Sevilha, Anderson Cassio (2016) ***Systematic conservation planning for the Paranã River Basin, Brazil, under climate change*****. PhD thesis, James Cook University.**

Systematic conservation planning for the Paranã River Basin, Brazil, under climate change

Thesis submitted by

Anderson Cassio Sevilha

Centre for Tropical Biodiversity and Climate Change, College of Marine and Environmental Sciences, James Cook University, Townsville, QLD, Australia

In fulfillment of requirements for the degree of Doctor of Philosophy in the College of Marine and Environmental Sciences, James Cook University, Townsville, QLD, Australia

Under supervision of the Professors:
Stephen E. Williams
Robert L. Pressey
Jeremy VanDerWal

March 2016

Statement on the contributions of others

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text, and a list of references is given. I declare that I have stated clearly and fully in the thesis the extent of any collaboration with others. To the best of my knowledge and belief, the thesis contains no material previously published by any other person except where due acknowledgment has been made. Author contributions to data chapters are as follows:

Chapter	Nature of contribution	Contribution	Names and affiliations of contributors
2	Intellectual	Chapter conception and design	Anderson C. Sevilha[1,2], Stephen E. Williams[1], Robert L. Pressey[3] and Jeremy VanDerWal[1,4]
		Conception and design of field biodiversity inventories	Anderson C. Sevilha[1,2], Aldicir Scariot[2], Bruno M. T. Walter[2], Taciana B. Cavalcanti[2], Guarino R. Colli[5], Reginaldo Constantino[5], Jader Marinho-Filho[5], Miguel Â. Marini[5], Ivone R. Diniz[5], Helena C. Morais[5] and Rosana Tidon[6]
		Overall data analysis	Anderson C. Sevilha[1,2], Stephen E. Williams[1], Robert L. Pressey[3], Jeremy VanDerWal[1,4] and Helga Correa Wiederhecker[1]
		Writing	Anderson C. Sevilha[1,2]
		Editing	Anderson C. Sevilha[1,2], Robert L. Pressey[3], Stephen E. Williams[1], Jeremy VanDerWal[1,4], Aldicir Scariot[2], Bruno M. T. Walter[2], Taciana B. Cavalcanti[2], Guarino R. Colli[5], Reginaldo Constantino[5], Jader Marinho-Filho[5], Miguel Â. Marini[5], Ivone R. Diniz[5] and Rosana Tidon[6]
	Data Collection	Overall data collection	Anderson C. Sevilha[1,2]
		Field biodiversity inventories	Anderson C. Sevilha[1,2], Helga Correa Wiederhecker[1], Aldicir Scariot[2], Bruno M. T. Walter[2], Taciana B. Cavalcanti[2], Guarino R. Colli[5], Reginaldo Constantino[5], Jader Marinho-Filho[5], Miguel Â. Marini[5], Ivone R. Diniz[5], Helena C. Morais[5], Rosana Tidon[6], Laurie J. Vitt[7] and Janalee P. Caldwell[7]

Chapter	Nature of contribution	Contribution	Names and affiliations of contributors
		Data preparation	Anderson C. Sevilha[1,2] and Helga Correa Wiederhecker[1]
	Funding	Research funding	Anderson C. Sevilha[1,2]
3	Intellectual	Chapter conception and design	Anderson C. Sevilha[1,2], Stephen E. Williams[1], Robert L. Pressey[3] and Jeremy VanDerWal[1,4]
		Overall data analysis	Anderson C. Sevilha[1,2], Stephen E. Williams[1], Robert L. Pressey[3] and Jeremy VanDerWal[1,4]
		Maxent data analysis	Anderson C. Sevilha[1,2], Jeremy VanDerWal[1,4] and Helga C. Wiederhecker[1]
		GDM data analysis	Anderson C. Sevilha[1,2], Kristen J. Williams[8] and Simon Ferrier[8]
		Writing	Anderson C. Sevilha[1,2]
		Editing	Anderson C. Sevilha[1,2], Stephen E. Williams[1], Robert L. Pressey[3] Jeremy VanDerWal[1,4]and Kristen J. Williams[8]
	Data Collection	Data preparation	Anderson C. Sevilha[1,2] and Helga C. Wiederhecker[1],
	Funding	Research funding	Anderson C. Sevilha[1,2]
4	Intellectual	Chapter conception and design	Anderson C. Sevilha[1,2], Stephen E. Williams[1], Robert L. Pressey[3] and Jeremy VanDerWal[1,4]
		Overall data analysis	Anderson C. Sevilha[1,2], Stephen E. Williams[1], Robert L. Pressey[3] and Jeremy VanDerWal[1,4]
		Zonation analysis	Anderson C. Sevilha[1,2], Helga C. Wiederhecker[1], Barbara J. Anderson[1,9] and Atte Moilanen[10]
		Writing	Anderson C. Sevilha[1,2]
		Editing	Anderson C. Sevilha[1,2], Stephen E. Williams[1], Robert L. Pressey[3] and Jeremy VanDerWal[1,4]
	Data Collection	Data preparation	Anderson C. Sevilha[1,2], Helga C. Wiederhecker[1], Barbara J. Anderson[1,9] and Atte Moilanen[10]
	Funding	Research funding	Anderson C. Sevilha[1,2]

Chapter	**Nature of contribution**	**Contribution**	**Names and affiliations of contributors**
5	Intellectual	Chapter conception and design	Anderson C. Sevilha[1,2], Stephen E. Williams[1], Robert L. Pressey[3] and Jeremy VanDerWal[1,4]
		Overall data analysis	Anderson C. Sevilha[1,2], Stephen E. Williams[1], Robert L. Pressey[3] and Jeremy VanDerWal[1,4]
		Maxent data analysis	Anderson C. Sevilha[1,2], Jeremy VanDerWal[1,4] and Helga C. Wiederhecker[1]
		Writing	Anderson C. Sevilha[1,2]
		Editing	Anderson C. Sevilha[1,2], Stephen E. Williams[1], Robert L. Pressey[3] and Jeremy VanDerWal[1,4]
	Data Collection	Data preparation	Anderson C. Sevilha[1,2] and Helga C. Wiederhecker[1]
	Funding	Research funding	Anderson C. Sevilha[1,2]

Affiliations:

[1] Centre for Tropical Biodiversity and Climate Change, School of Marine and Tropical Biology, James Cook University, Townsville, QLD, Australia;

[2] Embrapa Genetic Resources & Biotechnology, Brasília, DF, Brazil;

[3] Australian Research Council Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, QLD, Australia;

[4] eResearch Centre, James Cook University, Townsville, QLD, Australia;

[5] Department of Zoology, University of Brasília, Brasília, DF, Brazil;

[6] Department of Genetics and Morphology, University of Brasília, Brasília, DF, Brazil;

[7] Sam Noble Museum & Department of Biology, University of Oklahoma, Norman, OK, United States of America;

[8] CSIRO Ecosystem Sciences, Black Mountain Laboratories, Canberra, ACT, Australia;

[9] Rutherford Discovery Fellowship, Landcare Research, Dunedin, New Zealand, and

[10] Department of Biosciences, University of Helsinki, Helsinki, Finland.

Research funding

This study was carried out with financial support from the Brazilian Corporation of Agricultural Research (EMBRAPA); James Cook University (JCU); National Programme on Biological Diversity (PRONABIO); Project for Conservation and Sustainable Use of Brazilian Biological Diversity (PROBIO); National Environment Fund (FNMA); Conservation International-Brazil (CI-Brazil); International Bank for Reconstruction and Development/Global Environment Facility (BIRD/GEF), Ministry of Environment (MMA), and National Council of Technological and Scientific Development (CNPq). EMBRAPA also supported all expenses with tuition fees and scholarship.

Anderson Cassio Sevilha March 2016

Dedication

To the little boy and the little girls…

Acknowledgments

First of all, I wish to thank my supervisors Steve Williams, Bob Pressey and Jeremy VanDerWal, who guided me throughout the processes of this thesis conception, from data collection to data analysis. I am grateful for the time shared in intensive discussions and for their critical reviewing of all manuscripts.

From James Cook University (JCU), I thank Ben Phillips for his assistance with R, April Reside for her assistance with Maxent, Justin Welbergen for his advice, my colleagues from Bob's lab – Jorge, Steve, Natalie, Alana, Tatjana, Adrian, Georgina, Amelie, Mélanie, Ian, Vera, Vanessa, Morena, Mariana and Rafael, and colleagues from Steve's lab – John, Stewart, Collin, Nadiah, Scott, Lauren, Brett, Carina and Robert Puschendorf – for the time and discussions shared.

I am grateful to two anonymous examiners for their valuable and constructive comments on my manuscript.

I wish to acknowledge the many other people who helped and supported me during my PhD:

- Yvette Williams and Debbie Ford (JCU), Neuza Avelino and Antonio Craveiro (EMBRAPA) for their assistance with JCU and EMBRAPA paperwork, respectively.
- I thank Kristen J. Williams and Simon Ferrier (CSIRO Ecosystem Sciences, Australia) for their support and discussions on GDM. Glenn Manion (Office of the Environment and Heritage, Australia) for his prompt help in fixing bugs on .Net GD Modeller to generate community models.
- Barbara Anderson (Rutherford Discovery Fellowship, New Zealand) and Atte Molainen (University of Helsinki, Finland) assisted with advice on the application of Zonation.
- Aldicir Scariot, Bruno M. T. Walter and Taciana B. Cavalcanti (EMBRAPA), Guarino R. Colli, Reginaldo Constantino, Jader Marinho-Filho, Miguel Â. Marini, Rosana Tidon and Ivone R. Diniz (UnB), Júlio Dalponte (UNEMAT), and Laurie J. Vitt and Janalee P. Caldwell (University of Oklahoma), who helped me to coordinate, carry

out, identify species collected in the biological inventories and reviewed the manuscript referring to species occurrence.

- Bruno M. T. Walter, also as my ‘Walter’ ego and formal supervisor at EMBRAPA, Marcelo B. de Medeiros (EMBRAPA) for financial support to carry out physiographic inventories, Sergio E. Noronha (EMBRAPA) for his assistance on GIS and field work, Helena C. Morais (UnB) and Vitor O. Becker for their help in Saturnid-moths sampling design and specimens identification.
- The many Brazilian students at that time (Alexander Balbuíno, Alexandra Bezerra, Alexandre B. Sampaio, Alisson Gainsbury, Ana Hermínia, André Terra, Arthur Brant, Cledimara Sinigaglia, Daniel L. M. Vieira, Daniel Mesquita, Danilo L. Bernardo, Ednaldo Rocha, Ernestino Guarino, Fernanda Pinheiro, Fernanda Werneck, Flávia C. Martins, Flavio L. L. Souza, Gabriel C. Costa, Gustavo H. C. Vieira, Lemuel O. Leite, Leandro Godoy, Leonora T. Bastos, Luciana Zago, Luciana Silva, João B. Bringel, João M. Resende, Júlio César Roma, Renata A. Mata, Mariana Zatz and Patrícia Bueno), field assistants from Embrapa (Nilton B. Ferreira, Juarez P. Amaral, Glocimar P. Silva, João B. Pereira, Aécio A. Santos, Salim, Camarão, Zé Dias, Noel, Melão and Gledson Moreira), Santos Balbino and Joaquim from UnB, and Garça and Romilda, for their support and help in collecting for the biological data used here.
- Paula H. Valdujo and Daniela A. S. Oliveira (Brazilian Environment Ministry) for providing species occurrence information available at that Ministry.
- All people and Institutions listed in the sources of online Supporting Information 1.1 for providing data under their management.
- Lincoln Pereira and José Pereira (in memorian), Arnaldo Pereira, Carlos Pereira, Dona Isabel, Rogério Daia, Sebastião Nilson, Aldair Freire, Alfredo, Bira, Célio, Alagoanos, Pica-Pau and Regalito for the permits to work on their lands for such a long time.
- Cláudio, ZéMir (in memorian), André, Valdeci, Tião, André Macarrão, Chiquinho, Edson, Lirinho, Pixutinha Branco, Pixutinha Preto, Almir, Zézinho, Dida, Careca, Diu, Boda, Nirso and Lacerda, people from Monte Alto Village, who helped us to collect data through the years.

- Carmelito, Valdete, Aristeu and Agripina for their support to our team in Monte Alto Village during field work since we arrived there, in 1999.
- Luciano Adrián González (CSIRO Animal, Food and Health Sciences) for the long discussions trying to solve the problems of the world.
- Professors Alexandre Francisco da Silva (in memorian) and João Augusto Alves Meira Neto, who shaped my knowledge in plant ecology.
- Florence's and Young's families (Arthur, Melissa, Sebastian, Trinity, Doug, Connie and Janell), who adopted my family in Australia as part of theirs. I am really grateful.
- Newton and Hilda for their patient.
- Mom and Dad for their abnegation.
- Helga, as my partner, for her comprehension and for taking care of Olivia during my absences; and as a biologist, for her help in collecting, analyzing and discussing data.
- Olivia, the horse girl, for accepting (sometimes) my eternal absence during this period. Sorry for me, sorry for you!

This work was carried out with financial support from the Brazilian Corporation of Agricultural Research (EMBRAPA), James Cook University (JCU), National Programme on Biological Diversity (PRONABIO), Project for Conservation and Sustainable Use of Brazilian Biological Diversity (PROBIO), National Environment Fund (FNMA), Conservation International-Brazil (CI-Brazil), International Bank for Reconstruction and Development/Global Environment Facility (BIRD/GEF), Ministry of Environment (MMA), and National Council of Technological and Scientific Development (CNPq), all of which I thank. EMBRAPA also supported all expenses with tuition fees and scholarship.

Finally, as a public servant, I wish to thank the effort of all anonymous Brazilian taxpayers, hoping I would meet their expectations for valuable use of their tax contributions.

Note by the author

Writing a thesis is definitely a lonely journey. Sometimes, it makes you feel like Frodo Baggins – a Tolkien character – walking through windy paths and carrying all the weight of his own thoughts. Sometimes, it makes you feel like Gollum, so possessive of your own work: "my precious, mine". This thesis is my precious. I am responsible for every single word written here, all data, all analysis, all graphics and tables, that is, everything, including commas placed in the wrong spots. However, this thesis is not a result of a lonely work. It is based on 15 years of fieldwork. It has benefited from the efforts of many people interested in studying and preserving the natural complexity of a basin carved in a Gondwana landscape by the Paranã River, in the heart of Brazil.

I arrived in the Paranã River Basin for the first time in 1999 to study the effects of forest fragmentation in dry forest ecosystems. Later, between 2002 and 2005, I was also involved in reintroducing tree species threatened with extinction in the Basin. Meanwhile, during 2003 and 2004, I coordinated two projects with the primary objective of generating biodiversity information for the Basin, a region classified as poorly known in terms of its biological diversity by Brazilian environmental authorities. On these projects, I was responsible for the field inventories of plants and colleagues of mine, Professors from University of Brasília, for the field inventories of amphibians, birds, mammals, reptiles, termites, drosophilids and saturnids. On all of these projects, we were helped by many researchers, students and technicians from EMBRAPA (Brazilian Corporation of Agricultural Research), University of Brasília and State University of Mato Grosso. In total, there were more than 80 people involved. In 2006, after we have concluded the projects, my colleagues – from EMBRAPA and University of Brasília – and I were invited to participate in the validation meetings of the priority areas for biodiversity conservation in the Cerrado biome, selected by the Brazilian Government using information generated by our projects. At the time the Brazilian conservation planning was released, in 2007, we thought that our aims had been achieved. More than 80% of the Paranã River Basin was considered to have priority for conservation actions. However, no single fully protected area has been created in the region since then. Talking to stakeholders on different occasions, I heard numerous times that the original broadly delimited areas highlighted as priorities for biodiversity conservation were too large and roughly delineated to be translated directly into conservation actions.

Recognizing the importance of the Paranã River Basin for biodiversity and the massive effort made by the Brazilian Environmental Ministry in identifying priority areas for conservation – attending to the commitments made in different meetings of the Convention on Biological Diversity – I decided to act proactively to downscale broad priorities to resolutions matching those of units of management on the ground. At the end of 2009 and the beginning of 2010, I arrived at James Cook University to begin my PhD. I arrived without any background in systematic conservation planning, species and community modelling and the impacts of climate change on biodiversity. The only things I knew were the Paranã River Basin, its people, plants, animals, and the certainty of the importance of the region for worldwide biodiversity conservation. After long discussions with my supervisors, who guided me throughout the processes of a thesis conception, I returned to the Basin to search for more information on species occurrences. All the data that I had in my hands – the biodiversity inventories – was not sufficient to proceed with my work. Back in Brazil, I undertook over 10,000 km of ground travel to characterize the different environments present in the Basin, describing the landscape, type of soils, vegetation, and species occurrences. After this trip, I visited most of the museums, herbaria, libraries in different institutions and agencies that hold biological information from the region. Back to Australia, I processed and analyzed the data.

Therefore, this thesis is the story that I want to tell you. It is a story built by many hands, by many people. For this reason, as the Gollum that used to speak of himself as “we”, I beg your permission to write this thesis in the third person. The use of the third person is a recognition of all the effort of all the people that have been helping me for a long time to understand and protect the biodiversity of the Paranã River Basin.

This thesis consists of four core studies, each of which has been prepared as a manuscript for submission to a scientific journal and reformatted here as data chapters

(Chapters 2 to 5). The bibliography format adopted here follows that of the journal PLoS ONE, to which we have submitted the first manuscript (Chapter 2).

The data chapters are preceded by a general introduction (below) and followed by final considerations in a general discussion. In this introductory chapter are presented the problem, approaches, objectives and a general methods, organised to reduce repetition in the methods sections of the chapters that follow. In the general discussion are presented the main conclusions, problems faced during the execution of this thesis, their solutions, implications of the results, and also discussion of some future investigations that should take place to advance biodiversity conservation in the study region.

Systematic conservation planning for the Paranã River Basin, Brazil, under climate change

Abstract

Planning for effective conservation begins with precise information about species occurrences, accurate species identifications, knowledge of the distributions of species and communities and threats to biodiversity posed by anthropogenic activities and resultant climate change. In the Paranã River Basin, a key priority area for conservation within a global biodiversity hotspot – the Cerrado biome in Brazil – the areas previously highlighted as priorities for conservation by the Brazilian Government are too large and roughly delineated to be translated directly into actions on the ground**.** Which areas should be prioritized for biodiversity conservation within the Basin and why those areas should be prioritized are questions that are still not answered. These questions underpin the aims of this thesis. Answering those questions requires specific information, including: (a) identification of species occurrences within the Basin and gaps in biological information; (b) determination of patterns of species occurrences, species richness, and distribution of communities to support area prioritization for nature conservation; (c) evaluation of different biodiversity features as inputs for area prioritization, considering, at least, distributions of species and communities, richness patterns, threats and endemism of species, connectivity, complementarity, irreplaceability and vulnerability to anthropogenic activities; and (d) evaluation of the likelihood of changes in species ranges, habitat resilience and connectivity to identify priority areas for conservation considering the impacts of climate change on biodiversity.

To answer those questions, firstly, we compiled a database drawn from our intensive surveys of plants, birds, mammals, reptiles, amphibians, termites, drosophilids, and saturnids. Additionally, we exhaustively searched for biodiversity records in museums, herbaria, literature, and online databases. Having exhaustive biological information for the study area, we standardized taxonomic names and data formats compiled from diverse and disparate sources and applied an innovative method to automate identification of duplicate records. Secondly, we evaluated biodiversity patterns considering the distributions of species and communities. We analyzed richness and community patterns using, respectively, Maxent (a maximum entropy machine-learning algorithm) and GDM (Generalized Dissimilarity Modelling). We modelled distributions for 2159 species and modelled dissimilarities for each of the terrestrial vertebrate groups as well as insects and vascular plants. Thirdly, we evaluated the use of species and community distributions as biodiversity features for area prioritization using the Zonation conservation planning software. Here, we evaluated the separate contributions to area priorities of species distributions and community turnover as well as the influence on priorities of combined data on species and communities and the addition of other biodiversity features. We also compared the conservation solutions we generated with a previous prioritization of conservation units by the Brazilian Government and made inferences about the use of different approaches to area prioritization. Finally, we used Maxent to project species distributions for current and future climate and then applied Zonation again to identify areas for conservation that are resilient to the effects of climate change on species distributions.

Of ~140,000 records initially compiled for the entire Basin, only ~20,000 reliable records were retained. The 5,130 species currently known constitute between 35% and 54% of the larger Cerrado biome’s terrestrial vertebrate species, 34% of its plant species, and 11–43% of its endemics. Even with 200 years of data collection, spatial information gaps still exist for all taxa sampled in the Basin. The biodiversity information generated, however, was sufficient to build species and community distribution models for some important biological groups. Species distribution models showed differences in richness patterns between vertebrate groups (amphibians, birds, mammals, and reptiles) and plants. Plants are more diverse at higher elevations, contrasting with vertebrates that had higher richness at lower elevations. On the elevational gradient, environmental heterogeneity – represented by differences in climate and substrate – is related to more differentiation of vegetation types. At lower elevations, climate conditions are more stable, mean temperatures are higher, and the environment – dominated by dry forests and arboreal savanna – is vertically more structured, i.e. showing up to five strata or vertical layers (herbaceous, shrub, understory, canopy and emergent). Reliable models for community patterns could be

obtained only for birds, mammals and plants. With few exceptions, community groups based on birds and mammals followed the same pattern as those based on plants. The community patterns indicate three major bioregions within the Basin: the mountain ranges on the altitudinal gradient and the areas occupied by dry forest and savanna formations at lower elevations. Although the results produced by models of species distribution and community turnover are different in their details in depicting biodiversity patterns, they are complementary in highlighting areas of importance for biodiversity conservation. Additionally, we concluded that the use of richness patterns of some biological groups as proxies for other groups is unreliable. Different groups respond differently to variations in abiotic and biological heterogeneity across the Basin. These differences should be considered in conservation design.

The choice of whether species- or community-level patterns should be used in area prioritization using Zonation depends on the objectives of conservation and data availability. Considering the most important areas identified for biodiversity conservation (the top 20% fraction of the landscape), all Zonation solutions retained a similar amount of the total distributions of each biodiversity feature [4% on average (2.37 SD) across species, communities, and threatened ecosystems]. Additionally, the priority areas in these different solutions were strongly congruentities, whether the biodiversity features were populations, communities or ecosystems. Those most important areas were located along the mountain ranges of the Basin. Differences between species- and community-level approaches in allocating priorities are evident only in relatively small areas at lower elevations.

Prioritization of areas for the Paranã River Basin considering the impacts of climate change reveals some congruence between current and future areas indicated as important for biodiversity conservation. Besides the mountains ranges, the overlap of present and future priorities includes elevational gradients, some flat terrain that surrounds the ranges, and limestone outcrops. Additional important areas are those prioritized only for the future scenario, which emphasised the connecting zones between current and future priority areas in different elevation gradients of the Basin.

The broadly defined conservation planning done for the Basin in 2007 must be revisited to maintain resilience under climate change and facilitate species persistence into the future. There is a very low overlap between priority areas for biodiversity conservation revealed by this study and the areas identified for conservation action by Brazilian Government. Of the total area of the Basin officially identified for conservation measures in 2007 (some 82% of the Basin), the establishment of protected areas overlaps only 6.95% with the most important areas for biodiversity conservation identified by this study, considering the intersection of current and future priorities (~10% of the Basin). The enforcement of sustainable use overlaps 5.23%, while existing protected areas overlap 0.09%. However, a few small priority areas identified in this study (0.3%) are outside those identified in 2007. This thesis is a guide to the establishment of protected areas, spanning gradients from the flat areas of the lowlands to the tops of mountains, as a core part of a strategy to promote the persistence of species in the face of future shifts in climate.

Table of contents

List of tables

List of figures

List of plates

Chapter 1

Chapter 1 – General introduction

1.1. Introduction

Reconciling economic development and biological conservation is one of the biggest challenges facing human society. The current rate of species extinction is unprecedented in the history of mankind and is comparable in scale to prehistoric mass extinctions [1-4]. However, this is also the first time that any attempt has been made to preserve biodiversity globally. The Stockholm Conference in 1972 was the first major global environmental meeting arranged by the United Nations – which resulted in the creation of the United Nations Environment Programme (UNEP) – to tackle the problem. In 1992, at the Rio Earth Summit, the first Convention on Biological Diversity was held and provided a global legal framework for action on conservation biodiversity and sustainable use of biological diversity [5]. In 1998, the Kyoto Protocol was open for countries signature by countries to avoid the adverse effects of climate change on the physical environment, biodiversity, and human society. In 2002, world leaders committed, through the Convention on Biological Diversity, to achieve a significant reduction in the rate of biodiversity loss by 2010 [see 2]. However, on a global scale, it seems unlikely that the 2010 target has been reached. Despite some local successes and positive trends in some indicators – extent and biodiversity coverage of protected areas, sustainable forest management, policy responses to invasive alien species, and biodiversity-related aid – the rate of loss of biological diversity does not appear to be slowing [6]. Some species considered threatened by the IUCN Red List Index – an index developed by the International Union for Conservation of Nature and partners to measure trends in the extinction risk of species – have faced an increased risk of extinction during the past 40 years. At the same time, deforestation, pollution, invasive species and the impacts of climate change on organisms have also increased [6]. At the tenth meeting of the Conference of the Parties to the Convention on Biological Diversity, a revised and updated Strategic Plan for Biodiversity was adopted. This plan included the Aichi Biodiversity Targets [7] for the 2011-2020 period. Among others, by 2020, Aichi Biodiversity Target 11 aims to have at least 17% of terrestrial and inland waters of particular importance for biodiversity and ecosystem services conserved through effectively and equitably managed, ecologically representative and well-connected systems of protected areas. In this context, conservation actions have been more deeply debated within the Brazilian Government, introducing decision-makers to academic knowledge about conservation. This Thesis aims to contribute to the Brazilian Government's ability to achieve the targets of the Convention on Biological Diversity by prioritizing areas for biodiversity conservation in a global biodiversity hotspot.

Creating protected areas for biological conservation is not a new idea. It emerged in the mid-nineteenth century from the necessity of preserving scenic beauty, species rarity or even species of economic interest [8]. However, the process of identification and prioritisation has evolved towards the development of specific protocols through systematic conservation planning [9-12]. These protocols take into account not only the aims of conservation and information on biodiversity and/or some key species, but apply a process that includes: (a) estimates of costs of conservation, (b) identification of and consultation with key stakeholders; (c) reviewing policy, legal, institutional and socio-economic opportunities and constraints for conservation; (d) identifying comprehensive goals and strategies for conservation and its implementation; (e) gathering and evaluating spatially explicit data on biodiversity features and factors that determine their occurrence and threaten their persistence; (f) formulating objectives for biodiversity and other natural features to promote the persistence of biodiversity; (g) reviewing the effectiveness of existing conservation areas; (g) identifying and selecting additional conservation areas; (i) implementing conservation actions in selected areas, and (j) maintaining, monitoring and managing areas to promote biodiversity and other natural values [see 10, 11, 13, 14 for more details].

In addition to many other threats to biodiversity, global warming and its adverse effects are now widely accepted [15-17]. One of the challenges posed by climate change to conservation planning is the fact that fixed protected areas may have a different mix of species in the future than they have now. As a result, shifts in species' ranges into future suitable areas and/or local extinctions for species that cannot shift are expected [15, 17-24]. It is now necessary for conservation planning to be both temporally and spatially dynamic and move away from the static paradigms previously dominating thinking in conservation. Revisiting conservation plans or conservation areas is, therefore, necessary to monitor their efficiency and effectiveness on the ground for both the present and future. Tools to do this are now available, enabling planners to anticipate the ways that species and communities might change. Thus, assigning priority areas for conservation has advanced towards incorporating this new set of considerations into systematic conservation planning [13, 18, 21, 22, 25-27].

1.1.1. Conservation planning and climate change in Brazil

The Brazilian Government first began to identify priority areas for conservation in 2002 with PROBIO (Project of Conservation and Sustainable Use of Brazilian Biological Diversity), as a response to the Convention on Biological Diversity commitments adopted in 1992. Prioritization of areas was based on regional workshops of biodiversity programs organized by Conservation International [28]. These workshops drew on the expertise of some researchers to identify the most important areas for conservation. As a result, 390 areas were suggested as priorities for conservation, sustainable use, and benefit sharing.

In 2007, the Brazilian Government revised the priorities suggested in 2002 [29]. This revision followed the widespread acceptance of systematic conservation planning concepts, aggregating new data, and using new methods of analyzing data for area prioritization. In this new process, quantitative objectives were set for conservation features that embraced biodiversity, sustainable use, resilience and process. Biodiversity features consisted mainly of endemic and threatened species, habitats, and phytophysiognomies. Species of economic and medicinal importance, species necessary to promote conservation (flagship species), and keystone species were also considered. Attention focused also on areas necessary for maintaining environmental services, climatic refuges, connectivity and gene flow, or those with high endemism and diversification. As a result, the number of priority areas for conservation increased to 2684, encompassing ~54% of the Brazilian terrestrial environment.

Like the 2002 workshop, however, the 2007 process was dominated by subjectivity based on the personal experiences and knowledge of researchers invited to participate. This subjectivity prevailed over explicit criteria to define conservation features, establish their respective conservation objectives, and select areas to represent them (Bruno M. T. Walter, personal communication). An important reason for the approach taken was the lack of consistent information arising from the continental dimensions of the exercise, lack of information about species' absence/presence, and incomplete information on environmental units. The consultants extensively modified the maps resulting from the computational analyses used for the evaluation of areas during validation meetings, adding a considerable number of regions not prioritized by the explicit method. Moreover, potential climatic changes and their effects on biodiversity were not considered in the selection of priority areas for biodiversity conservation, sustainable use and ecological corridors [see 29].

1.1.2. Case study: the Paranã River Basin

Subjectivity bias is evident in the identified priority conservation areas in the Paranã River Basin, an area covering approximately 60,000 km^2 located in the middle of the Cerrado biome – a global biodiversity hotspot [30] – in central Brazil (Figure 1.1). The final plan in 2007 included several areas in addition to those previously indicated in 2002 (Figure 1.1). In one way, the result was beneficial because the total area of the Basin considered as priority increased considerably, from ~56%, in 2002 [28], to ~82%, in 2007 [29]. On the other hand, the results were questioned by stakeholders. According to them, the original broadly delimited areas highlighted as priorities for biodiversity conservation were too large and roughly delineated to be translated directly into conservation actions. Such an extensive coverage of priority areas still requires planners to decide exactly where conservation action will be applied and where to invest first. Importantly, these analyses have not resulted in a single additional fully protected area being established [29].

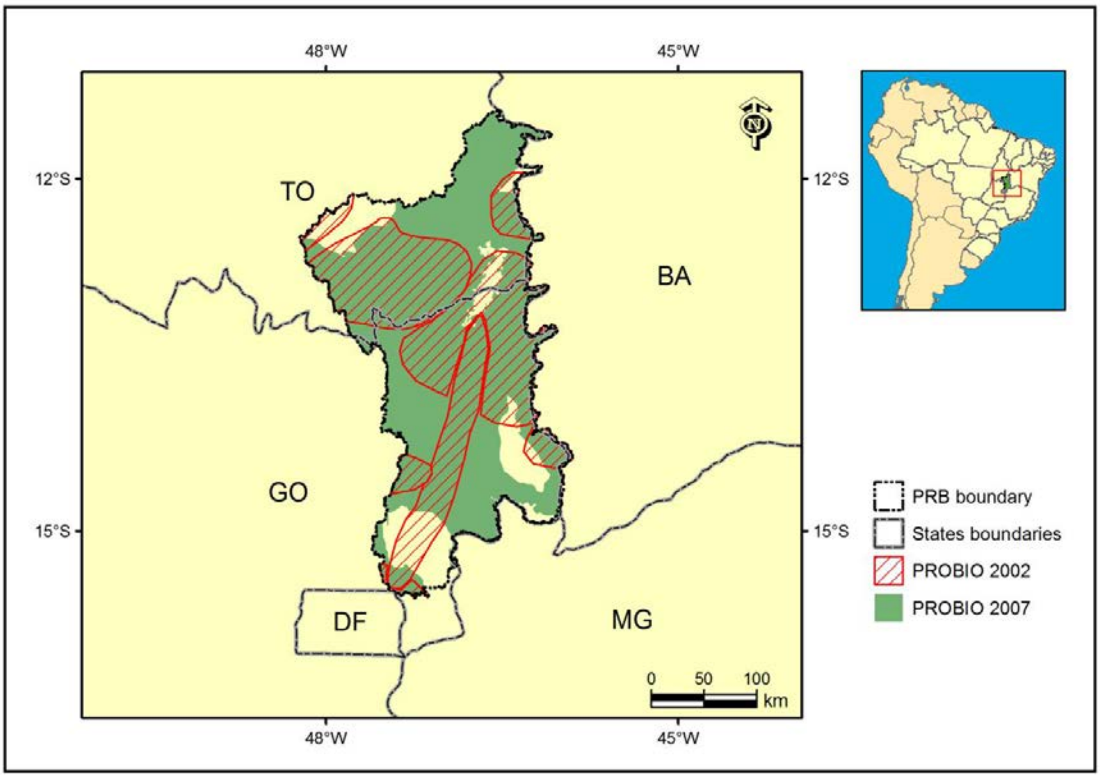


Figure 1.1. Areas indicated as priority for conservation, sustainable use and benefit sharing of biological diversity by PROBIO (Project of Conservation and Sustainable Use of Brazilian Biological Diversity) in 2002 (dashed red areas)[28], and in 2007 (green area)[29] across the Paranã River Basin.

For several reasons, the Paranã River Basin provides an excellent study region to develop a systematic conservation plan considering the future impacts of climate change. First, the Paranã River Basin has long been recognized as a priority for biodiversity conservation and stakeholders are mobilized and sensitive to the need for creating protected areas in the region [28, 29]. As part of a broad delineation of conservation priorities in the Cerrado, approximately 82% of the 61,856 km^2 of the Basin was identified as having national priority for conservation actions (Figure 1.2). Conservation priorities for the Basin include establishment of conservation units, biodiversity inventories, delineation of ecological corridors, restoration, and promotion of sustainable use [29]. Nevertheless, this plan does not provide any details of how these goals will be achieved. Therefore, it has not been successful in creating a single fully protected area within the Basin since 2007 when the last report on Brazilian priority areas for conservation was released. Up to the present, only 0.85% of the

Basin’s total area is fully protected within two conservation units: 3,674 ha in Chapada dos Veadeiros National Park, and 47,545 ha in Terra Ronca State Park (Figure 1.2). Second, the uniqueness of the Basin’s fauna, flora, vegetation, and land formations, high numbers of rare, endemic, and threatened species, and its significance as a center of diversity for several biological groups [31-41] are some aspects of its significance for conservation. Within the Basin is also located a globally endangered tropical ecosystem and one of the last remnants of one of the most threatened ecosystems in Brazil: the deciduous forest [31, 42-44]. Third, the Paranã River Basin is a mosaic of areas in different stages of fragmentation in the Cerrado biome but in a climatic transition zone between Amazonian and Caatinga biomes. The location of the Basin in this transition zone could mean that the region will serve as a refuge for species of different ecosystems and a corridor for species movement. Lastly, it is likely that the biodiversity of tropical deciduous forest and savanna formations of the Paranã River Basin is susceptible to the potential threat of climate change, like tropical formations in some other parts of the world.

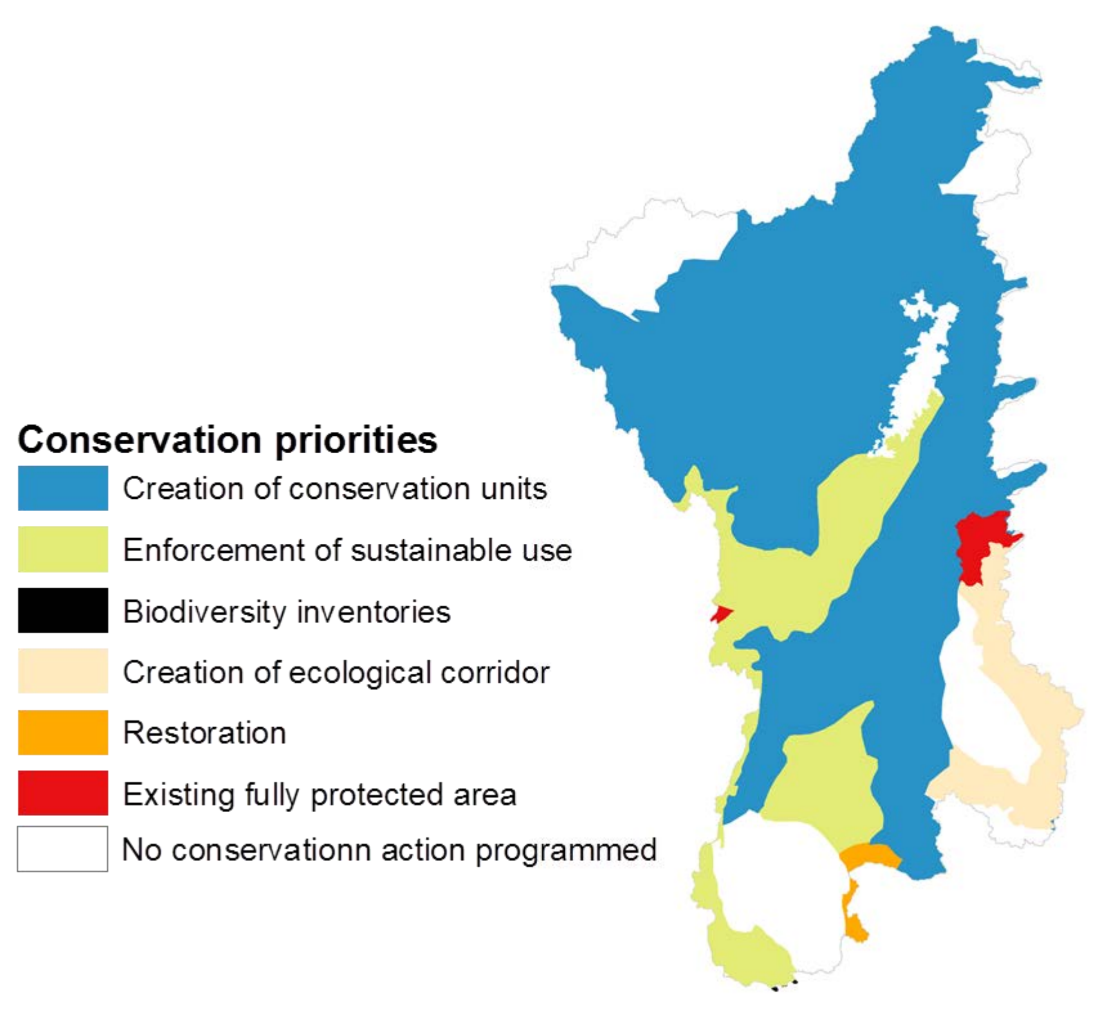


Figure 1.2. Conservation priorities identified by the Brazilian Environment Ministry for the Paranã River Basin in 2007.
Note that biological inventories were recommended for only two tiny areas on the southern margin of the Basin.

1.1.2.1. Human occupation

Cave paintings made by indigenous people about 10,000 BP are the first records of human occupation in the Paranã River Basin (see Plate 1.1). Naturalists and explorers such as Spix and Martius (1818) and Gardner (1839-1840), who visited the region in the 19th century [45, 46], were the first to make official records and detailed descriptions about Basin's natural history and human occupation. According to these explorers, during the first half of the 18th century, gold and diamond miners occupied the western and northern boundaries of the Basin. Their demand for food attracted farmers to the region, who took advantage of the natural pasturelands for cattle ranching. With the decline of mining activity in the late 18th century, cattle ranchers and fugitive African slaves remained in the area. They established themselves in the Basin in small villages, farms and communities of fugitive slaves named ‘Quilombos’. By the first half of the 19th century, cattle ranching was the most significant economic activity in the Basin [45-47]. Consequently, the negative impact of those occupations and activities on local indigenous populations were observed by naturalists. With many indigenous people exterminated or driven out by the new colonizers, only small remnants of these tribes remained in the region at that time [48].

Despite the lack of a government development plan for the region, until 1970 moderate production of cattle meat and leather kept the vegetation in the Paranã River Basin relatively well-preserved. However, after 1970 there was a large wave of human occupation, and farmers from southern and southeastern parts of Brazil arrived to extract timber and expand cattle ranching activity in the Basin [49], mainly in dry forest areas. Intensive timber extraction ended in the 1990s while cattle ranching has remained as the primary economic activity until the present day. Sporadic timber extraction for wood and charcoal production still occurs in dry forests. It is more intensive in savanna formations that are cleared for pastures. Moreover, in the past decade, limestone mining and eucalyptus plantations have increased in the region, threatening the last intact remnants of dry forests growing on limestone outcrops across the eastern boundary and the intact arboreal savanna formations of the northern and southern parts of the Basin [42, 50]. Currently at least 30% of the Basin’s total area is converted to human uses (see Figure 2.1).

At present, ~300,000 people live in the Paranã River Basin, of whom 69% live in urban areas. A few indigenous people remain from the Ava-Canoeiro tribe residing in Cavalcante county [48], on the western border of the Basin. The ‘Quilombola’ populations are more numerous, most of them living in small isolated groups in the well-preserved portions of western and northern parts of the Basin, unaware of their legal rights [50]. After the Brazilian Government recognized the remaining Quilombos as communities that have rights over their ancestral territories (Art. 68 of Brazilian Constitutional Law of 1988 and Decree-Law N. 4.887/2003), only one community (with approximately 4,500 people) had their land (253,000 hectares) legally delimited [51, 52].

The 33 counties that encompass the Paranã River Basin vary widely in development, with the Human Development Index (HDI) ranging from 0.603 to 0.750 [31]. The average HDI for the region (0.665) is much lower than that observed for the states of Goiás (0.770) and Tocantins (0.721), in which the Basin is embedded. Some 45% of rural properties in the Basin cover between 10 and 100 ha, most of these located in rugged terrain. On the other hand, larger farms (over 500 ha) are located in flat areas, with high-fertility soils. The regional economy is dominated by cattle ranching farms (69.4% of the rural areas), with ~1,300,000 cattle, followed by farms with both agriculture and cattle raising (15.5%), and farms exclusively used for agriculture (11.6%). Other rural economic activities are forestry, logging, charcoal production, fisheries and aquaculture. Calcareous rock mining and eucalyptus plantations, both destructive activities, are increasing in the region each year [31, 43, 50, 53].

