We are indebted to Fernando A.O. Silveira, Fabio R. Scarano and one anonymous referee for their valuable contributions to the manuscript.

References

Ab'sáber, A.N., 2003. Os Domínios de Natureza No Brasil: Potencialidades Paisagísticas. Ateliê Editorial, São Paulo.

Alves, R.J.V., Silva, N.G., Oliveira, J.A., Medeiros, D., 2014. Circumscribing campo rupestre – megadiverse Brazilian rocky montane savannas. Braz. J. Biol. 74, 355–362.

Bitencourt, C., Rapini, A., 2013. Centres of endemism in the Espinhaço Range: identifying cradles and museums of Asclepiadoideae (Apocynaceae). Syst. Biodivers. 11, 525–536.

Blanchet, F.G., Legendre, P., Borcard, D., 2008. Forward selection of explanatory variables. Ecology 89, 2623–2632.

Borcard, D., Legendre, P., Drapeau, P., 1992. Partialling out the spatial component of ecological variation. Ecology 73, 1045–1055.

Borcard, D., Legendre, P., Avois-Jacquet, C., Tuomisto, H., 2004. Dissecting the spatial structure of ecological data at multiple scales. Ecology 85, 1826–1832.

Dapporto, L., Ramazzotti, M., Fattorini, S., Talavera, G., Vila, R., Dennis, R.L.H., 2013. Recluster: an unbiased clustering procedure for beta-diversity turnover. Ecography 36, 1070–1075.

Dapporto, L., Ramazzotti, M., Fattorini, S., Vila, R., Talavera, G., Dennis, R.H.L., 2015. Recluster: ordination methods for the analysis of beta-diversity indices. In: R Package Version 2.8. http://CRAN.R-project.org/package=recluster.

Dray, S., 2010. SpacemakeR: spatial modelling. In: R Package Version 473 0.0-5/r101. http://R-Forge.R-project.org/projects/sedar/.

Echternacht, L., Sano, P.T., Trovó, M., Dubuisson, J., 2011a. Phylogenetic analysis of the Brazilian microendemic Paepalanthus subgenus Xeractis (Eriocaulaceae) inferred from morphology. Bot. J. Linn. Soc. 167, 137–152.

Echternacht, L., Trovó, M., Oliveira, C.T., Pirani, J.R., 2011b. Areas of endemism in the espinhaço range in Minas Gerais, Brazil. Flora 206, 782–791.

Echternacht, L., Sano, P.T., Bonillo, C., Cruaud, C., Couloux, A., Dubuisson, J.-Y., 2014. Phylogeny and taxonomy of Syngonanthus and Comanthera (Eriocaulaceae): evidence from expanded sampling. Taxon 63, 47–63.

Feres, F., Zucchi, M.I., Souza, A.P., Amaral, M.C.E., Bittrich, V., 2009. Phylogeographic studies of brazilian 'campo-rupestre' species: wunderlichia mirabilis riedel ex baker (Asteraceae). Biotemas 22, 17–26.

Fernandes, G.W., Barbosa, N.P.U., Negreiros, D., Paglia, A.P., 2014. Challenges for the conservation of vanishing megadiverse rupestrian grasslands. Nat. Conservacao 12, 162–165.

Frisby, S., Hind, D.J.N., 2014. Ichthyothere sasakiae, (compositae: Heliantheae: Milleriinae), a new species from the Amazonian campo rupestre of northern Mato Grosso State, Brazil. Kew Bull. 69, 9504.

Giulietti, A.M., Pirani, J.R., Harley, R.M., 1997. Espinhaço range region, eastern Brazil. In: Davis, S.D., Heywood, V.H., Herrera-Macbryde, O., Villa-Lobos, J., Hamilton, A.C. (Eds.), Centres of Plant Diversity: A Guide and Strategy for Their Conservation. IUCN Publication Unit, Cambridge, UK, pp. 397–404.

Gustafsson, A.L.S., Verola, C.F., Antonelli, A., 2010. Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus *Hoffmannseggella* (Orchidaceae: Epidendroideae). BMC Evol. Biol. 10, 177.

Hensold, N., 1988. Morphology and Systematics of Paepalanthus Subgenus Xeractis (Eriocaulaceae) Systematic Botany Monographs 23. The American Society of Plant Taxonomists, Michigan.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978.

Hopper, S.D., 2009. OCBIL theory: towards an integrated understanding of the evolution, ecology and conservation of biodiversity on old, climatically buffered, infertile landscapes. Plant Soil 322, 49–86.

Hughes, C.E., Pennington, R.T., Antonelli, A., 2013. Neotropical plant evolution: assembling the big picture. Bot. J. Linn. Soc. 171, 1–18.

IBGE, 1993. Mapa De Vegetação Do Brasil. Fundação Instituto Brasileiro de Geografia e Estatística, Ministério da Agricultura, Rio de Janeiro.

Jacobi, C.M., Carmo, F.F., 2011. Life-forms, pollination and seed dispersal syndromes in plant communities on ironstone outcrops, SE Brazil. Acta Bot. Bras. 25, 395–412.

Jacobi, C.M., Carmo, F.F., Vincent, R.C., Stehmann, J.R., 2007. Plant communities on ironstone outcrops – a diverse and endangered Brazilian ecosystem. Biodivers. Conserv. 16, 2185–2200.

Jacobi, C.M., Carmo, F.F., Campos, I.C., 2011. Soaring extinction threats to endemic plants in Brazilian metal-rich regions. AMBIO 40, 540–543.

Kreft, H., Jetz, W., 2010. A framework for delineating biogeographical regions based on species distributions. J. Biogeogr. 37, 2029–2053.

Legendre, P., Gallagher, E.D., 2001. Ecological meaningful transformations for ordination of species data. Oecologia 129, 271–280.

Legendre, P., Dale, M.R.T., Fortin, M.J., Gurevitch, J., Hohn, M., Myers, D., 2002. The consequences of spatial structure for the design and analysis of ecological field surveys. Ecography 25, 601–615.

Legendre, P., Mi, X., Ren, H., Ma, K., Yu, M., Sun, I., He, F., 2009. Partitioning beta diversity in a subtropical broad-leaved forest of China. Ecology 90, 663–674.

Legendre, P., Borcard, D., Roberts, D.W., 2012. Variation partitioning involving orthogonal spatial eigenfunction submodels. Ecology 93, 1234–1240.

Lepš, J., Šmilauer, J.P., 2003. Multivariate Analysis of Ecological Data Using CANOCO. Cambridge University Press, Cambridge, UK.

Lewin-Koh, N.J., Bivand, R., 2012. Maptools: tools for reading and handling spatial objects. In: R Package Version 0.8-17/r238, Available at: http://R-Forge.R project.org/projects/maptools/.

McCune, B., Grace, J.B., 2002. Analysis of Ecological Communities. MjM Software Design, Gleneden Beach, Oregon.

Mello-Silva, R., 1989. Velloziaceae de grão-mogol, minas gerais Brazil. In: MSc Thesis. Universidade de São Paulo.

Mews, H.A., Pinto, J.R.R., Eisenlohr, P.V., Lenza, E., 2014. Does size matter? Conservation implications of differing woody population sizes with equivalent occurrence and diversity of species for threatened savanna habitats. Biodivers. Conserv. 23, 1119–1131.

Nekola, J.C., White, P.S., 1999. The distance decay of similarity in biogeography and ecology. J. Biogeogr. 26, 867–878.

Neves, D.M., Damasceno-Junior, G.A., 2011. Post-fire phenology in a *campo sujo* vegetation in the Urucum plateau, Mato Grosso do Sul, Brazil. Braz. J. Biol. 71, 881–888.

Neves, D.M., Dexter, K.G., Pennington, R.T., Bueno, M.L., Oliveira-Filho, A.T., 2015. Environmental and historical controls of floristic composition across the South American Dry Diagonal. J. Biogeogr. 42, 1566–1576.

Nychka, D., Furrer, R., Paige, J., Sain, S., 2015. Fields: Tools for Spatial Data R Package Version 8., pp. 3–5 http://CRAN.R-project.org/package=fields.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2016. Vegan: Community Ecology Package R Package Version 2., pp. 0–3 http://CRAN.R-project.org/package=vegan.

Oliveira-Filho, A.T., 2015. Um Sistema de classificação fisionômico-ecológica da vegetação Neotropical. In: Eisenlohr, P.V., Felfili, J.M., Melo, M.M.R.F., Andrade, L.A., Meira-Neto, J.A.A. (Eds.), Fitossociologia no Brasil: métodos e Estudos de Casos, vol. 2. Editora UFV, Viçosa, Brazil, pp. 452–473.

Omland, K.E., Cook, L.G., Crisp, M.D., 2008. Tree thinking for all biology: the problem with reading phylogenies as ladders of progress. Bioessays 30, 854–867.

Peres-Neto, P.R., Legendre, P., Dray, S., Borcard, D., 2006. Variation partitioning of species data matrices: estimation and comparisons of fractions. Ecology 87, 2614–2625.

Pulliam, H.R., Danielson, B.J., 1991. Sources, Sinks, and Habitat Selection – a landscape perspective on population-dynamics. Am. Nat. 137, 50–66.

Quinn, G.P., Keough, M.J., 2002. Experimental Design and Data Analysis for Biologists. Cambridge University Press, Cambridge, UK.

R Core Team, 2016. R: A Language and Environment for Statistical Computing, Version 3.1.0. R Foundation for Statistical Computing, Vienna http://www.Rproject.org/.

REFLORA, 2016. Lista de Espécies da Flora do Brasil. Jardim Botânico do Rio de Janeiro, Brazil (Accessed 19 Dedember 2016) http://floradobrasil.jbrj.gov.br/.

Ribeiro, K.T., Medina, B.M.O., Scarano, F.R., 2007. Species composition and biogeographic relations of the rock outcrop flora on the high plateau of the Itatiaia, SE-Brazil. Rev. Brasil. Bot. 30, 623–639.

Safford, H.D., 2007. Brazilian Páramos IV: Phytogeography of the campos de altitude. J. Biogeogr. 34, 1701–1722.

Santos, H.G., Jacomine, P.K.T., Anjos, L.H.C., Oliveira, V.A., Lumbreras, J.F., Coelho, M.R., Almeida, J.A., Cunha, T.J.F., Oliveira, J.B., 2013. Sistema Brasileiro de Classificação de Solos, 3rd edn. Embrapa, Brasília.

Saravia, E.S.R., 2008. El costo de la conservación de los bosques tropicales. In: PhD Thesis. Universidad Autónoma G. René Moreno, Santa Cruz de la Sierra.

Silveira, F.A.O., Negreiros, D., Barbosa, N.P.U., Buisson, E., Carmo, F.F., Carstensen, D.W., Conceição, A.A., Cornelissen, T.G., Echternacht, L., Fernandes, G.W., Garcia, Q.S., Guerra, T.J., Jacobi, C.M., Lemos-Filho, J.P., Le Stradic, S., Morellato, L.P.C., Neves, F.S., Oliveira, R.S., Schaefer, C.E., Viana, P.L., Lambers, H., 2015. Ecology and evolution of plant diversity in the endangered campo rupestre: a neglected conservation priority. Plant Soil 403, 129–152.

Simon, M.F., Grether, R., Queiroz, L.P., Skema, C., Pennington, R.T., Hughes, C.E., 2009. Recenty assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. Proc. Nat. Acad. Sci. U. S. A. 106, 20359–20364.

Souza, E.R., Lewis, G.P., Forest, F., Schnadelbach, A.S., van den Berg, C., Queiroz, L.P., 2013. Phylogeny of Calliandra (Leguminosae: Mimosoideae) based on nuclear and plastid molecular markers. Taxon 62, 1200–1219.

Versieux, L.M., Barbará, T., Wanderley, M.G.L., Calvente, A., Fay, M.F., Lexer, C., 2012. Molecular phylogenetics of the Brazilian giant bromeliads (Alcantarea, Bromeliaceae): implications for morphological evolution and biogeography. Mol. Phylogenet. Evol. 64, 177–189.

Walter, H., 1985. Vegetation of the Earth and Ecological Systems of the Geo-Biosphere, 3rd edn. Springer-Verlag, Berlin.

Zomer, R.J., Bossio, D.A., Trabucco, A., Yuanjie, L., Gupta, D.C., Singh, V.P., 2007. Trees and Water: Smallholder Agroforestry on Irrigated Lands in Northern India IWMI Research Report 122. International Water Management Institute, Colombo, Sri Lanka.

Zomer, R.J., Trabucco, A., Bossio, D.A., van Straaten, O., Verchot, L.V., 2008. Climate change mitigation: a spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. Agric. Ecosyst. Environ. 126, 67–80.

Appendix C

Inserting Tropical Dry Forests into the Discussion on Biome Transitions in the Tropics

---

Reference:

Dexter, Kyle G., R. Toby Pennington, Ary T. Oliveira-Filho, Marcelo L. Bueno, Silva de Miranda, Pedro L, and Danilo M. Neves. 2018. "Inserting Tropical Dry Forests into the Discussion on Biome Transitions in the Tropics." *Frontiers in Ecology and Evolution* 6. https://doi.org/10.3389/fevo.2018.00104

frontiers
in Ecology and Evolution

PERSPECTIVE
published: 24 July 2018
doi: 10.3389/fevo.2018.00104

Inserting Tropical Dry Forests Into the Discussion on Biome Transitions in the Tropics

Kyle G. Dexter[1,2]*, *R. Toby Pennington*[2,3], *Ary T. Oliveira-Filho*[4], *Marcelo L. Bueno*[5], *Pedro L. Silva de Miranda*[1] *and Danilo M. Neves*[4,6]

[1] School of GeoSciences, University of Edinburgh, Edinburgh, United Kingdom, [2] Royal Botanic Garden Edinburgh, Edinburgh, United Kingdom, [3] Geography, University of Exeter, Exeter, United Kingdom, [4] Departamento de Botânica, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [5] Laboratório de Ecologia e Evolução de Plantas, Universidade Federal de Viçosa, Viçosa, Brazil, [6] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, United States

Tropical moist forests and savannas are iconic biomes. There is, however, a third principal biome in the lowland tropics that is less well known: tropical dry forest. Discussions on responses of vegetation in the tropics to climate and land-use change often focus on shifts between forests and savannas, but ignore dry forests. Tropical dry forests are distinct from moist forests in their seasonal drought stress and consequent deciduousness and differ from savannas in rarely experiencing fire. These factors lead tropical dry forests to have unique ecosystem function. Here, we discuss the underlying environmental drivers of transitions among tropical dry forests, moist forests and savannas, and demonstrate how incorporating tropical dry forests into our understanding of tropical biome transitions is critical to understanding the future of tropical vegetation under global environmental change.