The present scenario of land occupation within the Basin poses an extra challenge to the establishment of large protected areas for conservation, as initially proposed by the Brazilian Government [29]. Prioritization of feasible areas must now maintain, at least, the most valuable biodiversity features remaining in a diverse matrix of ownership, use rights, and commercial activities.

Plate 1.1. Cave paintings recorded in Terra Ronca State Park, Paranã River Basin.

1.1.2.2. Physiographic aspects

The Paranã River Basin lies within the Brazilian states of Goiás and Tocantins between coordinates -11.38 S to -15.57 S, and -45.90 W to -47.24 W (Figure 1.1), in a climatic transition zone between Amazon, Cerrado and Caatinga [26]. The Basin consists of a flat depression within the Brazilian central plateau oriented from south (at ~600 m asl in average) to north (~240 m). A mountain range surrounds the Basin on its western and southern borders – Serra Geral do Paranã (1200 m). A plateau – Chapadão Ocidental da Bahia (~900 m), with escarpments – Serra Geral de Goiás – extends throughout the eastern border. Small hills (450 m) complete the boundary of the Basin on its northern rim (Figure 1.3A and B and Plate 1.2). The predominant climate in the region is Aw (tropical semi-arid in Köppen’s classification), with dry winters and rainy summers. The average annual precipitation ranges from 1,164 to 2,012 mm and is highly seasonal with about 90% falling between October and March. The average temperature varies from 19.5 to 26.5 ^{o}C, and elevation ranges from 247 to 1,663 m asl [31, 43, 54].

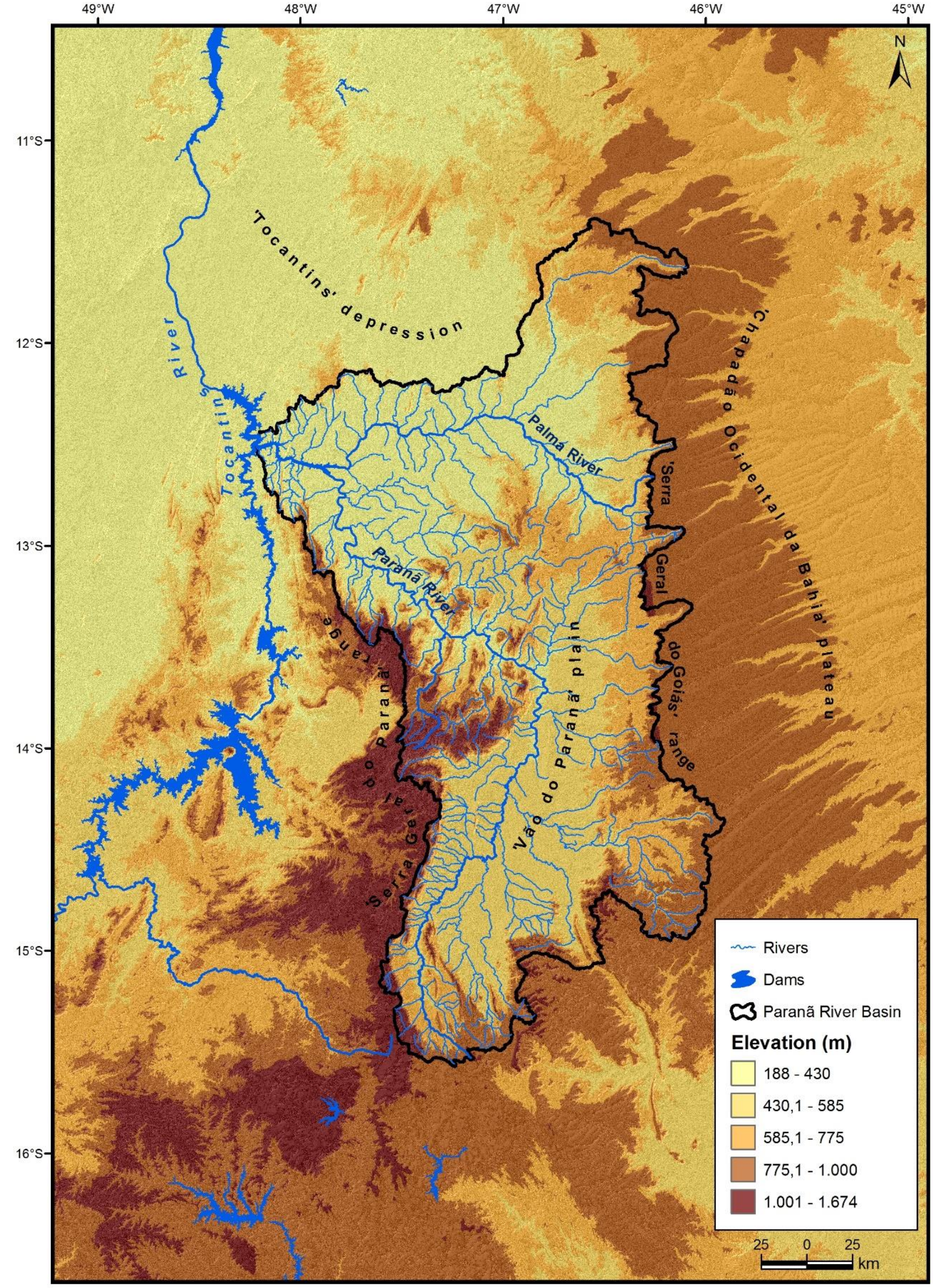


Figure 1.3A. The physical map of Paranã River Basin and surrounding terrains. The physical map shows classes of elevation [55], land relief (mountain ranges, plateaus, and plains [56]), and hydrology (dams, detailed hydrographic aspects of the Basin and main river for the surrounding terrain [57]).

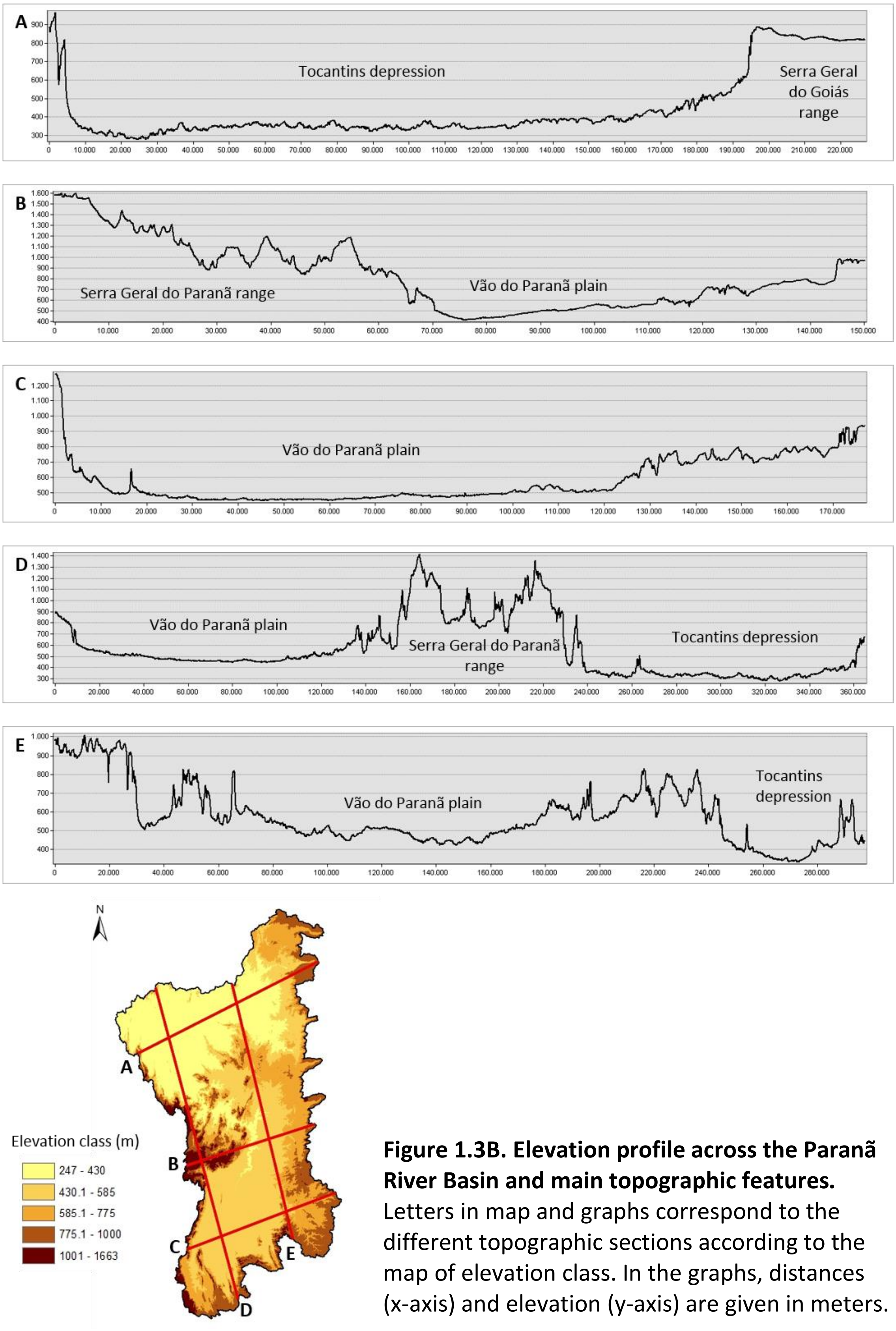


Figure 1.3B. Elevation profile across the Paranã River Basin and main topographic features. Letters in map and graphs correspond to the different topographic sections according to the map of elevation class. In the graphs, distances (x-axis) and elevation (y-axis) are given in meters.

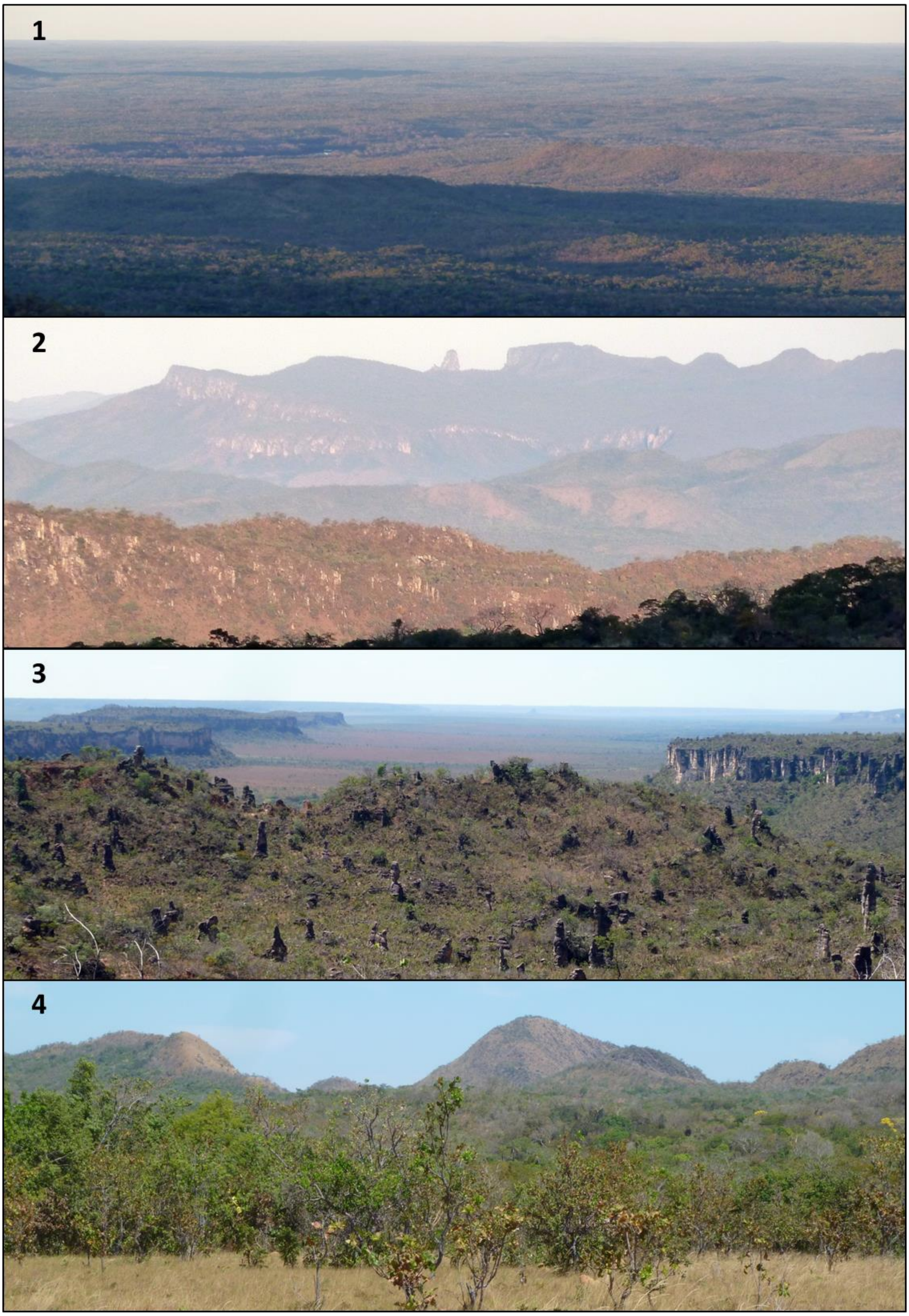


Plate 1.2. Landscape aspects of the Paranã River Basin.
1 = Flat area on the northern portion of the Basin's at low elevation Tocantins depression). 2 = Mountain range on the western border (Serra Geral do Paranã range). 3 = Eastern border (Serra Geral do Goiás range) formed by the Chapadão Ocidental da Bahia Plateau. 4 = Small hills that form the northern boundary of the Basin.

1.1.2.3. Vegetation

Inserted within one of the global hotspots, the Cerrado biome, the Basin is covered with a mosaic of distinct vegetation types. Savannas, grasslands, and seasonally dry forests are the most prominent vegetation forms (Figure 1.4 and Plate 1.3). Differences in vegetation types are mainly related to differences in bedrock. These differences are expressed in soil nutrient contents and soil physical properties that determine depth and permeability to water [31, 58].

According to tree canopy cover [59], the savanna formations of the Paranã River Basin vary from open (5 – 20% cover) to typical (20 – 50%), and from typical to arboreal or woodland (50 – 70%). These formations are found on acid soils that are relatively poor in nutrients. Open savannas are frequently associated with the mountainous areas and slopes on the eastern border of the Basin, the mountains and escarpments of the southern and eastern border, and the small hills of the northern border. Typical and woodland savannas are mainly in the relatively flat areas of the north and southern regions at low elevations. Typical savannas are also dominant in the mountainous regions of the western border and the plateau, escarpments and pediplains of the eastern border. On outcrops of sandstones or granite, that form the top of the mountains in the highest elevations of the western border (over 1011m high), a particular type of savanna – montane savanna – is also found.

Grasslands are mainly associated with shallow soils with high water tables at the top of the mountains of the western border and the sand soils of the pediplains on the eastern border.

Deciduous and semi-deciduous forests mainly constitute the dry forests of the Basin. Originally, dry deciduous forests occupied about 20% of the entire Basin [58]. These forests are one of the last remnants in Brazil [31, 43] and one of the most endangered tropical ecosystems in the world [44]. Within the Paranã River Basin, deciduous forest formations are mainly concentrated in the central and eastern portions. These forests are also found interspersed within areas dominated by savanna formations in the northern and southern regions of the Basin, in the foothills of the mountain range of the eastern border and along of the escarpments of the eastern border. Forest formations are always associated with relatively basic and fertile soils and limestone outcrops. Semi-deciduous forests are mainly associated with ravines in the mountains and hills of the Basin, run-on areas receiving moisture from surrounding terrain, and along the rivers and small water courses.

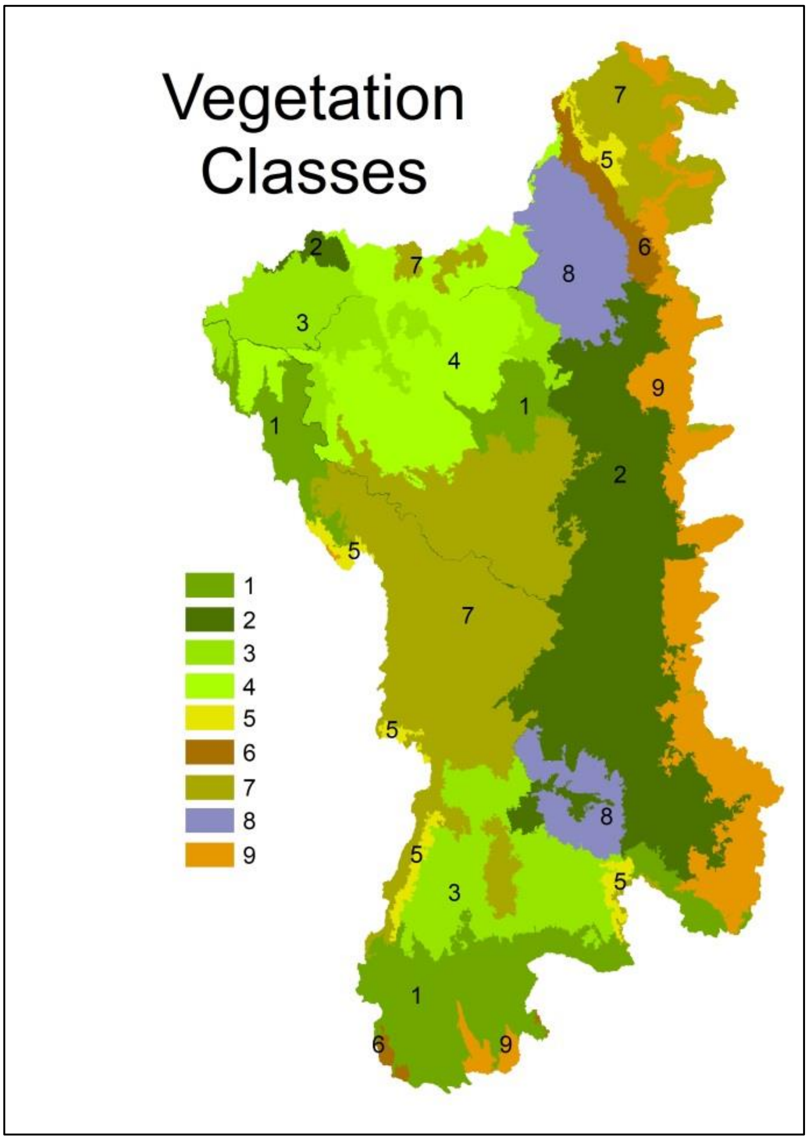


Figure 1.4. Spatial distribution of vegetation classes across the Paranã River Basin [adapted from 60].
1 = Savannas mixed with semi-deciduous forests. 2 = Deciduous forests. 3 = Savannas mixed with deciduous forests. 4 = Savannas mixed with grasslands and deciduous and semi-deciduous forests. 5 = Grasslands. 6 = Grasslands and semi-deciduous forests. 7 = Savannas and grasslands. 8 = Savannas in seasonally flooded lands. 9 = Savannas mixed with grasslands.

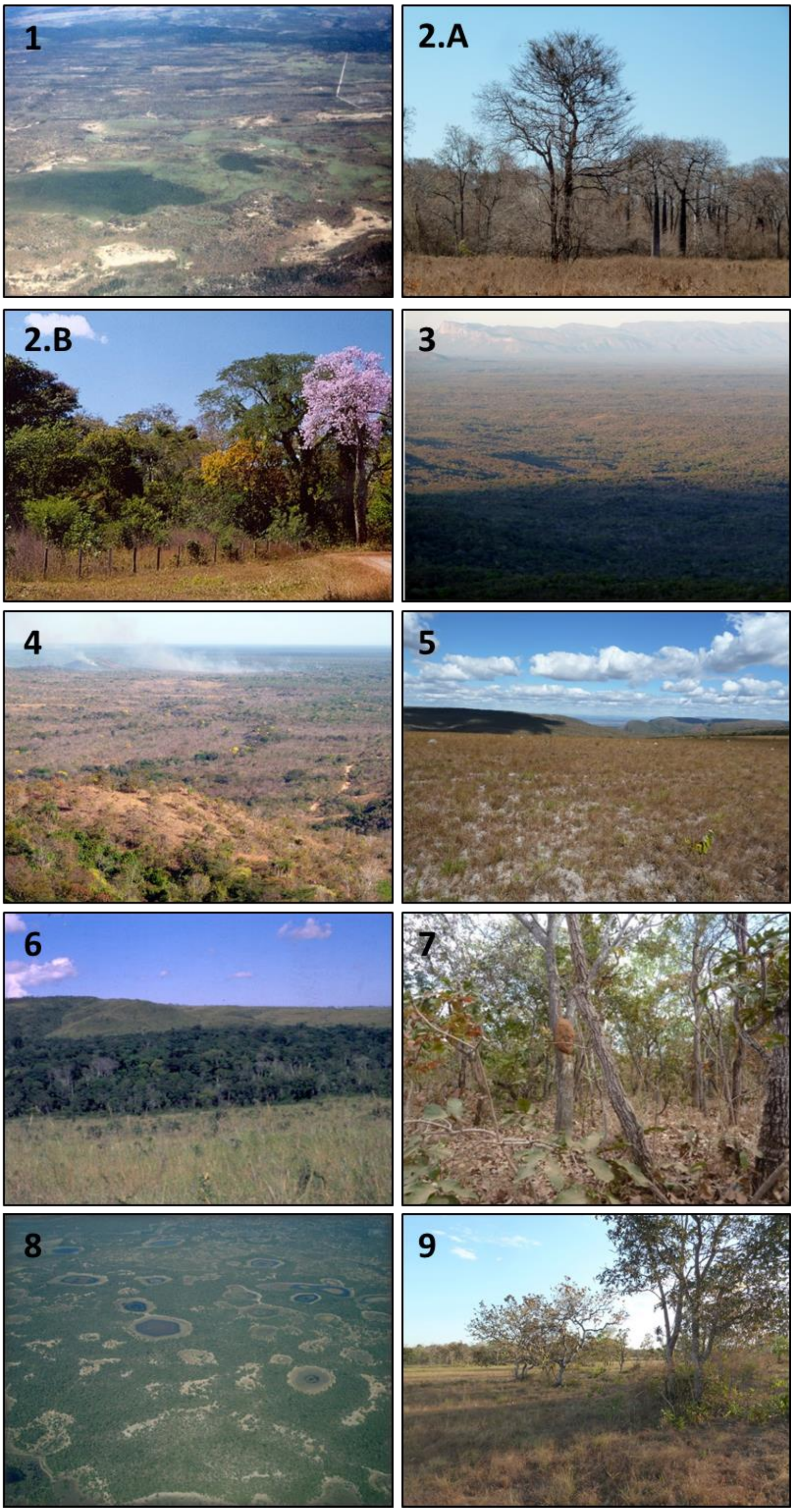


Plate 1.3. Vegetation classes across the Paranã River Basin.
1 = Savannas mixed with semi-deciduous forests. 2 = Deciduous forests in dry (A) and wet (B) season. 3 = Savannas mixed with deciduous forests. 4 = Savannas mixed with grasslands and deciduous and semi-deciduous forests. 5 = Grasslands. 6 = Grasslands and semi-deciduous forests. 7 = Savannas. 8 = Savannas in seasonally flooded lands. 9 = Savannas mixed with grasslands.

1.3.1.3. Conservation and knowledge gaps of biodiversity

Refining area prioritization for the Paranã River Basin poses several challenges. Prior to this study, no information about Basin's biodiversity was systematically compiled and available for direct use in conservation plans. Estimates of the total number of species already recorded within the Basin, including endemic and threatened species and trends in biodiversity loss, were completed unknown. Consequently, it was also unclear how species were distributed and arranged in different communities according to environmental gradients. This situation required a major effort in mining, collating and organizing information from disparate sources (herbaria, museums, literature and personal and online databases).

Just as collating biodiversity information in data-deficient regions is challenging, modelling the spatial distributions of species as a basis for area prioritization is still incipient in Brazil. Few studies have been done, and those are for single biological groups or significant species (threatened, rare or endemic) and generally at a national scale [35, 61-66]. These studies are difficult to use in planning for local or regional contexts, for which refined information is required. Furthermore, few previous studies have considered the possible impact of climate change on biodiversity, and none have anticipated these impacts specifically in the Paranã River Basin [67-71].

In previous studies, a myriad of methods and tools have been used to identify biodiversity patterns to support prioritization of areas for conservation. However, multi-taxonomic and community approaches remain important gaps in understanding biodiversity and prioritizing areas. The lack of a combined approach (species richness and community composition) constrains understanding of the differences and complementarities between species richness and species assemblages, how different biological groups respond to environmental variation, and whether some groups can be used as proxies for others in conservation planning.

The development of this thesis represents a first effort using systematic conservation planning considering a multi-taxonomic approach in the light of climate change for Brazil. Identifying and pointing out which and why such areas are priorities for conservation is likely to influence future government conservation policy and management decisions by stakeholders in the Paranã River Basin and other parts of the country.

1.1.3. Objectives

Given the importance of the Paranã River Basin for global biodiversity conservation and stakeholders' demands to define which areas should be prioritized and why those areas should be prioritized for conservation, the overall aim of this thesis is to define explicit criteria to determine which areas are important for biodiversity conservation within the Paranã River Basin, spatially identify these priority areas and clearly present arguments defining why those areas should be prioritized.

Answering those questions – which areas and why – is not a trivial task. For most regions of Brazil, the ideal information on biological occurrence and distribution is not known and/or not available for direct use in any process of conservation planning. The same is true for environmental variables that are determinants of species occurrence and distributions. All this information must be gathered from disparate sources or produced, processed and refined before it can be used. This process frequently demands years of work gathering sufficient information to proceed with area prioritization while simultaneously these natural areas are constantly being converted by people into modified ecosystems. The rapid land-cover change poses an additional challenge for biodiversity conservation in developing countries, that is, to produce the best solution with the information available for making informed conservation decisions in as short a time frame as possible. Therefore, the aims of this thesis are to:

a) Identify species occurrences within the Basin and gaps in biological information, and develop protocols to assure a reliable database for conservation planning.
b) Determine patterns of species richness and assemblage distribution to assist the identification of priority areas for conservation.
c) Evaluate different solutions in area prioritization, considering distributions of species and assemblages, richness, threats and endemism of species, complementarity, irreplaceability and vulnerability of habitats.

d) Evaluate likely future species movements and habitat resilience to identify priority areas for conservation considering the impacts of climate change on biodiversity.

1.1.4. General methods

This thesis is divided into four distinct but complementary studies (presented in chapters 2 to 5), namely:

I. Digging deep for biodiversity information in data-deficient areas: the case of the Paranã River Basin within a global biodiversity hotspot.
II. Biodiversity patterns within Paranã River Basin: what we can learn from species and community-level distributions.
III. Conservation in a hotspot: a multi-solution to conservation prioritization.
IV. Conservation in a hotspot: planning for a changing world.

In Study I, we were interested in knowing the “what” and “where” of biodiversity in the Paranã River Basin. Study II addressed biodiversity patterns and presents how species and communities are distributed in the Basin. In Study III, the relative effectiveness of species and community-level analyses in incorporating the maximum representation of biodiversity features into a reserve network is evaluated. Finally, in Study IV, we present a spatially explicit area prioritization for biodiversity conservation of the Basin considering the impacts of climate change on multiple taxa simultaneously.

Chapter 2

Chapter 2 – Digging deep for biodiversity information in data-deficient areas: the case of the Paranã River Basin within a global biodiversity hotspot

---

Keywords: biodiversity assessment, Cerrado hotspot, endangered species, RecordLinkage

Article type: Full-length article

Anderson C. Sevilha[1,2], Robert L. Pressey[3], Helga C. Wiederhecker[1,5], Jeremy VanDerWal[1,4], Stephen E. Williams[1], Guarino R. Colli[5], Jader Marinho-Filho[5], Miguel Â. Marini[5], Reginaldo Constantino[5], Rosana Tidon[6], Aldicir Scariot[2], Bruno M. T. Walter[2], Taciana B. Cavalcanti[2], Ivone R. Diniz[5], Laurie J. Vitt[7] and Janalee P. Caldwell[7].

[1] Centre for Tropical Biodiversity and Climate Change, School of Marine and Tropical Biology, James Cook University, Townsville, QLD, Australia
[2] Embrapa Genetic Resources & Biotechnology, Brasília, DF, Brazil
[3] Australian Research Council Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, QLD, Australia
[4] eResearch Centre, James Cook University, Townsville, QLD, Australia
[5] Department of Zoology, University of Brasília, Brasília, DF, Brazil
[6] Department of Genetic and Morphology, University of Brasília, Brasília, DF, Brazil
[7] Sam Noble Museum & Department of Biology, University of Oklahoma, Norman, OK, USA

Chapter 2 – Digging deep for biodiversity information in data-deficient areas: the case of the Paranã River Basin within a global biodiversity hotspot

2.1. Abstract

Planning for effective conservation begins with precise information about species occurrences and accurate species identifications. The Paranã River Basin is classified as data-deficient for biodiversity, although it is a key priority area for conservation within a global biodiversity hotspot, the Cerrado biome in Brazil. To compile an initial database for conservation planning in the Paranã River Basin, we: (a) surveyed plants, birds, mammals, reptiles, amphibians, termites, drosophilids, and saturnids; and (b) exhaustively searched for biodiversity records in museums, herbaria, literature, and online databases. We standardized taxonomic names and data formats compiled from diverse and disparate sources and applied an innovative method to automate identification of duplicate records. This approach allowed us to produce the best possible database to support conservation planning in the Basin. We analyzed geographical and taxonomic gaps in the database, characterized species in terms of threats and endemism and estimated the total number of species in major biological groups, considering potential additions from future sampling efforts. Of ~140,000 records initially compiled, only ~20,000 were retained after they had been checked for duplication, synonymies, misspelling, and other incongruities. Despite representing less than 3% of the area of the total Cerrado biome, the 5,130 species currently known from the Paranã River Basin overlap with between 35% and 54% of the biome’s terrestrial vertebrate species, 34% of the biome’s plant species, and 11–43% of the biome’s endemics. Even with 200 years of data collection, spatial information gaps still exist for all taxa in the Basin. Filling gaps for terrestrial vertebrates, plants, and insects would increase the total number of species sampled for these groups by 18–72%. Despite the data gaps, the biological information accumulated to date is sufficient to generate models of species distributions and community composition to guide fine-resolution conservation planning for the entire Basin. The Basin contains one of the most endangered ecosystems in the world, the seasonally dry tropical forest, where 70 threatened species and 927 Cerrado endemics are under increasing pressure from anthropogenic activities. Although our data and analyses are specific to the Paranã

Basin in the Brazilian Cerrado, we identify the myriad of problems faced when compiling basic information necessary to make conservation decisions. More importantly, we suggest methods that can be applied to make predictions of biodiversity in poorly known regions, especially those that are subject to rapid anthropogenic development.

2.2. Introduction

Accurately identified and spatially precise data on species are central to research and practice in conservation [72]. Several countries have responded to this need by creating spatially explicit databases on biodiversity across regional or national extents to support conservation initiatives. In Brazil, databases are not well developed, making the process of gathering biodiversity information for conservation planning a difficult task. Organizing biological data dispersed in museums, herbaria, literature, and some online databases is invaluable for conservation purposes, but this information needs to be integrated, refined, and checked for consistency before use; this task is not trivial.

The Brazilian government has invested significant resources to develop conservation plans and priorities at the scale of biomes [29]. However, implementation of conservation action demands that broad priorities are downscaled to resolutions matching those of units of management on the ground [73], and this process requires that spatial information for all aspects of planning, particularly biological data, are available at fine spatial resolutions. In the Cerrado biome—a global biodiversity hotspot [30]—the original 431 broadly delimited areas previously highlighted as priorities for conservation [29] are too large and roughly delineated to be translated directly into actions.

Within the Cerrado hotspot, the Paranã River Basin has long been recognized as a priority for conservation because of the uniqueness of its fauna, flora, vegetation, and land formations, high numbers of rare, endemic, and threatened species, and its significance as a centre of diversity for several biological groups [31-41]. However, estimates of total number of species already recorded within the Basin, including those of higher conservation value such as endemic and threatened species as well as trends in biodiversity loss, are completely unknown. Additionally, the co-occurrence of species from biomes adjacent to the Cerrado, including the Amazon, Caatinga, and Atlantic Forest [31, 33, 35, 43, 74], denotes the possible importance of this region as having current and historical refugial areas [75]. Furthermore, about 20% of the basin is seasonally dry tropical forest, the most threatened tropical ecosystem in the world due to its deforestation and fragmentation for agriculture and grazing [44]. These assessments confirm the Paranã River Basin as a region of outstanding global importance for conservation.

As part of the broad delineation of conservation priorities in the Cerrado, approximately 82% of the Paranã River Basin's ~60,000 km^2 was identified as a national priority for conservation action that included the creation of conservation units, biodiversity inventories, formation of ecological corridors, restoration, and promotion of sustainable use [29]. However, less than 1% of the entire basin is within fully protected areas. If protected areas are to be adequately enlarged and effectively located and configured in the Paranã River Basin, decisions about priorities for local areas will be needed within a regional context to achieve emergent properties of conservation systems such as complementarity and connectivity [10]. No currently identified conservation priorities for the Basin approach the fine resolution needed for effective action.

Despite the global significance of the Paranã River Basin, no fully protected area has been created there since 1989. Clearly, the Basin is a prime candidate for detailed conservation planning aimed at effective implementation of actions to specific local areas. Also obvious is the need to pull together the disparate sources of data on species as a basis for spatial modelling and identification of priority conservation areas. Before the study reported here, data on species were not in a form that could guide conservation planning: they were dispersed in public or private collections or databases or described in articles and reports that could not be easily accessed.

We describe an exhaustive and innovative process for compiling biodiversity data from diverse and disparate sources to produce the best possible database for conservation planning at a detailed scale using the Paranã River Basin as a model. We developed an approach to automatically link data sets and remove duplicate records. With the database compiled, we identified the full scope of biodiversity knowledge for our study region, geographic and taxonomic gaps in records, and levels of threat and endemism. This paper is the first stage of our work on conservation planning in the Basin. But our intent is that the compiled data and protocols for detecting and eliminating duplicates and continually updating information will be useful to all environmental decision-makers with interests not only in the Basin and surrounding parts of the Cerrado biome, but in other regions of the world.

2.3. Methods

2.3.1. Data collection

Biological data were gathered from primary and secondary sources. Secondary sources consisted of ad hoc information gathered from herbaria, museums, scientific collections, literature, and online species databases (see Supporting Information 1.1 for a complete list of data sources). Herbaria, museums, and scientific collections were chosen based on the size and representation of their collections and/or previous knowledge of material from the region. When data were not available online, visits to institutions were arranged and the collections searched directly.

Primary sources consisted of biological inventories designed and undertaken specifically for this study to fill spatial gaps in biological knowledge for the Paranã River Basin where biodiversity data were anecdotal or nonexistent and to target poorly known habitats or communities [28]. Our systematic inventories started in 1999 with the establishment of 375 permanent plots of 20 x 20 m, equally distributed between 15 sites (25 plots per site). Three sites were in savannas and 12 in dry forests under different regimes of anthropogenic disturbance, from heavily disturbed to intact. In these plots, annually, we evaluated temporal turnover of species and individual seedlings and saplings in savanna and dry forest formations and, every five years, measured the temporal turnover of tree species and individuals [31, 39]. Later, between 2003 and 2004, we extended the sampling to biological inventories of eight groups (plants, mammals, birds, amphibians, reptiles, termites, moths, and flies) in 110 sites distributed in four locations (26, 15, 42 and 27 sites, respectively – Figure 2.1), including the sites already established to sample plant populations and community dynamics.

Our biological inventories were based on a Rapid Assessment Program [76], carried out during two seasons (dry and wet), and encompassing the geographic diversity of the

Basin and its diversity of habitats and levels of disturbance. The inventories consisted of floristic and phytosociological surveys of plants; manual searching and pit-falls for reptiles and amphibians; nets, pit-falls, and tracking for mammals; nets and direct and acoustic observations of birds; fermented bait traps for drosophilids; manual searching in soils, vegetation, and litter for termites; and light traps for saturnids. In 80 days of effective field work we recorded 52,599 individuals of 1,352 species, many of them endemic to the Cerrado and endangered, as well as many species new to science (see Supporting Information 1.1).

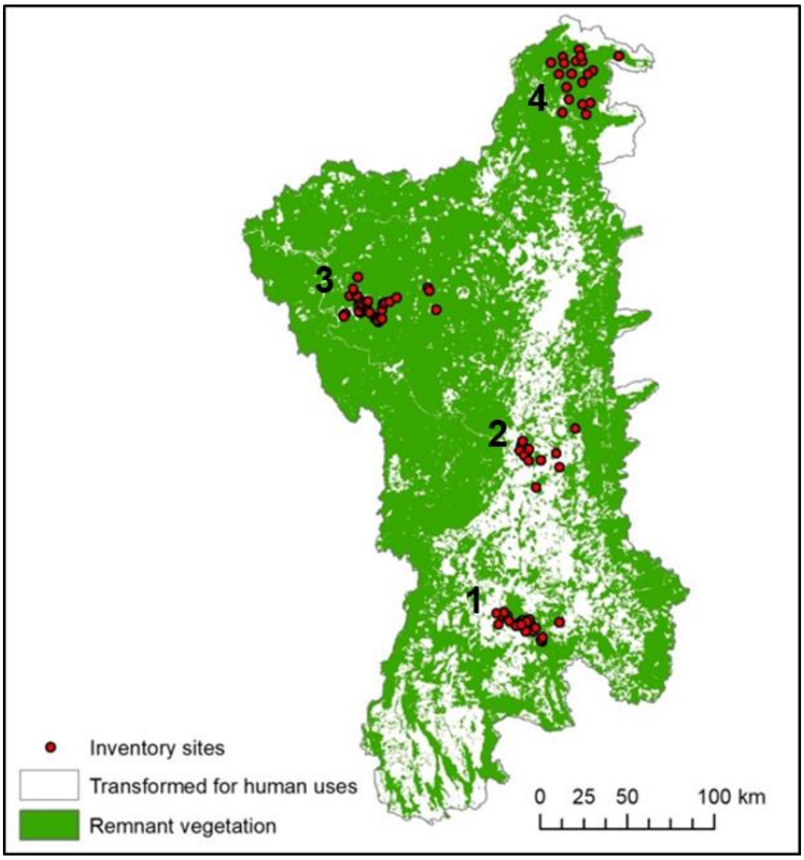


Figure 2.1. Remnant native vegetation [77] and locations of our biological inventories in the Paranã River Basin.

Permits and approvals for sampling, collecting and sacrificing animals specimens during field work on private land for each location were obtained from Brazilian Institute of Environment and Renewable Natural Resources – IBAMA (Licence Number:

007/2003-CGFAU/LIC) and with the consent of the landowners, respectively. According to the Brazilian Law, it is not required to have a written agreement with landowners to sample fauna and flora on private lands. In our case, future contact for collecting animals or plants must be done with ICMBio (the new Brazilian environmental agency that is responsible for giving permits), and future contact with landowners must be done in situ, since the localities sampled are in remote areas without telephone, mail or email access. A waiver to animal sacrifice for the purpose of this research was guaranteed by IBAMA Licence (Licence Number: 007/2003-CGFAU/LIC) since field work was conducted prior – before 2004 – to animal ethical approvals be requested by Brazilian legislation for studies not involving animal experimentation – after 2008. However, we performed all procedures manipulating and sacrificing animals according to international standard protocols and the subsequent (a) Brazilian Law [Law 11794/2008 (Ordinary Law) 08/10/2008 - http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/lei/l11794.htm] and (b) Resolution 301/2012 and Ordinance number 148/2012 of the Brazilian Biology Federal Council (http://www.cfbio.gov.br/resolucoes-cfbio/145-resolucao-no-301-de-8-de-dezembro-de-2012). The procedures adopted while in field work avoided or minimized animal pain and distress during their manipulation and ensured their wellbeing during care and management. Animal specimens were sacrificed only when they represented unknown taxa or a new record of the occurrence for the Cerrado biome or for the study region. A total of 415 individuals of 142 species were sacrificed. All animals sacrificed were humanely killed via the injection of a lethal dose of sodium pentobarbital performed by trained and qualified personal (Dr. Guarino R. Colli sacrificed herpetofauna species and Santos Fernandes Balbino and Joaquim Ribeiro da Silva – both Laboratory Technicians from Zoology Department of the University of Brasília – sacrificed birds and mammals species). No samples (blood or tissues) were taken from animals as part of this study. Zoological specimens were deposited in the following scientific collections, all hosted at the University of Brasilia: Coleção de Mamíferos (mammals), Coleção de Aves (birds), Coleção Herpetológica – CHUNB (reptiles and amphibians), Coleção de Cupins (termites), Laboratório de Biologia Evolutiva (fruit flies) and Coleção Entomológica (saturnids). Vouchers of plant specimens collected were deposited for reference at the Herbarium of Embrapa Genetic Resources and Biotechnology (Herbarium CEN).

We combined biological records from primary and secondary sources. We considered species as the basic units of analysis and filtered the original data set to keep only records with binomial species names. We therefore grouped varieties and subspecies within species, except for threatened taxa consistently identified in the data set at infra-specific levels. Records without coordinates or having suspicious coordinates (e.g. georeference inversions and records located outside of the grid domain) were discarded unless there was sufficient location description to correct errors with confidence.

2.3.2. Data preparation

Due to multiple data sources, some records were repeatedly obtained, meaning that the same observation or collecting event was reported by different databases, inflating the apparent number of records. Repeated records are not always identical because sources can organize data slightly differently. For example locality names, number format, misspellings can lead to duplication. In this context, and with thousands of species and many thousands of records, identifying duplicates is not trivial. We removed duplicate records using the RecordLinkage package in R [78]. Originally developed to determine duplicates in data stemming from a German cancer registry, this package uses stochastic and machine-learning methods to detect duplicates and linkages between datasets [79]. The software automates screening of duplicates by comparing all possible pairs of records in the dataset. Then each record pair is compared column by column and is transformed into comparison pattern (e.g. 0 0 0 1; where a match is 1 and zero indicates no match) and agreement and non-agreement of attributes is evaluated to determine whether the pair contains a duplicate.

To identify duplicates we considered eleven attributes: species binomial, geographic coordinates, collection date, collector’s name, species specialist identifier, date of species identification, geographic location, municipality of occurrence, collection code, name of the collection, and source of information. Prior to analysis with RecordLinkage, we standardized these attributes for each record. Standardization included spelling checks and removal of abbreviations in text fields, such as scientific names, municipality, and names of collections and collectors, to ensure consistency and to facilitate identification of duplicates (see Supporting Information 1.2 for a complementary description of data-attributes standardization process).

Within the RecordLinkage package, we used the function *RLBigDataDedup* to compare our standardized attributes and identify duplicates. This function returns comparison patterns between the attributes chosen as component pairs of the records. During this analysis, we applied a string comparator to identify and correct typographical errors. Then we used a blocking argument to compare records with equal values for a subset of attributes. The expectation maximization algorithm (EM algorithm) [80, 81] was used to estimate the weight resulting from the number of matched pairs of attributes and, consequently, to discern between match and non-match records. A weight value between a pair of records equal or close to one indicated a perfect duplicate or a high probability of duplication. A weight equal to or close to zero indicated a non-duplicate or a low probability of duplication. After inspecting the matching patterns between records, we identified a threshold weight of 0.7 as sufficient to identify possible duplicates. Values below this threshold were kept in our compiled dataset. For all pairs with weights above this threshold, we evaluated and decided for each case whether one of the records was redundant and should be removed.

Species synonymies were checked by taxonomic specialists using recent literature and online taxonomic indexes to assign the correct available name to each record (see Supporting Information 1.1 for a complete list of references used). Only valid names were retained for subsequent analyses, except for two cases. First, unresolved names of plants were retained (148 records of 32 species) if they were deemed valid taxonomically by appearing as individual entities in all source databases. Second, taxa identified only to genus based on the comprehensive new surveys for this study were retained if they represented new occurrences of genera for the region or new species, still unnamed by specialists.

All species in the vetted database were subsequently evaluated and classified according to their conservation importance in terms of threat status and endemism. For threat status, we used the Red List of threatened species from IUCN [82] and the lists of threatened species of Brazilian Fauna and Flora [83, 84]. To identify species endemic to the Cerrado that are present in the Paranã River Basin, we drew on recent literature that describes endemic species for the entire Cerrado biome [35, 62, 64, 74, 85-87]. The filtered dataset is available in Supporting Information 1.1 and will be kept and regularly updated in a web repository at http://eresearch.jcu.edu.au/tdh/datasets/parana_river_basin/species_occurrence/species_occurrence.xlsx.

For threatened species, we also recognized records within a one-degree buffer (~100 km) of the boundary of the Paranã River Basin (within a polygon formed by adding one degree to the latitudinal and longitudinal extremes of the basin). Although this procedure is by no means an accurate indication of extra-Basin occurrences, presence of species in the buffer indicates opportunities for conservation of the Basin's taxa additional to areas within the Basin.

2.3.3. Biodiversity patterns

To understand spatiotemporal patterns in species records, we plotted the number of surveys and species observed chronologically. To identify spatial and environmental gaps in collecting and systematic surveys, we plotted biodiversity records across elevational gradients and vegetation. We clipped the digital elevation model from SRTM data version 4.1 [55], applying Jenk's natural breaks in ArcGis 10.0 to generate elevation classes. The map of broad vegetation classes in the Basin, based on proportional occurrence of vegetation types, was derived from the map of natural vegetation cover in the Cerrado [60], allocating native vegetation remaining in 2009 [77, 88] to classes. To identify original vegetation types in presently transformed areas of this map, we used the vegetation maps of the Brazilian Army [89] and 2,664 field observations. At a resolution of 1:250,000, vegetation cover in the Army's maps was based on radar images taken in 1965, when natural vegetation of the Paranã River Basin was almost complete. The 2,664 field observations consisted of landscape descriptions at points placed across the Basin to represent most of its environmental diversity. Information recorded at these points included remnant vegetation or elements in the landscape (e.g. species of trees that remained on largely cleared pastures) that would give clues regarding previous vegetation types. The map of Cerrado remnants [77] was used to identify the parts of the landscape already converted to human uses.

To understand how additional sampling effort might contribute to increases in future numbers of species in different taxonomic groups, we produced species accumulation curves using EstimateS 9.0 software [90]. We defined each sampling event (hereafter "sample") as the list of species recorded on a particular date (defined by day, month, and year) in a particular locality within the Paranã River Basin. A sampling event could consist of one or more species observed one or more times in the specified day. We used bootstrap and jackknife2 richness estimators to correct the downward-biased number of observed species [91], that is, to reveal rare species not sampled because of the limited sampling effort. The two estimators differ in the way they assess the number of rare species not recorded in each simulated set of samples. Jackknife2 relies on the frequency of rare species (2 or less occurrences), whereas bootstrap uses the frequency of all species to estimate the total number of species that would be expected considering the samples taken. We obtained observed and estimated (using jackknife2 and bootstrap) species accumulation curves by plotting the mean and 95% confidence intervals from 1,000 random resamplings of species lists with replacement.

2.4. Results

Raw data gathered from different sources included 140,193 records. After cleaning the dataset, the total number of records was reduced to 20,415. Remaining valid records were obtained from 2,894 sites and 3,647 sampling events across the Basin. Sources of the cleaned data were: online databases (45.3% of records), our field inventories (26.3%), scientific collections (16.2%), and the literature (12.2%). Primary causes of error, and removal of more than 85% of records in the raw data, were duplicates, typographic errors, taxonomic synonymy, and lack of information on species names. Most of these errors were found in online databases. After cleaning the database, the total number of species decreased from 9,767 to 5,130 distributed among 21 biological groups (Table 2.1).

Apart from recently discovered species, for which conservation status is yet to be defined, the Paranã River Basin has at least 70 globally or nationally listed threatened species (Table 2.1). Of the threatened species recorded, 29 were found only within the Basin (and not within the one-degree buffer). Among these species were 17 plants, 7 birds, 4 mammals, and 1 reptile. Records within the Basin also included 927 species endemic to the Cerrado hotspot, comprising 872 plants, 23 amphibians, 13 birds, 11 reptiles and 8 mammals. Only 27 Cerrado endemics (25 plants, 1 bird, and 1 reptile) and 5 threatened species (3 plants and 2 birds) were found within the two fully protected conservation units in the Basin.

Table 2.1. Total species recorded in different biological groups across the Paranã River Basin, after data cleaning.

Target Taxa	Number of occurrences	Number of species	New species[a]	Cerrado endemic species	Threatened species[b]		
					CR	EN	VU
Plants							
Algae	372	216	1				
Vascular plants	15056	3467	185	872	22	3	15
Bacteria	1	1					
Fungi	24	23					
Excavates	3	2					
Animals							
Rotifera	9	9					
Arthropoda							
Arachnida	3	2					
Coleoptera	1	1					
Crustacea	25	19					
Diptera							
Drosophilid flies	183	40					
Phlebotomine sand flies	7	7	3				
Simuliidade black flies	19	16					
Hemiptera	2	1	1				
Hymenoptera	44	33	2				
Saturnid moths	364	291	32				
Termites	122	66	2				
Vertebrates							
Amphibians	206	63	2	23			
Birds	2945	430		13	9	3	4
Fishes	411	258	13				
Mammals	358	101		8	6	1	5
Reptiles	260	86	4	11	1		1
Total	20415	5130	243	927	38	7	25

[a] Species previously unknown to science and described using material collected in the Basin

[b] CR = Critically Endangered, EN = Endangered, and VU = Vulnerable, according to the IUCN classification of threatened species for national or global threat listings.

The first biological record for the Paranã River Basin dates from 1821. Following that, two conspicuous peaks of data collection and three peaks of new species discoveries occurred approximately in the periods 1830s to 40s, 1960s to 80s, and 1999 to 2005 (Figure 2.2A). The effort applied in surveys across the Basin through the years resulted in the discovery of 245 species new to science, that is, species originally described or being described from material collected in the Basin (Figure 2.2B). The locations of records of new species were more abundant where there were more surveys, indicating the spatial congruence among sampling effort and new records. Investment in collection effort is still leading to the discovery of new species from our systematic surveys between 1999 and 2005 (final period in Figure 2.2B).

The number and distribution of records have increased markedly since 1955 and especially after 1986 (Figure 2.2A). Records were not uniformly distributed across the Basin. The northern and southwestern regions of the Basin remain poorly sampled and can be considered knowledge gaps where additional species are likely to be detected.

For most taxa, records were concentrated within the flat areas of the Basin that have a gradual elevation gradient from north (lower) to south (higher) of 240 to 775 m (Figure 2.3A). The least sampled elevation class, from 776 to 1010 m, occurs around the margins of the Basin. All groups and particularly terrestrial vertebrates and insects, had fewer records in this class proportionally. In contrast, the highlands above 1011 m, despite being the least extensive elevation class, were relatively well sampled for all taxa. Mainly located within Chapada dos Veadeiros National Park, these high areas were intensively sampled in the past and are continually resampled to address the obligations for management planning of the Park, making it one of the best-known portions of the Basin and the Cerrado biome. From the total of 20,415 records, ~18% were obtained in this highest altitude class. The extent of the bias in sampling of the highest elevations can be illustrated with estimates of additional records needed to provide the same number of species records per unit area in the other elevation classes. Of a total of more than 82,000 extra records needed, 24,165 would be in the 240-430 m class, 27,803 in 431-585m, 17,535 in 586-775 m, and 12,986 in 776-1010 m.

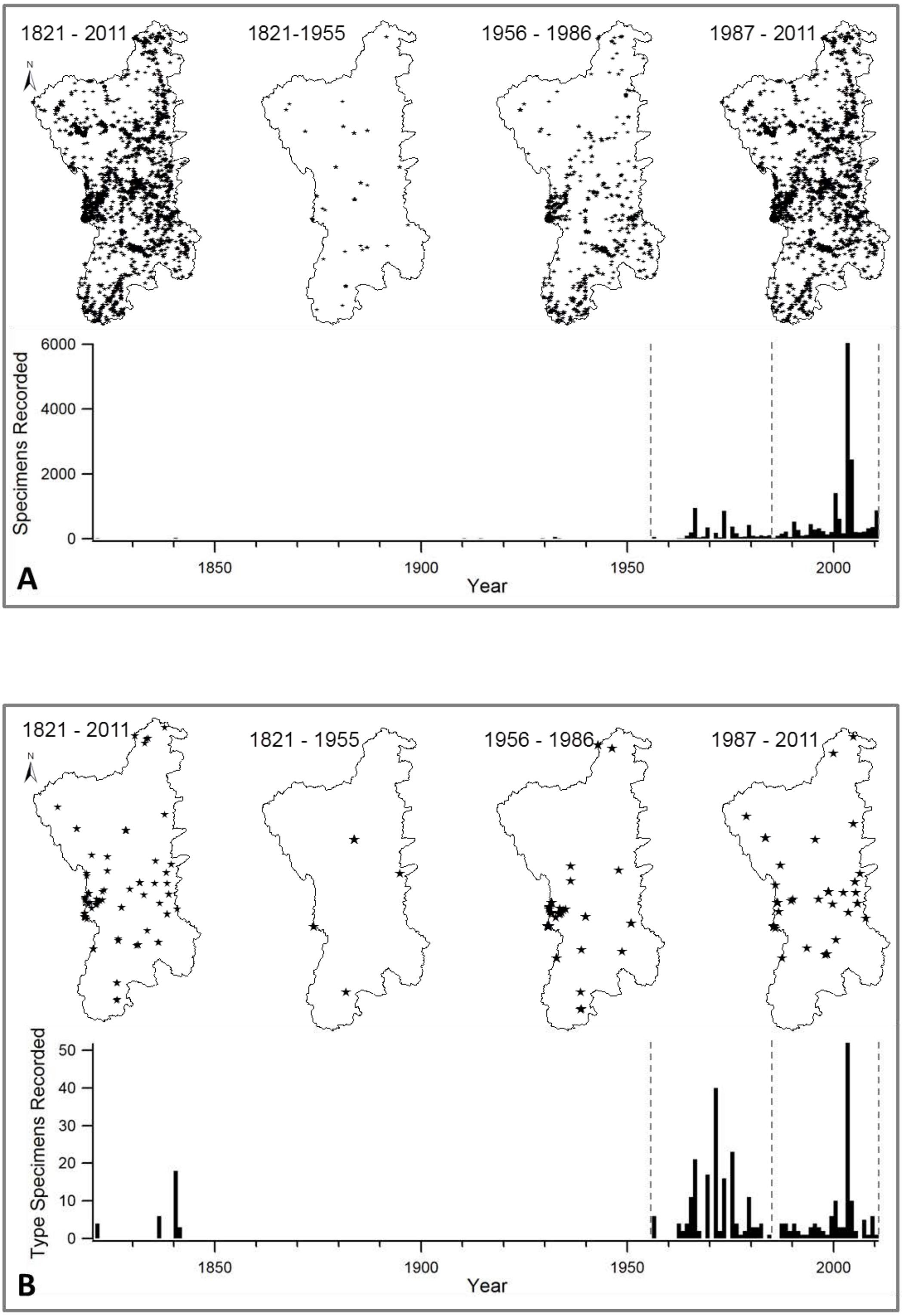


Figure 2.2. Chronological collection of data on species and their distributions within the Paranã River Basin from 1821 to 2011. A. All specimen records, B. All type specimens recorded, i.e. species previously unknown to science described using material collected in the Basin.
The left-hand map in each part of the figure shows the spatial distribution of all records. The other three maps in each figure show records collected during specific periods, delimited by vertical dashed lines in the frequency graphs.

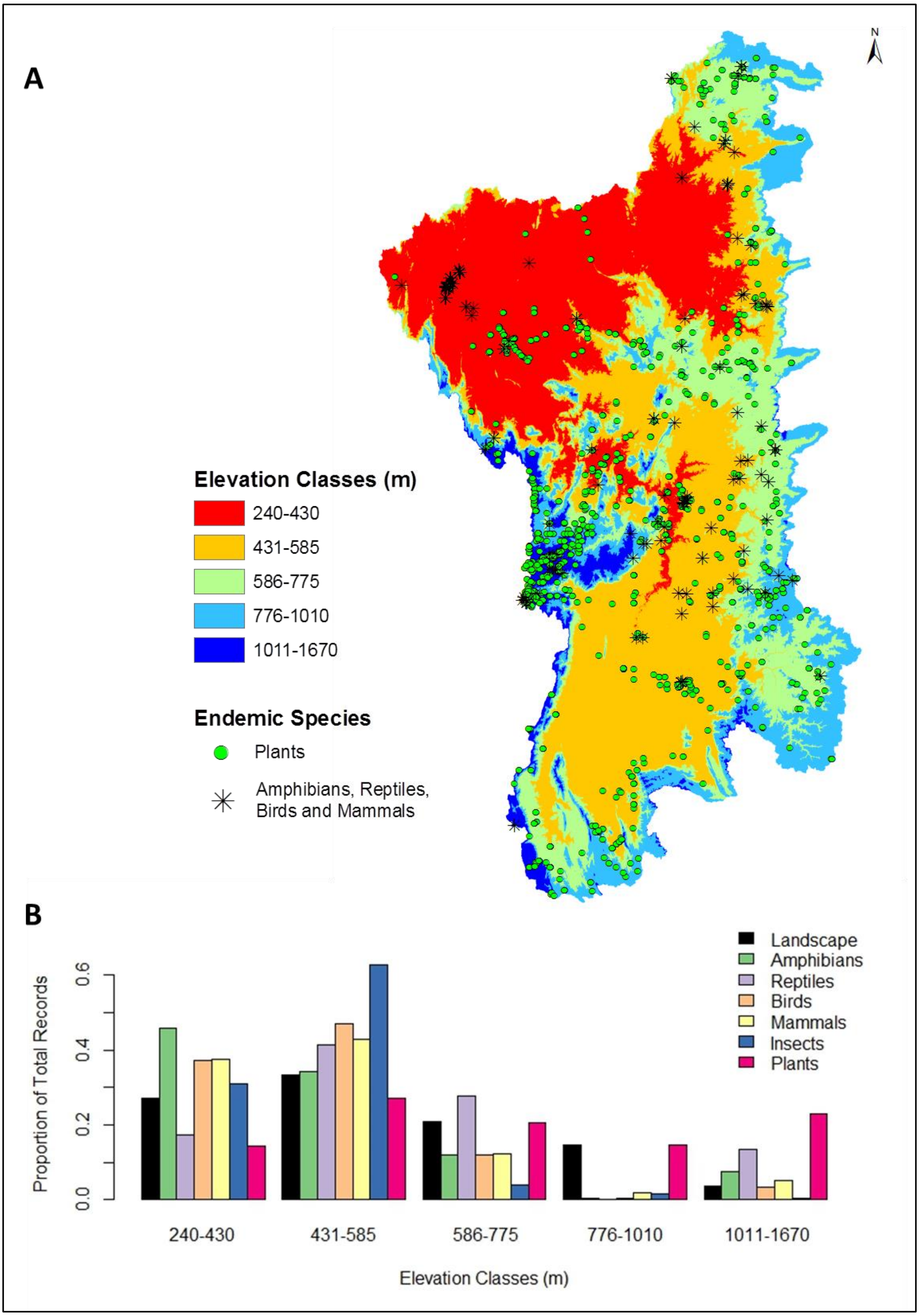


Figure 2.3. Records in relation to elevation within the Paranã River Basin. A. Spatial distribution of elevations and records of Cerrado endemic species. B. Proportions of records of each major taxon across elevation classes.

Each colored (not black) bar indicates the proportion of all records of each taxon in the Basin occurring within each elevation class. Landscape bars (black) refer to the proportion of the Basin's total extent occupied by the elevation class.

Records were also unevenly distributed across vegetation classes (Figure 2.4B). The deciduous forests and savannas (classes 2 and 7 in Figure 2.4A) are predominant across the flat areas of the valley, and are consequently better known biologically. These two vegetation classes are also predominant across the entire Basin, occupying around 44% of the total extent (19.8% and 24.8%, respectively), when distribution across presently cleared areas is considered. About 57% of the samples came from these classes. The least sampled classes – savannas / deciduous forests (areas of savannas intercepted by deciduous forests - class 3) and savannas / semideciduous forests (areas of savannas intercepted by semideciduous forests - class 1) – cover about 24% of the Basin (12.4% and 11.9%, respectively) but have only about 10% of all records (3.1% and 7.5%, respectively).

Considering the threatened species officially recognized and sampled within the Basin (Table 2.1 and Supporting Information 1.1), deciduous forests and savannas (the prevalent vegetation classes) had the highest concentration of records, independently of the biological group considered (Figure 2.4A). For endemic species, records were concentrated mainly in the highlands (above 1010 m) and distributed more sparsely in the lower areas (Figure 2.3A). This pattern coincides with the higher proportional sampling effort in the highland national park and our inventories at lower elevations.

Species accumulation curves estimated by bootstrap and jackknife2 indicated that, for most major taxa, the species list for the Paranã River Basin is still incomplete (Figure 2.5). According to the estimators, additional sampling effort might contribute between 1,092 and 4,182 additional species to the current list. By major groups, the additional species numbers are 852–3,270 for plants, 56–186 for birds, 16–51 for mammals, 17–60 for reptiles, 13–43 for amphibians, and 134–554 for insects. Only the curves for birds appear to be approaching asymptotes. For the other groups, even for plants, which is the best-documented biological group within the Basin, the species accumulation curves are still rising markedly.

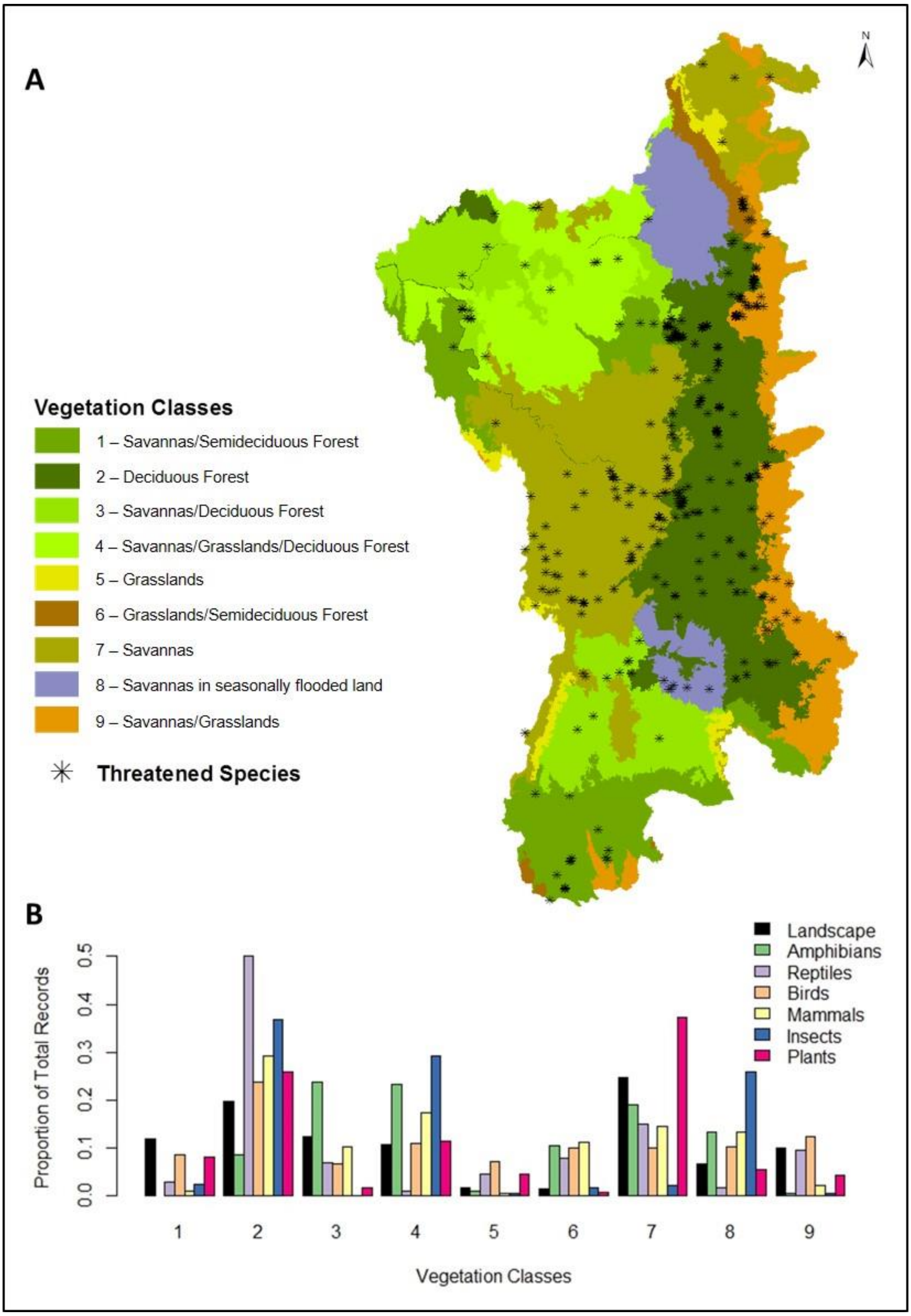


Figure 2.4. Records in relation to vegetation classes within the Paranã River Basin. A. Spatial distribution of vegetation and records of threatened species. B. Proportions of records of each major taxon across vegetation classes.
The vegetation map in A includes estimated cover of original vegetation classes (see Methods for details). Each colored (not black) bar indicates the proportion of all records of each taxon in the Basin occurring within each vegetation class. Landscape bars (black) refer to the proportion of the Basin's total extent occupied by the vegetation class.

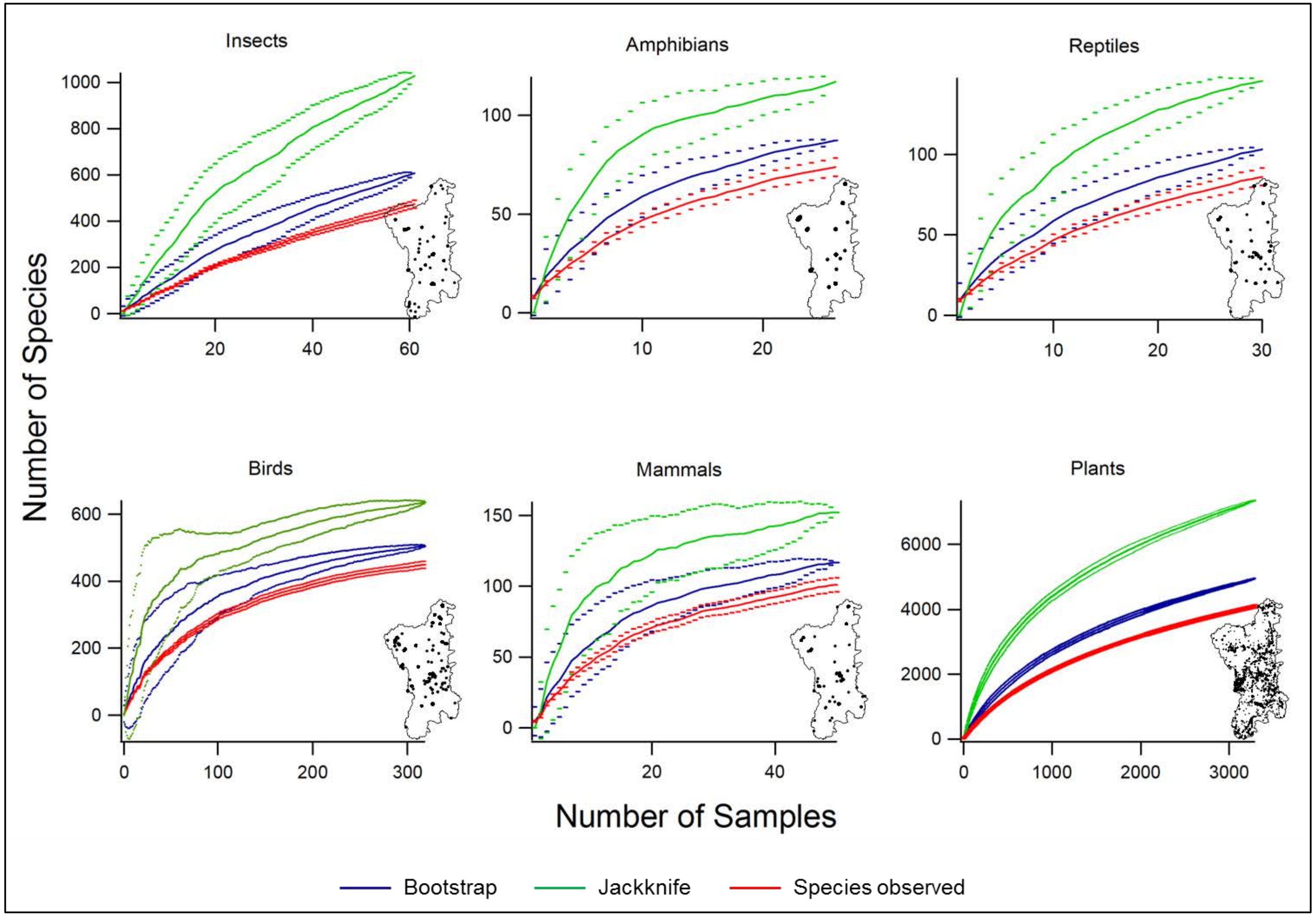


Figure 2.5. Species accumulation curves for the major biological groups sampled in the Paranã River Basin between 1821 and 2011. Dashed lines indicate the standard deviation of the average (continuous lines) richness observed or estimated by bootstrap and jackknife2, based on samples randomized 1,000 times with replacement. Number of samples refers to species records that occurred at a site per day.

2.5. Discussion

One of the constraints on applying scientific knowledge to guide conservation action is the lack of access to biodiversity information to inform decision makers [92]. In many cases, this information, based on the detection of species in the field, has already been generated. Collating existing biodiversity data is both challenging and time-consuming when information is dispersed and poorly organized. For our study on the Paranã River Basin, about 53% of final filtered records came directly from a recent biological inventory conducted by us and from biological collections and literature, while 47% were mainly scattered across online databases maintained by scientific collections or database compilers. Online information also represented the great majority of the raw data that was excluded as part of our data-cleaning project. Even with the help of RecordLinkage in detecting misspellings and data duplications, the elimination of about 85% of the raw data obtained shows that, unfortunately, the use of multiple sources of biological data for conservation planning must be followed by intensive exercises in organizing, standardizing, and filtering. For the Paranã River Basin, this work was essential to retain only high-quality data, and to eliminate data that would have been a distraction to conservation planning.

The chronology of biological records in the Paranã River Basin reveals peaks that illustrate the impact of isolated initiatives in producing information about biodiversity. These peaks closely match the occasional major efforts to collect biological information, with relatively little information added between them. The earliest peak corresponds to the first explorations of the interior of Brazil by European naturalists in the early 1800s [45, 46, 93, 94]. Despite the sparse collection points across the Basin at this time, these surveys revealed more than 30 species new to science. The second wave of collection followed the establishment of the University of Brasília in the country’s capital in 1962. Located only 50 km from the southern border of the Basin, the university’s opening attracted researchers from around the world. These researchers and their students were responsible for most of the specimen collections made around that time. The records following the early 1960s, however, were mainly concentrated along the few existing roads of the southern (close to the capital) and western (surrounding the Chapada dos Veadeiros National Park, created in 1961) sides of the Basin and so were limited in sampling biological diversity.

The third wave of collection occurred between 1999 and 2005 through our inventories. These inventories were systematically stratified to represent landscapes in which biological information was sparse or absent. Systematic sampling is a straightforward method and has the advantage of representing and providing comparability between environments, unlike ad hoc sampling. The recent inventories filled several of the previous gaps in biological information, mostly in the low-medium elevations of the Basin (below 775 m), in dry forests on flat lands, slopes, and limestone outcrops, in savannas on sandy, deep, and rocky soils, and in wetland savannas. The inventories also produced several new species records for the Paranã River Basin, the Cerrado biome, and to science. Nevertheless, these recently sampled areas are under high anthropogenic pressure, being targeted for agriculture, forestry, cattle farms, and limestone mining that have already converted around 30% of the entire Basin (see Figure 2.1C). Deforestation is concentrated on dry forest formations where few fragments remain, but savannas too are under threat, having been targeted recently for charcoal, sugar cane, and forestry with the expansion of *Eucalyptus* plantations.

Other types of vegetation on the elevational gradient from lowlands to highlands are less well explored biologically (savannas, grasslands, and semideciduous forests). The spatial distribution of all records shows that the savannas on lowlands of the northern and south-western portions of the Basin are still under-surveyed, as is the elevational class around the Basin’s rim with its distinctive types of savanna and forest formations. Future investments in biological inventories should prioritize these areas to fill gaps in information.

Despite the intensive effort in gathering data over almost 200 years, spatial information gaps remain for all taxa in the Paranã River Basin. Most species records are located near the roads that cross the Basin. Relatively few records are from less accessible areas. Roads are preferably established on flat areas of the tops of watersheds or on the bottoms of valleys to minimize costs in building and maintenance, resulting in a biased coverage of vegetation types and elevational classes. Sampling biases in the Basin are moderate for vegetation classes but strong for altitude classes. The very large number of additional records (about 82,000) needed to balance the concentration of records in the highest altitude class, highlights two requirements. First, extensive further inventories are required to fill taxonomic and elevational gaps. Second, models of the distributions of species and communities will be essential to make best use of information for conservation planning.

Regardless of representing less than 3% of the total Cerrado biome, the Basin contains large proportions of the biome's species. The Basin's 5,130 recorded species correspond to 35% of the Cerrado's amphibians, 54% of its birds, 53% of its mammals, 32% of its reptiles, 38% of its drosophilids, and 34% of its plants [41, 62, 64, 86, 95]. These records comprise 20% of the plants endemic to the Cerrado, 43% of endemic birds, 25% of endemic mammals, 21% of endemic amphibians, and 11% of endemic reptiles. Of the species sampled, 70 are included in official global or national threatened species lists, including 29 that are either absent or rare outside the Basin. The global conservation of these 29 species relies on conservation management within the Basin, mainly in dry forest formations that have been and still are targeted for agriculture, forestry, livestock and mining activities [31, 39, 50].

Among the lesser known groups in the Basin, fishes and insects stand out. Rivers crossing the karstic system in the dry forests of the Paranã River Basin are home to the richest and most abundant cave fish faunas in Brazil [36, 37]. Most of these fishes are endemic to the caves of Terra Ronca State Park and need to have their threat status assessed. A larger portion of this system remains unsurveyed and threatened by mining, likely putting at risk species yet unknown to science. Despite the recognized importance of insects in many ecological processes, such as pollination, predation, and decomposition, and their significance for conservation planning [41, 96], little is known about the composition and distribution of this group within the Basin. Our systematic inventories revealed significant numbers of new insect species or new species occurrences for the Cerrado biome. Independently of the groups sampled (i.e. termites, drosophilids, or saturnids), the numbers of species and new occurrences found during our field inventories were high when compared to other parts of the Cerrado sampled using similar methods [40, 41]. On the other hand, orders considered hyperdiverse in the Cerrado, such as Hemiptera and Coleoptera, are poorly represented in the Basin, indicating both the lack of surveys designed for these groups and easily accessible information for conservation planning.

Clearly, additional inventories of species are needed for this globally important region to narrow the Linnean shortfall of as-yet undescribed species and the Wallacean shortfall of under-sampled species [97]. Future biodiversity inventories cannot be restricted to the two tiny areas located in the south of the Basin recommended for surveys by the Brazilian Environmental Ministry [87] (see Figure 1.4). Sampling gaps for insects, mammals, amphibians, and reptiles are evident. Even for the best-surveyed groups—plants and birds—more species were expected to be sampled by both richness estimators (Figure 2.5). Performance of different estimators depends on sample size, sampling effort, and aggregation of species across samples [98, 99]. Generally, jackknife and bootstrap methods tend to converge as the total number of species is approached. Despite bootstrap and jackknife methods both having high accuracy in estimates of species richness, jackknife estimates seem to be less affected by grain size (sampling unit), providing a better picture of species richness patterns when incidence data (presence-only data) are used, compared to the bootstrap which is reported to underestimate species richness [99]. Except for insects, a poorly known group, estimates of the total species expected within the Basin seem to be reasonably reliable. Considering what is already known for the Cerrado biome (see numbers above), the curves for both methods are likely to be reasonable estimates of the total number of species in the Basin. To reveal rare and inconspicuous species, future inventories should cover areas and seasons not well surveyed previously. Although opportunistic sampling can provide valuable information for unsurveyed areas, systematic surveys should be prioritized. Systematic sampling allows the composition of sites to be compared in relation to species presence and absence, improving richness estimates. It also increases the likelihood of detecting ephemeral species by, for instance, sampling the same area in different seasons to improve detectability. Additionally, systematic surveys are designed to represent tenures, land uses, environments and biogeographical boundaries, adding records to poorly sampled parts of regions and improving the reliability of distribution modelling [100].

A much more complete system of fully protected areas within the Paranã River Basin is warranted. The present system encompasses only parts of the highlands (1011–1670 m) in the central west and limestone outcrops in middle elevations (586–775 m) of the central east (Figure 1.4). In total, less than 1% of the Basin’s area is under full protection, retaining merely 7% of the threatened and 3% of the Cerrado-endemic species already surveyed. While there is still limited information on smaller organisms such as insects, we have sufficient information, considering numbers of species and geographic representativeness, on vegetation and terrestrial vertebrates to support conservation planning. Species distribution models could fill many of the spatial gaps in data, thereby narrowing the Wallacean shortfall [97] to support conservation planning. Moreover, these models might be complemented by generalized dissimilarity modelling to analyse and predict spatial patterns of turnover in community composition [101].

Our work reinforces the global importance of the Paranã River Basin within the Cerrado hotspot. Our refined biological database is an important resource to downscale conservation planning for a region previously considered as data-deficient. Although data on the Basin’s biodiversity is far from complete, it is sufficient as a basis for species-distribution and community models for five major taxonomical groups. The pace of loss of native vegetation in the Basin indicates the urgency of conservation planning with available information, acknowledging that a plan for the Basin, like that for any other region, will be subject to adaptation as new information becomes available [73]. This study has also pointed to taxonomic and biological gaps to guide future biological inventories.

Despite our efforts to compile the available biological information for the Paranã River Basin, including site visits to several scientific collections, some unique records might have been missed because of constraints on the management of collections. However, this problem was partly circumvented by accessing species records through publications, theses and dissertations, gray literature, and personal databases. We interpreted the lack of response from some collections to our inquires as a lack of staff trained in preparing, organizing, and identifying specimens, and making material available. Additionally, some collections are regarded as private property, even though maintained with public investments, thereby denying or limiting access to deposited material. For these collections to achieve their full economic and cultural potential [102], it is clearly necessary to both sustain and strengthen public investments, in terms of infrastructure and taxonomic expertise. Frequently in Brazil, as elsewhere, specimens are lost from lack of maintenance, some still lacking scientific description. For conservation to be as effective as possible, it is also important to disseminate data on biodiversity as widely as possible through, for example, programs such as SpeciesLink, the Global Biodiversity Information Facility, and the Catalogue of Life [72].

Although our data and analyses are specific to the Paranã Basin in the Brazilian Cerrado, we identify the myriad of problems faced when compiling basic information necessary to make conservation decisions. More importantly, we suggest methods for compiling, detecting and eliminating duplicates in biodiversity data that can be applied to make predictions of biodiversity in poorly known regions, especially those that are subject to rapid anthropogenic development. All biodiversity information must be supplemented by metadata [72] to facilitate standardization and filtering. The minimum information necessary for data cleaning includes not only scientific names and precise coordinates of collection but also dates of collection, names of collectors, numbers of samples, descriptive locations, names of collections where material is housed, names of taxonomists making the identifications, and record (tag) numbers of specimens within collections. Furthermore, encouraging peer-reviewed publication of species lists for localities would enable access to reliable information that exists now only in reports with limited circulation. Thus, the integration of different strategies, from data collation to refinement and publication of biological data is essential to support decisions about conservation priorities, not only in Brazil but in all parts of the world.

Chapter 3

Chapter 3 – Biodiversity patterns within the Paranã River Basin: what we can learn from species and assemblage-level distributions?

Keywords: biodiversity patterns, species distribution, assemblage distribution, Maxent, GDM, Paranã, Cerrado

Article type: Full-length Article

Anderson C. Sevilha[1,2], Robert L. Pressey[3], Jeremy VanDerWal[1,4], Helga C. Wiederhecker[1], Kristen J. Williams[5], Simon Ferrier[5] and Stephen E. Williams[1]

[1]Centre for Tropical Biodiversity and Climate Change, College of Marine and Environmental Science, James Cook University, Townsville, QLD, Australia
[2]Embrapa Genetic Resources & Biotechnology, Brasília, DF, Brazil
[3]Australian Research Council Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, QLD, Australia
[4]eResearch Centre, James Cook University, Townsville, QLD, Australia
[5]CSIRO Ecosystem Sciences, Black Mountain Laboratories, Canberra, ACT, Australia.

Chapter 3 – Biodiversity patterns within the Paranã River Basin: what we can learn from species and assemblage-level distributions?

3.1. Abstract

Understanding biodiversity patterns and the ecological mechanisms that regulate them is fundamental to biological conservation. However, biological data are typically scarce and dispersed, limiting their direct use in conservation planning. Modelling the distributions of individual species or species assemblages in relation to environmental variables has been demonstrated to provide an effective complement to direct biological data, although it has been limited so far mainly to understanding species distributions. In the Paranã River Basin, a region of importance for nature conservation within the globally significant Cerrado biodiversity hotspot, biological patterns have until recently been largely unknown. After an intensive search for biological information in this region, we can now develop new insights into biodiversity patterns in relation to both species and assemblages. As a basis for conservation planning, we analysed species distributions and species richness using Maxent (Maximum Entropy species distribution modelling) and species assemblage patterns using Generalized Dissimilarity Modelling (GDM). We modelled distributions for 2159 species (1813 plants, 245 birds, 44 mammals, 30 reptiles and 27 amphibians). Species records were aggregated at 0.1 degree grid (~1 km) resolution across the Paranã River Basin. At the same resolution, we modelled geographic patterns in assemblage dissimilarity for each of the terrestrial vertebrate groups (63 species of amphibians, 86 reptiles, 430 birds, 101 mammals), vascular plants (3467) and insects (63 drosophilid flies, 33 hymenopterans, 66 termites and 289 saturniid moths). Analyses of species distribution models identified differences in richness patterns between animals and plants. Plants had higher diversity across the montane gradient in more elevated parts of the Basin. Diversity of vegetation types and plant species was related to environmental heterogeneity, represented by differences in climate (rain and temperature) and substrate (soil depth, permeability to water, and base saturation – the fraction of exchangeable cations present as Ca, Mg, K, and Na, which indicates differences in soil fertility). In contrast, vertebrates had higher richness in the lower-elevation sections of the Basin, apparently related to higher mean and more stable temperatures and higher

structural complexity of habitats in lowland environments dominated by dry forests and arboreal savanna. Reliable models for assemblage patterns were limited by data availability and could be obtained only for birds, mammals, and plants. These groups showed high spatial congruence in patterns of assemblage structure, characterizing three major bioregions important for biodiversity conservation within the Basin: the montane gradient and two vegetation formations – dry forests and savannas – in the relatively flat areas at the bottom of the Basin. The results of modelling species distributions and assemblages provide complementary perspectives on biodiversity patterns as a basis for conservation planning, depicting spatial differences in both richness and turnover in composition.

3.2. Introduction

The study of biological variation within and between species and species assemblages, and their relationships with environmental conditions, provides insights for mapping, planning, and prioritizing conservation actions [103, 104]. Conservation actions require information on the spatial distribution of biodiversity, often across large regions [10, 105]. Direct field sampling of such regions is typically sparse, with biological surveys or collection sites covering very small patches within extensive unsurveyed areas [101]. This scarcity of direct biological data applies particularly to Brazil's Paranã River Basin, located in the centre of the Cerrado global biodiversity hotspot [58], and considered to be of extreme importance for biodiversity conservation [31-41, 58].

Despite efforts in acquiring biological information in the Paranã River Basin over the last 200 years, patterns of species richness and the distribution of assemblages are still poorly known [58]. Direct biological records take time and money for sampling and identification, and the Paranã River Basin exemplifies regions with high biodiversity but where biological inventories are not a high priority [106], the race for economic growth is driving rapid environmental transformation, and there is an urgent need for systematic conservation planning and subsequent implementation of conservation actions on the ground.