Keywords: tropical dry forest, tropical moist forest, savanna, biomes, fire, soil fertility, water stress, deciduousness

Edited by:
Colin Osborne,
University of Sheffield,
United Kingdom

Reviewed by:
Imma Oliveras,
University of Oxford, United Kingdom
Liam Jude Langan,
Senckenberg Biodiversität und Klima
Forschungszentrum (SBiK-F),
Germany

***Correspondence:**
Kyle G. Dexter
kyle.dexter@ed.ac.uk

Specialty section:
This article was submitted to
Biogeography and Macroecology,
a section of the journal
Frontiers in Ecology and Evolution

Received: *01 May 2018*
Accepted: *29 June 2018*
Published: *24 July 2018*

Citation:
Dexter KG, Pennington RT, Oliveira-Filho AT, Bueno ML, Silva de Miranda PL and Neves DM (2018) Inserting Tropical Dry Forests Into the Discussion on Biome Transitions in the Tropics. Front. Ecol. Evol. 6:104. doi: 10.3389/fevo.2018.00104

INTRODUCTION

Predicting vegetation change in the tropics depends on understanding the drivers of transitions among major vegetation types, or biomes. Climatic factors, such as mean annual precipitation (MAP) and its seasonality are of obvious importance, but edaphic factors, and disturbance via fire, humans and herbivores also play key roles. Recent large-scale studies across the tropics have focused on transitions between forest and savanna (Hirota et al., 2011; Staver et al., 2011; Oliveras and Malhi, 2016; Xu et al., 2016; Langan et al., 2017). While there is value to simplifying vegetation concepts in the tropics, we believe the simplification used by these authors in defining "forest" goes perhaps one step too far. There are two principal kinds of forest in the lowland tropics, moist forests and dry forests. With very few exceptions (e.g., Hirota et al., 2010; Lehmann et al., 2011), studies of biome transitions in the tropics have either failed to distinguish them, or have completely ignored dry forests, focusing solely on moist forests when using the term "forest." The aim of this perspective is to discuss biome transitions in the tropics and their underlying drivers, while including dry forests in the discussion. We focus on transitions among savanna, moist forest and dry forest, the three biomes in the lowland tropics with a substantial tree component.

Tropical moist forest and savanna are relatively well understood at a global scale compared to tropical dry forests (Pennington et al., 2018). Moist forest is tall, multi-stratal, and with a closed canopy. Tropical moist forests include tropical rain forests, as well as forests with lower rainfall where soil moisture is maintained throughout the year, via edaphic factors such as proximity to rivers, or water recycling, allowing most trees to be evergreen (Guan et al., 2015). The understory is often dominated by saplings of taller-statured tree species, although small tree and shrub species are present. Terrestrial forbs and grasses are a minor component of diversity and biomass. Savanna is a more open environment, where tree species are present, but individuals do not form a closed canopy. There is a significant understory grass component, which is flammable, and fires are common. Tree species that occur in savannas are adapted to these recurring fires (Simon and Pennington, 2012), and regular fires are necessary for the maintenance of savanna biodiversity (Parr et al., 2014; Durigan and Ratter, 2016; Abreu et al., 2017). Some savannas in the paleotropics (e.g., miombo woodlands in Africa, deciduous dipterocarp forests in southeast Asia) are generally referred to as dry forests, but we consider them as savannas given that they have a grassy understory and experience regular fire (Ratnam et al., 2011; Dexter et al., 2015; Pennington et al., 2018).

Tropical dry forests vary greatly in structure, from tall, closed canopy forest to short scrub vegetation, occasionally not forming a closed canopy, especially in drier areas (Pennington et al., 2000). They are distinct from savanna in not having a significant grass component and not experiencing regular fires (Murphy and Lugo, 1986; Gentry, 1995). In fact, regular fires would be lethal for many of the characteristic life forms and taxa of tropical dry forest (e.g., cacti; Mooney et al., 1995). This is not to say that dry forests never experience fires. Even moist forests can experience fire under extreme drought conditions (Aragão et al., 2016). Rather, damaging fire is sufficiently rare in dry forests such that fire-intolerant species can persist in the landscape as metapopulations (Hanski, 1998). The exact threshold of fire return interval or intensity involved in the tropical dry forest - savanna transition is poorly understood and likely to vary with the broader environmental context (e.g., soil fertility and annual precipitation; Hoffmann et al., 2012; Murphy and Bowman, 2012), and we suggest that this should be a priority for further study. Tropical dry forest is distinct from moist forest in its seasonal drought stress, which leads many tree species to lose their leaves in the dry season (Reich and Borchert, 1984; Murphy and Lugo, 1986). The combination of seasonal drought stress and lack of fire leads to ecosystem function in dry forests that is markedly different from savannas or moist forests, which justifies their distinction as a unique biome.

BIOMES IN LOWLAND TROPICAL SOUTH AMERICA

We focus this review on continental lowland tropical South America (LTSA), where we have conducted most of our research. In a recent study (Silva de Miranda et al., in press), we used an unsupervised classification, or hierarchical clustering, of sites based on their tree species composition (see inset in **Figure 1**), followed by interpretation of the resulting cluster using site information on vegetation physiognomy (savanna vs. forest) and leaf flush regime (evergreen vs. semideciduous vs. deciduous) to delimit and map biomes across LTSA east of the Andes (**Figure 1**). Moist forests fell in two major groups in the cluster and occurred in two large geographic blocks, one in the Amazon basin and another along the Atlantic coast of Brazil. Semideciduous forests did not form a distinct group in the cluster and were instead mixed with evergreen, moist forest sites. Semideciduous forests are often found in drier regions, where they occur along rivers, lake margins and submontane areas with orographic precipitation. Savanna formed a single group in the cluster and was most prevalent in central Brazil, in an area commonly termed the Cerrado.

Tropical dry forest also formed a single group in the cluster, which was comprised almost entirely of forest sites with deciduous phenology. In LTSA, the largest block of dry forest occurs in the Caatinga region of northeast Brazil. The Caatinga has been referred to as a biome (Hirota et al., 2010; Instituto Brasileiro de Geografia e Estatística, 2012), although as a region it contains non-dry forest habitat (e.g., patches of savanna). Further, tropical dry forest is found outside of the Caatinga, in patches throughout the Cerrado, in an area spanning the Pantanal and Chiquitania in Bolivia (**Figure 1**), and scattered more widely across the Neotropics (DRYFLOR, 2016). While Silva de Miranda et al. (in press) broadly assessed climatic overlaps amongst biomes, they did not focus on the environmental drivers of transitions between individual biomes. That is the goal of the present manuscript.