In such regions, direct biological information can be complemented by modelling the distributions of species and assemblages in relation to environmental variables, based on biological data already available [107]. Most information on species occurrences is limited to species records from herbaria and museums and much of it is now widely available in online repositories. Very few data come from systematic studies developed specifically for modelling purposes [108]. However, the ready availability of online databases for species records and environmental data has been accompanied by a long debate about which tools should be used to analyse and assess biodiversity patterns, and the pros and cons of each [101, 104, 108-111]. There are three key aspects to decisions about modelling techniques: (i) an ecological framework or conceptual model, which includes the theory used to link environmental predictors to biodiversity distribution; (ii) the data model, which considers data availability, resolution, and selection; and (iii) the statistical model, which includes the modelling method and the choice of explanatory variables during model building and evaluation [111-113].

With these three considerations in mind, we used data collected from project-specific biological inventories and a large volume of additional data collated and filtered from herbaria, museums, literature, and online databases [58] to model the distribution of species and assemblages within the Paranã River Basin. We developed an ecological conceptual framework to preselect variables likely to be related to the distributions of species and assemblages. We then applied and evaluated two different methods to model distinct dimensions of biodiversity, namely species distribution models and assemblage dissimilarity modelling.

The reliability of these modelling methods depends on the type and spatial bias of direct biological information available and the quality of data on potential predictors used for model building [111, 112]. The limited availability of species occurrence data poses a challenge to the creation of accurate species distribution models, especially for endemic and threatened species. However, reasonable models can be obtained even with a few sampling points [105], although lacking the accuracy of those made with larger datasets and potentially not appropriate for all applications. Among the benefits of species distribution models are insights into environmental relationships, help with design of additional field inventories, and improved information for selection of areas for conservation, management, and reintroduction [105, 114].

Species assemblage models include all species in the data set, regardless of the number of records per species [115]. This approach is useful for groups of species that are poorly sampled, or that exhibit rapid spatial turnover in composition relative to the distribution of sampling sites. Both of these characteristics limit the accuracy of species distribution models [116]. Species assemblage models predict spatial patterns in assemblage composition across the study area based on environmental variables, extrapolating patterns in compositional turnover beyond sampled localities [115, 116]. The resulting predicted patterns of beta diversity serve as a biodiversity surrogate for conservation assessment that has the potential to provide information complementary to species distribution models [117]. However, species assemblage models assume that, to some extent, all species have been surveyed (i.e. recorded as present or absent) at the same set of sites [115]. Using presence-only data in assemblage modelling therefore requires adjustments to the method [101].

To build on the field inventories and data compilation described in Chapter 2, and to provide a basis for conservation planning in the Paranã River Basin described in Chapter 5, this chapter has three main aims:

1. To use species distribution modelling to understand the spatial patterns of composition and richness of taxonomic groups, and whether these patterns are congruent between groups;
2. To use assemblage dissimilarity models to understand how spatial compositional turnover shapes assemblages of species, and whether these assemblages are congruent between groups;
3. To compare the results of models of species distribution and assemblage dissimilarity to understand their potential to provide complementary inputs for conservation planning software.

3.3. Methods

3.3.1. Input data and models

We modelled species environmental niches and compositional dissimilarity surfaces using data on the terrestrial taxonomic groups previously compiled for the region [58]. We used field records of amphibians, reptiles, birds, mammals, insects, and plants. Climatic variables (precipitation and temperature), landscape characteristics, and data on substrate were combined with species data to create correlative models of species distributions, using Maxent [118, 119], and compositional dissimilarities, using Generalized Dissimilarity Modelling [101, 117, 120, 121]. With these methods, it was possible to evaluate spatial patterns of species composition and richness and assemblages related to spatial turnover of composition within the Basin.

3.3.2. Selection of predictor variables for species distribution modelling

Selecting predictor variables is a critical step in modelling distributions of species and assemblages. Particular attention needs to be given to the explanatory power of those variables, their ecological basis, and physiological relevance for the taxa being considered [111, 112, 122, 123]. We preselected a comprehensive set of 38 physiographic environmental variables at 0.01 degree grid (~1 km) resolution, which potentially influence the distribution of species in the study area (see Table 2 in Supporting Information 2 for variable descriptions, their ecological rationale, and sources). Each variable was then assigned to one of four categories reflecting its ecological relevance (precipitation, temperature, landscape, and substrate). Variables within categories were assessed in terms of redundancy and relevance to the different taxonomic groups.

To avoid model over-parameterization, we removed redundant variables within each ecological category that were highly correlated with others in that category (Pearson's correlation index above 0.70), retaining those with more physiological importance in controlling growth, reproduction, morphology, and behaviour [104, 124]. These variables were *a priori* considered more ecologically relevant as influences on distribution patterns for each taxonomic group [as per 112].

The remaining, relatively independent, variables were then submitted to manual selection strategies, including forward selection, backward elimination, and stepwise procedures using Maxent [see 112]. The variables contributing the least information based on their permutation importance (less than 1% via jackknife tests for overall species within two groups – terrestrial vertebrates and plants) were successively dropped [as per 112]. The final model retained seven variables for fauna and nine for plants that were relatively independent (Table 3.1). Insects were not included because all had less than the minimum of five locality records considered adequate for modelling.

Table 3.1. Predictor variables selected for species distribution models. "Fauna" corresponds to mammals, birds, reptiles, and amphibians.

Variables	Type of variable	Fauna	Plants
Annual mean temperature	Temperature	X	X
Temperature seasonality	Temperature	X	X
Annual precipitation	Precipitation	X	X
Precipitation seasonality	Precipitation	X	X
Precipitation of the driest quarter	Precipitation	X	X
Surface roughness	Landscape	X	X
Vegetation class	Substrate	X	
Soil depth	Substrate		X
Soil permeability	Substrate		X
Soil base saturation	Substrate		X

Spatial variability in the selected variables can be considered as a proxy for habitat structure and its spatial heterogeneity. We interpret habitat structure here as vertical habitat stratification, that is, the number of strata or layers in a given vegetation type, and assume, based on previous studies, that stratification is relevant to niche-space partitioning [125-127]. A typical forest formation of the Paranã River Basin has at least five layers: herbaceous, shrub, understory, canopy, and emergent. Temperature, humidity, and light can vary vertically, decreasing from the canopy to the bottom of the forest. At the other extreme are the grasslands, with just a herbaceous layer. Savanna formations vary from open to arboreal or woodland, according to the densities of herbaceous plants, shrubs, and trees. Open savannas have herbaceous and shrub layers with scattered trees. In a typical savanna, the densities of the herbaceous and shrub layers are lower, and the trees are taller but without forming a canopy. In an arboreal savanna, the tree density is higher still and can form a canopy, and the densities of the herbaceous and shrubs layers are lower than in open savannas [59, 128, 129]. Therefore, in structural complexity, forested habitats are highest, followed by savannas (from arboreal to open) and grasslands.

Habitat heterogeneity in this study represents a measure of horizontal variation in the environmental variables across the landscape. For example, variation across the landscape in classes of soils or vegetation types and values for landscape roughness, temperature, and precipitation determine spatial heterogeneity that might also influence niche-space partitioning [125, 126, 130].

Our choice of predictor variables influenced both habitat structure and heterogeneity, allowing us to capture complementary information relevant to species distributions across the environmental gradients, both localized and region-wide, in the Paranã River Basin.

3.3.3. Species-level modelling

For species distribution models (SDMs) we used a total of 2159 species (1813 plants, 245 birds, 44 mammals, 30 reptiles, and 27 amphibians) occurring in five or more geographically unique localities. We gridded the study region at 0.01 degree (~1 km) resolution, so “unique” refers to localities after they were aggregated to this resolution across the Paranã River Basin extended area. The extended area consisted of a buffer (~100 km) outside the Basin’s boundary, within a polygon formed by adding one degree to the latitudinal and longitudinal extremes of the Basin. Our aim of defining the extended area was to reduce the arbitrary effects of ecological truncation associated with the Basin’s perimeter by including surrounding environmental gradients and related species distributions.

We are aware of the criticisms of using small numbers of occurrence points in SDMs [105, 131]. However, using a larger minimum number of records would have reduced the number of species included in the analyses significantly, compromising the understanding of biodiversity patterns within the Basin. Furthermore, reliable results in modelling species distributions with Maxent can be achieved even when few records are used to estimate species distributions [105, 131], as long as the points cover a reasonable proportion of the species’ environmental space (see Discussion for further consideration of the limitations of our SDM modelling).

We estimated environmental suitability for each species as a function of predictor variables in the gridded study domain using Maxent 3.3.3k [118, 132]. Where there were 20 or more gridded geographical records, individual species models were fitted using a maximum of 500 iterations, a convergence threshold set to 0.00001, and 0.5 as prevalence of the species. For each model, a random sample of 10,000 sites from the study area, in addition to the sampled sites, defined “absences”, which might or might not be true absences (random background). The regularization parameter was set to linear, quadratic, product, threshold, and hinge feature classes using Maxent default parameters for number of samples at which features were first used [see 118]. Additionally, we used a 10-fold cross-validation partitioning species occurrences in 90% for training data and 10% for testing models.

For other species (those with 5 to 19 gridded occurrences), we removed the threshold feature from the regularization parameter to smooth the response curve and obtain a general tendency of variable contribution. For background sites (“absences”), we used a set of all gridded occurrence records surveyed for other species in the same taxonomic group (target group background). There were two reasons for this background approach: first, to avoid inflation of AUC values due to the contrast between few species records and a large number of background points [108]; and, second, to minimize the effect of biased sampling [118, 133], which is inherent in incidence species records [110, 118, 133, 134]. The number of background sites used for species in each taxonomic group is presented in Table 3.1.

Model performance was evaluated by the AUC – area under the receiver operator characteristic curve [110], automatically calculated by Maxent. Maxent´s AUC measures how well the predicted distribution ranks test presences above background test absences. Its value is equal to the probability that a randomly chosen presence will be ranked above a randomly chosen absence [132]. The AUC is considered an adequate indicator of overall model performance (ability to discriminate suitable and unsuitable areas) [118, 135, 136] and is often used to assess the consistency of model predictions of species distributions [137]. AUC can also be used to compare performance between different models based on the same presence-only data [108]. Generally, an AUC ≥ 0.7 is taken as a threshold above which to retain models for further analysis, while values ≥ 0.9 are considered indicative of high accuracy [138]. Values < 0.5 indicate performance no better than random [132].

To analyse richness patterns from SDMs, we needed criteria for retaining species based on the reliability of their models. For species with 20 or more locality records, we retained those with mean AUC ≥0.70 in the 10-fold cross-validation. For each species not evaluated through cross-validation (those with less than 20 records), we retained species only if their AUCs were ≥0.70.

For retained species, we estimated species richness per grid cell using two separate sets of stacked models to identify differences and similarities between them. One set consisted of models for all retained species, regardless of the number of localities. The second set consisted of only those models for species with 20 or more localities. Models were stacked by combining the predicted distributions of species in each of the five taxonomic groups evaluated: plants, birds, mammals, reptiles, and amphibians. For stacking models, we converted the continuous logistic output for each species into a Boolean map (binary 0/1 absence/presence map). The threshold marking the difference between suitable and unsuitable was defined by the 'equate entropy of thresholded and original distributions logistic'. This threshold finds the cutoff value between absence and presence that most closely reflects the distribution of suitability values in Maxent's continuous logistic output [139]. The final species richness maps therefore represented the sum of potential presences (or relative suitability) across all retained species in each taxonomic group in each pixel. These values were combined using the SDMTools R package [140]. Additionally, we tested for relationships between patterns of species richness and the environmental layers using Pearson's correlation coefficient.

3.3.4. Assemblage-level analysis

We also examined the spatial pattern of biodiversity with a quantitative, spatially explicit analysis of changes in assemblage structure using generalized dissimilarity modelling (GDM) [101, 117, 120, 121]. GDM is essentially a nonlinear extension of Mantel correlation analysis [141]. It uses GLMs (generalized linear models) to model observed dissimilarity in biological composition between site pairs, as a function of explanatory variables [101, 116]. The pair-wise dissimilarities between multiple species presences and absences were calculated with the Bray-Curtis index [101].

The mathematical description of GDM assumes that environmental predictors are continuous variables or, at least, ordered categories [101, 112]. We transformed classes of vegetation and soils into ordinal scales relevant to species occurrences. The original vegetation types were converted into ordinal classes by mean values of percent canopy cover, height, and number of strata to reflect structural complexity. Similarly, we extracted the mean values of depth, base saturation, and permeability for soil classes. All transformed variables were standardized to values from 0 to 1 (see Table 1 in Supporting Information 2 for a complete description).

We calculated dissimilarities between pairs of sites for each of the terrestrial vertebrate groups (63 species of amphibians, 86 reptiles, 430 birds, 101 mammals), vascular plants (3467) and insects (63 drosophilid flies, 33 hymenopterans, 66 termites, and 291 saturniid moths) using .NET GD Modeller software version 2.70 [142]. We grouped all insects together, attempting to identify patterns of insect distributions as a whole, given that small numbers of records precluded SDMs.

The GDM models developed for the terrestrial biota selected a total of 20 relatively independent predictors (Table 3.2) from the 38 candidate predictors tested for inclusion (the same 38 first considered for the SDMs; see Table 1 in Supporting Information 2). The nonlinear monotonic functions fitted for all variables were derived from three *I*-spline basis functions. The predictors included in the model, and the number of *I*-spline basis functions employed for each predictor, were determined using manual and automated selection strategies, including forward selection, backward elimination, and stepwise procedures [112] over the variables selected (see Table 1 in Supporting Information 2).

Table 3.2. GDM variables selected for each taxonomic group modeled in the Paranã River Basin.
A = Amphibians; B = Birds; M = Mammals; R = Reptiles; I = Insects; and P = Plants

Variables	Type of variable	A	B	M	R	I	P
Annual mean temperature	Temperature					X	
Mean temperature of coldest month	Temperature		X		X		
Mean Temperature of driest Quarter	Temperature						
Mean temperature of wettest quarter	Temperature			X			X
Mean temperature of warmest quarter	Temperature	X					
Temperature seasonality	Temperature	X		X		X	
Isothermality	Temperature	X	X				
Precipitation of coldest quarter	Precipitation				X		X
Precipitation of driest month	Precipitation			X			
Precipitation of wettest month	Precipitation	X		X			X
Precipitation of warmest quarter	Precipitation		X		X	X	X
Precipitation seasonality	Precipitation		X				
Surface roughness	Landscape			X	X		
Soil base saturation	Substrate			X			
Soil permeability	Substrate						X
Vegetation complexity	Substrate	X			X		X
Vegetation coverage	Substrate		X				
Vegetation height	Substrate			X			
Distance to stream	Terrain	X				X	
Geographic distance	Spatial	X	X	X	X	X	X

To limit model complexity, we used a threshold of 0.05% improvement in minimum partial deviance explained from including each predictor or additional *I*-spline in each model. After selecting the environmental variables, we tested the effect of adding geographic distance between locations of site pairs. Geographic distance was included if it independently explained additional variation, i.e. not at the expense of existing environmental predictors, and if it contributed at least the threshold deviance increment used in variable selection. Geographic distance is hypothesised to relate to latent variables, evolutionary history, dispersal barriers, and non-equilibrium conditions [112]**.** Each final model used a set of relatively independent predictors for its taxonomic group (see Table 2 in Supporting Information 2) that minimized the number of variables included without compromising deviance explanation.

When data are aggregated from multiple sources, as in this study, spatial and environmental bias in sampling is likely. GDM performs best when the aggregated sample begins to approach a representative set of presence/absence samples for the ecological region of interest. In this study, we found relatively small deviance explained by GDM models for some taxonomic groups. However, we considered that the signal variation could be robust, albeit masked by variation in data quality. Previous workers have used a minimum number of species per site as an indicator of sampling adequacy for sites to be included in the analysis [112], or have applied a weight to the regression based on the number of species recorded at each site pair [101]. Therefore, we developed an exploratory analysis to test for signal robustness in the models as an indicator of their utility in subsequent conservation assessments.

The different model comparisons aimed to test the robustness of the GDM models developed using the complete dataset with the GDM models developed for ten subsets of purported adequately inventoried sites. The subsets of sites came from systematic biodiversity inventories carried out in the study area [58]. Ignoring the model fit statistics, and as a test of signal robustness, we hypothesized that the dissimilarities predicted by the model derived using all the data and applied to the adequately inventoried sites [58] should approximate the predicted dissimilarities obtained from a model based only on the adequately inventoried sites. We also compared the observed and predicted dissimilarities between just the adequately inventoried sites with the predicted dissimilarities for the complete dataset using simple linear regression. The parameters of the straight line fit (intercept and slope) and the r-squared values were used to measure concordance. Additionally, to confirm that the improvement in deviance explained by the adequately inventoried sites was in fact due mostly to the thoroughness of surveys at these locations, rather than being simply an artifact of smaller numbers of sites allowing better-fitting models, the deviances explained by those models were compared with what was explained, on average, by a set of 10 models fitted to randomly selected sub-samples of sites, each with the same number of sites as the adequately inventoried sub-samples).

Finally, the predicted dissimilarities were visualized by clustering all grid cells in the study area into ten classes based on UPGMA hierarchical classification [101]. Each class was colored using a multi-dimensional scaling of the predicted similarities using a sample size of 10000 cells, such that similar colors indicated compositional similarity [143]. For clustering the grid cells based on records of plant species, we set the number of groups to 10, based on our prior knowledge of the vegetation assemblages in the Paranã River Basin. We assumed that animal assemblages would align roughly according to those of the plants, so again set the number of groups to 10.

3.4. Results

3.4.1. Species-level modelling

From the total number of species used to model distributions (2159), 513 species with 20 or more records had AUC values of ≥ 0.70 (456 plants, 51 birds, 4 mammals and 7 reptiles) and were retained for further analysis. For the 513 retained models, no taxonomic group had an average AUC smaller than 0.70 (Table 3.3).

Of the species with < 20 records (models not evaluated with cross-validation), we retained models for 1640 species with acceptable AUCs (1357 plants, 194 birds, 40 mammals, 23 reptiles, 26 amphibians). For these species, the vast majority of AUCs were higher than 0.9, and average AUCs for taxonomic groups were 0.85 or higher.

From this set of models, only one species (the amphibian *Elachistocleis ovalis*) was removed from the analysis following the AUC threshold filter. This species is a habitat generalist, occurring from forests to grasslands, with a range extending beyond the Cerrado biome [144]. Ecological truncation of its distribution and, therefore, inadequate correlation with the available environmental predictors, might explain, in part, its low AUC score (0.64). The high AUC values for the retained species with < 20 records might be inflated by the relatively large number of background points [108]. Nonetheless, these models were considered reliable, at least as a basis for analysis of biodiversity patterns.

Table 3.3. Maxent performance in relation to mean values (AUC) and standard deviations (SD) for each taxonomic group.
Spp = number of species with retained models per taxonomic group. Models for species with ≥ 20 records used 10,000 random background sites, added to sampled sites, as assumed absences. Models for species with < 20 records used background points (PO indicates number) based on the target group background (a set of all gridded occurrence records surveyed for other species in the same taxonomic group).

Taxonomic Group	**Retained species with ≥ 20 occurrences (513)**			**Retained species with < 20 occurrences (1640)**			
	AUC	SD	# Spp	AUC	SD	# Spp	PO
Amphibians	-	-	-	0.89	0.08	26	37
Birds	0.85	0.07	51	0.92	0.04	194	135
Mammals	0.81	0.08	4	0.90	0.07	40	115
Reptiles	0.76	0.01	7	0.85	0.07	23	62
Plants	0.83	0.07	456	0.94	0.05	1357	2523

The species richness maps derived from the all retained models, regardless of number of records of species, were similar spatially across the four faunal groups, but different from the richness pattern of plants (Figure 3.1). Plant richness was higher in mountain regions while that of birds, mammals, amphibians and reptiles was greater in the lower parts of the Basin.

Plants were more diverse across the montane gradient of the ‘Serra Geral do Paranã’ (775 to 1663 m – see Figure 1.3) that surrounds the western and southern portions of the Basin. This pattern extends through the central-western highlands that divide the northern part of the Basin (in the Tocantins depression – see Figure 1.2) from southern parts. The montane gradient is dominated by open savannas and grasslands interspersed with dry forest formations along water courses and ravines (semi-deciduous forests) and occasional deciduous forests on basic soils.

Faunal richness was highest in the relatively flat terrain of the northern part of the Basin. This area is occupied by arboreal savanna formations intermixed with dry forest, where elevations vary from 247 to 430 m [58]. A secondary region of high richness was across the northeastern-southwestern axis in the central-southern part of the Basin (the ‘Vão do Paranã’ plain in Figure 1.3). This region is occupied by dry forests and ecotone zones with arboreal savannas, with elevations that range from 430 to 585 m [58].

Similar richness patterns to those derived from all retained species (Figure 3.1) were obtained with models based on species with at least 20 records (Figure 3.2). Amphibians were not included in these models given that small numbers of records precluded SDMs. Again, richness for fauna was highest in the lower parts of the Basin and highest for plants on the montane gradient. However, there were some slight differences for birds and plants between the two sets of models, and more pronounced differences for reptiles and mammals. For models based on species with at least 20 records, the richness pattern of plants was less sharply limited to high altitudes, grading somewhat into lower areas. High richness for birds expanded further to the northeastern and southern parts and into some parts of the montane gradient in the west. High richness for reptiles diffused south. High richness for mammals shifted from north to south. However, there is a caveat on interpreting the differences in richness patterns for reptiles and mammals between the two sets of models. The models used in Figure 3.2 built for only a few species (7 reptiles, 4 mammals), so might not be reflect the overall patterns of richness for these groups.

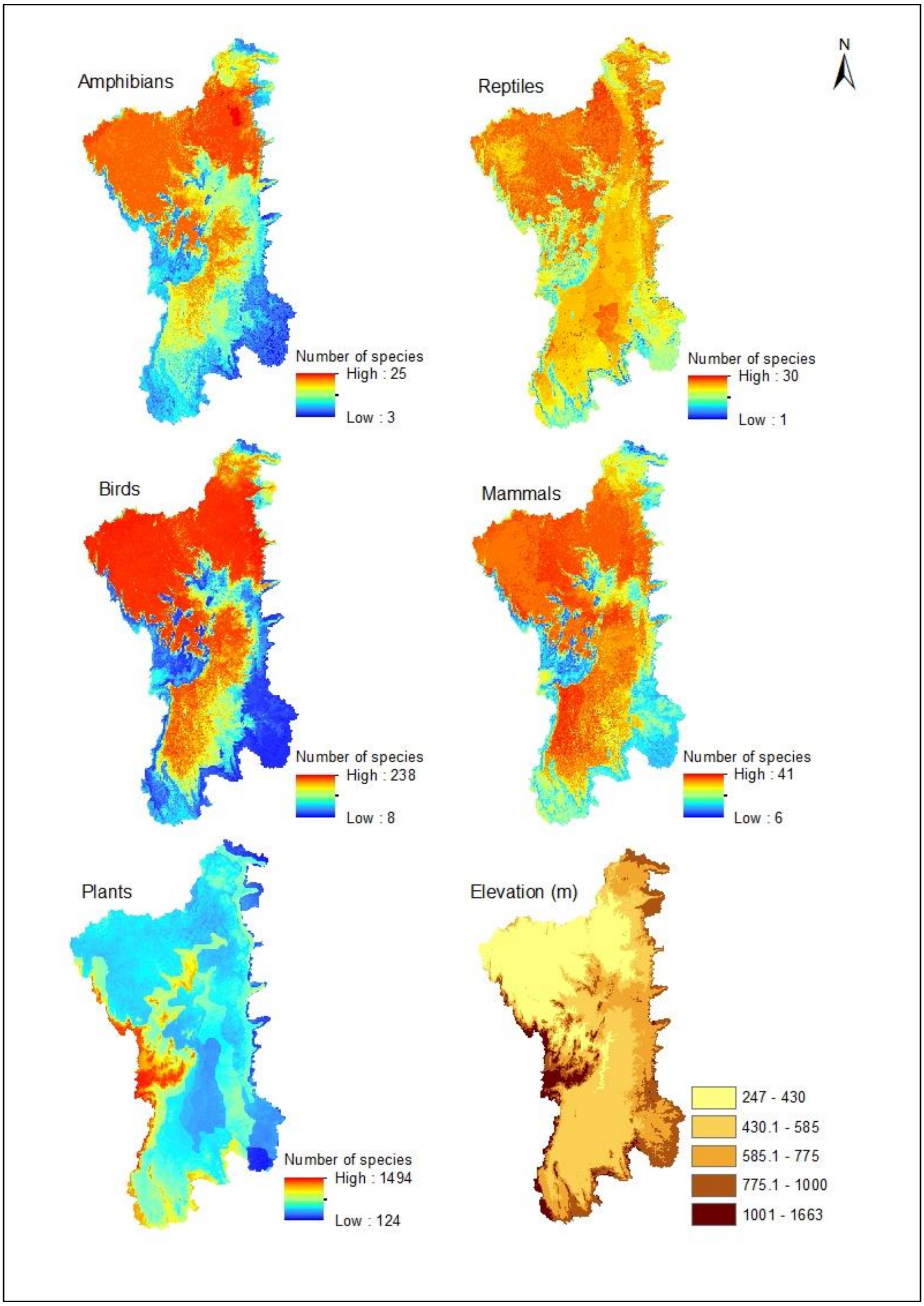


Figure 3.1. Spatial species richness maps for all models combined (stacked models of all retained species, regardless of number of records) and elevation classes within the Paranã River Basin.

Warmer (reddish) and colder (bluish) colors indicate higher and lower richness values (or relative habitat suitability), respectively.

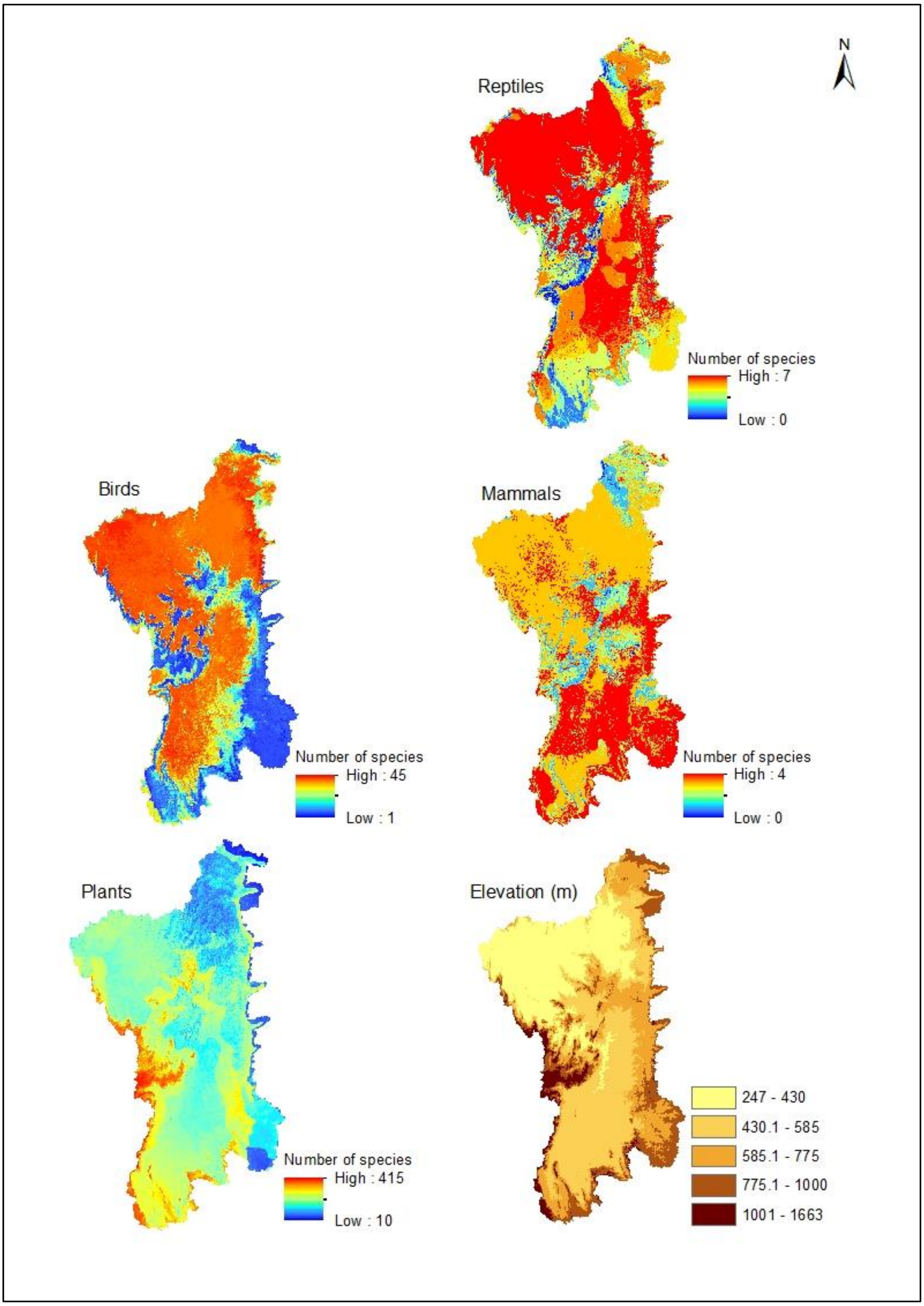


Figure 3.2. Spatial species richness maps for species recorded in 20 or more localities and elevation classes within the Paranã River Basin.

Warmer (reddish) and colder (bluish) colors indicate higher and lower richness values (or relative habitat suitability), respectively. Amphibians were not included in these models because all species had few records.

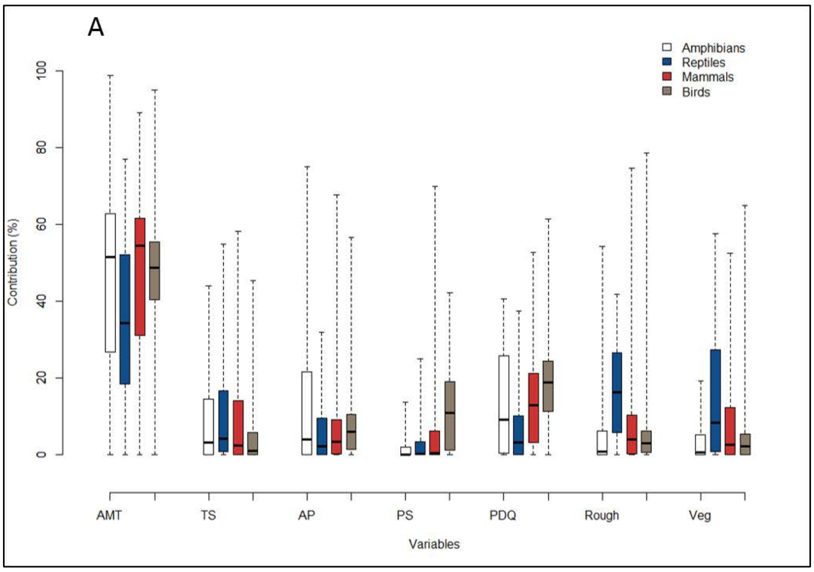
Aspects of temperature and precipitation are the most important correlates of the richness patterns observed in the Paranã River Basin (see Table 3.4). Vertebrate richness was primarily correlated with higher and stable temperatures found in the low parts of the Basin while plant richness was correlated with higher precipitation, especially during the dry season, and less variable precipitation found in rugged terrains.

Table 3.4. Pearson correlation coefficients among estimated richness and environmental variables used to model species distributions within the Paranã River Basin.

AMT = annual mean temperature, TS = temperature seasonality, AP = annual precipitation, PS = precipitation seasonality, PDQ = precipitation of the driest quarter, Rough = surface roughness, SD = soil depth, SP = soil permeability, SB = soil saturation of bases. Shaded cells highlight coefficients with absolute values larger than 0.60.

Variable	Amphibians	Birds	Mammals	Reptiles	Plants
Birds	0.90				
Mammals	0.90	0.91			
Reptiles	0.67	0.66	0.69		
Plants	-0.11	-0.12	-0.03	-0.10	
AMT	0.75	0.72	0.65	0.65	-0.35
TS	-0.69	-0.56	-0.61	-0.69	-0.10
AP	0.28	0.26	0.23	0.44	0.45
PS	-0.09	0.05	-0.09	-0.21	-0.44
PDQ	-0.42	-0.42	-0.36	-0.22	0.62
Rough	-0.13	-0.17	-0.17	-0.50	0.43
SD	0.03	0.02	0.01	0.03	-0.22
SP	-0.14	-0.16	-0.16	-0.01	-0.53
SB	-0.04	-0.02	-0.01	-0.10	-0.07

Despite variation between species in taxonomic groups in the relative contribution of each variable to the models, annual mean temperature was consistently the most important variable related to species' distributions for plants and fauna. Temperature was followed in importance by water (precipitation), substrate, and landscape structure, which contributed to different degrees depending on the taxonomic group (see Figure 3.3 A, B).



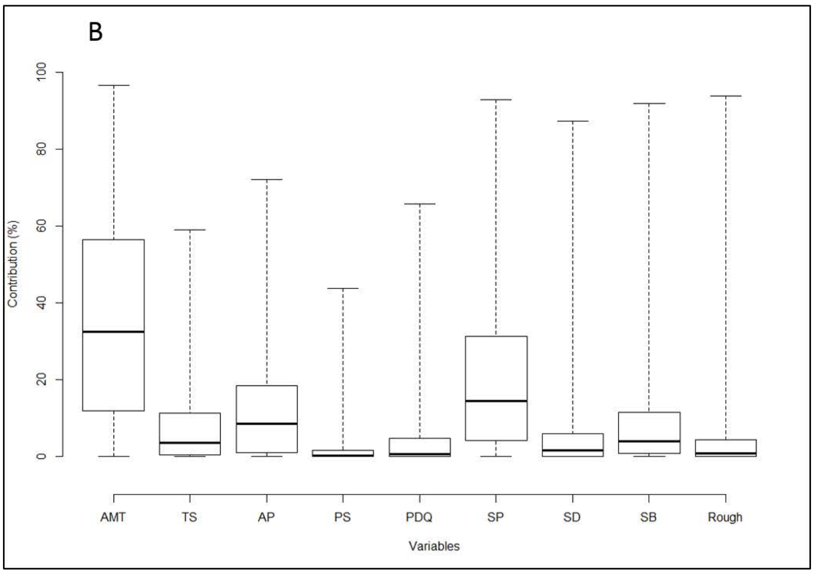


Figure 3.3. Boxplots showing the relative contribution of each environmental variable (Jackknife test of variable importance) to the Maxent distribution models for species in faunal groups and plants. A. Contributions to models for vertebrate species. B. Contributions to models for plants.

Contribution values indicate percentage contributions of each variable to estimated suitability for each species in each group. Boxes represent 50% of the values (second and third quartiles), whiskers represent first and fourth quartiles, and darker bars within boxes indicate medians.

AMT = annual mean temperature; TS = temperature seasonality; AP = annual precipitation; PS = precipitation seasonality; PDQ = precipitation of driest quarter; Rough = surface roughness; Veg = vegetation classes; SD = soil depth; SP = soil permeability; SB = soils base saturation.

3.4.2. Assemblage-level analysis

The comparative results of models derived from all samples for each taxonomic group and models based solely on comprehensively surveyed sites (inventoried sub-samples) showed that only models for birds, mammals, and plants were unaffected by the adequacy of field sampling (see values of 'R' and 'adjusted R-squared' columns in Table 3.5). For reptiles, dissimilarities of site pairs based on all data were also significantly related to dissimilarities based only on fully inventoried sites, but models for this taxonomic group appeared to be affected by the smaller number of species in the sub-samples. The deviance explained by the sub-samples of the inventoried sites for reptiles (51.42%) was not different from that obtained by models fitted to the filtered data – the randomly selected sub-samples of sites (48.77%; SD = 8.9). The results for birds (25.64%; SD = 7.4), mammals (26.08%; SD = 7.2) and plants (18.08%; SD = 1.7) validated the robustness of their models and the sampling adequacy for all sites to be included in the analysis.

Therefore, the correlation between all samples and inventoried sub-samples indicates that the low deviance explained by the inclusion of all samples in the models for plants, mammals, and birds (4.53; 9.44; and 10.98%, respectively – see 'percentage deviance explained' for 'all samples' columns in Table 3.5) was, in fact, an effect of the high number of species recorded in few locations. In consistent surveys, or filtered site data, where total deviance explained by the models was high, the number of sites with few species was small. These models – for birds, mammals, and plants – were then retained for the subsequent analyses of biodiversity patterns.

All predictors included in the retained GDM models (those for birds, mammals, and plants) were related to climatic variables (water and temperature), substrate, and geographic distance. Geographic distance influenced other predictor coefficients by decreasing their relative contribution to the models and the intercepts, but increased the total deviance explained by the models by at least 5%. Differences in mean temperatures (in winter for birds, and in summer for mammals and plants) and mean precipitation (in summer for birds, and spring and winter for mammals and plants) were the main environmental correlates of the spatial patterns of assemblage turnover (indicated by the maximum “y” values in Figure 3.4). The maximum height reached by each fitted function for each variable indicates the relative contribution of each variable to the model. The maximum height also reflects the total amount of compositional turnover (beta diversity) associated with the environmental gradient concerned, and by extension, the relative importance of that variable in explaining beta diversity, holding all other variables constant [101].

The highest rates of species turnover between site-pairs (indicated by the slope of the fitted curves in Figure 3.4 for birds, mammals, and plants) were associated with low geographic distance (birds, mammals, and plants), precipitation seasonallity (birds), precipitation in summer (mammals and plants), and precipitation in winter (plants). Higher values of species turnover were also found in areas with higher values of temperature in winter (birds), temperature in summer (mammals and plants), precipitation in summer (birds), precipitation in winter (mammals and plants), surface roughness (mammals), vegetation cover (for birds), and vegetation complexity (for plants).

Table 3.5. GDM parameters and regression results between all samples and sub-samples of sites with comprehensive inventories for each taxonomic group.

Taxonomic groups evaluated	Percentage deviance explained by the models		Number of site pairs (number of sites)		Model intercept		Number of predictors included	Model comparison statistics (samples versus sub-samples)						
	all samples	Inventoried sub-samples	all samples	sub-samples	all samples	sub-samples		R	Residual SE	df	Multiple R-squared	Adjusted R-squared	F-statistic	p-value
Amphibians	5.54	51.52	2080 (65)	66 (12)	1.76	0.54	8	0.40	0.18	64	0.16	0.14	11.85 on 1 and 64 DF	0.001021
Reptiles	10.59	51.42	5356 (99)	66 (12)	1.42	0.43	7	0.81	0.11	64	0.66	0.65	122.8 on 1 and 64 DF	$< 2.2e^{-16}$
Birds	10.98	76.04	38,503 (250)	300 (25)	1.21	0.39	6	0.67	0.14	298	0.45	0.45	243.4 on 1 and 298 DF	$< 2.2e^{-16}$
Mammals	9.44	70.24	17,205 (186)	66 (12)	1.38	0.08	9	0.76	0.12	64	0.58	0.58	90.26 on 1 and 64 DF	$7.63e^{-14}$
Insects	18.5	49.59	3741 (73)	66 (12)	0.90	0.07	8	0.71	0.22	64	0.19	0.18	15.07 on 1 and 64 DF	0.000248
Plants	4.53	62.40	500,000 (3039)	2346 (69)	3.67	0.66	7	0.84	0.14	311	0.72	0.72	791.9 on 1 and 311 DF	$< 2.2e^{-16}$

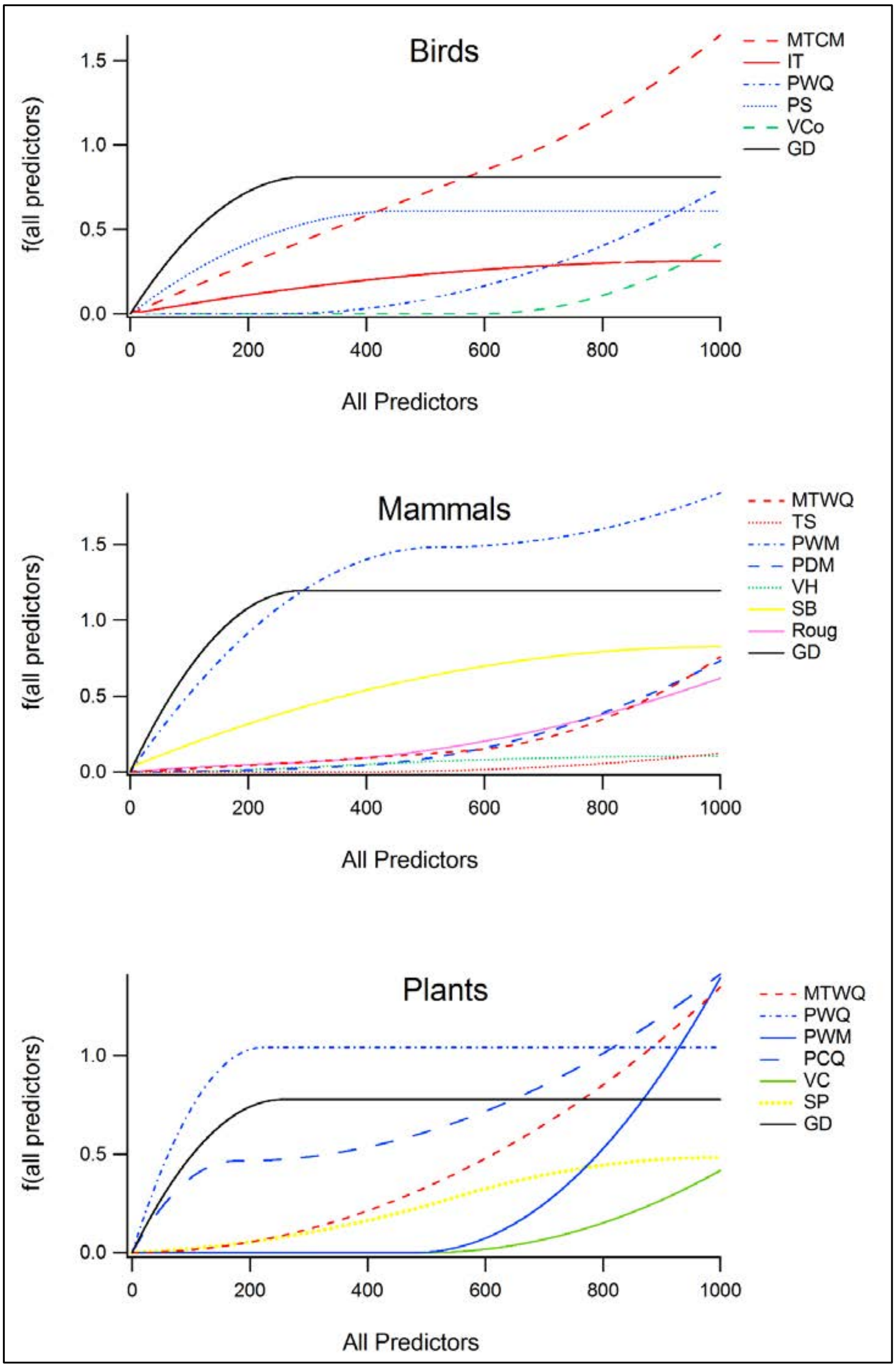


Figure 3.4. GDM fitted functions for birds, mammals, and plants. X-axes show the range-standardized values of environmental variables (environment gradients). Y-axes show values of the GDM-transformed ecological distance (degree of compositional dissimilarity).

Predictors: GD = geographic distance; IT = isothermality; MTCM = mean temperature of coldest month; MTWQ = mean temperature of wettest quarter; PCQ = precipitation of coldest quarter; PDM = precipitation of driest month; PWM = precipitation of wettest month; PWQ = precipitation of warmest quarter; PS = precipitation seasonality; Roug = roughness; SB = soil base saturation; SP = soil permeability; TS = temperature seasonality; VC = vegetation complexity; VCo = vegetation coverage; VH = vegetation height.

At least three very different species assemblages were observed in the spatial representation of the GDM full models for plants, birds, and mammals (Figure 3.5). For plants, differences in depicted assemblages are represented by greenish, reddish, and bluish colours (Figure 3.5A). Greenish areas indicate the assemblages located on the montane gradient (775 to 1663 m) of the western portion of the Basin (‘Serra Geral do Paranã’ range in Figure 1.3 and Plate 1.2.2). This region is covered by a myriad of vegetation types that replace one another along the gradient. Grasslands and savanna formations (from open to dense) are the dominant vegetation in the highest parts of the mountain range and constitute the matrix that surrounds the patchy forest formations (mainly semi-deciduous forests along the watercourses and ravines) (see Plates 1.3.5, 1.3.6, and 1.3.9). Savanna formations (on acid soils) and dry forests (deciduous forests on basic soils and semi-deciduous forests in watercourses and ravines) make up the main vegetation types found along the montane gradient [58] (see Plate 1.2.2). This vegetation heterogeneity in the western elevated areas resulted in the highest species turnover and richness values estimated for plants in the Basin (1470 at the highest elevations and 1012 to 1268 species across the montane gradient).

The vegetation in the low parts of the Basin, the foothills of the mountain range in its western border, and the foothills and the watershed scarps on the southern and eastern borders is predominantly arboreal (reddish colours in Figure 3.5A). The foothills (~ 654 species) are occupied mainly by savannas on acid soils that are relatively poor in nutrients, deciduous forests on basic and more fertile soils, and semi-deciduous forests in the ravines and along watercourses. In the low parts of the Basin, there are three predominant vegetation types: the arboreal savanna formations interspersed with dry forest (~ 566 species) in the north and south; the dry forests (~529 species) in the central-eastern and some portions of the northern and north-western borders; and the ecotones among arboreal savannas and dry forests in the flat lands and limestone outcrops (~ 601 species) located along the eastern border.

A distinct region on the north-eastern boundary of the Basin, represented on the map by bluish colours (Figure 3.5A; ~646 and 485 species), is dominated by savannas and grasslands on sandy soils and semi-deciduous forests in gallery forests in plateaus and plains of the stepped terrain, and deciduous forests in limestone outcrops. The major part of this mapped blue area corresponds to an extension of the ‘Chapadão Ocidental da Bahia’ plateau (Figure 1.3 and Plate 1.2.3) into the Basin. This plateau extends along the eastern side of the Basin (oriented about 600 km north-south and 150 km east-west) at elevations around 1000 m. The original vegetation there is now almost completed removed for intensive agriculture, including soybeans, cotton, and corn.

For birds, the mountain assemblages are related to the highest elevations of the Basin on its western border and the highest elevations and montane gradient of the south-western border (~ 125 species) (bluish colours in Figure 3.5B). This assemblage is also found at the higher elevations along the south-eastern border (~ 60 species).

At low elevations, two distinct bird assemblages are represented. The reddish colours (Figure 3.5.B; ~183 to 228 species) indicate bird assemblages in the lowest elevations of the Basin (in the Tocantins depression – Figure 1.3), a region dominated by savanna formations interspersed with dry forests along the watercourse of the Paranã River (on the axis southeast to northwest – Figure 1.3) and affluents of the Palma River (oriented on an axis from east to west of the north portion of the Basin – Figure 1.3). Additionally, these regions are characterized by permanent and intermittent lagoons along the Palma and Paranã Rivers (see Figure 1.3 and Plate 1.3.8), formed by rivers that overflow during the wet season or by superficial accumulation of rain water above underground limestone (see Plate 1.3.8). Also at low elevations, greenish colours represent the core bird assemblages within dry forests and arboreal savannas (Figure 3.5B). In the north-west are also assemblages at the lowest elevations (~ 247 to 300 m) in the Basin (~ 99 species), a region dominated by arboreal savannas interspersed by dry forests.

In the axis south-west to north-east (~ 40 to 60 species) – in the ‘Vão do Paranã’ plain (Figure 1.3 and Plates 1.3.2A and B) – are located the bird assemblages at intermediate elevations (430 to 585 m), dominated by dry forests interspersed with savannas. Transitional assemblages (brown colour in Figure 3.5B; ~ 183 species) are in the areas of the montane gradient in the central-western region that extends in an arc towards the south, following the watershed scarps and the highest elevations of the valley - on plains and low-relief hills adjacent to nearby mountains - and north again, following the scarps and the those plains and low relief hills of the eastern border of the Basin.

For mammals, montane and transitional assemblages are found on montane gradients (represented by reddish colours in Figure 3.5C). Pink areas in this Figure (~ 21 and 23 species) are the assemblages located at the highest elevations of the Basin. Transitional assemblages on the montane gradient are represented by yellow (~ 28 species), orange (~ 15 species), and brown (~ 22 species). Greenish colours represent the assemblages of the low parts of the Basin in areas dominated by savannas interspersed with dry forests in the northern and southern regions of the Basin (~ 37 species) and areas dominated by dry forests in the central-eastern region (~ 31 and 38 species). Bluish colours in the northeastern and southeastern portions of the Basin (~ 15 and 22 species) represent the assemblages associated with savannas and grasslands over acid sand soils and semi-deciduous forests in gallery forests.

Across all three taxonomic groups, the study area exhibits relatively high compositional turnover (beta diversity), that is, site-pairs share relatively few species across environmental gradients. Composition turnover is indicated by a non-linear function of environmental variables and geographical separation for birds, mammals, and plants' models based on all samples and sub-samples of well-surveyed sites (Figure 3.6). The fit between observed and predicted dissimilarity between site pairs is represented by the curved lines, which are the inverse-link functions of the models. This figure provides an overview of the degree of scattering in the data, that is, how well the predicted compositional dissimilarity between site pairs (x-axes) matches the actual compositional dissimilarity (y-axes) [101]. For the full models for birds, mammals, and plants, for which the number of species per site varies from a few to many, the wide scatter of points indicates a low concordance between predicted and observed compositional dissimilarities. In the sub-samples, the species' number per site is relatively even distributed, resulting in better-fitted models.

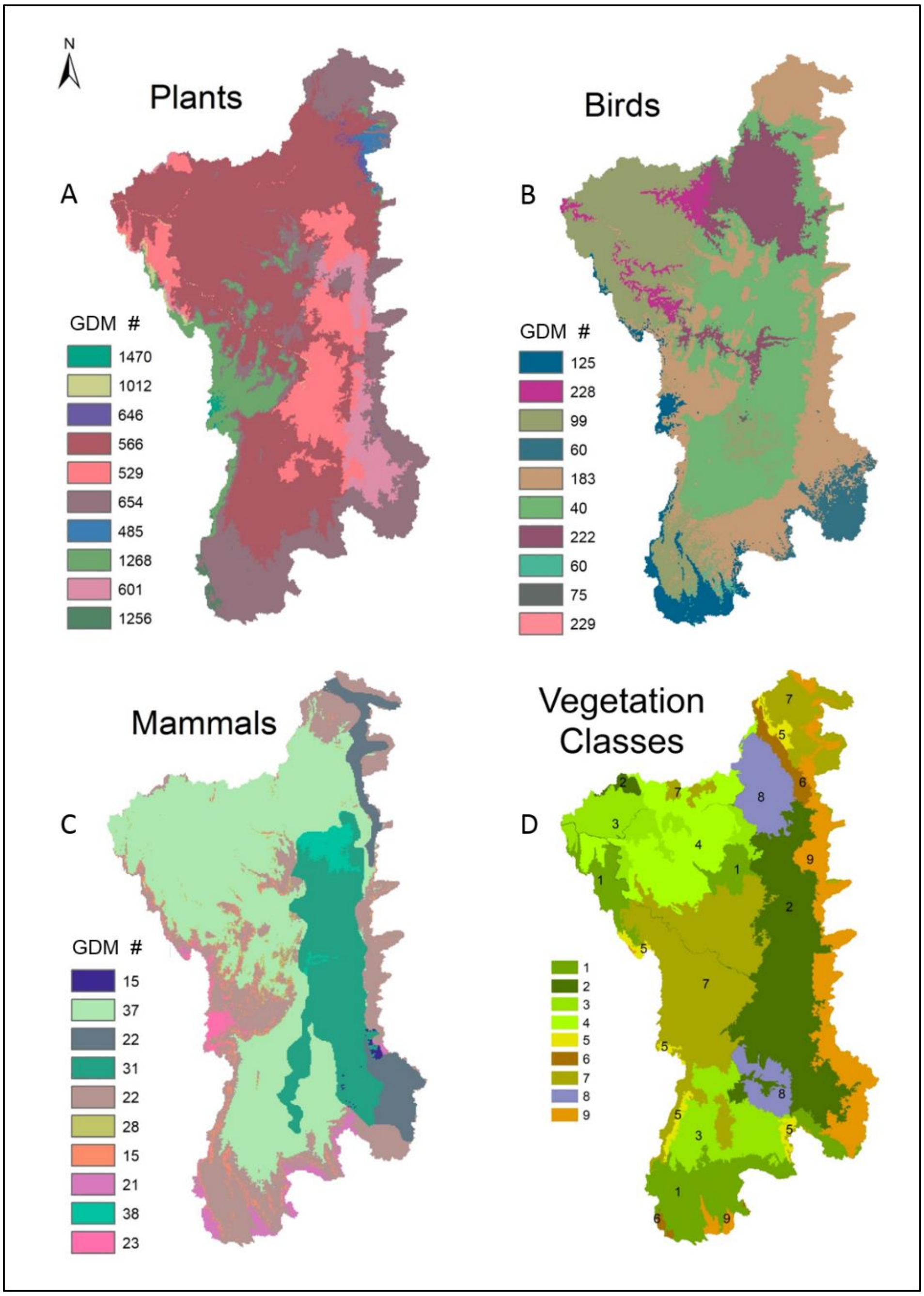


Figure 3.5. Spatial pattern of assemblage composition for birds (A), mammals (B) and plants (C) within the Paranã River Basin, derived by applying metric multidimensional scaling to compositional dissimilarities predicted by a GDM model fitted to all samples. Previously mapped vegetation classes are in D.

In A, B and C, grid cells mapped in a similar colour are predicted to have similar species composition, while cells mapped in a very different colour are predicted to be highly dissimilar in composition.

Numbers on legends (#) represent the average richness estimated in each of the classes. In D, vegetation classes are: 1 = savannas/semideciduous forest, 2 = deciduous forest, 3 = savannas/deciduous forest, 4 = savannas/grasslands/deciduous and semideciduous forests, 5 = grasslands, 6 = grasslands/semideciduous forests, 7 = savannas and grasslands, 8 = savannas in seasonally flooded land and 9= savannas/grasslands.

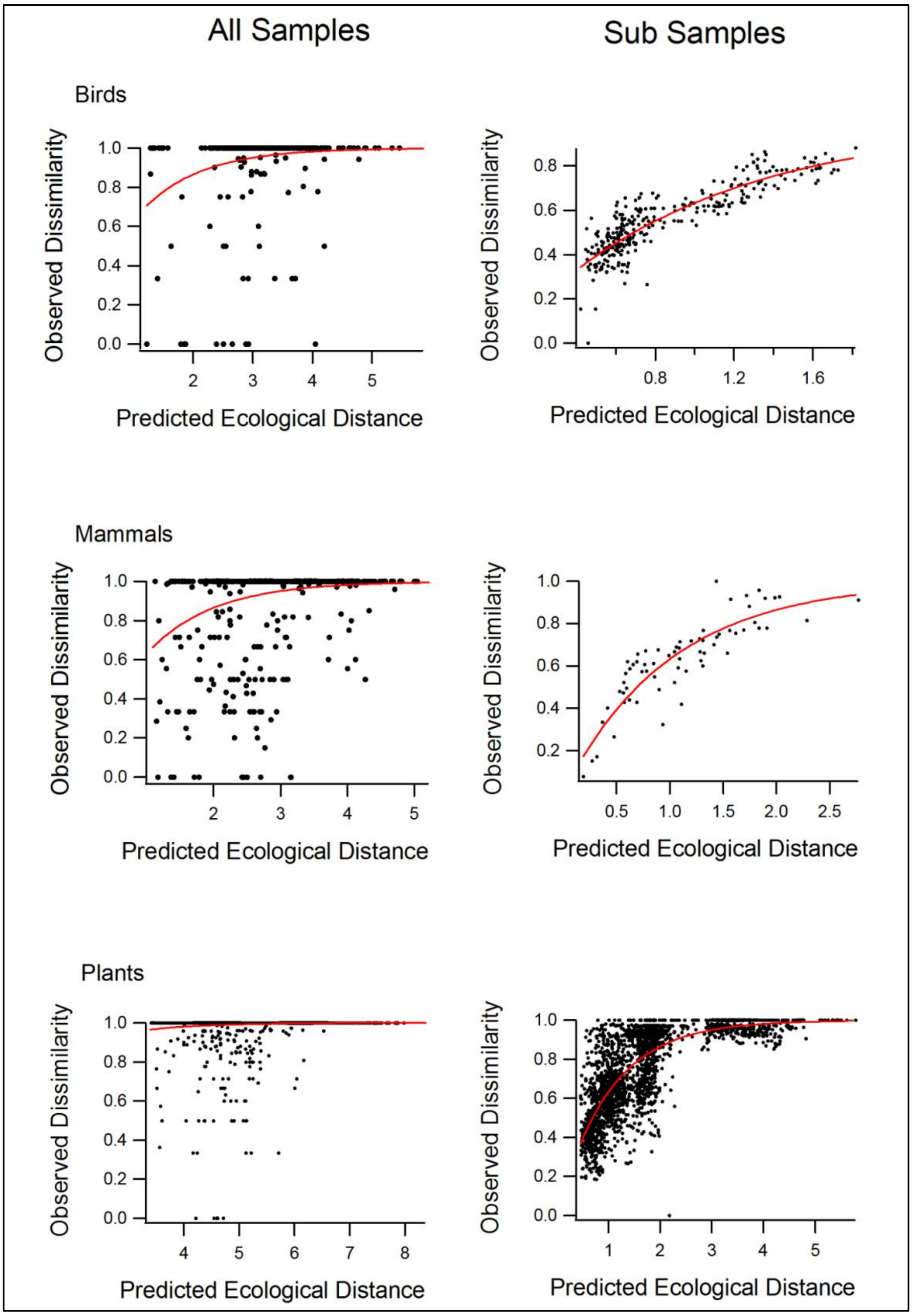


Figure 3.6. GDM link functions relating predicted ecological distance and observed dissimilarity for birds, mammals, and plants considering all samples and the sub-samples (adequately inventoried sites used to test the robustness of the GDM models developed using the complete dataset).
The scatterplots depict the relationship between environmental/ecological distance as predicted by the models and observed species dissimilarities between site pairs based on the Bray-Curtis measure. Each dot represents a site-pair, and red lines represent a perfect fit for the transformed ecological distance, i.e., the non-linear link function relating dissimilarity to environmental/ecological distance (predicted compositional dissimilarity).

3.5. Discussion

Based on species richness, assemblage similarities and turnover of plants and terrestrial vertebrates, at least three distinct regions could be identified across the Paranã River Basin: the montane gradient in the mountain range of the western side, and the relatively flat terrains of the low parts of the Basin in the north and the central-eastern regions. The mountain range has the highest richness values for plant species (Figures 3.1 and 3.2), unique species assemblages of plants, mammals, and birds (Figure 3.4), and high species turnover of plants. The highest richness values for all vertebrates (Figures 3.1 and 3.2) and unique species assemblages for birds (Figure 3.4B) are concentrated in the depression of the northern portion of the Basin. In the central-eastern region of the Basin, at intermediate elevations (430 to 585 m) of the 'Vão do Paranã' plain (See Figure 1.3), there is high richness of vertebrates (Figure 3.1 and 3.2) and unique assemblages of plants, birds and mammals (Figure 3.4). Recognition of these these three major regions in conservation planning is important if the biodiversity of the Basin is to be adequately represented in conservation areas.

The combined use of SDMs and GDMs in the Paranã River Basin identified different dimensions of biodiversity that could not have been captured if one of these methods had been used alone. Conservation planners need to include areas of high diversity (α diversity), critical habitats for species, areas that have unique assemblages (probably with local restricted endemism), and zones of high species turnover to adequately represent the biodiversity pattern of a region, as well as promote the persistence of ongoing evolutionary processes (β diversity) [13, 100, 145, 146]. In this study, the SDMs revealed two general contrasting patterns between richness distributions of plants and vertebrates, irrespective of the accuracy of the models (Figures 3.1 and 3.2). The GDMs showed some degree of congruence between the distributions of assemblages in the different species groups and, within the same taxonomic groups, highlighted regions of similar species composition, unique species assemblages, and regions of high species turnover.

3.5.1. Species-level modelling

Despite the longstanding debate within the ecological literature about the determinant factors of geographical richness patterns [125, 147], there still is little consensus about the ecological mechanisms that generate and maintain them [148], since no single mechanism adequately explains all examples of observed patterns [125, 148-151]. Furthermore, richness patterns can vary with spatial scale. Processes operating at a regional level have a nested influence upon patterns found at the local level. Therefore, the relative balance of causal mechanisms means that there will invariably be differences in and exceptions to any proposed causal relationship. Overall, though, the main factors proposed to be associated with richness patterns are interactions between climate, habitat heterogeneity and structure, time for diversification, and the size and shape of habitat areas [125, 130, 148].

In the Paranã River Basin, contrasting richness patterns between plants and vertebrates seem to be regulated by different processes. Richness of plants (see Figures 3.1 and 3.2) might suggest a regional pattern that includes the Cerrado biome as a whole. In the mountain range, there are unique combinations of humidity, temperature, light exposure, and different types of substrate (Figure 1.2 and Plate 1.2) that are not present in the lower parts of the Basin (see Table 1 in Supporting Information 2). The combination of these factors produces high levels of abiotic heterogeneity, resulting in a patchy, heterogeneous, and highly diverse range of vegetation types. This all contributes to the highest values of plant richness observed in the Basin. This pattern, however, is not reflected in richness of vertebrates. The prevalence of structurally simple vegetation (grasslands and open savannas) in the mountain landscape, as well as possible environmental limitations, might be responsible for this pattern. Additionally, portions of the montane gradient (775 – 1000 m) of the mountain range are still relatively poorly known for vertebrates compared to the low parts of the Basin [58], so sparse field sampling might also be influencing the observed pattern.

From the environmental variables evaluated, consistent water availability across the year, represented by the higher precipitation (mainly during the dry season) in the mountain region, is suggested as the main driver of high richness for plants in that region (see Table 3.4). The mountain range acts as a barrier to the northwesterly winds that bring humidity and rains from the Amazon Basin. The range retains more water on the windward side and at the highest elevations; warming and drying are higher on the lee side of the range and in the valleys. Rains during the dry season lead to less stressful conditions in water availability for plants during the winter period.

In contrast, the combination of climatic stability, unique topography, and a biogeographic mixing of three distinct faunas (Amazon, Cerrado and Caatinga) [58] are the suggested drivers of the observed richness patterns of vertebrates. Thus, this pattern could represent the local variation of a fauna adapted to the conditions of the Basin’s bottom that is not represented elsewhere in the Basin or the Cerrado biome as a whole. Huge incrusted depressions within the Brazilian central plateau, like the Paranã River Basin, are nonexistent elsewhere in the entire biome. In this depression, higher and stable temperatures found in the low parts of the Basin are correlated with high richness of vertebrates (see Table 3.4). The depression combines environmental variables that might provide a unique environment, related to the high number of endemic species observed [32, 34, 36, 40, 58, 74, 152, 153]. Higher average annual temperature (27.1 ^{0}C) and less pronounced temperature seasonality are found at the Basin’s bottom. This contrasts with the cooler (19.5 ^{0}C) and more seasonal temperature regime of the mountains (see Table 1 in Supporting Information 2), which could pose a constraint on fauna distributions. Additionally, the presence of more structured environments represented by a predominance of dry forest formations (deciduous and semi-deciduous) and arboreal savannas at the Basin’s bottom, interspersed with small patches of open areas, favours the co-occurrence of species from disparate biomes (Caatinga, Amazon and Atlantic) [33, 35, 58, 74]. In a broader context, this biogeographic mix contributes to increasing the number of species, even compared to other parts of the Cerrado biome in Brazil [58].

3.5.2. Assemblage-level analysis

Patterns of assemblage dissimilarity and species turnover (GDMs) highlighted aspects of biodiversity not detected in species-level modelling (SDMs) of birds, mammals, and plants. In contrast to SDMs, GDMs followed the same general spatial pattern for all groups evaluated, with only slight differences between major groups. Assemblages located in the mountain range in the western portion of the Basin (Serra Geral do Paranã range – Figure 1.3) are dissimilar from those of the valley, irrespective of the taxonomic group (Figure 3.5). Additionally, this mountain region was identified as one of the centres of species endemism within the Basin [58] and in the Cerrado biome [62, 95, 154]. These assemblage differences extend in an arc shape towards the south following the watershed scarps of the range, and north again following the higher terrain of the eastern border (Serra Geral de Goiás range – Figure 1.3). The difference between assemblages in mountain areas and valleys marks the major contours of species turnover within the Basin.

These compositional patterns are influenced mainly by the same variables that appear to regulate differences in richness: precipitation, temperature, and habitat heterogeneity and structure (Figure 3.4). The difference here, when compared to SDMs, is that geographic distance was included as an explanatory variable, and contributed significantly to explain dissimilarities and higher rates of species turnover within all groups (birds, mammals and plants) in the montane gradient of the central-western mountains and the southern boundary of the Basin, for example. Geographic distance provides supplementary information about the responses of composition to latent variables, evolutionary history, dispersal barriers, and non-equilibrium conditions [112] not considered directly in the analyses.

Overall, variations in precipitation across the Basin seem to be primarily related to differences observed in species assemblages, followed by changes in other variables (see Figure 3.4). The highest areas of the Basin that mark the major contour for species turnover receive more precipitation during different seasons of the year (see Table 1 in Supporting Information 2). In these areas, differences in composition primarily follow the montane gradient in climate – availability of water and temperature – and secondly the habitats – differences in substrate for plants and in substrate and vegetation structure for animals (Figure 3.4).

Distinct assemblages for all groups are delimited clearly across the montane gradient that forms part of the Basin's perimeter in an open arc shape towards the north. The typical assemblages of the highlands were mapped for plants (in green areas in Figure 3.5A), birds (in dark blue areas in Figure 3.5B), and mammals (in pink areas in Figure 3.5C) in regions that receive the highest precipitation during the year. The transitional assemblages between the valley and the highlands were mapped for plants, birds, and mammals in brownish areas in Figure 3.5A, B, and C (dark brown for plants and light brown for birds and mammals) across the slopes of the mountain ranges and on adjacent plains and low-relief hills. The transitional assemblages are associated with an increase in temperature and decrease in precipitation from the highlands to the bottom of mountains, with vegetation types replacing one another across with differences in soil content and permeability to water [58].

Assemblage dissimilarities within the biota in the relatively flat terrain at low elevations (Tocantins depression in Figure 1.3) and intermediate elevations (Vão do Paranã plain region in Figure 1.3) are lower (reddish colours for plants and greenish colours for birds, and mammals in Figures 3.5A, B, and C, respectively) than those in mountain regions for all taxonomic groups. This pattern is explained by the lower environmental complexity in the valleys. However, species turnover is still evident in these regions and is consistent for the groups evaluated. Higher precipitation in summer contributes to the formation of permanent and intermittent lagoons along the Paranã and Palma Rivers and their affluents in the region of the Tocantins depression (see Figure 1.3). This environment provides suitable conditions for migration and nesting of Amazon and Atlantic birds species and coexistence of terrestrial and aquatic birds species [152]. Migratory species and those strictly related to aquatic environments add to the avifauna of the savannas and dry forests, resulting in the high number of species estimated and a highly distinct bird assemblage (reddish colours in Figure 3.5B).

Higher precipitation in the stair-stepped terrain of the eastern border of the Basin is also related to singular assemblages of mammals and plants (Figure 3.5.B). Distinct assemblages of mammmals are located in the plains and low relief hills of the eastern border of the Basin (bluish colours in Figure 3.5B) where typical and open savannas on sandy soils predominate. Singular assemblages of plants are located on the extension of the 'Chapadão Ocidental da Bahia' plateau into the Basin (Figure 1.3) and in the plains and low limestone hills of the northeastern border of the Basin (blues area in the Figure 3.5A) that are covered by open to typical savannas (on plateaux and plains) and dry forests (on limestones outcrops).

Finally, distinct assemblages of birds (represented by light green areas in Figure 3.5B), mammals (dark green areas in Figure 3.5C), and plants (light red and pink areas in Figure 3.4A) are associated with the predominance of dry forests on flat terrains and limestone outcrops of the Vão do Paranã plain region of the valley (Figure 1.3). This region is recognized by the significant endemism of flora and fauna and occurrence of species from the drier Caatinga biome [31-41]. Additionally, the slightly different colours (light red and pink areas in Figure 3.5A) mapped for plant assemblages in the same region is mainly related to the predominance of dry forest in flat areas (light red) and on limestone outcrops (pink areas). A study comparing tree species in those environments has already pointed out differences in species composition and forest structure between the dry forests sampled in the flat areas and limestone outcrops [31].

3.5.3. Uncertainties in models

Even with high values of AUC, richness maps based on all SDMs (Figure 3.1) might contain errors. AUCs can be inflated by many ways [108, 155]. Species recorded in few locations and biases in species collections, both characteristic of the dataset used to build these models [58], might have inflated AUC values [108, 155] observed in the completed models (Table 3.3). Methods using only presence data as input for modelling species distributions (such as Maxent) assume that samples are unbiased, a premise easily violated when using presence-only data [118, 133]. Most of presence-only data come from incidence records derived from herbaria and museum specimens that might be biased towards accessible locations or better-surveyed areas [118], and are rich in species sampled in few locations. In models based on species recorded in few locations, the true niche-based signal might not be correctly established, leading to inflated AUC values [155]. For models based on many localities, the environmental signal of species’ niches is more likely to be captured [155]. In this sense, the main problem for modelling species distributions is not the spatial bias of the records but a bias in sampling of environmental conditions associated with species distributions [155].

Several methods have been proposed to correct or, at least, minimize the sampling bias effect: (1) choosing a reasonably large number of random background sites (where there are a large number of presence records that might cover environmental variation); or (2) manipulating the background to fit the original distribution of the species sites sampled by restricting background samples to the known habitat of the species (for example, by limiting background samples to distance from water for riparian forest species), or by creating species-specific bias files representing the relative sampling effort or record density, or by using target-group backgrounds [see, for instance, 108, 118, 133, 134, 155, 156-159]. However, background manipulation can be impractical for a large number of species from multiple taxa, for which habitat specificity is unknown.

In this study, the effect of sampling bias might still be present in models due to an uneven collection effort in the montane gradient (775 – 1000 m) [58], possibly underestimating faunal richness in those locations. Across the whole Basin, records for plants and birds are well distributed (see Figure 2.5), but those for mammals, reptiles, and amphibians less so. Sampling bias arising from incidence records from herbaria and museums (~ 60% of records present in the dataset) is partially compensated by adequately sampled sites inventoried across the Basin (~ 40% of records) [58], most of them designed to target poorly known habitats and species assemblages, or to assess the impacts of development projects (such as dams, railways, and roads) [58]. Additionally, most of the species sampled in few locations in the dataset are, in fact, species with restricted spatial distributions that were targeted and searched for in unusual habitats, while more common species were more easily detected in an opportunistic way.

The sampling bias and number of samples are of least concern for the models using a large number of species-site records and tested for accuracy. For other models, the target-group background adopted in this study might not completely circumvent the problems of: (1) AUC inflation [108, 155]; and (2) mistakes in detection of relevant predictors, which is influenced by the sampling process [155] and often biased towards accessible locations (near roads or towns) [155, 157]. The contrast of using species well sampled or under-sampled to estimate richness patterns is evident for mammals and reptiles (Figures 3.1 and 3.2). However, the low number of species used to build the tested models might also explain the differences in predicted richness of mammals. Of the four species used for tested models, three (maned wolf, giant anteater, and three-banded armadillo) live primarily in grasslands and open savannas, which are the dominant vegetation types in the areas of higher predicted richness. The fourth species (the South American tapir) lives mainly in forests following the watercourses. The same problem might apply for reptiles, of which only seven species were used to build the tested models. In this case, the pattern of richness was relatively flat, but higher towards the Basin´s southeastern regions (see Figure 3.2). This pattern might be explained by the presence of generalist species, such as the giant Ameiva and the Amazon lava lizard, or species more restricted to rocky environments as *Tropiduros torquatos* (calango), a species frequently found in limestone outcrops, but that is distributed widely along the eastern portion of the Basin. Additionally, higher overall richness predicted by the tested models in the southern part of the Basin is not corroborated by richness observed and estimated in adequately sampled sites in four different locations (Location 1 in Figure 2.1). Sites located in the central-northern portion of the Basin (Location 3 in Figure 2.1) had high numbers of species, irrespective of the group considered (insects, amphibians, birds, mammals, reptiles, and plants), compared to sites sampled in the other three locations [42]. In general, the richness in adequately sampled sites corroborates the fauna richness patterns predicted for the low altitudes of the Basin using the full models, which are probably more realistic than the tested models for these faunal groups.

Conversely, for plants and birds, tested models and full models presented much the same richness patterns with only slight variations. The more evenly distributed richness values observed in the tested models for plants can be explained by the inclusion in the full models of a large number of species with restricted distributions in the distinct environment of the higher elevations of the Basin. The same applies to the tested models for birds, resulting in somewhat even richness towards the northeastern and southern regions, with highlighting the montane gradients of the western parts. A high number of vagile species occasionally sampled in perennial and intermittent lagoons along the Paranã and Palmeiras rivers (Figure 1.3) that are not found in others parts of the Basin were used to build the full models and also contributed to the more evenly distributed richness values observed in the tested models for birds. Considering the above, we can make some assumptions about spatial richness patterns. First, both tested and non-tested models for plants and birds seem to be informative about actual richness patterns (Figures 3.1 and 3.2, respectively). Additionally, faunal groups (birds, mammals, reptiles, and amphibians) seem to have similar patterns of richness (see Figure 3.1). In this sense, it is possible that, for amphibians, reptiles, and mammals, the richness patterns between tested and non-tested models would be similar to that of birds if they had have been subject to the same sampling effort.

Notably, though, stacking species distribution models by thresholding continuous distribution maps to presences and absences has been criticized recently for over-predicting species richness in species-poor cells, and underestimating richness in species-rich cells [155, 160]. However, this bias is likely to be more a statistical than an ecological issue [160]. In this study, the use of a threshold to stack the SDMs using the cutoff results in models that give the most similar results to the continuous model outputs [139] may not offer a complete understanding of richness in particular cells; but it is acceptable to identify the general *n* bests cells (or generally species-rich cells) [155] across species. Consequently, the general spatial richness patterns in this study are probably realistic.

In an exploratory analysis, differences in depicting species-rich cells between stacked SDMs based on continuous values (see additional Figures 1 and 2 in the Supporting Information 2) versus thresholded binary values (see Figures 3.1 and 3.2) can be observed. Summing continuous values seems to be more accurate in pointing out species-rich cells (red colour in additional Figures 1 and 2 in Supporting Information 2), but can produce patterns resulting from minor relative probabilities across species, which might not be necessarily indicate high species richness. Nevertheless, if the the intermediate and moderately high values of the relative probability of species occurrence derived from the continuous values are considered a good signal for species-rich cells (represented by yellow colour in additional Figures 1 and 2 in the Supporting Information 2), richness patterns between binary (red cells) and continuous (yellow and red cells) methods are quite similar. Thus, the spatial patterns of richness across the Basin are preserved to some extent for all taxonomic groups, irrespective of the method used. However, further evaluations of the similarities and dissimilarities of richness patterns between continuous and binary methods are needed for a thorough understanding of their respective strengths and weaknesses.

Regarding the GDMs, the significant scatter observed in the full models of birds, mammals, and plants denotes the influence on the analysis of a large number of sites with few species per site, resulting in relatively small deviance explained by the GDM full models. Aggregated data from multiple sources is likely to have substantial spatial and environmental bias, masking robust signal variation as found in analyses restricted to adequately sampling sites [112]. Two methods have been proposed to circumvent this problem: (1) applying a weight to the regression based on the number of species recorded at each site pair; or (2) using a minimum number of species per site as an indicator of sampling adequacy for sites to be included in the analysis [101, 112]. We subsetted the dataset and compared the results of models derived from all samples and sub-samples (as described in the Methods section), concluding that the GDM models for birds, mammals, and plants were sufficiently reliable for further biodiversity assessment.

Aware of the challenges of working in highly diverse but poorly known areas and the limitations and uncertainties of modelling under these circumstances, we consider nonetheless that the full models of species and assemblages are reliable for preliminary conservation planning at a fine resolution. Most of all, the overall spatial patterns depicted by the models matched our previous empirical information on biodiversity patterns in terms of both species richness and assemblages.

3.5.4. Species and assemblage-level approaches: pros and cons

Planning for the conservation of multiple taxa requires, at minimum, reliable information on their distributions. While, in detail, the SDMs and GDMs produce complementary results in assessing biodiversity patterns, it was possible to find congruencies between them in depicting regions across the Basin that are with both high richness and distinctive assemblages. The montane gradient of the mountain range on the western border, the savannas and dry forests in the northern section of the Basin, and the dry forests and the transition zone with savanna formations in the central-eastern region are some examples. All these areas contain distinct associations of species, high richness values, high spatial turnover of species, besides the presence of a large number of endemic and threatened species [58]. The congruence in results from disparate approaches in identifying these biodiversity patterns increases confidence in considering these regions 'a priori' as important for conservation of biodiversity within the Paranã River Basin.

These results also reveal the vulnerabilities of both approaches in assessing biodiversity patterns for conservation when they are used alone. Although SDMs are useful to evaluate spatial richness patterns within a region, they do not offer any information about assemblages. Additionally, using taxonomic groups as surrogates for others in conservation should be done with caution [161]. Within the Paranã River Basin, for example, areas with high richness of plants cannot be used as proxies for areas of high richness for birds, mammals, reptiles, or amphibians, and vice-versa (see Figure 3.1) since different environmental variables regulate patterns of richess (see Table 3.1). However, vertebrate groups showed a high correlation between richness patterns (Table 3.1), especially between birds, mammals, and amphibians. In any case, richness patterns alone are of more importance for understanding biogeographic patterns and environmental relationships than for conservation planning, which relies more heavily on composition.

The sensitivity of the GDMs in identifying differences in species assemblages, even within the same phytophysiognomies, indicates the importance of an assemblage-level approach in defining sub-regions for biodiversity conservation. Accounting only with this method, however, would imply selecting areas for conservation that did not consider the distributions of individual species.

3.5.5. What have we learnt?

The montane gradient on the western side of the Basin, and the northern and central-eastern flat terrains of the lower elevations in the Basin are the most important regions for biodiversity conservation. Given the landscape features in the mountain range, these areas are naturally protected making intensive use by agriculture difficult. In contrast, the relatively flat areas of the northern and central-eastern portions of the Basin are the most vulnerable to transformation, which demands urgent conservation efforts. The north of the region encompasses the most pristine areas of the entire Basin, while in the central-eastern region are concentrated the last remnants of one of the most threatened ecosystems in Brazil, the dry forests [31, 58]. Additionally, the entire northern region of the Basin is now covered by the new Brazilian land development program for agriculture – the MATOPIBA Project – that aims to extend the frontiers and facilities for farming activities in the Cerrado Biome occupation [162]. However, any established or proposed conservation actions are considered in this project, making the northern region of the Basin suitable for investigating new conservation areas.

This study is the first substantial effort to understand the general biodiversity patterns of a region considered a global priority for biodiversity conservation [28, 29]. Investments in data collection on fauna and flora still necessary, in particular for groups that could not be properly evaluated in this study due to lack of species records, especially insects, reptiles, and amphibians. The poor state of biological information for some groups contributes limitations and uncertainties to modelling.

Chapter 4

Chapter 4 – Conservation in a biodiversity hotspot: a multi-solution approach to conservation prioritization

Keywords: SDM, Maxent, GDM, Zonation, Paranã River Basin, Conservation Planning

Article type: Full-length Article

Anderson C. Sevilha[1,2], Robert L. Pressey[3], Stephen E. Williams[1], Jeremy VanDerWal[1,4], Helga C. Wiederhecker[1], Barbara J. Anderson[1,5] and Atte Moilanen[6]

[1] Centre for Tropical Biodiversity and Climate Change, School of Marine and Tropical Biology, James Cook University, Townsville, QLD, Australia;
[2] Embrapa Genetic Resources & Biotechnology, Brasília, DF, Brazil;
[3] Australian Research Council Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, QLD, Australia;
[4] eResearch Centre, James Cook University, Townsville, QLD, Australia;
[5] Rutherford Discovery Fellowship, Landcare Research, Dunedin, New Zealand;
[6] Dept. of Biosciences, University of Helsinki, Helsinki, Finland.

Chapter 4 – Conservation in a biodiversity hotspot: a multi- solution approach to conservation prioritization

4.1 Abstract

Our main goal was to produce a spatial prioritization based on biodiversity value to inform conservation planning in the Paranã River Basin, an important region for protection within a global biodiversity hotspot in the Cerrado Biome, or Brazilian savanna. We formulated area prioritization solutions using species distribution models (SDMs) from Maxent and models of assemblages from Generalized Dissimilarity Modelling (GDM). The objective was to maximize the value of biodiversity features in areas selected for conservation actions, considering endemic and threatened species, unique and rich species assemblages, and threatened ecosystems. We weighted biodiversity features using the premise that: (i) the persistence of rare, threatened, and endemic species relies more on conservation actions than the persistence of common species; (ii) greater weights should go to areas with higher species richness and uniqueness; and (iii) threatened ecosystems require more conservation efforts than non-threatened ecosystems. Using these weighted biodiversity measures as input, we developed six different conservation-prioritization planning solutions in the Zonation software. To understand the contributions of species and assemblage distributions to prioritization of areas, we used three solutions to evaluate the separate contributions of species distributions (SDMs), assemblage turnover (GDMs), and combining species and assemblages (SDMs and GDMs). We primarily considered the better-sampled taxonomic groups (birds, mammals, and plants). However, we also evaluated the additional contributions to the analyses when we incorporated the poorer information on the species distributions of amphibians and reptiles, and information on threatened ecosystems. We also compared the prioritization solutions generated with a previous prioritization of protected areas by the Brazilian Government and made inferences about the use of different approaches to area prioritization.

4.2. Introduction

One of the biggest challenges of conservation planning in tropical areas is to maximize the representation of biodiversity within a heterogeneous, species-rich landscape and a complex political and social land-use matrix. Planning for conservation in such regions confronts the lack of field data for planning and demands appropriate procedures to identify conservation area networks with limited data [9, 14, 100]. Well-considered planning procedures are essential to achieve representativeness, complementarity, connectivity, and persistence of biodiversity features and other natural values in the planning region [9, 10, 73, 163]. Given the multiple land-use pressures often present in these landscapes, a spatially-explicit, transparent planning approach is crucial to both technical effectiveness and real-world applicability.

A sensible approach to face the lack of biodiversity data is to use a combination of biodiversity surrogates. Surrogates include, for instance, species distribution models, species assemblage models, patterns of richness and species assemblages, habitat types, taxonomic surrogates such as endemic and threatened species, environmental variables, and spatial representations of ecological processes [9-11, 100, 164]. Conservation planning is primarily based on a spatial prioritization approach. It determines priority areas in the landscape based on defined goals such as the maintenance of maximum biodiversity values [165]. However, which biodiversity features to use and how to weight them during the planning process depends not only on the goals of the conservation program but on the quantity and quality of available data.

Prioritization of areas is just one step in the systematic conservation planning process [see (14) for a complete description of conservation planning process]. It requires the development of knowledge that takes into account the distributions of species and assemblages and how threatening processes might affect these distributions [9]. One of the pervasive constraints in planning for biodiversity conservation is the limited availability of good data on distributions of species and assemblages, particularly in tropical regions [166]. In the absence of such information, species distributions models (SDMs) are an alternative to filling the gap in spatial information on biodiversity (the

Wallacean shortfall [97, 167]) to support conservation planning. SDMs also give information about the suitability of areas for species as a function of climate, landscape, and substrate variables and allow estimation of patterns of species richness across the study domain. Complementary to understanding spatial variation of biodiversity at the species-level are analyses that evaluate and predict spatial patterns of turnover in assemblage composition [101]. These analyses can identify assemblages that are both rich in species and distinctive in composition, and consequently, regions where evolutionary processes are potentially more significant [168].

However, all species or other biodiversity features are not of equal concern in conservation planning, so should be weighted depending on the identified goals of the particular planning program [4, 164]. Endemic and threatened species and threatened ecosystems, vegetation types, and land units are some examples of biodiversity features that have received more attention in conservation prioritization across the world [29, 164, 169-171]. Nevertheless, the specifics of weighting cannot be generalized for all study regions, and the weighting of biodiversity features varies according to the aims of the conservation plan and the planning context shaped by region-specific information on biodiversity and threats [172].