TRANSITIONS BETWEEN TROPICAL SAVANNA AND DRY FORESTS

A common view of dry forests in the tropics is that they are transitional between savannas and moist forests along precipitation gradients (e.g., Whittaker, 1970; Malhi et al., 2009). If, however, we examine how the sites featured in **Figure 1** are distributed over variation in MAP (**Figure 2**), a more complex picture emerges. Moist forests do occur under wetter conditions than savanna, but dry forests are largely found under drier conditions. Below 1,000 mm MAP, savanna quickly disappears and dry forest becomes the only tree-dominated vegetation type. The largest area of these arid dry forests is found in the Caatinga region of northeast Brazil. As discussed above, a key distinction between savanna and dry forest is the regularity of fire, and in these dry conditions, there is not sufficient biomass build-up to sustain regular fires (Van Der Werf et al., 2008). In particular, this reflects the relative lack of grasses. The tree species that are present are able to tolerate severe drought, but they do not invest in adaptations for fire, such as thick bark or underground stems, characteristic of savanna species (Simon and Pennington, 2012).

There is also extensive occurrence of the dry forest biome under the same precipitation conditions as savanna. These are dry forests found in the Cerrado region and around the Pantanal and Chiquitania regions of Brazil and Bolivia. Within the Cerrado, dry forests are known to occur on and around calcareous outcrops, where soils have higher phosphorus and base cation concentrations (Ratter et al., 1978; Furley and Ratter, 1988; Oliveira-Filho and Ratter, 2002; Neves et al., 2015). On these soils, trees can grow more quickly, have better chances of escaping the "fire trap" and are more likely to

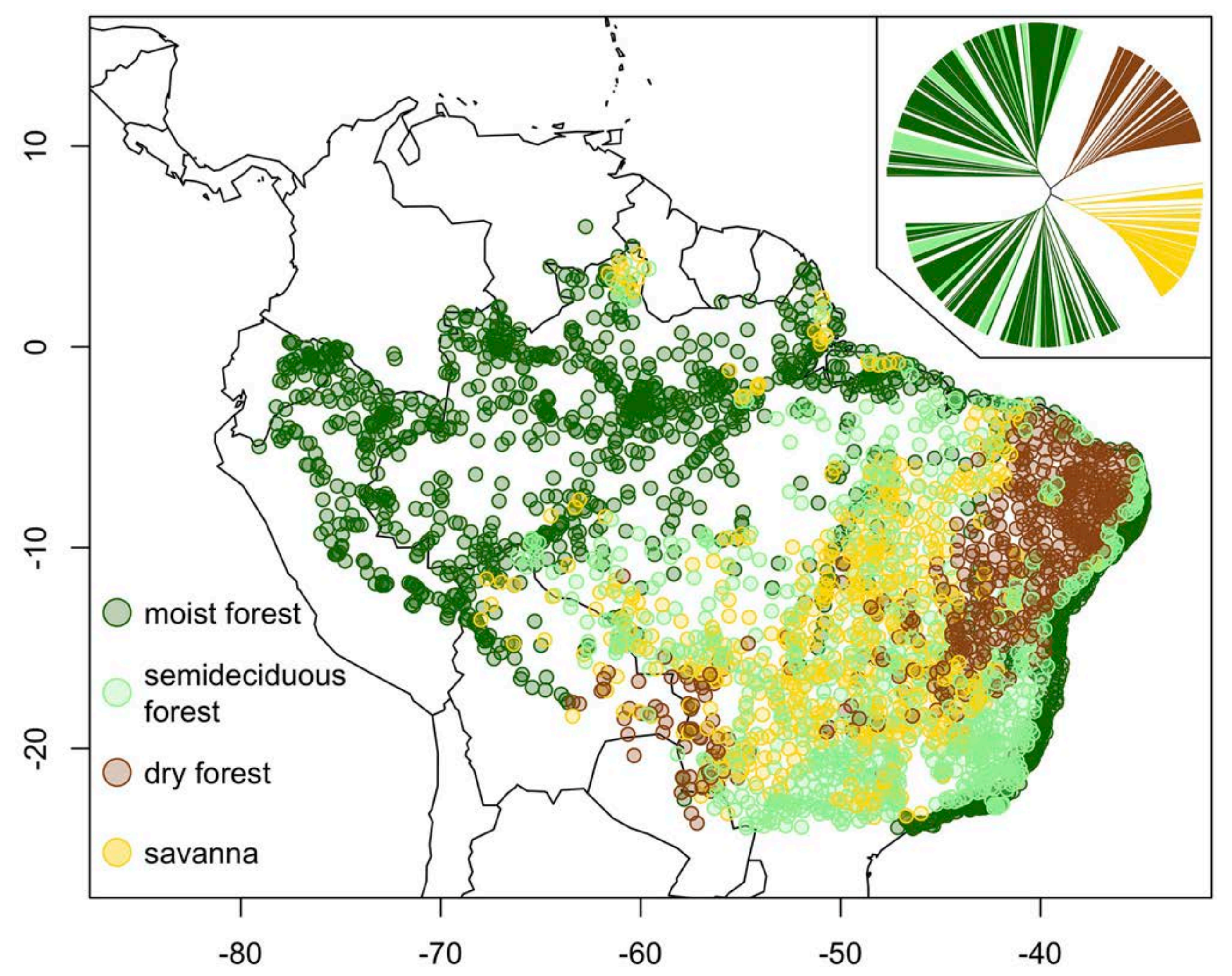

FIGURE 1 | Map of biomes across lowland tropical South America (after Silva de Miranda et al., in press). Inset shows results of unsupervised clustering of 3,331 georeferenced sites based on tree species composition. The biomes of the major groups in the cluster were determined based on vegetation physiognomy and leaf flush regime. Semideciduous forests were placed in the moist forest biome based on similarities in tree species composition evident in the cluster by Silva de Miranda et al. (in press), but are distinguished here. We exclude sites south of 23°S latitude or above 1,000 m elevation.

form a closed canopy (Hoffmann et al., 2009). There can be positive feedbacks between tree growth, grass exclusion and fire mitigation that leads to a forested vegetation (Hoffmann et al., 2012; Silva et al., 2013; Pausas and Dantas, 2017). Calcareous outcrops in the Cerrado also have poorly developed, shallow soils, and vegetation occurring on them may experience greater drought stress than surrounding vegetation, thus making them similar to the arid dry forests, which lack fire because of insufficient fuel build-up. However, it is likely that soil fertility is relevant for the presence of dry forests around calcareous outcrops as a different vegetation, cerrado rupestre, which is floristically related to savanna vegetation, is found on non-calcareous outcrops in the Cerrado (Ribeiro and Walter, 1998). Whichever factor is more important (soils with high fertility or low water-holding capacity), it is evident that the same drought-tolerant, fire-intolerant tree species and lineages that dominate vegetation in the arid Caatinga are also found in dry forest patches in the moister Cerrado (Prado and Gibbs, 1993; Neves et al., 2015; DRYFLOR, 2016; Silva de Miranda et al., in press).