The use of both species- and assemblage-level data in conservation prioritization has been examined previously for both simulated and real data [173, 174]. However, in these studies SDMs were used to calibrate assemblage models for richness and uniqueness during the prioritization, that is, to determine total species and unique species present within each assemblage. There is no published study evaluating and comparing the use of both data types separately or simultaneously in spatial prioritization. In this study, the main goal is to assess the relative effectiveness of species and assemblage-level analyses in incorporating biodiversity features into a proposed reserve network, giving particular attention to endemic and threatened species, unique and richer species assemblages, and more threatened ecosystems. To promote the retention of features of greater conservation concern, we weighted species based on their regional and national rarity and taxonomic distinctiveness, assemblages based on their richness and uniqueness, and ecosystems based on their threats.

To understand how species- and assemblage-level data shape prioritization of areas, we evaluated in separate prioritizations the contributions of species, assemblages, and both data types in combination for the same taxonomic groups. For each prioritization, we compared the retention of biodiversity features in high-priority areas. For comparisons of species and assemblage data, we used only taxonomic groups that could be modelled by both approaches (birds, mammals, and plants). Additionally, to assess the information gain in area prioritization by adding extra features to the prioritization based on species- and assemblage-level data, we included information on species distributions from other taxa, which could not be evaluated by assemblage models due to the lack of data, and for ecosystems. Finally, we compared our conservation prioritizations with the areas identified by the Brazilian Government as priorities for the establishment of protected areas in the Paranã River Basin [29]. We also measured the overlap between priority areas from this study and the existing protected area system.

4.3. Methods

4.3.1. Overview of data and sources

In this study, we used three sets of biodiversity data features in area prioritization: a set of individual species-level distributions; a set of species assemblages; and a set of ecosystem types. Methods used for modelling species and assemblage distribution are briefly described in the next sections [for a complete methods' description see (58), 175]. The models were based on an intensive process of data gathering and filtering biological information available for the region.

For the first data set, we fitted habitat suitability models to existing species data, and we created spatial predictions of biodiversity features across the planning region [175]. For prioritization, species were weighted according to three values: 1. their AUC (area under the curve) values to reflect the discrimination ability [134] of their models; 2. level of endemism; and level of threat.

For the second data set, we used assemblage-level analysis [175] to maximize complementary representation of assemblage types, weighted by their richness and uniqueness. The complementarity approach utilized information about similarity, i.e. shared and unshared species across assemblage classes [172].

For the third data set, we used ecosystem types, weighted by identified threats posed by anthropogenic activities. The identification of ecosystems and threats was carried out by specialists during the Workshop ‘Priority areas for the conservation, sustainable use, and benefits sharing of Brazilian biological diversity’ [29], organized by the Brazilian Ministry of the Environment in 2006, to evaluate priority conservation areas in the Cerrado and Pantanal biomes. The aims of this workshop were to validate the priority areas previously assigned and identify, for each area highlighted as a priority, its conservation importance, type of conservation priority, opportunities for conservation, and actions needed to maintain biodiversity. For each area, threats and necessity of protected areas of varying categories (fully protected, sustainable use, or undefined – a type that requires further assessment to be categorized) were also addressed. Three ecosystem types were identified during the workshop based on vegetation maps (http://mapas.mma.gov.br/i3geo/datadownload.htm) and satellite images for each priority area. The ecosystems identified were then rated according to their vulnerability to threats (in the case of the Paranã River Basin: dry forests > savannas mixed with dry forests > savannas). Secondly, types of threats within each ecosystem and their coverage of each ecosystem were identified. These included mining, intensive or extensive cattle-ranching, intensive or extensive agriculture, and contamination from fertilizers and agrotoxins.

To assign threat values to ecosystems for our prioritization analyses, we used the ‘threats’ attribute of the shapefile of the workshop’s priority areas maps (http://www.mma.gov.br/biodiversidade/biodiversidade-brasileira/%C3%A1reas-priorit%C3%A1rias/item/489) to identify threatened areas within each ecosystem. Threat values were then assigned according to pressure on each 0.01 degree grid (~1km) resolution area within each ecosystem in categories from ‘1 to 5’ (low, low-intermediate, intermediate, high-intermediate, and high). For areas within the Basin but outside the scope of the workshop (~ 18% of the ~ 60,000 km^2 of the Basin), the degree of threat was unknown, and we assumed the lowest value of ‘1’. Finally, we assigned different weights to areas covered by each ecosystem according to the ecosystem’s vulnerability to threats (see Section 4.3.2 below).

4.3.1.1. Species distribution modelling (SDMs)

We modelled species distributions using a total of 2159 species (1813 plants, 245 birds, 44 mammals, 30 reptiles, and 27 amphibians) occurring in five or more geographically unique localities, aggregated at 0.01 degree grid (~1km) resolution. Of 70 threatened and 927 Cerrado endemic species sampled within the Parana River Basin [58], we modelled the distributions of 30 threatened and 208 endemic species. We estimated environmental suitability for each species as a function of climate, habitat structure, and substrate variables in the gridded study domain using Maxent [118, 132]. For area prioritization, we used two different sets of models that had AUC values ≥ 0.7: (1) models tested through a 10-fold cross-validation (for species sampled in 20 or more localities), and non-tested models (for species sampled in less than 20 localities).

4.3.1.2. Assemblage distribution modelling (GDMs)

We modelled compositional dissimilarities of assemblages and extracted a dissimilarity matrix between grid cells based on GDM – Generalized Dissimilarity Model – [101] using .NET software version 2.70 [142]. In assemblage analyses, we considered species records of 430 birds, 101 mammals, and 3467 vascular plants. Based on our knowledge of the vegetation community patterns in the Paranã River Basin, we set 10 classes for clustering the grid-cells of plants and fauna groups (birds and mammals). Additionally, each class was attributed values of average richness and uniqueness (number of species exclusive to each GDM class) based on richness maps derived from SDMs. We modelled assemblages of reptiles and amphibians too, but discarded them for conservation assessment due to the weakness of the data and their low power in explained deviance of the GDM models (see Chapter 3).

4.3.2. Area prioritization

Priority areas were selected using Zonation [172], a reserve selection software designed to identify networks of areas necessary for retaining high habitat quality for multiple biodiversity features [165, 172, 176]. The Zonation algorithm [177, 178] produces a hierarchical prioritization of the conservation value of a landscape. It determines the relative contribution of the total amount of each biodiversity feature in a given cell and then iteratively discards grid cells of lowest proportional value across all features, maximizing the retention of more highly weighted features. We used the Additive Benefit Function, as a rule, to remove cells [177], a function that has a higher performance on average over all species [172] (see Supporting Information 3 for a brief description on Zonation and the Additive Benefit Function as a cell removal rule).

For our multi-solution approach, we formulated different solutions to an area-prioritization problem using SDMs, GDM-derived assemblages, and ecosystems. In total, we developed six different area-prioritization solutions:

A. SDMs for birds, mammals, and plants
B. GDMs for birds, mammals, and plants
C. GDMs and SDMs for birds, mammals, and plants
D. SDMs for birds, mammals, amphibians, reptiles and plants
E. GDMs for birds, mammals, and plants, and SDMs for birds, mammals, amphibians, reptiles, and plants
F. GDMs for birds, mammals, and plants, SDMs for birds, mammals, reptiles, amphibians, and plants, and ecosystems defined by dry forest, savanna-dry forest complex, and savanna.

The first three solutions were intended to indicate the relative contributions of species and assemblages to prioritization of areas. For these comparisons, we considered only better-sampled taxonomic groups (birds, mammals, and plants). Solutions D-F were intended to evaluate the contribution of additional features to prioritization. For these, we incorporated information on the distributions of species in other taxonomic groups (amphibians and reptiles) and ecosystems mapped from the Government workshop [29].

Weighting biodiversity features is a critical component of the Zonation algorithm [172, 179]. By default, weights are equal, but features can be assigned differential weights based on factors such as perceived threat, endemism, rarity, taxonomic importance, economic value, or population trend [172]. Weighting features affects the order in which cells are removed from the landscape. Cells that include the distributions of valuable features (high weight) remain longer in the iterative cell removal process than cells only containing low-weight features, all else equal. Weighting influences the fraction of a feature's distribution retained at any point of the cell removal. Highly weighted features retain a relatively higher proportion of their distributions, all else equal. Weighting can also balances the amount of information available for different groups of features. In our case, weighting avoided area prioritization dominated by information on plants, given the larger number of species.

We weighted biodiversity features based on three considerations: (1) the persistence of rare, threatened, and endemic species relies more on conservation actions than the persistence of other species, (2) greater weights should go to assemblages with higher richness and uniqueness, and (3) threatened ecosystems require more conservation efforts than non-threatened ecosystems. Accounting for these differences, we set different weights according to which groups of features were included in the prioritizations, as summarized in Figures 4.1A and 4.1B. To evaluate the contributions of SDMs and GDMs on conservation reserve selection (Figure 4.1A), we established an overall weight of 50% for each of these groups. Within SDMs, weights were distributed among the taxonomic groups considering the number of species in each taxon (high number of species, high weight), quality of the models (based on models tested and non-tested for accuracy), and taxonomic groups as surrogates for other groups, in this case, plants as structural habitat for animal groups. Thus, plants, birds, and mammals received respectively 50, 30 and 20% of the weight given to SDMs. Within taxonomic groups, SDMs for species were weighted according to the formula:

$$\mathrm{W}_{\mathrm{sp}i} = \left\{\left[\left(\sum\left(\mathrm{IUCN}i_{\mathrm{sp}i} + \mathrm{EB}i_{\mathrm{sp}i} + \mathrm{EC}i_{\mathrm{sp}i}\right)\right) * \mathrm{AUC}i_{\mathrm{sp}i}\right] + 1\right\}$$

where:

$\mathrm{W}_{\mathrm{sp}i}$ = the weight of species *i*

$\mathrm{IUCN}i_{\mathrm{sp}i}$ = IUCN category of species *i* assuming values: 1, for no threat listing
2, for vulnerable species
3, for endangered species
4, for critically endangered species

$\mathrm{EB}i_{\mathrm{sp}i}$ = Endemism for species *i* considering the percentage of species *i* modelled distribution restricted to within the Paranã River Basin boundaries in relation to a buffer area of 1 degree beyond the Basin extremities

$\mathrm{EC}i_{\mathrm{sp}i}$ = Endemic species *i* of Cerrado biome occurring within the Basin assuming values: 1, if present
0, otherwise

$\mathrm{AUC}i_{\mathrm{sp}i}$ = 1 * (AUC value of model of species *i* distribution tested for accuracy), or
0.5 * (AUC value of model of species *i* distribution non-tested for accuracy)

Within GDMs, weights were distributed among taxonomic groups following the same percentage distribution used in SDMs, i.e., 50% for plants, 30% for birds, and 20% for mammals. Within each taxonomic assemblage group, the weights assigned to each assemblage were 85% for average richness and 15% for uniqueness (species exclusive to each assemblage).

To evaluate the contributions of the additional features to area prioritization, weights were distributed following the same criteria established to assign weights to the previous scenarios (Figure 4.1B). Considering that GDMs have, in total, fewer taxonomic groups than the larger set of SDMs including the additional features, we attributed more relative value to the SDMs. We established a weight of 30% for all GDMs, 30% for all ecosystems, and 40% for all SDMs (this time with five taxonomic groups - birds, mammals, reptiles, amphibians, and plants). Within these broad data types, weights were assigned to groups of features in this way: for GDMs, 50% for plants, 30% for birds, and 20% for mammals; for SDMs, 38% for plants, 23% for birds, 15% for mammals, 12% for reptiles, and 12% for amphibians; and for ecosystems, 50% for dry forest, 30% for savannas mixed with dry forests, and 20% for savannas. The higher weight for dry forests reflects their status as one of the most threatened ecosystems in the world, and in Brazil, especially in the Paranã River Basin [29, 31, 39, 43, 44].

Priority areas for each solution were presenteded in maps of hierarchical prioritization of the conservation values across the Basin. The hierarchical nature of the results means, for example, that the most valuable 2% of the Basin is within the most valuable 5%, which is within the top 10%, and so on. Additionally, we evaluated (1) the incidental representativeness in the prioritizations of threatened and endemic species of the Cerrado biome that were not modelled with SDMs, and (2) we measured the overlap of the priority areas from our analysis with the areas recommended for conservation in protected areas by the Brazilian Government [29] and the existing fully protected area system. We assumed the value of 20% coverage of protection as the minimum amount to be achieved in the Basin. This value also represents the percentage determined by the new Forest Code for conservation in the Cerrado biome, approved recently by the Brazilian Government [180].

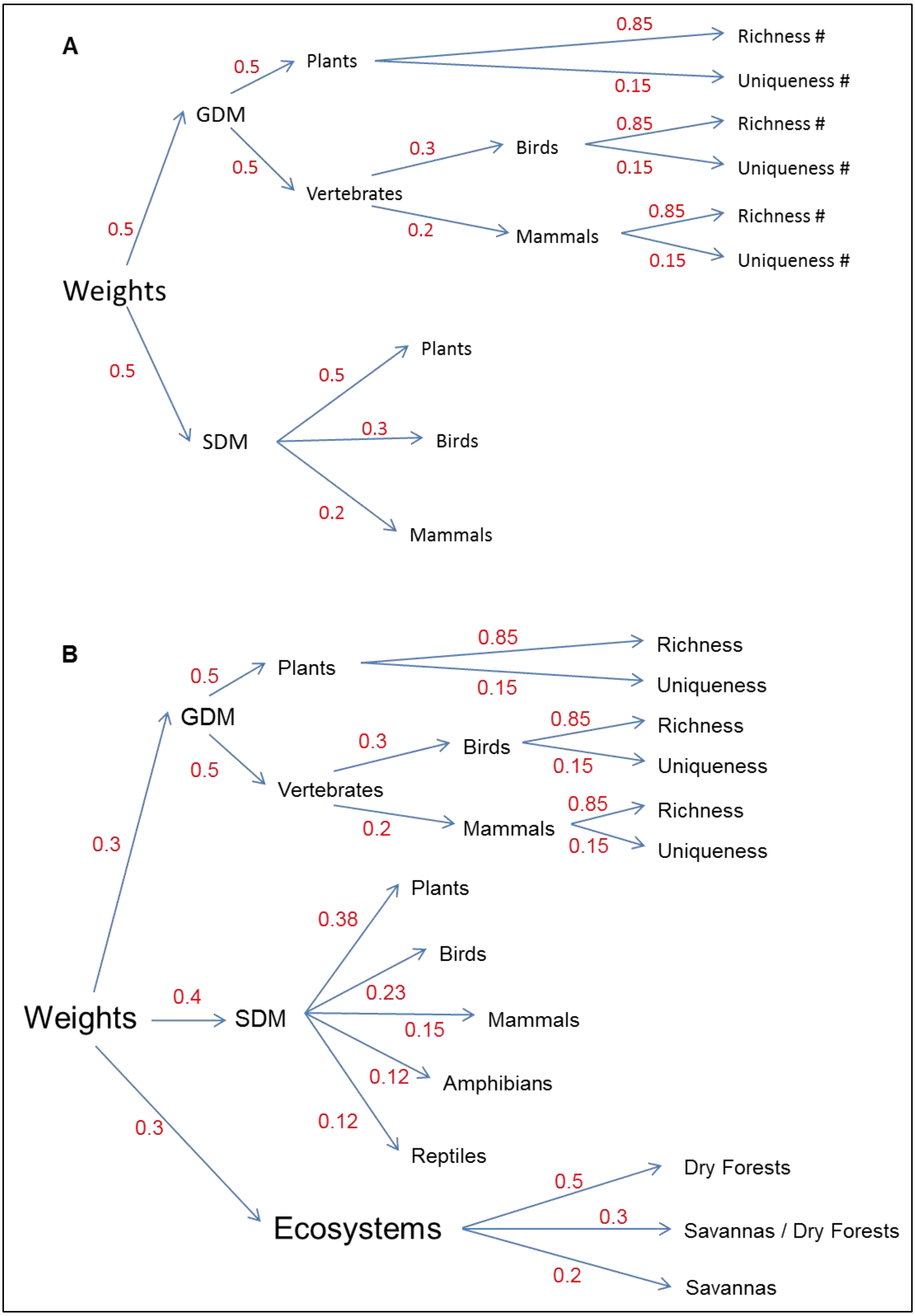


Figure 4.1. Schematic representation of differential weights for each group of biodiversity features used to prioritize areas for conservation in the Paranã River Basin. A. Weights for prioritizations based only on SDMs and GDMs. B. Weights for prioritizations based on SDMs and GDMs from A, plus additional features.

4.4. Results

Regardless of the sets of biodiversity features used in the prioritizations – species, species assemblages, ecosystems, and their combinations – all converged in highlighting the top 2 to 5% fraction of the landscape in the mountain range on the western border of the Basin (Figure 4.2A-F). The mountain range, with an elevational gradient varying from 430 – 1663 m, corresponds to the richest area for plants and the highest species turnover and distinctive assemblages for birds, mammals, and plants [175]. Differences in priorities between solutions were evident at lower elevations where the proportion and extent of the landscape in priority classes varied according to the biodiversity features used in the analysis.

Solutions using SDMs (Figures 4.2A, C-F) highlighted, basically, the richest areas for vertebrate groups in the north-west, north-east, and central-east regions of the Basin [175]. Areas of higher priority in the north-west are the lowest elevation areas of the Basin (247 – 430 m) where arboreal savannas mixed with deciduous forests are dominant in the landscape [58]. In the north-east, higher-priority areas are at intermediate elevations (430 to 775 m) and dominated by savannas and deciduous forests on limestone outcrops, and also on the montane gradient (776 – 1010 m) that corresponds to the cliffs on the rim of the Basin. On these cliffs, the dominant vegetation is arboreal savanna in concave faces of the landscape and semideciduous forests in ravines of convex faces. In the central-eastern region, higher-priority areas are deciduous forests on relatively flat terrain and limestone outcrops. Additionally, scattered areas along the Paranã River (Figure 1.3), which connects the central-east and north-west regions, were also in the 20% top of the landscape. However, those areas were more evident in solutions that used only SDMs (Figures 4.2A, D) and in solutions that combined SDMs with assemblages (Figures 4.2C, E). In this section of the Paranã River Basin, deciduous and semideciduous forests constitute the riparian forests in a landscape matrix formed by arboreal savanna interspersed with patches of deciduous forest. Unique richness areas formed by marginal temporary lagoons along the river are also present and contribute to the high number of vertebrate species found in this part of the Basin.

Landscape prioritization based only on assemblages (Figure 4.2B) for birds, mammals, and plants is the most dissimilar to the others. Basically, it gave high priority to regions with distinctive species composition, higher species turnover, and higher richness, overlapping to some extent with solutions using SDMs. The most important areas selected for assemblages (red in Figure 4.2.B) correspond to the montane gradients of the same mountain range in the west and south-western borders of the Basin highlighted solutions including SDMs. Additionally, however, the montane gradient located in the southern and eastern boundaries of the Basin were also given high priority for assemblages. At intermediate elevations (430 to 775 m), areas with higher priority correspond to the transition zones between deciduous forests and savannas across the flat areas of the eastern region (yellow in Figure 4.2B). In the lowest-elevation part of the northern region (230 – 430 m), some areas with higher priority correspond to the ecosystems formed by arboreal savanna interspersed with patches of deciduous forest. In these locations, unique and rich areas for birds are found in a landscape formed by marginal temporary lagoons along the main river (Paranã River in Figure 1.3), in the central north-western region (orange and yellow line in Figure 4.2B), and in permanent and intermittent lagoons along the main river (Palma River in Figure 1.3) and its affluents of the north-eastern region [175] (orange area and yellow line in Figure 4.2B). These areas are also the richest areas for mammals, amphibians, and reptiles [175]. Singular areas in species composition were similarly highlighted across the eastern boundaries of the Basin (orange, pink, and red patches in Figure 4.2B) where unique communities of birds, mammals, and plants were also identified [175].

Finally, the incorporation of ecosystems into the prioritization contributed basically to highlight four distinct regions of the Basin where threatened ecosystems – mainly dry forests – are present in patches mixed with savannas (in the central-west mountain range and north-west regions) or are dominant (in the central-east and north-east regions) in the landscape (Figure 4.2F).

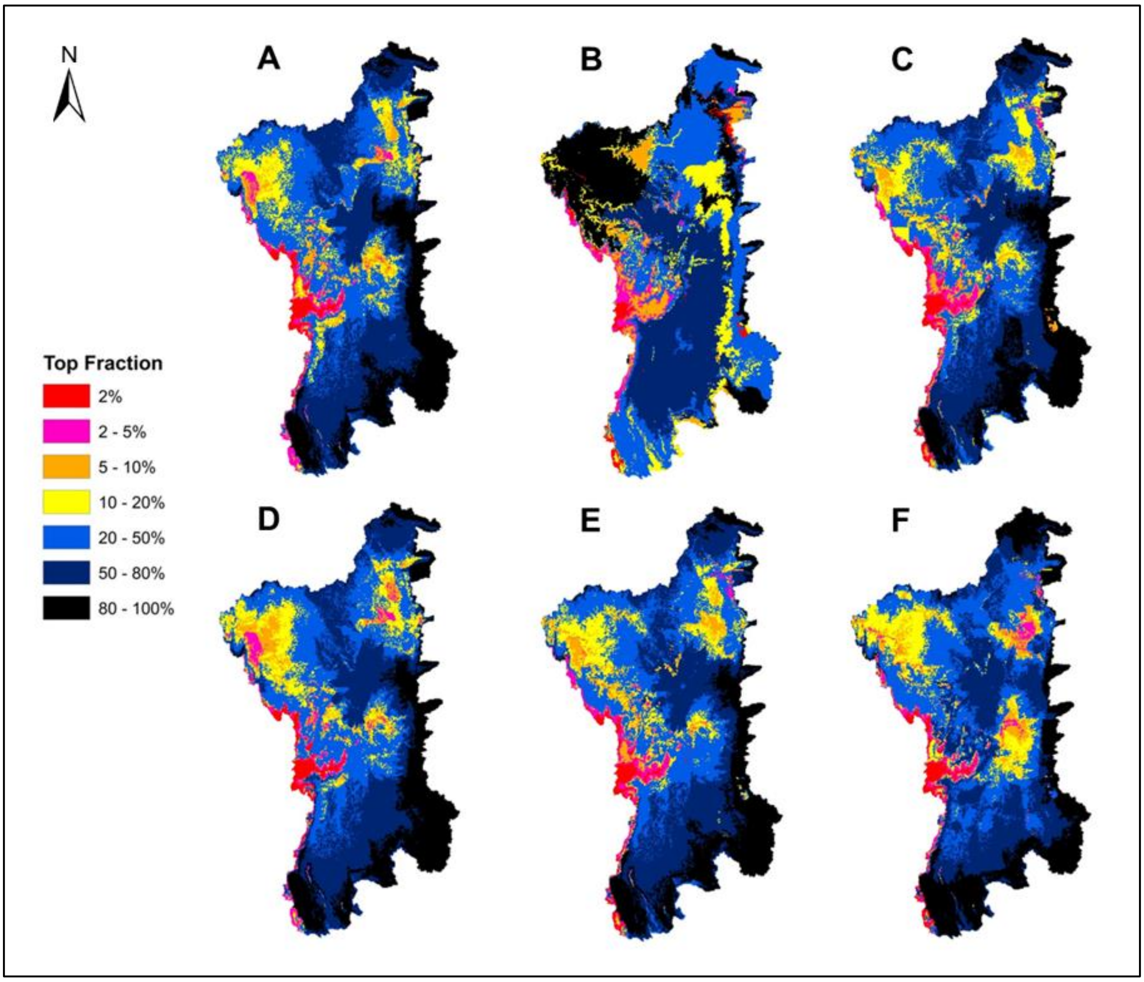


Figure 4.2. Alternative area prioritizations for conservation in the Paranã River Basin. In this figure, different colours represent different priorities of grid cells. Warmer colours (from yellow to red) represent the highest priority areas for conservation considering the top 20% most important areas (yellow), the top 10% (orange), 5% (pink), and 2% (red). A. SDMs for birds, mammals, and plants. B. GDMs for birds, mammals, and plants. C. GDMs and SDMs for birds, mammals, and plants. D. SDMs for birds, mammals, amphibians, reptiles and plants. E. GDMs for birds, mammals, and plants, and SDMs for birds, mammals, amphibians, reptiles, and plants. F. GDMs for birds, mammals, and plants, SDMs for birds, mammals, reptiles, amphibians, and plants, and ecosystems defined by dry forest, savanna-dry forest complex, and savanna.

Across groups of features, the variation in the average proportion of each group's feature distributions that remained as cell removal progressed was similar in solutions using only SDMs (Figures 4.3A and D and Table 4.1). In contrast, average proportions varied strongly between groups where GDMs were used in prioritization (Figure 4.3B, C, E and F and Table 4.1). However, all solutions retained – on average – similar amounts of the average distributions of biodiversity features (between 26 and 30.8% of feature distributions) until 80% of the landscape was lost within the area prioritization that is, until only the highest 20% priority cells remained (Table 4.1). Nevertheless, some biodiversity features were lost (reduced to 0% of their initial distributions in the prioritization) while larger percentages of cells remained (Table 4.1). For example, the first features were lost in two prioritizations when 50-80% of cells remained. Other solutions were better at retaining biodiversity features, including features not considered directly in the prioritization, with no outright losses until less than 2% of the landscape remained (represented by red cells in Table 4.1).

Assemblages defined by plants, followed by those based on birds and mammals, tended to have larger proportions remaining when 80% of cells had been removed in most of the prioritizations (Figure 4.3 and Table 4.1). For species distributions, the proportions remaining of the original distributions in the landscape (amphibians, birds, mammals, reptiles, and plants) was generally smaller than that for assemblage features. The exception was the plants group, a biodiversity feature highly represented in all solutions (Figure 4.3 and Table 4.1).

Retention of additional features (SDMs for amphibians and reptiles and the three ecosystems) was variable across the prioritizations. The most threatened ecosystem within the Basin – dry forests – was represented well only when ecosystem feature was included in the prioritization. In contrast, additional information on amphibians and reptiles did not noticeably increase the representation of these features (see Table 4.1).

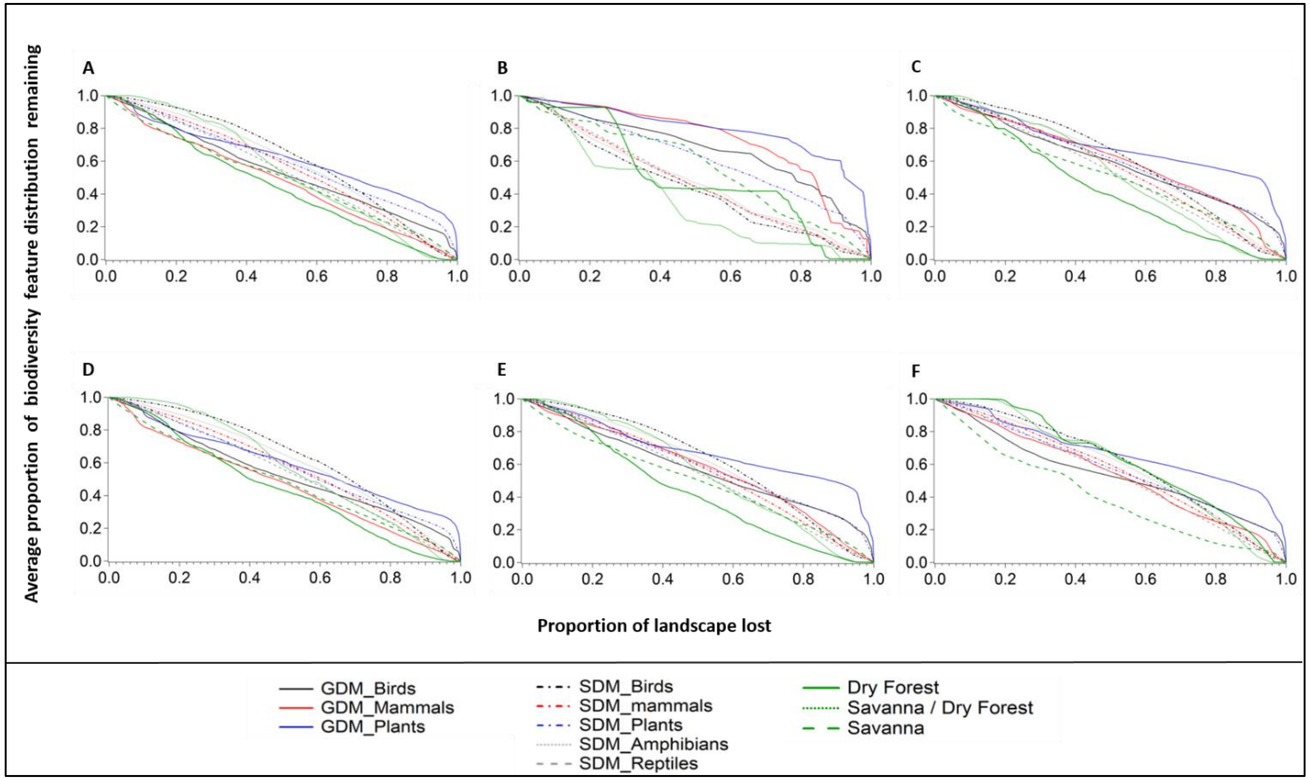


Figure 4.3. The average proportion of biodiversity features remaining (in eight groups, with ecosystems considered individually) according to the percentage of the Basin removed from the prioritization.

X-axis represents the fraction of the landscape lost for conservation during the process of cell removal. *Y-axis* represents the average proportion of the original extent of biodiversity features remaining in the analysis. GDM = assemblage-level approach for three groups. SDM = species-level approach for five groups.

A. Prioritization based on SDMs (birds, mammals, and plants). B. Prioritization based on GDM (birds, mammals, and plants). C. Prioritization based on GDM and SDMs (birds, mammals, and plants). D. Prioritization based on SDMs (birds, mammals, amphibians, reptiles, and plants). E. Prioritization based on GDM (birds, mammals, and plants), and SDMs (birds, mammals, amphibians, reptiles, and plants). F. Prioritization based on GDM (birds, mammals, and plants), SDMs (birds, mammals, reptiles, amphibians, and plants), and ecosystems.

Table 4.1. Average feature retention in groups of features retained by each prioritization in the top 20% of priorities (when 80% of the total area of the Paranã River Basin is lost).
Colours in this table follow the same colouring pattern of the area prioritization presented in Figure 4.2. Colours indicate in which top fraction of the landscape the first biodiversity feature is reduced to zero: navy blue (50–80%), blue (20–50%), yellow (10–20%), orange (5–10%), pink (2–5%), and red (2%).
Letters correspond to taxonomic groups with models included in the prioritization (b = birds, m = mammals, p = plants, a = amphibians, r = reptiles, Eco = ecosystems, DF = dry forest, Sav = savanna). Ave = average, and SD = standard deviation.

Prioritization solutions	**Biodiversity features**											**Ave (SD)**
	GDM b	**GDM m**	**GDM p**	**SDM b**	**SDM m**	**SDM p**	**SDM a**	**SDM r**	**Eco DF**	**Eco S/DF**	**Eco S**	
A. SDM (bmp)	**29.2**	**18.7**	**42.6**	**29.2**	**25.5**	**35.4**	**28.6**	**22.1**	**13.5**	**18.8**	**22.6**	**26.0 (8.2)**
B. GDM (bmp)	**46.8**	**54.9**	**69.1**	**16.3**	**18.5**	**38.0**	**20.6**	**17.2**	**24.9**	**9.3**	**22.7**	**30.8 (18.9)**
C. GDM (bmp) and SDM (bmp)	**36.5**	**37.6**	**55.9**	**26.4**	**23.6**	**37.1**	**27.1**	**20.7**	**11.8**	**14.5**	**25.2**	**28.8 (12.4)**
D. SDM (bmpar)	**30.2**	**18.2**	**37.5**	**32.2**	**26.8**	**32.3**	**31.0**	**23.1**	**11.3**	**23.9**	**21.0**	**26.1 (7.5)**
E. GDM (bmp) and SDM (bmpar)	**35.3**	**30.8**	**53.7**	**28.6**	**24.5**	**35.7**	**28.9**	**21.5**	**10.3**	**19.9**	**23.2**	**28.4 (11.1)**
F. GDM (bmp), SDM (bmpar), and Eco	**33.4**	**25.1**	**51.7**	**28.3**	**25.3**	**33.2**	**28.2**	**22.3**	**33.2**	**26.8**	**11.8**	**29.0 (9.7)**
Ave (SD)	**35.23 (6.3)**	**30.88 (13.9)**	**51.75 (11.0)**	**26.83 (5.5)**	**24.03 (2.9)**	**35.28 (2.2)**	**27.40 (3.6)**	**21.15 (2.1)**	**17.50 (9.4)**	**18.87 (6.3)**	**21.08 (4.7)**	**28.18**

Regardless of the solution considered, all threatened and endemic species modelled were represented in the top 2% fraction. Of the threatened and endemic species not modelled and not considered in the prioritizations, all solutions were similar in terms of the number of species represented incidentally (Table 4.2), although the identities of those species differed between solutions. Of a total of 719 endemic species without species distributions models, more than 70% had at least one point of occurrence included in the top 20% of the different prioritization solutions. Of the 40 threatened species not modelled, an average, across solutions, of 65% were included in the top 20% of cells (Table 4.2).

Table 4.2. Total of endemic and threatened species represented incidentally in each prioritization in the top 20% of priorities (when 80% of the total area of the Paranã River Basin is lost).
´Within´ columns indicate species that had at least one point of occurrence included in the top 20% of priorities and ´out´ columns, otherwise.

Prioritization solutions	Endemic		Threatened	
	within	out	within	out
A. SDM (bmp)	531	188	26	14
B. GDM (bmp)	514	205	25	15
C. GDM (bmp) and SDM (bmp)	536	183	27	13
D. SDM (bmpar)	534	185	26	14
E. GDM (bmp), SDM (bmpra)	535	184	25	15
F. GDM (bmp), SDM (bmpra), and Eco	526	193	27	13

About 37,500 km^2 (~61%) of the total area of the Basin was suggested for the establishment of protected areas by the Brazilian Government [29, 87]. These areas overlapped partially (between approximately 40 and 66%) with the top 20% fractions of the different solutions (Figure 4.4 and Table 4.3). Higher overlaps are mainly related to the areas with management categories still undefined, whilst areas intended for full protection have smaller overlaps. However, overall, there are only small overlaps between proposed protected areas and high-priority areas (the top 2%) from Zonation (between approximately 6 and 30%). Similarly, existing fully protected areas within the

Basin only partially overlap the most important areas for conservation selected by Zonation. Located at the top of the mountain range of the western border of the Basin, the Chapada dos Veadeiros National Park has ~ 54 km^2 of its extent within the top 2% of Zonation priorities. The Terra Ronca State Park, located at the opposite border (eastern side), overlaps only the top 20% fraction, but to different degrees (from 1 to 81 km^2), according to the solution considered (Figure 4.4 and Table 4.3).

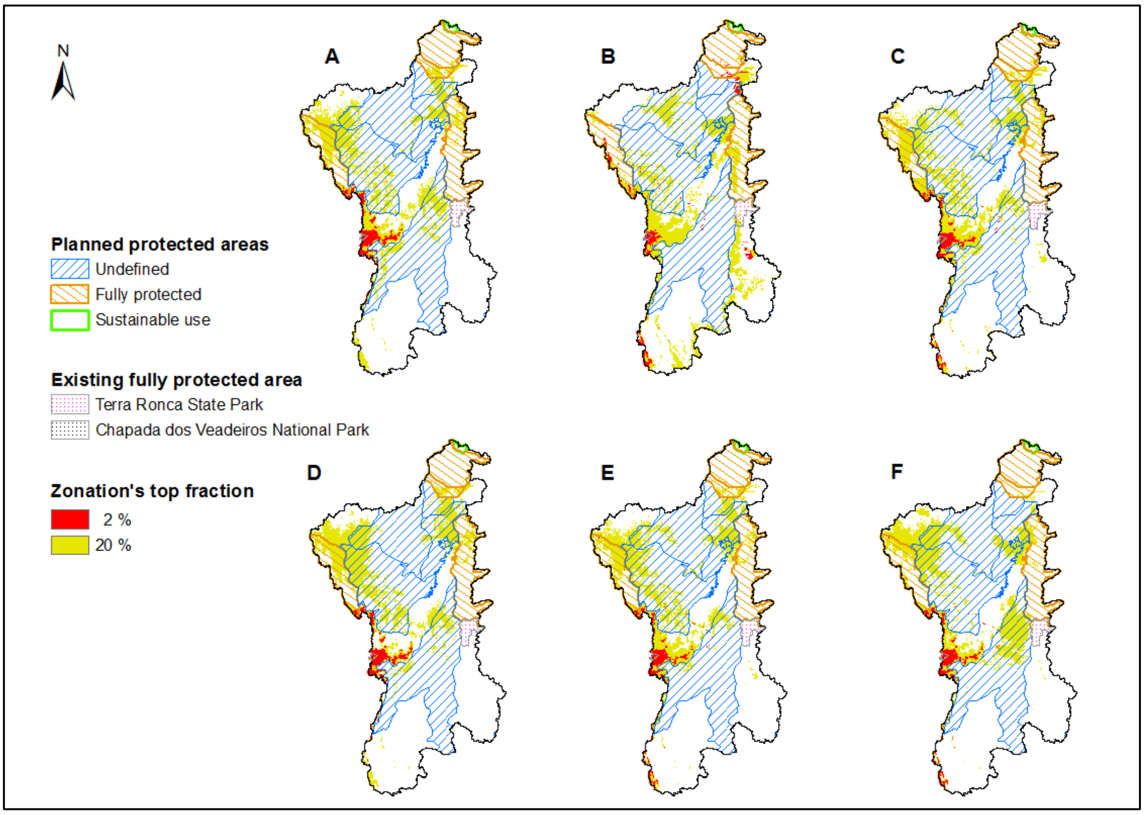


Figure 4.4. Priority areas for conservation in the Paranã River Basin derived from the six different Zonation solutions and their overlap with the priority areas recommended by the Brazilian Government [29, 87] and with the existing fully protected areas.

In this figure, A. SDMs for birds, mammals, and plants. B. GDMs for birds, mammals, and plants. C. GDMs and SDMs for birds, mammals and plants. D. SDMs for birds, mammals, amphibians, reptiles and plants. E. GDMs for birds, mammals and plants, and SDMs for birds, mammals, amphibians, reptiles and plants. F. GDM for birds, mammals and plants, SDM for birds, mammals, reptiles, amphibians, and plants, and ecosystems.

Zonation top fractions refer to the highest 20% and 2% of cells in the Basin.

Note that the Chapada dos Veadeiros National Park is represented by a tiny area on the western border of the Basin

Table 4.3. Percentage of overlap between the top 2% and top 20% fraction of the landscape indicated by each Zonation prioritization solution and the priority areas recommended by the Brazilian Government [29, 87] and existing fully protected areas in the Paranã River Basin.

Overlap percentages are presented in relation to the protected areas (in two categories) recommended by the Brazilian Government [29, 87] and in relation to existing protected areas. According to the Brazilian Government, areas in the undefined (u = ~28,000 km^2) category need further investigation to determine whether they will eventually be fully protected (fp = ~9,000 km^2) or managed for sustainable use. Areas recommended for sustainable use by the Brazilian Government (~157 km^2) do not overlap Zonation´s top fractions (2% and 20%) and are not represented in the table.

Letters correspond to taxonomic groups with models included in the prioritizations (b = birds, m = mammals, p = plants, a = amphibians, r = reptiles). Eco represents ecosystems.

BPA = Brazilian Government protected areas. ZTF = Zonation´s top fraction.

Prioritization solutions	Brazilian priority areas for conservation in the Paranã River Basin												Existing fully protected areas	
	Overlapping with top 2% (%)				Overlapping with top 20% (%)				Total overlapping (%)				Overlapping (%)	
	BPA (u)	ZTF	BPA (fp)	ZTF	BPA (u)	ZTF	BPA (fp)	ZTF	BPA	ZTF 2%	BPA	ZTF 20%	ZTF 2%	ZTF 20%
A. SDM (bmp)	0.03	0.69	0.73	5.54	20.58	46.74	23.33	17.61	0.20	6.23	21.18	64.35	3.74	3.78
B. GDM (bmp)	0.48	10.80	2.69	20.28	13.18	29.92	14.25	10.75	1.02	31.07	13.39	40.68	3.74	4.30
C. GDM (bmp) and SDM (bmp)	0.12	2.84	0.71	5.33	18.81	42.72	23.78	17.94	0.27	8.17	19.96	60.66	3.74	3.75
D. SDM (bmpra)	0.03	0.69	0.72	5.40	21.76	49.43	22.72	17.15	0.20	6.09	21.91	66.58	3.74	3.75
E. GDM (bmp) and SDM (bmpra)	0.13	3.04	0.68	5.12	19.94	45.29	20.70	15.63	0.27	8.17	20.05	60.92	3.74	3.74
F. GDM (bmp), SDM (bmpra), and Eco	0.13	3.04	0.73	5.54	21.66	49.19	15.84	11.96	0.28	8.58	20.12	61.15	3.74	3.90

4.5. Discussion

Regardless of the biodiversity features considered (species, species assemblages, and ecosystems), all six solutions aligned with their highest 5% of cell priorities. This alignment highlights patterns that are shared across the different dimensions of biodiversity evaluated in these analyses. This convergence of highest priorities also indicates the importance of these areas for future conservation efforts.

Areas prioritised with species-level data centred mainly in the highlands of the Basin (Figures 4.2A and 4.2D) which are the richest areas for plants [175]. This region also has a high concentration of threatened and endemic species of plants and animals (see Figure 2.3 and 2.4) [58]. Moreover, the montane gradient that surrounds this mountain range is associated with high species turnover, independent of the taxonomic group evaluated (birds, mammals, and plants). The species turnover reflected in the assemblage-level approach [175] denotes a specialized biota occupying different habitats, which contributes to this region's uniqueness and richness (Figure 4.2B).