Dry forest and savanna vegetation also intermingle in the Chiquitania region of Bolivia, but here dry forest predominates and savanna occurs in patches, which may be because soils in the Chiquitania are more fertile on average than in the Cerrado (Silva de Miranda et al., in press). Some of the savannas that are present in the Chiquitania region may represent dry forest that has been degraded by logging or anthropogenic fire (Devisscher et al., 2016), highlighting that human land-use patterns can readily drive transitions between dry forest and savanna. However, "old-growth savannas" that would exist independent of anthropogenic

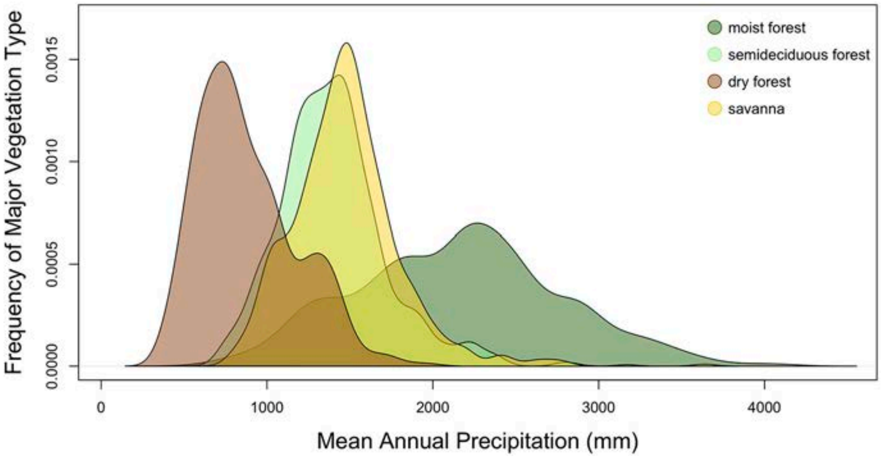

FIGURE 2 | Frequency distribution of sites in **Figure 1** over a gradient in mean annual precipitation. For heuristic purposes, semideciduous and moist forests are distinguished although they may represent the same biome.

influence are also clearly present in Bolivia (Power et al., 2016; Veldman, 2016).

If fire is excluded from savanna vegetation, it first converts to a formation with a higher percentage of tree canopy cover termed "cerradão" (Durigan and Ratter, 2006, 2016), which translates from the Portuguese as "big cerrado" and is generally considered as a forest. Cerradão shares some tree species with tropical dry forest and comparatively few with semideciduous and evergreen moist forests (Bueno et al., 2018), even though the latter are present near savanna vegetation along river courses and lake margins that have year-round water availability (Ribeiro and Walter, 1998). Tree species from semideciduous and evergreen forests may be less likely to immigrate into cerradão than typical dry forest tree species because they are not adapted to seasonal drought. While cerradão is initially comprised of fire-adapted tree species from the cerrado, dry forest tree species colonize this environment if propagules are available, fire remains absent and soils are sufficiently fertile. These dry forest tree species may eventually outcompete cerrado tree species, since they do not invest in fire defense (Ratajczak et al., 2017), and in the prolonged absence of fire, cerradão may transition to a dry forest if there are positive feedback cycles between forest vegetation, lack of fire and soil fertility (Silva et al., 2013; Pellegrini et al., 2014, 2018). However, if the underlying soils remain poor and/or if there are high aluminum concentrations in the soil that do not attenuate over time, then cerrado tree species, which are adapted to infertile, aluminum-rich soils may continue to dominate the vegetation.

Overall, the savanna-dry forest transition is distinct from the savanna-moist forest transition in two key ways: (1) the contrasting role of water availability (lowest in dry forest, intermediate in savanna and highest in moist forest) and (2) the potentially critical importance of soil fertility for the savanna-dry forest transition (savanna and moist forests are similar in generally having infertile, acidic soils).

TRANSITIONS BETWEEN MOIST FOREST AND DRY FOREST

Both of these biomes are forest, but they function in distinct ways. Dry forests are found in areas with marked precipitation seasonality, which leads most species to lose their leaves during the dry season and has significant implications for nutrient cycling (Reich and Borchert, 1984; Murphy and Lugo, 1986). Many moist forests also experience seasonality in water availability (e.g., in the southern and eastern Amazon), but the dry season is three months or less and subsurface water remains available to trees (Guan et al., 2015). The systems also differ in the rate at which they accrue and cycle carbon, with trees in moist forests growing more quickly and storing more total carbon (Murphy and Lugo, 1986; Poorter et al., 2017). There are often differences in soil fertility, with dry forests occurring on more fertile soils, which facilitates their ability to shed their leaves as they can readily afford to grow new ones. However, high rainfall in moist forests results in nutrient leaching, and this correlation between soil fertility and biome identity may be due to overriding climatic factors (Webb, 1968; Hall and Swaine, 1976).

In LTSA, there are multiple areas of contact between the moist and dry forest biomes (**Figure 1**). One transition zone is in northeastern Brazil, where evergreen Atlantic forest on the coast transitions to dry forest in the arid Caatinga. In between the two lies a band of semideciduous forests. Another transition zone is found in the Chiquitania region of eastern Bolivia and adjacent areas of Brazil, where there is a gradual transition over 200+ km of geographic distance, largely covered by semideciduous forest. We suggest that transitions between dry forest and moist forest are primarily mediated by water availability and that intermediate states are possible in zones of intermediate water availability (Oliveira-Filho and Fontes, 2000; Oliveira-Filho et al., 2006). This contrasts with transitions between savanna and moist forest that can be more abrupt and may represent alternative stable states (although see Lloyd and Veenendaal, 2016).

SEMIDECIDUOUS FORESTS

Previous studies have variously grouped semideciduous forests with the dry forest biome (Murphy and Lugo, 1986; Pennington et al., 2000) or the moist forest biome (DRYFLOR, 2016; Silva de Miranda et al., in press). In fact, as discussed above, these forests may be transitional between the two. Semideciduous forests in LTSA have few endemic tree species and instead contain tree species associated with the dry or moist forest biomes (Oliveira-Filho and Fontes, 2000). As moist forests contain many more tree species than dry forests (Esquivel-Muelbert et al., 2017), we suggest that they may contribute more species to semideciduous forests simply via mass effects (Shmida and Wilson, 1985), and this may be why they group with moist forests in clustering analyses based on presence versus absence of tree species (as in Silva de Miranda et al., in press). If abundance information were to be taken into account (e.g., via inventory plot data), we hypothesize that semideciduous forests may cluster with moist or dry forest based on the proportion of individuals belonging to moist versus dry forest tree species. It is clear that future comparative studies across dry, moist and semideciduous forests are needed to understand their origins and how they compare in terms of ecosystem function. Their geographically variable species composition and lack of endemic species suggests that semideciduous forests may have been independently and recently assembled in different ecotonal areas.