Differences between the species- and assemblage-level approaches for identifying priority areas are evident in the areas selected in the Basin's lower elevations. Prioritization based on species (Figures 4.2A and 4.2D) highlighted the richest areas for birds, mammals, amphibians, and reptiles in areas with more vertically structured habitats [175], specifically areas dominated by dry forests (in the central-eastern parts) and arboreal savannas mixed with dry forest (in north-eastern and north-western parts). In contrast, prioritization based solely on assemblages highlighted transitional zones where species richness is higher and areas of singular assemblage composition. Transitional areas between dry forests and savannas are evident along the eastern border of the Basin, represented by yellow areas selected at the top 20% in Figure 4.2B.

Areas of singular assemblage composition (Figure 4.2B) were found along the scarps and plateaus (Figure 1.3) of the eastern border and in the central-northern parts at low elevations. The scarps of this region are occupied mainly by arboreal and typical savannas (see Plate 1.3). Sandy soils on the plateau are occupied by formations that vary from grasslands to typical savannas. Semideciduous forests are interspersed with those formations, forming gallery forests along the water courses. In these regions, orange (top 10%) and red (top 2%) indicate singular assemblages of plants and mammals. Bird assemblages are indicated by orange (top 10%) in the central-north, along with the lagoons in savannas interspersed with dry forest along the Palma River (see Figure 1.3).

Differences observed in prioritizations based on species and assemblages are removed in solutions that combine both approaches (Figures 4.2C and 4.2E). In these cases, high-priority areas are highly complementary in terms of the features they contain (see Table 4.1), including species-rich areas, singular species associations, and transitional zones.

The efficacy of using additional features in conservation prioritization can be assessed in terms of the complementarity of the information added to the solutions. The addition of reptile and amphibian distributions, did not alter the general pattern of area prioritization (see Figures 4.2A and 4.2D and Table 4.1). The distributions of the additional species overlapped those of the vertebrate groups already considered [175], thereby adding little complementary information in defining priorities. However, the analyses that included ecosystems highlighted priority areas in the Basin's central-eastern section that were only partially prioritized by species and assemblage approaches (Figure 4.2 F). Including ecosystems mainly highlighted areas where one of the most threatened ecosystems of the Basin – the dry forests – is dominant in the landscape. Including ecosystems also led to higher priorities allocated incidentally to endemic and threatened species of the dry forest ecosystem (see Figures 2.3 and 2.4), without substantially reducing representation of other biodiversity features (see Table 4.1).

4.5.1. Comparison of species and assemblage-level prioritization

The choice of which approach (species or assemblage) to use in area prioritization using Zonation depends on the objectives of conservation and data availability. In the species-level analysis, the value of a gridded cell was the summation of all of its biodiversity features, weighted according to endemism, threat, and quality of the models. In analyses based on assemblages and ecosystems, prioritization was determined by classes of landscape, weighted by richness and/or uniqueness of species composition (for assemblages) or threats (for ecosystems). The simultaneous use of species, assemblages, and ecosystems prioritizes multiple spatially distinct regions important for one or more data types (see Figure 4.2F prioritization pattern compared to the other solutions 4.2A, B, C, D, and E).

Modelling the distributions of individual species is a relatively common strategy [110] to provide spatially explicit biodiversity data for prioritization [100, 113, 173]. Depending on the method used, the models can be derived from continuous or categorical environmental variables. Additionally, prioritization based on these models can produce relatively good results in representing other biodiversity features incidentally (see Table 4.1 and 4.2). However, SDMs have high data demands when applied to many species. It is necessary to have a minimum number of records per species to produce a useful model [105, 131], so rare species need to be excluded from the analysis due to insufficient observational data. Often, these are the species of the greatest concern in conservation planning [173]. This study was no exception. The majority of threatened and endemic species could not be modelled due to insufficient occurrence data and the models for those species with sufficient field records often came with high uncertainty [175]. Despite this limitation, most of the biologically important species not modelled were represented incidentally in the top 20% fractions of the landscape in all solutions (see Table 4.2).

An alternative approach would be to include these species as records directly in the prioritization of gridded cells [172]. However, their inclusion could result in a scattered pattern of unique cells (1 Km^2) of high importance in which the feasibility of conservation and the persistence of species are questionable. Additionally, small patches in fragmented landscapes are more susceptible to the invasion of species from the surrounding matrix due to an edge effect [181] and sampling bias in species records would focus reserves on areas near to roads and towns where anthropogenic impacts are more frequent and might compromise species persistence.

Assemblage modelling with GDMs is an alternative basis for conservation planning in data-poor regions [101, 115]. Because GDMs deal with assemblage composition as a whole, they include rare species by assuming they have similar responses to environmental variables as more common species. In this sense, GDMs incorporate interactions between species, rather than relying on individual species' records [173, 174] as do SDMs. Additionally, the use of GDMs for Zonation produced useful results in prioritizing areas of high species turnover commonly found in transitional areas or ecotones. In ecotones, the active evolutionary processes responsible for biodiversity generation and maintenance are often found [13, 145, 182]. Given the importance of ecological gradients and transitional habitats [13], their conservation warrants greater attention in tropical regions. However, most protected areas are planned to select biodiversity patterns in the reserve design and do not include ongoing evolutionary processes [13, 145]. Moreover, GDMs were also successful in capturing the overall richness patterns – considering birds, mammals, and plants – and, complementary, areas where species composition was unique. On the other hand, data limitations prevented assemblages being defined for amphibians and reptiles [175].

Moreover, GDMs operate only on continuous or ordinal variables, which is a limitation in areas that are poorly known with respect to environmental variables related to patterns of species occurrence and turnover. Most of the environmental data available for tropical regions in developing countries is available solely as classes of soils, geology, and vegetation, for example. Converting categorical environmental data into continuous or ordinal variables is a significant and time-demanding challenge. Depending on the type of meta-information associated with those classes, this conversion might be unmanageable. Additionally, fitting useful GDM models requires an even distribution of records across the entire region of interest as shown in the Chapter 3 [175]. Finally, the use of GDMs in Zonation is not as straightforward as SDMs, requiring a greater level of expertise and time. An assemblage similarity matrix must be generated before being used in Zonation [see (172) for a complete description of the method] and additional analytical steps are also required before applying Zonation.

Overall, though, multi-taxonomic species and assemblage models in conservation prioritization are strongly recommended due to the complementarity of these approaches. Both performed well in retaining the biodiversity features considered in this study, even when species- and assemblage-level models were used alone for prioritization (see Table 4.1). However, each approach tended to retain better the features on which they were focused; that is, the SDM solution had better retention of species distributions to the detriment of assemblages, and the GDM solution retained assemblages at the expense of species distributions. Such differences are balanced, though, when both approaches are used in the same solution.

4.5.2. Uncertainties in area prioritization

Uncertainties in area prioritization might be present in solutions where the species-level approach was used given the known shortcomings of modelling species distributions from very few records and a relatively large number of variables and background points [108]. However, considering the similarities of the signal of species distributions between tested and non-tested models and the similarities in species distributions across vertebrate groups [175], uncertainties in species models seem to have been partially circumvented or minimized by the weighting framework adopted in this study. Giving much less weight for models not tested for accuracy, the potentially highly inaccurate distribution maps had a relatively small influence on prioritization. However, further investigations must be done using, for example, only species tested for model accuracy to check similarities and dissimilarities between area prioritizations to understand whether species recorded in few locations were adequately protected in high-priority areas. The high incidental representation of endemic and threatened species in each prioritization in the top 20% of priorities (see Table 4.2), though, might indicate that species recorded in few locations are likely to be protected, even using models not tested for accuracy in prioritization.

It is possible, however, that the weighting framework used in all solutions might be so strongly focused on endemic and threatened species, unique and species-rich assemblages, and threatened ecosystems that high-priority areas do not represent other features well. That is, the general representation of overall biodiversity features until the 20% top fraction of the landscape (see Table 4.1) does not guarantee their optimal representation. However, if we consider a general average of 20% of the original biodiversity features as an adequate level of biodiversity representation, the different solutions seem to be acceptable for area prioritization within the Basin.

4.5.3. Implications for biodiversity conservation within the Paranã River Basin

Macro-scale approaches at the biome level for conservation planning, as proposed by Brazilian government [29], are essential to highlight the importance of some regions for biodiversity conservation. However, the areas indicated as a priority on a national scale are rarely converted in their totality into fully protected areas. Territorial dimensions of the areas involved, lack of biological data, lack of stakeholder engagement, conflicts between human communities and, particularly in Brazil, land property disputes, are some of the reasons for this gap between high-level design and on-ground implementation. Additionally, reviewing and updating regional conservation designs is part of the dynamism intrinsic to conservation planning [73]. Updates are necessary to accommodate, for example, new data on biodiversity, loss of areas of interest, and new methods and approaches to analysis [14].

More than 60% of the Paranã River Basin has been identified by the Brazilian Government as important for establishment of protected areas [58]. This extensive and coarse-resolution analysis requires fine-resolution conservation design to identify which of these areas should be prioritized and why, thereby providing stakeholders with clear arguments to act locally. This refined design can be done with our Zonation analyses, Irreplaceable areas of extreme value for biodiversity (the top 2% of the landscape), such as the mountain range in the Basin's central-west, are under-represented in the priority areas identified by the Brazilian government [29] and the current fully protected area system (see Figure 4.4 and Table 4.3).

Additionally, areas indicated as the top priority (2–5 %) by different solutions (see Figures 4.2A-F) are mainly located in the western mountains, which have low values for agricultural proposes due to their steepness. Furthermore, those areas are possibly climatic refugia for species from lower elevations, which would contribute positively to long-term biodiversity conservation outcomes. Equally important are the complementary areas highlighted in the flatter, lower parts of the Basin. Deforestation threatens most of the areas located in the eastern region, due to pressures from agriculture, cattle ranging, and calcareous mining on the limestone outcrops [29, 31, 39, 50]. Additionally, those areas embrace distinct endemic and threatened species and are considered a centre of endemism for species from different taxonomic groups [31-41, 58]. Also in the top 2–5%, the north-west region is the most pristine environment in the Basin with the highest richness values for vertebrates [58]. Its preservation has been due to terrain constraints – relatively low fertility and difficulty of mechanization posed by stony soils. However, afforestation projects with *Eucalyptus* spp. and sugarcane plantations are already being implemented in the region with loss of native vegetation.

According to the new Brazilian Forest Code, 20% of the lands within Cerrado biome must be left as Legal Reserves by landowners. Additionally, the new Forest Code resolution allows landholders who have already cleared their lands to allocate these obligatory reserves in areas outside of their properties as long as they are located at the same biome. The new resolution will offers an opportunity to use the prioritizations developed for this study, in focusing offsets and also limited future clearing.

The minimum area necessary to represent all species and promote their persistence within a region is highly variable and depends upon the diversity and endemism of the taxa of concern and on the size of the selection units considered [183]. Although reserving 20% of the landscape is not necessarily an optimal target for biodiversity conservation, it is realistic in the Paranã River Basin since the law now requires it. Additionally, most of the species, assemblages, and ecosystems were represented in the top 20% fractions identified in all prioritizations (see Table 4.1). Prioritizing these areas for conservation will make an important contribution to maintaining areas of high biodiversity interest within the Basin.

Chapter 5

Chapter 5 – Conservation in a biodiversity hotspot: planning for a changing world

---

Keywords: conservation planning, conservation prioritization, climate change, Maxent, Zonation, Paranã River Basin

Article type: Full-length Article

Anderson C. Sevilha[1,2], Stephen E. Williams[1], Robert L. Pressey[3], Jeremy VanDerWal[1,4] & Helga C. Wiederhecker[1,5]

[1] Centre for Tropical Biodiversity and Climate Change, School of Marine and Tropical Biology, James Cook University, Townsville, QLD, Australia

[2] Embrapa Genetic Resources & Biotechnology, Brasília, DF, Brazil

[3] Australian Research Council Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, QLD, Australia

[4] eResearch Centre, James Cook University, Townsville, QLD, Australia

[5] Department of Zoology, University of Brasília, Brasília, DF, Brazil

Chapter 5 – Conservation in a biodiversity hotspot: planning for a changing world

5.1. Abstract

Designing conservation reserves for biodiversity conservation under a changing climate requires careful consideration of future projections of species displacement, movement patterns and corridors, the potential for localized population extinction, and landscape resilience. Identifying refugia and/or resilient areas under climate change is essential to minimize adverse impacts on biodiversity. Most studies around the world have focused on optimizing the resilience of reserve networks to climate change based on selected species or multiple species from the same biological group, while only a few have considered multiple taxonomic groups, with none of these in Brazil. We used the Paranã River Basin, a key area for biodiversity conservation within the global Cerrado hotspot, as a case study to identify areas for conservation that are resilient to effects of climate change on the distributions of species from multiple taxonomic groups**.** We used Maxent to project species distributions for current and future climate data for 1813 plants, 245 birds, 44 mammals, 30 reptiles, and 27 amphibians. We then spatially prioritized areas based on these distributions under current and future climates using Zonation. Climatic space favourable for high terrestrial vertebrate species richness is currently at the lower elevations of the Basin but is predicted to expand upwards in elevation under future climate projections. High plant species richness is currently in the highlands of the Basin and is predicted to shrink further upward. Congruence between priority areas across taxa and between current and future distributions corresponds to the top 10% fraction of the most important areas for biodiversity conservation. These areas are of the highest priority now and into the future and thus represent the most important areas for optimizing resilience to climate change. We also identified potential climate refugia and corridors between important areas for current species distributions and those that will be important in the future. These high-priority areas are mostly under-represented in current Brazilian Government planning, putting most of the endemic and threatened species at risk. The broadly defined conservation priorities identified in previous studies must be revisited to accommodate the effects of climate change.

5.2. Introduction

Global warming and its effects on ecosystems and biodiversity are now widely accepted [15-17] and pose an additional challenge to conservation planning and management [13, 25]. Protected areas might have a different mix of species in the future than they once had [25]. Migration into suitable areas and/or local extinctions are expected for many species [15, 17-24]. Climate change acts simultaneously as a driver of biodiversity processes, promoting adjustments of ranges over time, while also posing a dynamic threat to biodiversity [13]. An additional expected impact of climate change is the increasing likelihood of extreme climatic events.

The field of systematic conservation planning aims to prioritize areas for conservation spatially [13]. Currently, planning is faced with the additional challenge of incorporating new sets of considerations around the future dynamics driven by a changing climate [13, 18, 21, 22, 25]. Planning for long-term persistence of biodiversity features has now also to consider changes such as future shifts in species distributions. Among the necessary aspects of planning are scheduling conservation actions with expanding threats, minimum population sizes to maintain genetic variability for adaptation to natural and anthropogenic threats, and corridors and refugia to enable connectivity between current and future species distributions [13, 184, 185].

Different schemes and strategies have been proposed to ensure future persistence of individual species, species assemblages, and ecosystem services under climate change [22, 186-188]. Basically, these schemes and strategies are aimed at promoting landscape and species resilience via an understanding of: (a) resistance (ability of a system to remain unchanged); (b) resilience (system ability to recover from disturbance or to adapt to new climatic conditions without undergoing transitions to new states, which would imply in loss of ecosystem functioning, populations, and species); and (c) managing changes from one system state to another. In this context, the identification of resilient areas to climate change is essential to map regions that are either likely to be more stable or will offer some relative level of resilience within a given landscape. In these resilient areas, many actions will take place to promote changes – by restoration, translocation, and adaptive management and monitoring of target populations, communities and ecosystems, for example – and/or increase resistance of ecosystems to threats that would be potentially magnified by climate change, for example by reducing undesirable or extreme effects of fires, insect attacks, diseases, and invasive species.

Consequently, using future species distributions in area prioritization per se as targets for biodiversity conservation must be accompanied by an understanding of resilient areas and paths and connectivity between current and future species distributions [22, 189]. Otherwise, some populations will be trapped by natural or anthropogenic barriers to their movements, preventing them from reaching suitable climates in the future, thereby increasing the risk of species extinction [190]. Additionally, the presence of an appropriate climate niche in the future does not guarantee species persistence if other environmental requirements are not also available or if obligate biotic/abiotic interactions are lost. For example, specialisation of plant species to particular fertile soils can prevent the establishment and development of many dry forest tree species in soils that are relatively poor in nutrients.

Species distribution modelling (SDM) is a widespread technique used to assess species' exposure to future climate change [17, 191, 192] and to identify areas for future prioritization [67, 193]. In Brazil, SDMs have been used to estimate future species distributions for different groups of organisms to identify climate-induced shifts in distributions and to prioritize areas under future scenarios of climate change [67-71]. However, all of these studies are based only on single taxonomic groups. None of them has considered the prioritization of future areas for multiple taxonomic groups. This gap impedes the identification of resilient areas for biodiversity in area prioritization, and compromises the cost-effectiveness of future conservation and the persistence of species.

Species are shifting and will continue to shift under a changing climate. Adapting to these changes requires identification of important biodiversity areas now, where they are likely to be in the future, and the pathways between current and future ranges. Given the uncertainty in future projections, the most robust approach is to consider all three aspects and to look for sets of areas that will cover a good representation of each aspect in one solution. Our aim is, therefore, to identify a set of priority areas for biodiversity conservation that are important for current distributions and resilient to the effects of climate change on species distributions of multiple taxonomic groups, considering endemic, threatened, and declining species. The Paranã River Basin – a critical area for biodiversity conservation within the globally significant Cerrado biodiversity hotspot [30] – was used as a case study to evaluate species representation in area selection in both current and future climatic scenarios. We compared our results with the areas in the study region already proposed for biodiversity conservation by the Brazilian government [29] and also identified congruities and gaps between both approaches. We which areas should be prioritized for biodiversity conservation under climate change within a global hotspot, offering to stakeholders a clear rationale for area prioritization based on explicit criteria.

5.3. Methods

5.3.1. Data sources

Species distribution modelling (SDM) – We fitted habitat suitability models to existing species data [58] to create spatial predictions of biodiversity features across the planning region for current [175] and future climates. A total of 2159 species (1813 plants, 245 birds, 44 mammals, 30 reptiles, and 27 amphibians), of which 30 are threatened and 208 are endemic to the Cerrado biome, were modelled at 0.1 degree grid (~1km) resolution. We estimated environmental suitability for each species as a function of climate, landscape, and substrate variables in the gridded study domain using Maxent [118, 132]. We used the same bioclimate variables (annual mean temperature, temperature seasonality, annual precipitation, precipitation seasonality, and precipitation of the driest quarter) used for predicting current distributions [see 175] to predict future distributions.

Multiple atmosphere–ocean general circulation models (GCMs) are recommended for predicting species distributions in the future [172], but the selection of GCMs remains problematic [194]. The GCMs generally converge well at the global level, but show a wide range of climatic conditions for individual regions. Factors underlying regional variation include land cover, elevation, climate zones, and chaotic behaviours embedded in multi-decadal simulations of variability [194]. These differences among model simulations are recognized as one of the major sources of uncertainty for regional-level assessment [194, 195].

The performance of different GCMs has been tested according to their skill scores to predict temperature and precipitation based on the inverse root mean square error (RMSE) of the model simulation relative to observed variables [194]. The skill scores were defined for 17 different GCMs globally at the resolution of 2^0 x 2^0 grid cells over the same time period. Inverse RMSE values were normalized to values between 0 and 1 to represent the relative skill scores of each GCM (by dividing the inverse RMSE of one GCM by the sum of the inverse RMSE values of the 17 GCMs). A threshold of 0.06 was used to define performance above the average [see 194 for methods details]. For the region of interest of this study, several GCMs had performance above average [see 194]: the NCAR PCM1 and UKMO-HADCM3 for the wet an dry tropical areas that include the Amazon and part of the Cerrado biomes; the MRI CGCM2.3.2 and UKMO-HADCM3 for the dry semi-arid areas, which include Caatinga biome and part of the Cerrado; and the MIROC3.2 (hires), MPI ECHAM5, and UKMO-HADCM3 for the temperate sub-tropical areas, which include parts of the southern Cerrado and Atlantic, Pantanal, and Pampas biomes in Brazil. However, the UKMO-HADCM3 GCM consistently presented highest skill scores (0,17 – 0,22) for temperature and precipitation for the study region which, besides variations in climate, elevation, and land cover typical of the Cerrado biome, also encompasses parts of the climatic influences of the Amazon and Caatinga domains.

Based on these results, for the exploratory evaluation of the possible impacts of climate change on biodiversity in the study region, spatial predictions of future climate used the UKMO-HADCM3 Atmosphere–Ocean General Circulation Models - GCM (available by International Centre for Tropical Agriculture in http://ccafs-climate.org). The selected climate scenario was (SRES) A1B from the Intergovernmental Panel on Climate Change Fourth Assessment Report, which reflects an intermediate severity of projected carbon emissions [196] to the year 2050.

We also restricted future distribution models by habitat structure and substrate as we did for the current models. The performance of SDMs was evaluated in the same way as in previous chapters [168, 175], using values of the AUC. For conservation assessment, we retained species with models having: (a) mean AUC greater than 0.70 in a 10-fold cross-validation (for those species recorded in 20 or more localities); and (b) AUC greater than 0.70 for species with less than 20 records that were not tested in the cross-validation [for a complete description of the methods used see 175]. The final maps of species richness for current and future climates were obtained by combining the logistic output for each species and converting it into a binary 0/1 (absence/presence) using the values of the "equate entropy of thresholded and original distributions" logistic threshold. These maps represent the sum of potential presences across all species of each taxonomic group by 0.1 degree grid cell.

5.3.2. Conservation prioritization

To identify priority areas for conservation of species under present and future climates, we used the Zonation software [172, 177], with the additive benefit function applied to remove cells in reverse order of conservation priority [177]. As in Chapter 4, Zonation basically produces a hierarchical ranking of conservation priority over the entire landscape [177, 178], while the additive benefit function gives more weight to locations with high relative values of biodiversity features [172] during the process of cell removal (see Supporting Information 3 for a brief description on Zonation and the additive benefit function as a cell removal rule). However, in this study, besides given more weight to rare, threatened, and endemic species, considering that these species relies more on conservation actions than that of common species, we additionally considered the impact of predicted climate change on species distributions. The assumption in doing so was that species with distributions that are reduced by climate change should be weighted more than species with future ranges that are unchanged or larger in extent than present.

Across SDMs, weights were distributed among the taxonomic groups considering: the number of species in each group (higher number of species, greater weight); quality of the models (based on AUC and total number of occurrences per species); and the potential for plants to serve as structural habitat for animals. Weights were therefore distributed 38% for plants, 23% for birds, 15% for mammals, 12% for reptiles, and 12% for amphibians. Within taxonomic groups, species were weighted according to the formula:

$$\mathrm{W}_{\mathrm{sp}i} = \left\{\left[\left(\sum\left((\mathrm{IUCN}i_{\mathrm{sp}i} + \mathrm{EB}i_{\mathrm{sp}i} + \mathrm{EC}i_{\mathrm{sp}i}) + (\mathrm{Ccc}fi_{\mathrm{sp}i)})\right)\right) * \mathrm{AUC}i_{\mathrm{sp}i}\right] + 1\right\}$$

where:

$\mathrm{W}_{\mathrm{sp}i}$ = the weight of species *i*

$\mathrm{IUCN}i_{\mathrm{sp}i}$ = IUCN category of species *i* assuming values: 1, for no threatened species
2, for vulnerable species
3, for endangered species; and
4, for critically endangered species

$\mathrm{EB}i_{\mathrm{sp}i}$ = Endemism for species *i* considering the percentage of species *i* distribution modelled restricted to the Paranã River Basin boundaries and a buffer area of 1 degree beyond the Basin's extremities

$\mathrm{EC}i_{\mathrm{sp}i}$ = Endemic species *i* of the Cerrado biome occurring within the Basin assuming values: 1, if present
0, otherwise

$\mathrm{Ccc}fi_{\mathrm{sp}i}$ Constraints on the distribution of species *i* posed by climate change, expressed as a function of proportional reduction in predicted niche space in the future, relative to the current modelled distribution (increases not considered)

$\mathrm{AUC}i_{\mathrm{sp}i}$ = 1 * (AUC value of model of species *i* distribution tested for accuracy with cross validation), or
0.5 * (AUC value of model of species *i* distribution not tested for accuracy with cross-validation)

Priority areas for current and future solutions are presented by Zonation in maps of nested values. So the most valuable 2% of the landscape lies within the most valuable 5%, which lies within the top 10%, and so on. The top 20% of priority areas for future conservation based on future species distributions was compared with top 20% of areas selected based on current distributions. We used the 20% figure because it is the standard for conservation extent set out in the new Brazilian Forest Code [180]. This code aims to preserve 20% of each rural property within the Cerrado biome as Legal Reserves, with the goal of good representation of biodiversity features [168].

Zonation also evaluates habitat quality and connectivity between time steps in future projections of species distributions and takes into consideration species-specific dispersal abilities [172]. By prioritizing areas based on connectivity between current and future distributions, it is possible to identify priority areas that are not only currently valuable for the species but which also facilitate dispersal to expected future areas of value [193]. However, estimation of species-specific dispersal distances is difficult due to the inadequate biological knowledge about almost all species present in the study area. Consequently, in this study, we consider only the overlap between the top 20% of areas identified in current and future climates. We compared the overlap between priority areas identified in this study and areas proposed as priorities for the establishment of protected areas by the Brazilian government [29]. Finally, we also evaluated the representativeness of the areas selected in the current protected areas network, that is, the overlap of identified priority areas and the existing priority areas in the Basin .

5.4. Results

Differences in species richness patterns between current and future climate conditions depend on the taxonomic group evaluated (Figure 5.1). For the current climate, vertebrates are more diverse at lower elevations while plants are more diverse in mountainous areas. Across the valleys, habitats formed by arboreal savanna and dry forest are more structured vertically, with more strata or layers [175]. This higher vertical structure allows greater habitat compartmentalization of vertebrate species. In contrast, horizontal heterogeneity of vegetation is higher across the mountain range that surrounds the Basin on its western and southern borders and extends from the central-west towards the north-east, in a landscape dominated by grasslands [175].

For future climatic conditions, climate niches for most of the plant species are further constrained to the high elevations of the mountain range. In contrast, climatic niches favourable for vertebrates expanded across the valleys and over the lower slopes of the mountain ranges. This expansion is related to a general pattern of homogenization and increase in richness in many parts of the Basin. Even in areas predicted to be less rich in the future (with dark blue colours in figure 5.1), the total number of species predicted is higher in future than in current scenarios, except for plants. However, a small reduction in the total number of future species of birds, mammals, and reptiles was predicted (from 238 species to 235 species for birds, from 41 to 40 for mammals, and from 30 to 29 for reptiles) (Figure 5.1). Additionally, future niches for 32 species (1 amphibian, 16 birds, 9 mammals, and 6 reptiles) show reductions in distributions (ranging from 11 to 70%) into the future, including 13 threatened and 7 endemic species.

Bioclimatic variation between current and future scenarios is mainly related to the warmer average annual mean temperatures predicted for different classes of elevation in future (3.2 ^{o}C between 240–430 m; 3.0 ^{o}C between 430–585 m; 2.9 ^{o}C between 585–775 m; 2.8 ^{o}C between 775–1000 m, and 2.7 ^{o}C between 1000–1663 m). Other bioclimatic variables used to model species distribution do not differ significantly from current scenarios, except for a reduction in average annual precipitation of 333 mm*year^{-1} in elevations between 240–430 m (see Table 5.1).

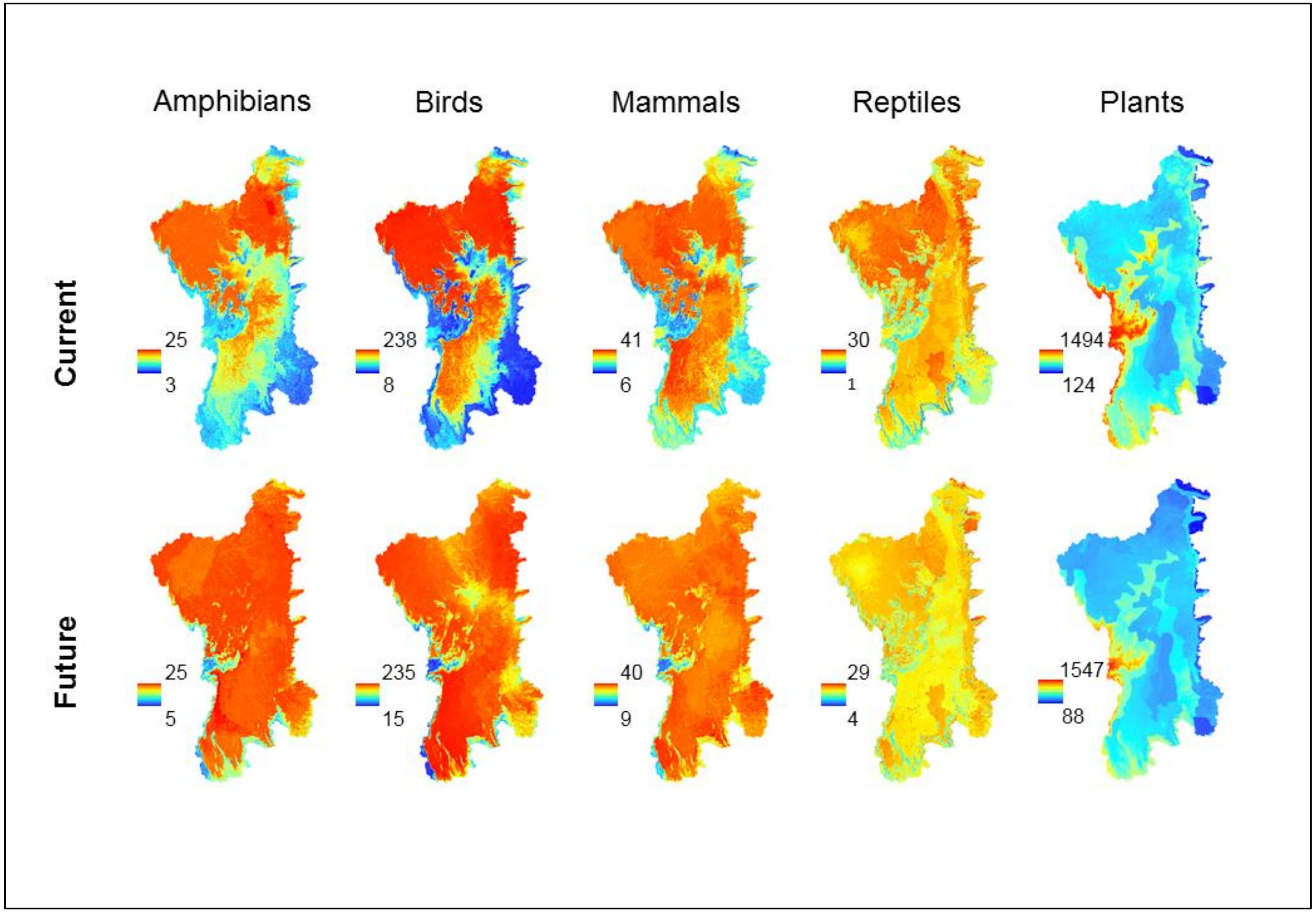


Figure 5.1. Species richness patterns for current and future climates in the Paranã River Basin. The scales are normalised as percentages of maximum values of species number within each group.

Table 5.1. Bioclimatic variables for current and future climates of the Paranã River Basin by 2050, considering the intermediate emission scenario (SRES4 - A1B) and using the global climate model UKMO-HADCM3.

Bioclimatic variables	**Time**	**Values**	**Elevation (m)**				
			240-430	**430-585**	**585-775**	**775-1000**	**1000-1663**
Bio_1 Annual mean temperature (°C)	current	mean	25.98	25.04	23.90	22.90	21.81
		SD	0.53	0.91	0.93	0.98	0.95
	future	mean	29.18	28.02	26.84	25.72	24.49
		SD	0.61	1.04	1.07	1.13	0.97
Bio_4 Temperature seasonality	current	mean	722.25	902.71	971.45	973.13	920.91
		SD	63.25	300.46	239.71	207.12	123.44
	future	mean	728.07	896.70	953.16	932.08	870.28
		SD	89.10	255.17	202.88	165.61	87.03
Bio_12 Annual precipitation (mm)	current	mean	1526.09	1425.35	1325.60	1434.61	1556.01
		SD	122.58	235.98	227.37	162.33	130.04
	future	mean	1193.18	1257.34	1193.80	1419.06	1620.21
		SD	110.26	241.08	280.03	231.61	95.10
Bio_15 Precipitation seasonality	current	mean	81.67	84.47	84.75	81.00	78.09
		SD	3.33	3.48	3.86	3.78	2.69
	future	mean	79.88	84.55	86.20	83.79	81.49
		SD	3.77	5.52	5.90	4.02	2.27
Bio_17 Precipitation for driest quarter (mm)	current	mean	13.71	12.58	11.34	18.63	24.04
		SD	3.95	4.03	6.23	8.79	5.51
	future	mean	11.92	11.18	10.03	17.33	22.64
		SD	3.86	3.91	6.20	8.88	5.51

The top 20% priority areas selected to preserve current and future species distributions are shown in Figure 5.2A and 5.2B, respectively. These priority areas give particular attention to threatened and endemic species and to species which current distributions that are predicted to contract under climate change.

For current species distributions (Figure 5.2A), the most important areas (top 20%) are divided between the Basin´s valleys and mountain ranges. Within the valleys, the most important areas are in the north-western and north-eastern portions (in the Tocantins depression – see Figure 1.3) and in the central-east (at the ‘Vão do Paranã’ region – see Figure 1.3). The mountain ranges were assigned the very highest priorities, with the most important areas on the western rim of the Basin (at the ‘Serra Geral do Paranã’ range – see Figure 1.3) and, to a lesser extent, on the eastern rim (at the ‘Serra Geral do Goiás’ range – see Figure 1.3).

For future species distributions, the most important areas are centred on the mountain ranges, and on plains and small hills surrounding these ranges. The moutain range selected as the highest priority for future scenarios of climate change is located in the western rim of the Basin (the ‘Serra Geral do Paranã’ range – see Figure 1.3), which is oriented from south to north. In the central-western portion of the highest priority area, the range splits in two, extending towards the north and north-east, where high-priority areas for future were also selected. Additional priority areas in most elevated areas are apparent in the far north-east of the Basin, along the ‘Serra Geral do Goiás’ range (see Figure 1.3). In the Basin´s valleys, the most important areas are located in the central-east, in the ‘Vão do Paranã’ region (see Figure 1.3).

Considering the overlap between the top 20% of priorities based on both current and future species distributions (Figure 5.2C), highly elevated areas were generally identified as the most important parts of the landscape to be protected. Most of this overlap is located in the elevated terrains of the central-western rim of the Basin, extending towards the north and north-east, and on the north-eastern rim. Equally important are the overlaps in current and future priorities identified at lower elevations in the central-eastern portion of the Basin.

Interestingly, the areas selected as priorities under future climate conditions on the montane gradients and surrounding plains and small hills (Figures 5.2B and C) lie between two sets of areas identified as priorities under the current climate: those in the Basin's valleys and those on the mountain ranges. The areas selected for the future therefore contribute to establishing spatial links between the most important areas indicated for current climate selected in the north-western, north-eastern, and central-eastern parts of the Basin (Figure 5.2C).

On average, current prioritization solutions retain a higher proportion of biodiversity features (Figure 5.3) when considering the top 20% fraction of the landscape (29.1% – see Table 5.2) or the intersected areas between current and future scenarios (15.7% – Table 5.2). This overall difference is reflected in most of the taxonomic groups, except for plants which had a higher average retention in the solution for future climate (Table 5.2). Higher retention of plants in the future solution was favoured by the general contraction of plant distributions to higher elevantions on the mountain ranges.

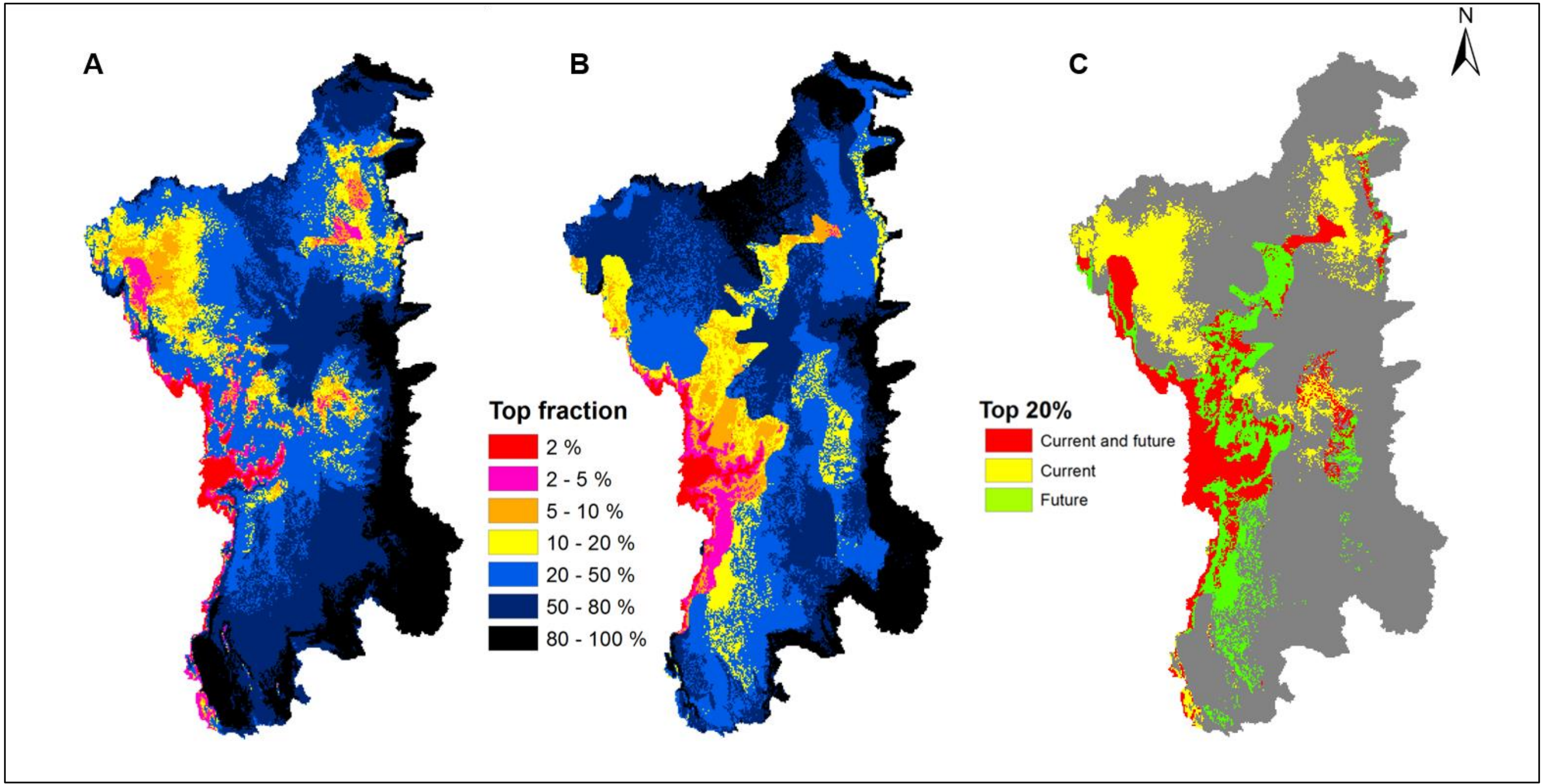


Figure 5.2. Spatial prioritization maps for biodiversity conservation in the Paranã River Basin, considering species distributions of amphibians, birds, mammals, reptiles and plants. A. Area prioritization for the current climate. B. Area prioritization for future climate. C. Intersection between area prioritization for current and future climate (areas in grey are below 20% priority in both current and future climates).

In ‘A’ and ‘B’, warmer colours (from yellow to red) indicate the highest priority areas for conservation, considering the top 20% fraction of the landscape (yellow), which includes the top 10% (orange), top 5% (pink) and top 2% (red).

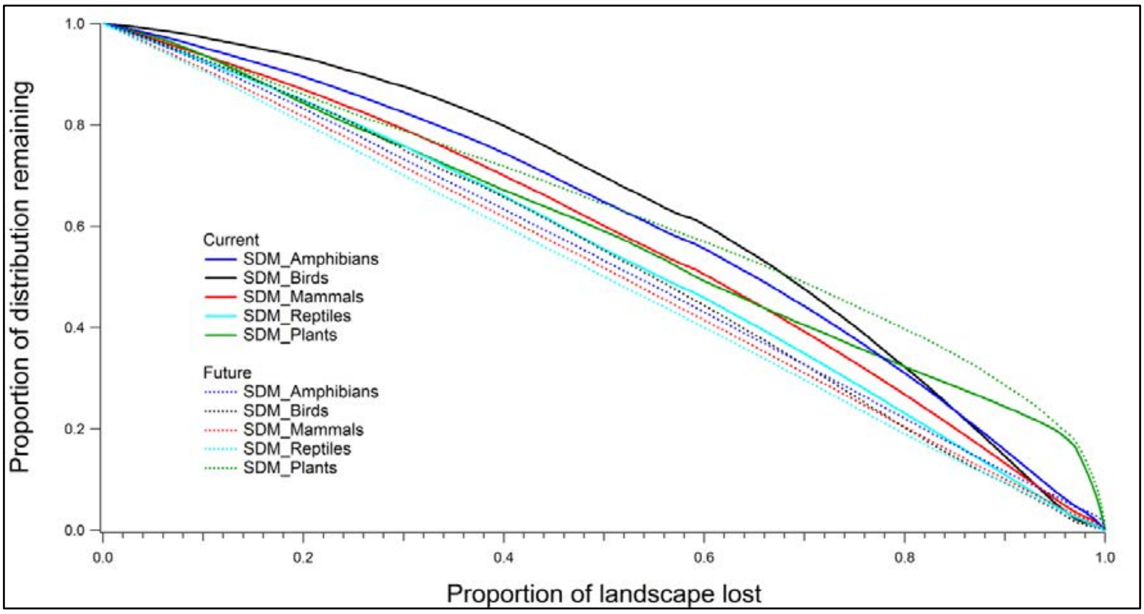


Figure 5.3. Proportion of current and future species distribution of amphibians, birds, mammals, reptiles and plants remaining in area prioritization according to the proportion of the landscape lost during the process of cell removal analysis in the Paranã River Basin.
X-axis shows the fraction of the landscape lost for conservation during the process of progressive cell removal. *Y-axis* shows the average proportion of initial distributions of species in each taxonomic group remaining in priority areas.

Table 5.2. The average proportions of initial distributions of species in each taxonomic group in the highest 20% priority cells, for current and future climates in the Paranã River Basin.
Cu ∩ Fu = remaining species distributions in the intersected area between current and future scenarios of climate change, covering about 10% of the total ~60.000 km^2 of the Basin´s area.