BIOME TRANSITIONS TO DRY FOREST OUTSIDE THE NEOTROPICS

As in South America, moist forest—savanna transitions have been studied extensively on other continents, but transitions to dry forest have received less attention. This is partly because it is unclear where dry forest exists outside of the Neotropics (Lock, 2006; Dexter et al., 2015; Pennington et al., 2018). In a recent study, Linder (2014) delimited and described the main "floras" of Africa, which are large-scale units of vegetation that have a distinct evolutionary and biogeographic history and differ in their present-day plant taxonomic composition. Linder did not assign the term "biome" to these vegetation units, although his "floras" correspond to several previously defined biomes. There is a "savanna flora," which readily corresponds to the savanna biome, and a "lowland forest flora" that largely corresponds to the moist forest biome. Linder postulated an "arid flora" that is most evident in the Horn of Africa (Somalia, Ethiopia and northern Kenya), but is also present in arid regions of Angola and Namibia. The distribution of this "arid flora" largely overlaps the distribution of the "succulent biome" in Africa, as proposed by Schrire et al. (2005). The "arid flora" or "succulent biome" is similar to dry forest in arid regions of the Neotropics in that there is not adequate water availability to allow for sufficient biomass build-up to sustain regular fires. Thus, as in the Neotropics, water availability may be one environmental factor that underlies transitions between savanna and dry forest in Africa.

Soil fertility is another significant factor that has been shown to underlie savanna-dry forest transitions in the Neotropics. An important question for future research in Africa is whether, amongst its great expanses of savanna, there is distinct vegetation that does not regularly burn, is found on more fertile soils and shows greater floristic similarity with the "arid flora" of Linder (2014) than it does with surrounding savanna vegetation. It may be that in Africa, a higher abundance of large herbivores, including elephants, favors grasses over trees, leading to a more open savanna vegetation with more frequent fires, even in areas of higher soil fertility (Charles-Dominique et al., 2016; Pellegrini et al., 2017). If dry forests are not found on fertile soils in more mesic areas of Africa, there may not be moist forest-dry forest transitions on this continent, because the areas mapped as belonging to the "arid flora" or "succulent biome" are completely separated from moist forest regions by large areas of savanna (Schrire et al., 2005; Linder, 2014).

In the tropical regions of continental Asia and in Malesia, moist forest is the predominant vegetation type, although drier forest formations are present (e.g., deciduous forests in the Western Ghats and dry dipterocarp forests in Indochina). As we have discussed elsewhere (Dexter et al., 2015; Pennington et al., 2018), the majority of these drier forest formations have a significant grassy component in the understory, burn regularly and may be better considered as savannas (Ratnam et al., 2011). The succulent biome of Schrire et al. (2005) is mapped as present in arid regions of northwest India and extending across the coast of Pakistan and Iran to the Arabian Peninsula. Thus, an arid form of the dry forest biome may occur in Asia, as in Africa, and water availability may underlie transitions between vegetation in seasonally dry areas that regularly burns (what we term savanna) and that does not regularly burn (what we term dry forest). As with Africa, future research in Asia should assess if there are vegetation formations in seasonally dry, yet not arid, areas that: (1) are found on fertile soils, (2) do not regularly burn and (3) show greater floristic similarity with arid areas than with surrounding vegetation that does regularly burn. This will help determine if soil fertility is also important in understanding savanna-dry forest transitions in Asia, as it is in the Neotropics.

CONCLUSIONS

The aim of this perspective has been to bring the tropical dry forest biome into discussions of biome transitions in the tropics. Previous studies of tropical biome transitions have largely focused on forest-savanna transitions, with all forests being considered as a single biome. In fact, there are many kinds of forests in the tropics, some of which are distinct from each other in species composition and ecosystem function and represent different biomes (i.e., dry vs. moist forest) and others which are more difficult to classify (e.g., semideciduous forests, cerradão). Water availability is a key factor underlying tropical biome transitions. While forests are often thought to occur under wetter conditions than savannas, tropical dry forest is actually more prevalent in areas of lower water availability (<1,000 mm MAP). Meanwhile, soil fertility, which has received limited attention in studies of biome transitions, is also critical in the Neotropics, and merits future research on other continents. More generally, recognizing tropical dry forest as a distinct biome within the tropics should improve the accuracy of modeling studies that aim to predict the future of tropical vegetation and ecosystem function under global environmental change.

AUTHOR CONTRIBUTIONS

KD and RP wrote the first draft of the manuscript and contributed to revising the manuscript. AO-F, MB, PS and DN contributed to revising and improving the manuscript.

FUNDING

KD, RP, and DN thank the Natural Environment Research Council (UK) for funding via NE/I028122/1. KD thanks the Leverhulme Trust for supporting him during the time this study was completed via an International Academic Fellowship. PS thanks the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior–Brazil (CAPES) for support via the Science without Borders Programme (grant 99999.013197/2013-04).

REFERENCES

Abreu, R. C. R., Hoffmann, W. A., Vasconcelos, H. L., Pilon, N. A., Rossatto, D. R., and Durigan, G. (2017). The biodiversity cost of carbon sequestration in tropical savanna. *Sci. Adv.* 3:e1701284. doi: 10.1126/sciadv.1701284

Aragão, L. E., Anderson, L. O., Lima, A., and Arai, E. (2016). "Fires in Amazonia," in *Interactions Between Biosphere, Atmosphere and Human Land Use in the Amazon Basin*, eds L. Nagy, B. R. Forsberg, and P. Artaxo (Heidelberg: Springer), 301–329.

Bueno, M. L., Dexter, K. G., Pennington, R. T., Pontara, V., Neves, D. M., Ratter, J. A., et al. (2018). The environmental triangle of the Cerrado domain: ecological factors driving shifts in tree species composition between forests and savannas. *J. Ecol*. doi: 10.1111/1365-2745.12969

Charles-Dominique, T., Davies, T. J., Hempson, G. P., Bezeng, B. S., Daru, B. H., Kabongo, R. M., et al. (2016). Spiny plants, mammal browsers, and the origin of African savannas. *Proc. Natl. Acad. Sci. U.S.A.* 113, E5572–E5579. doi: 10.1073/pnas.1607493113

Devisscher, T., Anderson, L. O., Aragão, L. E., Galván, L., and Malhi, Y. (2016). Increased Wildfire Risk Driven by Climate and Development Interactions in the Bolivian Chiquitania, Southern Amazonia. *PLoS ONE* 11:e0161323. doi: 10.1371/journal.pone.0161323

Dexter, K. G., Smart, B., Baldauf, C., Baker, T. R., Balinga, M. B., Brienen, R. J. W., et al. (2015). Floristics and biogeography of vegetation in seasonally dry tropical regions. *Int. For. Rev.* 17, 10–32. doi: 10.1505/146554815815834859

DRYFLOR, Banda, R. K., Delgado-Salinas, A., Dexter, K. G., Linares-Palomino, R., Oliveira-Filho, A., et al. (2016). Plant diversity patterns in neotropical dry forests and their conservation implications. *Science* 353, 1383–1387. doi: 10.1126/science.aaf5080

Durigan, G., and Ratter, J. A. (2006). Successional changes in cerrado and cerrado/forest ecotonal vegetation in western São Paulo State, Brazil, 1962–2000. *Edinburgh J. Bot*. 63, 119–130. doi: 10.1017/S0960428606000357

Durigan, G., and Ratter, J. A. (2016). The need for a consistent fire policy for Cerrado conservation. *J. Appl. Ecol.* 53, 11–15. doi: 10.1111/1365-2664.12559

Esquivel-Muelbert, A., Baker, T. R., Dexter, K. G., Lewis, S. L., Steege, H., Lopez-Gonzalez, G., et al. (2017). Seasonal drought limits tree species across the Neotropics. *Ecography* 40, 618–629. doi: 10.1111/ecog.01904

Furley, P. A., and Ratter, J. A. (1988). Soil resources and plant communities of the central Brazilian cerrado and their development. *J. Biogeogr.* 97–108. doi: 10.2307/2845050

Gentry, A. H. (1995). "Diversity and floristic composition of neotropical dry forests,"in *Seasonally dry tropical forests*, eds S.H. Bullock, H.A. Mooney, and E. Medina (Cambridge: Cambridge University Press), 146–194.

Guan, K., Pan, M., Li, H., Wolf, A., Wu, J., Medvigy, D., et al. (2015). Photosynthetic seasonality of global tropical forests constrained by hydroclimate. *Nat. Geosci.* 8, 284–289. doi: 10.1038/ngeo2382

Hall, J. B., and Swaine, M. (1976). Classification and ecology of closed-canopy forest in Ghana. *J. Ecol.* 64, 913–951. doi: 10.2307/2258816

Hanski, I. (1998). Metapopulation dynamics. *Nature* 396, 41–49. doi: 10.1038/23876

Hirota, M., Holmgren, M., Van Nes, E. H., and Scheffer, M. (2011). Global resilience of tropical forest and savanna to critical transitions. *Science* 334, 232–235. doi: 10.1126/science.1210657

Hirota, M., Nobre, C., Oyama, M. D., and Bustamante, M. (2010). The climatic sensitivity of the forest, savanna and forest–savanna transition in tropical South America. *New Phytol.* 187, 707–719. doi: 10.1111/j.1469-8137.2010.03352.x

Hoffmann, W. A., Adasme, R., Haridasan, M. T., de Carvalho, M., Geiger, E. L., Pereira, M. A., et al. (2009). Tree topkill, not mortality, governs the dynamics of savanna–forest boundaries under frequent fire in central Brazil. *Ecology* 90, 1326–1337. doi: 10.1890/08-0741.1

Hoffmann, W. A., Geiger, E. L., Gotsch, S. G., Rossatto, D. R., Silva, L. C., Lau, O. L., et al. (2012). Ecological thresholds at the savanna-forest boundary: how plant traits, resources and fire govern the distribution of tropical biomes. *Ecol. Lett.* 15, 759–768. doi: 10.1111/j.1461-0248.2012.01789.x

Instituto Brasileiro de Geografia e Estatística (ed.). (2012). *Manual Técnico da Getveação Brasileira (2a Edição Revista e Ampliada)*. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística-IBGE.

Langan, L., Higgins, S. I., and Scheiter, S. (2017). Climate-biomes, pedo-biomes or pyro-biomes: which world view explains the tropical forest–savanna boundary in South America? *J. Biogeogr.* 44, 2319–2330. doi: 10.1111/jbi.13018

Lehmann, C. E., Archibald, S. A., Hoffmann, W. A., and Bond, W. J. (2011). Deciphering the distribution of the savanna biome. *New Phytol.* 191, 197–209. doi: 10.1111/j.1469-8137.2011.03689.x

Linder, H. P. (2014). The evolution of African plant diversity. *Front. Ecol. Evol.* 2:38. doi: 10.3389/fevo.2014.00038

Lloyd, J., and Veenendaal, E. M. (2016). Are fire mediated feedbacks burning out of control. *Biogeosci. Discuss.* 2016, 1–20. doi: 10.5194/bg-2015-660

Lock, J. M. (2006). "The seasonally dry vegetation," in *Neotropical Savannas and Seasonally Dry Forests: Plant Diversity, Biogeography, and Conservation* (Boca Raton, FL: Taylor and Francis), 449.

Malhi, Y., Aragão, L. E., Galbraith, D., Huntingford, C., Fisher, R., Zelazowski, P., et al. (2009). Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest. *Proc. Nat. Acad. Sci. U.S.A.* 106, 20610–20615. doi: 10.1073/pnas.0804619106

Mooney, H. A., Bullock, S. H., and Medina, E. (1995). "Introduction," in *Seasonally Dry Tropical Forests*, eds S. H. Bullock, H. A. Mooney, and E. Medina (Cambridge: Cambridge University Press), 1–8.

Murphy, B. T., and Bowman, D. M. (2012). What controls the distribution of tropical forest and savanna? *Ecol. Lett.* 15, 748–758. doi: 10.1111/j.1461-0248.2012.01771.x

Murphy, P. G., and Lugo, A. E. (1986). Ecology of tropical dry forest. *Annu. Rev. Ecol. Syst.* 17, 67–88. doi: 10.1146/annurev.es.17.110186.000435

Neves, D. M., Dexter, K. G., Pennington, R. T., Bueno, M. L., and Oliveira Filho, A. T. (2015). Environmental and historical controls of floristic composition across the South American Dry Diagonal. *J. Biogeogr.* 42, 1566–1576. doi: 10.1111/jbi.12529

Oliveira-Filho, A. T., and Fontes, M. A. L. (2000). Patterns of floristic differentiation among Atlantic forests in southeastern Brazil and the influence of climate. *Biotropica* 32, 793–810. doi: 10.1111/j.1744-7429.2000.tb00619.x

Oliveira-Filho, A. T., Jarenkow, J. A., and Rodal, M. J. N. (2006). "Floristic relationships of Seasonally Dry Forest of eastern South American based on tree species distribution patterns," in *Neotropical Savannas and Dry Forests: Plant Diversity, Biogeography and Conservation,* eds R. T. Lewis, G. P. Ratter, and J. A. (Orgs.) (Pennington, NJ: Boca Raton: CRC Press), 159–192.

Oliveira-Filho, A. T., and Ratter, J. A. (2002). "Vegetation physiognomies and woody flora of the cerrado biome: ecology and natural history of a neotropical

Savanna," in *The Cerrados of Brazil,* eds P. S. Oliveira and R. J. Marquis (New York, NY: Columbia University Press), 91–120.

Oliveras, I., and Malhi, Y. (2016). Many shades of green: the dynamic tropical forest–savannah transition zones. *Phil. Trans. R. Soc. B. Biol. Sci.* 371:20150308. doi: 10.1098/rstb.2015.0308

Parr, C. L., Lehmann, C. E., Bond, W. J., Hoffmann, W. A., and Andersen, A. N. (2014). Tropical grassy biomes: misunderstood, neglected, and under threat. *Trends Ecol. Evol.* 29, 205–213. doi: 10.1016/j.tree.2014.02.004

Pausas, J. G., and Dantas, V. D. L. (2017). Scale matters: fire–vegetation feedbacks are needed to explain tropical tree cover at the local scale. *Glob. Ecol. Biogeogr.* 26, 395–399. doi: 10.1111/geb.12562

Pellegrini, A. F., Ahlström, A., Hobbie, S. E., Reich, P. B., Nieradzik, L. P., Staver, A. C., et al. (2018). Fire frequency drives decadal changes in soil carbon and nitrogen and ecosystem productivity. *Nature* 553, 194–198. doi: 10.1038/nature24668

Pellegrini, A. F., Hoffmann, W. A., and Franco, A. C. (2014). Carbon accumulation and nitrogen pool recovery during transitions from savanna to forest in central Brazil. *Ecology* 95, 342–352. doi: 10.1890/13-0290.1

Pellegrini, A. F., Pringle, R. M., Govender, N., and Hedin, L. (2017). Woody plant biomass and carbon exchange depend on elephant-fire interactions across a productivity gradient in African savanna. *J. Ecol.* 105, 111–121. doi: 10.1111/1365-2745.12668

Pennington, R. T., Lehmann, C. E. R., and Rowland, L. M. (2018). Tropical savannas and dry forests. *Curr. Biol.* 28, R541–R545. doi: 10.1016/j.cub.2018.03.014

Pennington, R. T., Prado, D. E., and Pendry, C. A. (2000). Neotropical seasonally dry forests and Quaternary vegetation changes. *J. Biogeogr.* 27, 261–273. doi: 10.1046/j.1365-2699.2000.00397.x

Poorter, L., van der Sande, M. T., Arets, E. J., Ascarrunz, N., Enquist, B., Finegan, B., et al. (2017). Biodiversity and climate determine the functioning of Neotropical forests. *Glob. Ecol. Biogeogr.* 26, 1423–1434. doi: 10.1111/geb.12668

Power, M. J., Whitney, B. S., Mayle, F. E., Neves, D. M., de Boer, E., and MacLean, K. S. (2016). Fire, climate and vegetation linkages in the Bolivian Chiquitano Seasonally Dry Tropical Forest. *Philos. Trans. R. Soc. B. Biol. Sci.* 371:20150165. doi: 10.1098/rstb.2015.0165

Prado, D. E., and Gibbs, P. E. (1993). Patterns of species distributions in the dry seasonal forests of South America. *Annu. Mol. Bot. Gard.* 80, 902–927. doi: 10.2307/2399937

Ratajczak, Z., D'Odorico, P., and Yu, K. (2017). The Enemy of My Enemy Hypothesis: Why Coexisting with Grasses May Be an Adaptive Strategy for Savanna Trees. *Ecosystems* 20, 1278–1295. doi: 10.1007/s10021-017-0110-7

Ratnam, J., Bond, W. J., Fensham, R. J., Hoffmann, W. A., Archibald, S., Lehmann, C. E., et al. (2011). When is a 'forest' a savanna, and why does it matter? *Glob. Ecol. Biogeogr.* 20, 653–660. doi: 10.1111/j.1466-8238.2010.00634.x

Ratter, J. A., Askew, G. P., Montgomery, R. F., and Gifford, D. R. (1978). Observations on forests of some mesotrophic soils in Cental Brazil.(Observacoes sobre florestas de alguns solos mesotroficos no Brasil Central.). *Rev. Brasil. Bot.* 1, 47–58.

Reich, P. B., and Borchert, R. (1984). Water stress and tree phenology in a tropical dry forest in the lowlands of Costa Rica. *J. Ecol.* 61–74. doi: 10.2307/2260006

Ribeiro, J. F., and Walter, B. M. T. (1998). "Fitofisionomias do bioma cerrado," in *Cerrado: Ambiente e Flora*, eds S. M. Sano and S. P. Almeida (Planaltina: Embrapa-CPAC), 89–166.

Schrire, B. D., Lavin, M., and Lewis, G. P. (2005). Global distribution patterns of the Leguminosae: insights from recent phylogenies. *Biologiske Skrifter - Det Kongelige Danske Videnskabernes Selskab*, 375–422.

Shmida, A. V. I., and Wilson, M. V. (1985). Biological determinants of species diversity. *J. Biogeogr.* 20, 1–20. doi: 10.2307/2845026

Silva, L. C., Hoffmann, W. A., Rossatto, D. R., Haridasan, M., Franco, A. C., and Horwath, W. R. (2013). Can savannas become forests? A coupled analysis of nutrient stocks and fire thresholds in central Brazil. *Plant Soil* 373, 829–842. doi: 10.1007/s11104-013-1822-x

Silva de Miranda, P. L., Oliveira-Filho, A., Pennington, R. T., Neves, D. M., Baker, T. R., and Dexter, K. G. (in press). Using tree species inventories to map biomes and assess their climatic overlaps in lowland tropical South America. *Global Ecology and Biogeography.*

Simon, M. F., and Pennington, T. (2012). Evidence for adaptation to fire regimes in the tropical savannas of the Brazilian Cerrado. *Int. J. Plant Sci.* 173, 711–723. doi: 10.1086/665973

Staver, A. C., Archibald, S., and Levin, S. A. (2011). The global extent and determinants of savanna and forest as alternative biome states. *Science* 334, 230–232. doi: 10.1126/science.1210465

Van Der Werf, G. R., Randerson, J. T., Giglio, L., Gobron, N., and Dolman, A. J. (2008). Climate controls on the variability of fires in the tropics and subtropics. *Glob. Biogeochem. Cycles* 22:GB3028. doi: 10.1029/2007GB003122

Veldman, J. W. (2016). Clarifying the confusion: old-growth savannahs and tropical ecosystem degradation. *Phil. Trans. R. Soc. B.Biol. Sci.* 371:20150306. doi: 10.1098/rstb.2015.0306

Webb, L. J. (1968). Environmental relationships of the structural types of Australian rain forest vegetation. *Ecology* 49, 296–311. doi: 10.2307/1934459

Whittaker, R. H. (1970). *Communities and Ecosystems*. New York, NY: Macmillan, 162.

Xu, C., Hantson, S., Holmgren, M., van Nes, E. H., Staal, A., and Scheffer, M. (2016). Remotely sensed canopy height reveals three pantropical ecosystem states. *Ecology* 97, 2518–2521. doi: 10.1002/ecy.1470

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

…"If I were myself" seems to represent our greatest risk of living, seems to be a new entrance into the unknown. However, I have the intuition that, past the first so called highs of the party that it would be, we would have the world's experience. We would, at last, taste in fullness the pain of the world. And our pain, the one we learned not to feel. But we would also, at certain times, be taken by a rapture of joy so pure and legitimate that I can barely conceive it. Actually, I think I am already conceiving it, because I felt myself smiling and I also felt a sort of coyness one feels before all things that are larger than us."

Clarice Lispector ("Se eu fosse eu" in: **A Descoberta do Mundo**. Rio de Janeiro: Rocco, 1999. Freely translated from Portuguese into English)