Amp = amphibians, Bir = Birds, Mam = mammals, Rep = reptiles, Pla = plants, and SD = standard deviation.

Area prioritization	**Remaining average distributions across species in each taxonomic group (%)**						
	Amp	**Bir**	**Mam**	**Rep**	**Pla**	**Average**	**SD**
Current climate (top 20%)	31.0	32.2	26.8	23.1	32.3	29.1	4.0
Future climate (top 20%)	22.0	20.3	20.4	19.0	39.7	24.3	8.7
Cu ∩ Fu considering current species distribution	15.7	14.5	13.2	10.9	24.4	15.7	5.2
Cu ∩ Fu considering future species distributions	11.6	9.3	10.0	9.3	28.5	13.7	8.3

With only ~10% of the total area of the Basin, the intersection between highest 20% priority areas in current and future solutions retains disproportionately large portions of species distributions (see Table 5.2). This is likely to reflect the strong weightings used to maximize retention of species with smaller ranges: endemic and threatened species, and those with reduced ranges under future climatic conditions. Additionally, from the total of 2159 species modelled, only seven species of plants and one bird species are not represented at all in the intersected area, and all these missing species are found commonly in other parts of the Cerrado biome. The high species retention in the final solution indicates the intersected area as important for biodiversity conservation in the Basin considering the effects of climate change.

The overlap between priorities for biodiversity conservation revealed by this study (for current and future scenarios of species distributions) and areas indicated for conservation actions by Brazilian Government [29, 87] is very low (Figure 5.4). Of ~82% of the total area of the Basin officially proposed by the Government for conservation actions, only 7.0% of the intended protected areas (conservation units) overlap our current and future intersected areas (covering 55% of the intersected area). About 5.2% of proposed sustainable use areas overlap our intersected priorities (covering 41% of the intersected area). About 0.1% of the existing protected areas overlap our intersected priorities (covering about 0.7% of the intersected area).

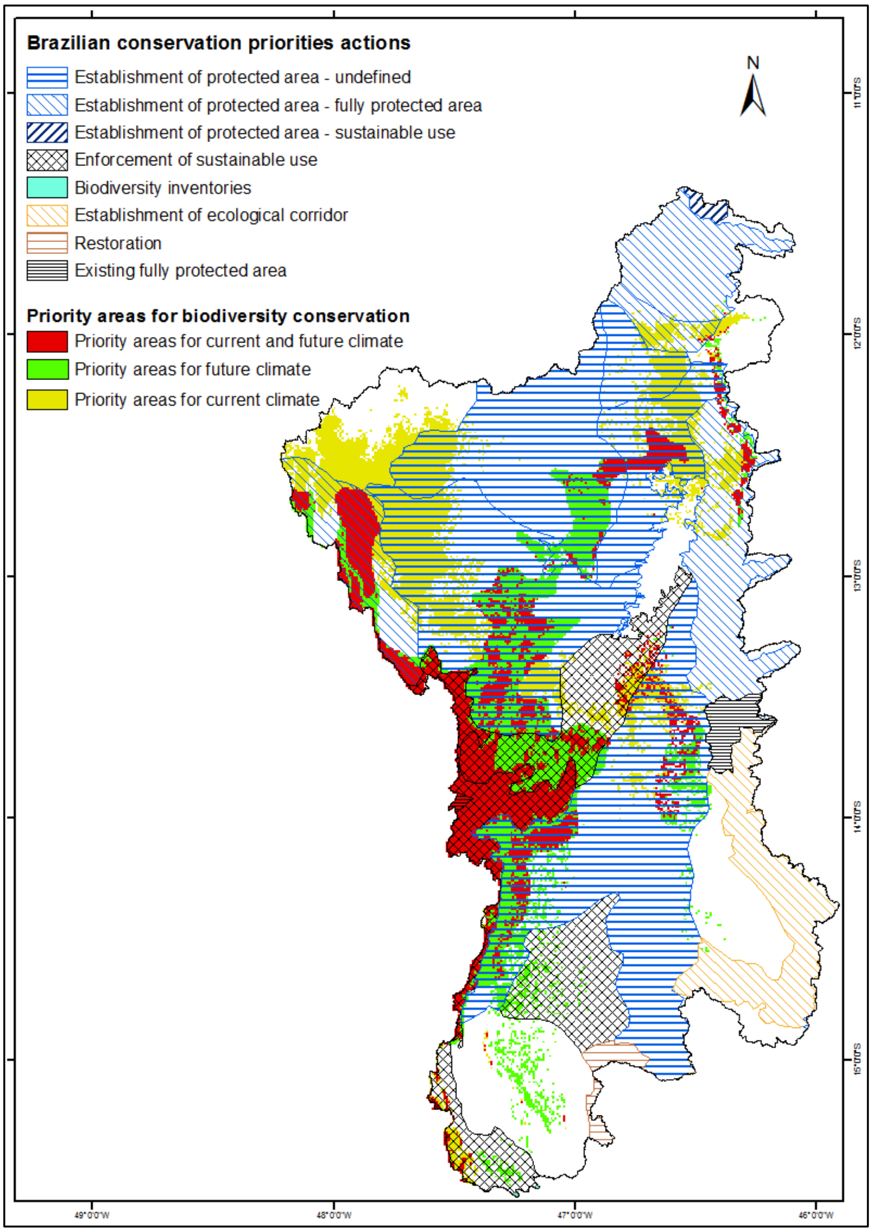


Figure 5.4. Priority areas for biodiversity conservation identified in this study, and the Brazilian Government's conservation priorities for the Paranã River Basin. Priority areas for biodiversity conservation identified by this study are indicated in three colours: red – top 20% areas for both current and future climates; yellow – top 20% areas for current climate; and green – top 20% areas for future climate. Conservation priority actions identified by the Brazilian Government are represented by proposed fully protected areas, protected areas where sustainable uses are allowed, and areas that need to have their category of protection to be defined (fully protected or sustainable use). Enforcement of sustainable use, establishment of ecological corridors and restoration categories are conservation actions to be promoted on private land. Biodiversity inventories identify priority areas for biodiversity sampling. Existing fully protected areas are Chapada dos Veadeiros National Park on the western border of the Basin and Terra Ronca State Park, on the eastern border.

5.5. Discussion

The existing fully protected areas within the Paranã River Basin are not sufficient to protect important areas important for species under future climatic conditions. Low overlap between current protected areas and the top 20% priority areas identified for the intersection of present and future species distributions (0.7% of the intersected area) indicates the need for a considerable increase in size and number of protected areas to adequately represent the biological diversity of the Basin, considering future climatic conditions.

Additional protected areas proposed by the Brazilian Government [29] overlap only partly with the most important areas for biodiversity under both current and future climates. The poor overlap is noticeable at lower elevations (see Figure 5.4), but also on the montane gradients and the top of the mountain range in the west (Figure 5.4). On this range are most of the threatened and endemic species present within the Basin, as well as a concentration of plant species and unique assemblages of plants, birds, and mammals [58, 168, 175]. Furthermore, under future climate change, this mountain region seems to be the most resilient part of the Basin, and is likely to be a refuge for species from the range high elevations as well as those presently on the montane gradient and the Basin's valleys.

5.5.1. Shifts in biodiversity patterns

Predicted shifts of species with climate change from lower elevations towards higher elevations and from lower latitudes towards higher latitudes in the Paranã River Basin are similar to predictions in other parts of the globe [197-200] and the Cerrado biome [67, 68]. However, in this study, the shift towards higher latitudes (see Figure 5.1) seems to be mostly due to the increase in elevation towards the south (see Figure 1.3), so the shift is mediated by temperature. The landscape of the Basin might also interfere with the velocity and direction of future changes of species distributions. Generally, the velocity of climate change is low on mountains and higher on flat areas [191, 201]. Assuming this is true in the Basin, vertebrates and plants of the flat areas of the valleys might face the need for faster geographic changes than those on elevational gradients, while the interaction between changes in temperature and precipitation might result in multi-directional shifts according to the organisms' position in the landscape and climatic tolerances [199]. Organisms from the Basin’s flat valleys might also be indirectly affected by changes in vegetation structure over time, which are related to current patterns of vertebrate richness [175].

Species’ adjustments to changing climate are also constrained by barriers posed by land cover change [202, 203]. Ongoing fragmentation of landscapes by cattle ranching and agriculture in the flat areas of the Basin are likely to limit connectivity across altitudinal gradients. Keeping connection between current and future areas is therefore an important strategy for organisms adaptation [18, 22, 189, 190] and should be considered in any planning process of the region.

Even without land use change, range adjustment by some species in response to climate change might be difficult. Although many vertebrates can move relatively quickly across a landscape, plants move slowly, generation after generation. A phylogenetic niche conservatism and stable geographically structured phylogenies of dry forest tree species suggest, for example, a limited potential for range adjustment of associated woody plants [204]. Additionally, changes in patterns of pollinators, dispersers, and even exposure to new predators due to climate change might increase the vulnerability of many species to local extinction or reduce distributional ranges.

Shifts in contemporary climatic patterns of the Basin might affect the expansion of forested systems during this interglacial period. Disregarding impacts promoted by human activities, forested systems are advancing from the richest soils and limestone outcrops on the flat areas of the Basin towards the mountains. These forests occupy ravines and the richest juvenile soils exposed by continuous intemperism processes, whether on calcareous rocks or other parent material rich in nutrients. Although average precipitation, precipitation in the driest quarter, and rain seasonality will not differ significantly from current to future climates for most of the Basin (see Table 5.1), these variables indicate a trend towards less rain annually and in the dry season. Additionally, a trend of more concentrated rains could also contribute to increased extreme climate events in the future. For forested environments of the Basin, less rain in the wet season could reduce seed germination [205, 206] and establishment and growth of seedlings, thereby reducing their survival during the subsequent dry season. Changes to establishment and development patterns of arboreal species imply rapid changes in the structure and composition of forests and arboreal savannas. These changes, in turn, will affect vertebrate species that rely on forested ecosystems, and consequently, the richness patterns now observed for vertebrates (see Figure 5.1).

The pattern of richness expansion of vertebrates under climate change towards the south and uplands (Figure 5.1) might be an indicative of one of the multiple possible routes of arrival of 'new' species currently occurring in regions outside the Basin. The Basin's aperture in the intersection of the Paranã and Tocantins Rivers, and the relatively low elevation of the hills on the northern rim of the Basin might offer an open route for species dispersal from the northern part of the Cerrado. Some species of the Amazon biomes could move into the Basin through the Tocantins depression into the Paranã Valley (see Figure 1.3). Additionally, if the historic Quaternary climatic narrow corridor in Central Brazil [75] was reestablished under future changes in climate, a possible biogeographic link for fauna and flora dispersion routes would also be restored, bringing to the Basin potential 'new' species from the drier areas of the north-eastern regions of the Cerrado and Caatinga biomes. Moreover, considering the upslope species shifts on the western side of the 'Serra Geral do Paranã' range, more species could arrive in the Basin due to migration from the western, southern, and northern neighbouring regions of the most elevated parts of the Basin.

Some plant and animal species might locally have their populations reduced or even extirpated given future changes in climate niches. Some animal species are predicted to have significant reductions of their spatial niches (~70%). For plants, this pattern is more frequent, given the overall contraction of species distributions to the mountain ranges with climate change. For species with current distributions restricted to the tops of the ranges, the favourable current climate niche might not be present in future.

Although reductions in species distributions are a key driver of species decline, cases of long-term species persistence in remarkably small ranges demonstrate that the link between proportional reduction in extent and extinction likelihood cannot be assumed simplistically [207]. However, extinction risk of reduced ranges of the Basin's species could be exacerbated by immigration of species from disparate areas. In the mountain regions of the Basin's western rim, existing ecological assemblages could be strongly modified due to rapid spatial turnover of species and possible local extinctions. Additional invasive plants, animals, pathogens, and altered fire regimes might interact with climate change at local and regional scales [186] to put some of the Basin's species at risk. Those are some of the multiple possibilities arising from climate change in the Basin.

5.5.2. Area prioritization

Prioritization of areas for the Paranã River Basin reveals congruence between current and future areas indicated as priorities for biodiversity conservation (red areas in Figure 5C). These areas are of extreme importance for the Basin's biodiversity conservation considering both present and future climate conditions. Besides mountains ranges, the areas include elevational gradients and parts of plains and small hills that surround those ranges and limestone outcrops. These areas together form an escape route for species living in the valley bottoms of the Basin that will have to shift their ranges under climate change. Most of the vegetation types present at low elevations – grasslands, savannas, and dry forests – are found within the areas selected [58, 168, 175]. Additionally, these areas already encompass parts of all species distributions for current and future climate predictions (see Figure 5.3 and Table 5.2). Additional important areas are those prioritized for the current scenario (yellow areas in Figure 5.2C). The maintenance of these areas might avoid eventual reduction of species ranges by land-use change and promote species persistence long enough for them to disperse to future sites.

Less emphasis should probably be given to the areas identified only in future scenario of climate change due to the uncertainties involved in modelling future species distributions as discussed below (see section 5.5.3). However, the priority areas identified for the future that are congruent with areas identified to be current priorities, together with priorities identified only in the current climate scenario, might also be necessary for promoting resilience and facilitating species persistence in the future. Furthermore, future priority areas might also support the valuable spatial biodiversity links between areas selected in the Basin´s low elevations in the central-eastern and north-western regions and the mountain regions on the Basin´s western rim (see Figure 5.2C). Additionally, the corridor formed by future priority areas linking the eastern and western borders of the Basin might offer a climate corridor for species movement.

Models of historic dry forest distribution during Quaternary climatic fluctuations have indicate a potential narrow corridor in Central Brazil connecting the extremes of the predicted distribution range of the core areas of dry forest formations in South America [75]. This corridor could have served as a biogeographic link for fauna and flora dispersion routes through the Cerrado biome across the Paranã River Basin region, connecting the Brazilian Caatinga to the Andean dry forests [75]. The presence in the Basin of species like the Brazilian baobab (*Cavanillesia arborea*), a tree typical of the Caatinga biome, or the climbing cavy (*Kerodon acrobata*), endemic to the limestone outcrops of the Basin and with a sister species, *K. rupestris,* widely distributed in the Caatinga domain, illustrates past expansion of the Caatinga drier system over the region. The same climate corridor pattern seems to be reproduced for the future climatic scenario, making the conservation of these potential drier corridors a priority.

Selection of areas for the future appears to be heavily influenced by the high number of plant species and the weights given to species for reductions in range under climate change. One result is that areas selected for future priorities overlap substantially with current priorities for plant species (see Figure 5.1). However, a relatively good representation of animal species is retained in the area prioritization for future conditions (Figure 5.3 and Table 5.2).

The maintenance of protected areas that form spatial link between the flat areas of the Basin’s bottom to the top of mountains should be considered a strategy to help species to move and adapt to future climate shifts. Corridors, stepping-stones, and matrix management are necessary to facilitate species movement [18, 189, 190], mainly for those for which mobility relies only on the movement of propagules. Additionally, the maintenance of elevational gradients in reserve designs could be critical to protect places where evolutionary processes are prominent [13, 145], although no systematic studies have investigated patterns of diversification in the Basin. The high number of endemic species recorded for different taxa in this region [58, 95, 154] might be indicative, however. Edaphic interfaces, ecotonal zones, inter-basin riverine corridors, macroclimatic gradients, and upland–lowland gradients present in the mountain ranges of the Basin could be considered potential surrogates for biodiversity processes such as diversification and contraction to refugia during harsh climates [18, 145, 164, 182, 183, 208, 209].

5.5.3. Uncertainties in area prioritization and shifts in biodiversity patterns

The uncertainties in this study are mainly related to the input data used in the correlative models for estimating species distributions in current and future scenarios. As discussed in the previous chapter [168], prioritization of areas based on species distribution models considering current scenarios must be interpreted with care due to the limitations imposed by scarce records for many species. These limitations might be exacerbated by extrapolating modelled relationships between species occurrences and environmental variables into the future [137]. Additionally, even for the best models, we cannot be confident that models for current climate conditions will be the best to estimate future distributions [210]. In respect of the statistical models used to predict future distributions, most of them poorly represent multiple ecological and evolutionary processes operating at different spatial and temporal scales that are expected to determine contemporary distributions of most species [211]. Future projections of species distributions come with a variety of other uncertainties from disparate sources [71, 137, 207, 210, 212].

Uncertainties in future species distributions based on climate projections are mainly related to the inherent uncertainties contained in the initial dataset used for input to the modelling software, the statistical method used for niche modelling, the methodological differences in general circulation models (GCM) and their ability to depict climate change in different regions, and differences between gas emission scenarios (GES) [71, 137, 207, 210, 212]. Of these, differences among SDM predictions have been suggested as the major source of uncertainties by increasing variation in future projections [71, 207, 212]. SDMs are, of course, dependent on the the selection of GCMs and GES to provide climatic variables with which to model [71, 212]. One study has shown that the interaction of SDMs with GCMs is as important as the effects of GCMs alone [71].

The ensemble forecasting framework considering several SDMs, climate projections from different GCMs, and GESs has been proposed to allow quantitative evaluation of the relative importance of different sources of variation in predicting future species distributions [71, 212]. The quantitative approach in conjunction with maps of uncertainty and maps of projected impacts on species could lead to a more realistic understanding of the future effects of climate change [71, 212] allowing conservation decisions to be taken with an awareness of the inherent uncertainty involved in predictions [212].

However, we consider that the predicted future distributions presented by this study are valuable as an exploratory analysis of the possible general trends in future climate shifts and species turnover for the Paranã River Basin. SDMs contributed the most to uncertainty in the range of predictions, when compared with GCMs and GESs [71]; that is, different SDMs can result in distinct predictions for future species distributions, compared to different GCMs and GESs under the same dataset in different parts of the world. However, different SDMs tended to give similar results and their differences are of minor concern for predictions of species turnover rates in some few regions in the New World [71], which seems to include the Paranã River Basin [see 71].

Different climate projections (GCMs) and their interaction with SDMs show high variations, mainly in the dry regions of the eastern part of South America, in the Caatinga biome [see 71]. While GCMs converge relatively well at the global scale, individual outcomes may vary significantly at the regional scale in different parts of the globe. GCM model structures, parameterization, and model validation might be affected by the distribution of geographic variables (such as land cover, earth surface elevation, and climate zones) [194], resulting in the differences among GCMs for the same regions in different parts of the world [see 71]. A regional mapping of GCM performance might be useful for nesting of regional climate models and guide the choice of a specific or set of GCMs for a certain region [194]. Given that, a more systematic evaluation of the uncertainties related to the different approaches to niche modelling, global circulation models, and different gas emission scenarios should be undertaken to properly assess different possible ways of reducing uncertainties in predicting the impacts of climate change on biodiversity of the Basin.

5.5.4. Representativeness of current and future species distribution in current reserve network

The broad-scale conservation planning done for the Basin by the Brazilian Government [29] must be revisited to maintain resilience under climate change. Alteration of the status of some conservation actions planned under that strategy and/or application of supplementary measures in other areas are necessary to achieve long-term biodiversity conservation within the region. The priority areas identified in our study for the current climate combined with the priority areas for both current and future climate conditions should be considered for the establishment of fully protected areas. For the mountain ranges in the Basin, for example, the expansion of the existing fully protected reserve network is strongly recommended. The enforcement of sustainable use, as envisaged by Brazilian government [29] for this mountain range, might not guarantee biodiversity persistence under climate change. Removing or reducing non-climate-related threats to biodiversity or ecosystems present nowadays in sustainable-use areas would increase system resilience. Deforestation, habitat loss, habitat fragmentation, expansion of invasive species, and overharvest are examples of additional stresses that occur in sustainable-use areas but can be avoided by the establishment of fully protected areas, thereby increasing the resilience of these ecosystems to climate change [22]. Enforcement of sustainable practices in soils, terrains, and use of biodiversity in private and sustainable-use areas must be seen as relevant alternatives to maximize ecosystem resilience, but not the only strategies to be adopted for this mountain region.

The establishment of ecological corridors should be prioritized to increase the spatial link between current and future species distributions and facilitate the movement of species generally throughout the Cerrado biome. The current areas indicated by the Brazilian Government to establish ecological corridors [29] do not promote climate-related connectivity and do not guarantee the movement of species through the biome in a changing climate. Although ecological corridors do not constitute protected areas under the Brazilian national protected-area system [213], they might support some conservation actions to promote species persistence. Ecological corridors should be prioritized in the elevational arc that connects the mountain range of the central-western highlands to the plateau in the north-east of the Basin. In this arc, both current and future priority areas were identified (Figure 5.2C). The relative protection offered by the ecological corridors proposed by the Brazilian Government [29] would facilitate species movement from low to high elevations in the Basin and the movement of species through the biome generally by favouring the maintenance of biogeographical links established in past climate changes [75].

Furthermore, restoration of areas should not be restricted only to the Basin’s southern portion, as indicated by the Brazilian Government [29] (Figure 5.4). Restoration must be extended to others parts of the Basin identified here as priorities for biodiversity conservation in a changing climate. Restoration should include areas of dry forests on flat soils and limestone outcrops of the central-eastern portion of the Basin. These areas are under pressure from cattle ranging, agriculture, and mining activities [31, 50, 58]. Furthermore, these areas were identified in our study as having the highest priority for biodiversity conservation [168]. Additional restoration efforts must be implemented in dry forests and savannas in regions across the mountain ranges of the central-west and in the arc that connects this range to the plateau of the north-eastern rim of the Basin. In this arc, intensive cattle grazing is the main threat to biodiversity. Restoration in these regions, therefore, must be seen as an opportunity to increase ecosystem resilience to climate change and maintain connectivity within the Basin and throughout the Cerrado biome.

Finally, as well as needing modification to incorporate the results of our study, the conservation actions planned by the Brazilian Government for the Paranã River Basin must be supported by clear definition of conservation objectives, and followed by implementation. Engagement of diverse stakeholders is necessary to achieve full community commitment to biodiversity conservation [14, 73]. This engagement avoids discontinuities between the planning process and its implementation on the ground. The lack of stakeholder involvement in the past planning process (B.M.T Walter, personal communication) is becoming more evident in the Paranã River Basin. Since the Brazilian Government released the Cerrado priority areas in 2007, not a single fully protected area has been established in the Basin. Despite the uncertainties involved in species modelling and area prioritization, we believe we have produced solid data, analyses, and constructive ways forward to support managers in future decision-making by identifying which areas should be prioritized and why.

Chapter 6

Chapter 6 – Final considerations

In this last Chapter are discussed the main results and conclusions of this study and their implications for biodiversity conservation in the Paranã River Basin and the Cerrado Biome as a whole. We revisit here the central questions and objectives that initiated this study, discuss the problems faced during the execution of this work, and describe how we circumvented these problems to reach the most reliable possible results. Finally, we suggest future investigations that should take place to advance the knowledge of biodiversity conservation in this important region. We also discuss further investigations in the light of the knowledge already accumulated, gaps in biodiversity information, present and future scenarios of land use and climate change, and policies and tools for biodiversity conservation in Brazil.

6.1. Introduction

One of the biggest challenges to developing effective conservation plans for megadiverse regions in developing countries is finding reliable and sufficient information on species distributions. The biodiversity of many of these regions is poorly known. As we have seen in Chapter 2, the Paranã River Basin is not an exception to this rule. Addressing the first questions of this thesis about the Basin's biodiversity – what?, where?, and where not? - required carrying out biological inventories in the field and an intensive search for information on online databases, literature, museums, and herbaria. A large effort was also required to detect and remove duplicates in the dataset to obtain the most reliable information on species presence and gaps in occurrences.

Using the best available biological information gathered in Chapter 2, in Chapter 3, we modelled the distributions of individual species and species assemblages as a function of environmental variables. This procedure allowed us to predict species and assemblage distributions across the entire region of interest. In doing so, we answered our second major question that was to understand how species and assemblages are distributed in the Basin.

In Chapter 4, we used the knowledge about species and assemblage distributions developed in the previous Chapter to answer the question of primary stakeholders – which areas are the highest priorities for biodiversity conservation in the Paranã River Basin and why are these areas important? In this Chapter, we also compared and evaluated the use of species and assemblages distributions in spatial prioritization to inform conservation planning for the Basin.

Finally, in Chapter 5 we developed a comprehensive assessment of current conservation priorities and those under projected climate change, and made recommendations about interpreting this information for conservation actions that take a view of likely future conditions in the Basin.

6.2. Detailed considerations of the data chapters

The intensive search for biodiversity information developed in Chapter 2 resulted in the collation of a biodiversity database that we have used for developing an area prioritization for a region that was previously classified by Brazilian environmental authorities as poorly known for biodiversity [see 28]. During the process of data compilation, we were able to gather a large volume of useful data from different sources (online databases, literature, museums, and herbaria), as well as our specially designed field inventories. In total, more than 140,000 records from 12,126 species were compiled. This compilation revealed that some of the gaps in biodiversity information that impeded its utilisation in conservation planning were related more to the lack of organisation of this information into a unified database than to an actual absence of data. Additionally, it is worth highlighting the importance of data searches in literature. Most of the information available in the online database derived from herbaria and museums. However, reliable information that existed previously only in reports with limited circulation, such as theses, dissertations, monographs, and studies of environmental impacts, contributed significantly to the understanding of the biodiversity patterns of the region. This information was not available in herbaria and museums, or in online databases. Encouraging peer-reviewed publication of species lists for localities would enable a broader access to such data arising from future studies.

An analysis of the records gathered revealed a high rate of duplication and incongruities in information within data records. Data duplication was evident particularly in plants since duplicates of vouchers might be deposited in different collections around the world. Accessing this information frequently resulted in uncertainties related to the assigned taxonomic names and geographic coordinates associated with the same record in various herbaria. Solving these problems was critical for gathering reliable data and required the use of an innovative method to detect duplicates in registers, reducing the total number of records to 20,415 and species to 5,130 (see a complete description of the procedures adopted in data cleaning in the Methods section of Chapter 2 and Supporting Information 1.2). This method could be of use in many places around the world for which available data involve similar problems.

Most of the duplication in species occurrence data originated in the online databases. In Brazil, online biological databases are still not widely used or available, and only during the course of this thesis have they started to become widely accessible. Nevertheless, like many biodiversity databases around the world, online repositories of data on species in Brazil are not thoroughly checked or verified. In contrast, the database generated in this study now has all its information checked and verified. Errors could still present, but they were vastly reduced by the cleaning procedures adopted. Other errors, such as species misidentification in the field, or in herbaria and museums, and mistakes in transcription of field notes are unavoidable and impossible to detect in most cases.

Although the effort put into preparing a reliable database significantly improved the information available, there still is a lack of species occurrence information for most of the taxonomic groups in the Paranã River Basin (see Figures 2.3 and 2.4). Gaps in biological information for the Basin go beyond the two small areas highlighted by Brazilian environmental authorities as priorities for biodiversity inventories (see Figure 1.4). Future inventories should contemplate gaps in species records in all elevational and vegetation classes as pointed out by this study (see Figure 2.3 and 2.4).

Besides an accurate database and information on gaps on species occurrence, this study has also substantially improved information on total species occurrence within the Basin, including endemic and threatened species. The Paranã River Basin has long been recognized as a priority for conservation due to the uniqueness of its fauna, flora, vegetation, landforms, large numbers of rare, endemic, and threatened species, and its significance as a centre of diversity for several taxonomic groups [31-41]. However, the total number of species, including endemic and threatened species, and its level of representativeness of the Cerrado biome, were unknown until our study comprehensively collated, verified, and analysed the available data. The information generated here will make a significant contribution in informing stakeholders' decisions about conservation management in the Basin.

The resultant database we are making available to the scientific community and stakeholders is the most up-to-date and complete information available for the Basin. The database will continue to serve future needs as long as it is maintained and updated when possible. Additionally, after a presentation of the partial results of this thesis at the 51st Annual Meeting of the Association for Tropical Biology and Conservation, held in Cairns in 2014, we were invited to contribute the knowledge developed in this thesis to the SIBBr, a future Brazilian System of Information on Biodiversity. The SIBBr has a mission to provide better support for decision-making in conservation and sustainable use of biodiversity through an interdisciplinary government initiative based on integration and publication of high-quality data. In this sense, a future collaboration with SIBBr, sharing knowledge and expertise, will provide better biological information for the Paranã River Basin and for Brazilian biodiversity more generally. A Brazilian national repository of biological information would help to facilitate access to unique biological records, avoiding the difficulties we have faced over the course of this study and discussed in Chapter 2.

While gaps in species occurrence persist in the Basin for most taxonomic groups, this study has provided sufficient information to develop primary recommendations on area prioritization for some of the terrestrial species groups (amphibians, birds, mammals, reptiles, and plants), with relative reliability. Using the best possible taxonomic information gathered in Chapter 2, in Chapter 3 we modelled the distributions of individual species and species assemblages across the entire Basin as a function of environmental variables. In doing so, we have answered our second major question that was to understand how species and assemblages are distributed in the Basin.

In Chapter 3 we faced a common problem when modelling species or assemblage distributions using data on species occurrences that were collected for purposes other than modelling. Such data carry with them information only on species presence, but no indication about where species were looked for but found to be absent. Additionally, many species in our database still had inadequate numbers of records to be used for modelling, reinforcing the necessity to collect new data in this region, as pointed in Chapter 2. We used Maxent [132] to model species distributions since it performs well in predicting distributions compared to other available approaches, especially when the number of samples is small [104, 110, 132]. However, statistical modelling methods such as Maxent assume that samples are unbiased, a premise easily violated when using presence-only data [118, 133]. Species with few and environmentally biased records are common in the dataset used to build these models [58], and might have inflated the indicator of overall model performance (AUC values) [108, 155]. Accurate environmental profiles therefore might not be correctly established for those species [155].

As a complement to species distribution models, we used generalized dissimilarity modelling (GDM) [101, 117, 120, 121] to maximize the utilization of all species records gathered in our database, regardless of the number of records per species. GDM deals with assemblage composition as a whole, incorporating interactions between species and environmental variables, rather than relying on individual species. In this sense, GDM includes rare species with scarce records by assuming they have similar responses to environmental variables as the common species [13,14].

The results of Chapter 3 show the importance of using models of both species and assemblage distributions for assessing biodiversity patterns, considering both similarities and dissimilarities between the results produced by each method. From species distribution models (see Figures 3.1 and 3.2), we were able to identify centres of richness for different biological groups, how the distributions of each group changed across the landscape, and species suitability as a function of climate, landscape, and substrate variables. Plants showed higher richness values on the highlands of the Basin, where habitats are more diverse, humid, and cooler. Vertebrates (amphibians, birds, mammals, and reptiles) were richer at the low elevations of the Basin, where habitats are more structured, and climate is warmer and more stable. These results illustrate the problems associated with using taxonomic groups or other groups of species as proxies for other groups in conservation planning, potentially leading to under- or over-estimates of species presence in areas considered priorities.

Assemblage models (see Figure 3.5) provided information about how species are organized in the landscape, making it possible to identify different bioregions across the Basin. Through these models, we were able to identify rich and unique areas in species composition and predict spatial patterns of species turnover in assemblage composition where evolutionary process are potentially more significant. Additionally, congruence between the results of species and assemblage modelling reinforced the importance of the mountain range in the western portion of the Basin and some parts of the valley bottom for biodiversity conservation. In those regions are high values of species richness and/or singularities in species composition. The understanding of these different dimensions of biodiversity patterns proved to be useful in conservation planning.

Once we had solved the problem of estimating how species and assemblages were distributed across the Basin, we proceeded in Chapter 4 with an area prioritization to identify which areas are most important for biodiversity conservation. There are some different methods available for systematic area prioritization [172, 214, 215], each one with its pros and cons [216, 217]. We used Zonation because we wanted to identify continuous values of priority across the Basin for our complete picture of biodiversity [178], considering the following assumptions: the persistence of rare, threatened, and endemic species relies more on conservation actions than that of common species;

greater weights should go to assemblages with higher richness and uniqueness; and threatened ecosystems require more conservation efforts than non-threatened.

Using some protocols that had already been developed to deal with species and assemblage models in prioritization [see 172], we compared the advantages and disadvantages of using SDMs and GDMs separately and, when they were used together, also considered only taxa shared between them (birds, mammals, and plants). Moreover, we evaluated the additional contribution of species distributions of amphibians and reptiles and threatened ecosystems to area prioritization and discussed the implications of their use. GDMs are presented as an alternative input for area prioritization when there is not sufficient biological information on species occurrences to produce SDMs [101, 115]. However, there is no systematic comparison in the literature between GDMs and SDMs in area prioritization using real data. Since we have developed models only for species distributions of amphibians, birds, mammals, reptiles and plants and assemblage models for birds, mammals, and plants, we worked on different solutions for area prioritization and compared the results.

In this fourth Chapter, we were interested in the identification of priority areas by evaluating different solutions considering the distributions of species and assemblages, richness, threats, and endemism of species, complementarity, and irreplaceability. So, why should we be concerned about these aspects of planning? A broad delineation of priorities areas for conservation has already been made for the entire country by the Brazilian government [29], which includes the Basin. However, this broad delineation is of limited use due to its low spatial resolution. The priorities identified in that exercise do not approach the fine resolution needed for practical actions. Assuming a similar situation in the biological state of knowledge for priority areas already identified nationally, the evaluation of the contribution of species and assemblage-level analyses would be useful, not only for the Basin, but for a fast response to prioritize conservation actions at fine resolution across Brazil [see 29].

Ideally – where data availability, expertise, and time are available – the results of this study suggest that the simultaneous use of species and assemblage-level models in area prioritization is the best approach for broad-scale prioritization (for instance, to the 20% top fraction of the landscape) (see Table 4.1 and Figure 4.2). This broad approach takes advantage of complementary strengths and limitations of the two approaches. For more focused prioritizations, say in the top 2-5% of the landscape, SDMs and GDMs produce similar results in representing biodiversity features, and can be used without any significant prejudice to their respective outputs. SDMs have the advantage of being easily incorporated into area prioritization, with no pre-processing required. However, SDMs require more data on species occurrence to achieve reliable models. Consequently, most of the possible useful data from species sampled in just a few localities might not be utilized via SDMs. GDMs, on the other hand, can accommodate all available species records, even from locally rare species. However, the data preparation and the expertise required to prepare and analyse GDM models are more demanding than for SDMs.

Considering area prioritization produced by using both SDMs and GDMs, the analyses suggest some congruence in the results. Within the top 5% of priority areas, SDM and GDM prioritizations converged in selecting similar areas. More pronounced differences were observed for priorities outside the top 5%, and when threatened ecosystems were included in the analysis. Not surprisingly, including threatened ecosystems increased their retention in the highest-priority areas, especially for dry forests (see Table 4.1). This result illustrates the need for managers and policy planners to consider carefully what biodiversity features are important so that these can inform the analyses.

Although some spatial differences in area selection are observed in each solution (see Figure 4.2), all biodiversity features used in the analysis were relatively well represented in the final solutions (see Table 4.1 and Figure 4.3). Additionally, the congruities between different solutions highlighted patterns that are shared across distinct dimensions of biodiversity evaluated in these analyses. These patterns were captured irrespective of the biodiversity features used as input, giving more confidence that conservation decisions can be taken based on any of the solutions examined here.

In this sense, consistent planning can be obtained for the highest fraction of priorities irrespective of the model inputs, whether species or assemblages.

Having a better understanding of the pros and cons of using different conceptual models in area prioritization, in Chapter 5 we explored the possible impacts of climate change on biodiversity and identified priority areas considering these effects. From Chapter 4, we knew the differences in area prioritization between SDMs and GDMs are relatively minor, at least for the highest fraction of priorities. Therefore, given the SDM approach is easier, less time-consuming, and more widely available to practitioners around the world, we focused on SDM techniques for an exploratory analysis of the possible future climate-change impacts on biodiversity and their implications for area prioritization.

Comprehensive assessment of conservation planning and biodiversity should consider the potential impacts of future climate change, and this is what was done in Chapter 5. The Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report, on which these analyses were based, presents six SRES (Special Report Emission Scenario) marker scenarios. Although another set of scenarios (RCPs) is currently available, there is little qualitative or quantitative difference for their use in spatial conservation planning because the range of futures described are not significantly different [218]. We therefore decided to work with scenarios from the Fourth Assessment Report.

The IPCC scenarios are based on distinct development pathways, covering a broad range of demographic, economic and technological driving forces and resultant greenhouse gas emissions [196]. For an exploratory evaluation of the impact of climate change on biodiversity and its influence in spatial prioritization, we used the SRES A1B) for the year 2050 as an intermediate scenario of emission levels and time-frame. When limiting evaluations of species distributions to future projections of greenhouse gas emissions for 2050, we considered that countries' measures to reduce emissions are still uncertain, but that longer time-frames, say to 2080, involve more uncertainty [212]. This choice of time frame also recognized that the choice of emission scenarios has little effect when projecting future species distributions for the first few decades of the 21st century [212].

Even using an intermediate scenario for emissions and a relatively short time interval, changes in biodiversity patterns were predicted in the Paranã River Basin (Figure 5.1). These changes follow the same general patterns already reported for other parts of the world: a tendency for distributions to shift from lower to higher elevations and from lower to higher latitudes [200], although there might be local variations in the direction of movement [199].

The existing fully protected areas system within the Paranã River Basin is not sufficient to meet biodiversity conservation objectives either now or in the future. The present system requires a considerable increase in size and/or number of areas to better represent areas important for current and future biodiversity conservation. To some extent, the areas identified as high priorities for both current and future species distributions might contribute to species persistence over time, at least by reducing the need for populations to shift spatially. These overlap areas can therefore be seen as important for conservation by expanding protected areas. Notably, the broad-scale Brazilian conservation planning that covers the region [29] does not recommend substantial protection where current and future priorities overlaps (see Figure 5.4).

Apart from fully protected areas and reorientation of biodiversity inventories to areas where the gaps in biological information are present – discussed in Chapter 2 – other priorities indicated in the national plan [29] should be revised. The great importance of facilitating species movements to track shifts in their climate space means that ecological corridors assume a high priority. The ecological corridors proposed in the national plan do not match the areas likely to be resilient under climate change. This mismatch could compromise linkages between current and future species distributions within the Basin. In the same way, environmental restoration efforts proposed in the national plan do not align with priority areas identified by our study or account for opportunities to maximize future resilience under climate change. In addition, areas proposed for management of sustainable use are some of those identified in the present study as most important for conservation under current and future climatic conditions. Relying only on landholders' self-enforcement of sustainable use in these areas risks the persistence of biodiversity. In these areas, there are advantages in integrating extractive reserves into territories inhabited by the 'Quilombolas' communities. These traditional slave-descendants are recognized by the low impact of their activities on biodiversity, although their territories have still to be officially recognized [50-52]. Regularizing these territories to incorporate the cultural management practices of these communities will, however, require information and consultation [50] to reconcile the interests of people and biodiversity.

We recognize the Brazilian Government's effort in attending to the commitments of the Convention on Biological Diversity. Investments in biodiversity conservation and related studies have been increased considerably in Brazil in recent decades, but these efforts are still insufficient to counteract the high rates of conversion and loss of ecosystems. There are still vast natural areas with little to no information about their biodiversity, indicating the need for basic studies to identify, at least, the species that are there. Time, money, and expertise are needed for biodiversity inventories, species identification, data preparation, and data analysis to identify which areas are priorities for biodiversity conservation. The basic biodiversity inventory that we conducted in the Paranã River Basin, for example, took almost three years, and we sampled only a small fraction of the entire Basin. The additional data gathering, preparation, and processing for spatial modelling and prioritization encompassed more than four years of research. In countries such as Brazil, this time is priceless. Frequently, while decisions to preserve and protect areas of importance for biodiversity are being considered, irreplaceable biodiversity values were already lost or degraded. The natural environment of the Cerrado biome in Brazil in which the Paranã River Basin is embedded, for example, has been converted into anthropogenic landscapes at a rate of 1.23 ha per minute; and that rate, measured between 2009 and 2010, reflected deforestation rates that had been reduced by more than 50% compared to previous periods (2002 to 2008) [219].

The research and planning framework developed in this study and presented in the form of data-chapters can be used in other study regions, thereby minimizing the time spent in processing, filtering, and analyzing data. In Brazil, similar problems with data availability are likely to apply to many of the other 1,560 broad areas indicated as national priorities for biodiversity conservation [29]. Macro-scale approaches, as proposed by the Brazilian government [29], are essential to provide broad perspectives on important places for conservation. Nevertheless, revising conservation designs is part of the dynamism that is intrinsic to conservation planning [73]. The revision of these national priority areas is necessary to accommodate, for example, new data on biodiversity, loss of areas of interest, and new methods and approaches of analysis [14], as well as the kind of fine-resolution insights generated by this study. Additionally, such plans should consider biodiversity as broadly as possible, or at least incorporate information on the distributions of different taxonomic groups to understand how biodiversity patterns are organized and distributed. Basing conservation planning on just a few species considered of primary interest, as is commonly done, can result in lack of protection for many species, as well as for unique and richer species assemblages.

Incorporating knowledge and informing new decisions in conservation also depends on efficient communication with stakeholders. For instance, the Chico Mendes Institute for Biodiversity Conservation (ICMBio), a Brazilian organization responsible for the establishment of protected areas, is currently planning to increase the total area of the Chapada dos Veadeiros National Park. The park is located in the highlands of the centre western border of the Basin (see Figure 1.2), a region identified in this study as among the most important for biodiversity conservation in the Basin. The presentation of our results to that organization was considered of great significance to inform and justify their decisions in the process of expanding the park. In this communication process, however, scientific knowledge must be demystified for a complete understanding of its strengths and weaknesses.

In conclusion, in the development of this study, we attempted to address two different types of expectations: those of stakeholders and those of the scientific community. Stakeholders, such as the technicians of the Chico Mendes Institute for Biodiversity Conservation (ICMBio), who are responsible for determining the boundaries and categories of protected areas proposed, are interested in the practical approach of this study to inform decisions; that is, their focus is on the identification of priority areas for biodiversity conservation and the reasons why these areas are priorities. The scientific community is more interested in the theoretical approach involved in the study.

Considering stakeholder's expectations, we have pointed out the importance of the region for biodiversity conservation and its representativeness in relation to the Cerrado biome. We also indicated the total number of species occurrence already recorded, and the presence of threatened and endemic species within the Basin. We additionally suggested areas that should be prioritized for conservation action, taking into account the distributions of species and assemblages, endemic and threatened species, and threatened ecosystems. We also provided information on area prioritization considering current and future species distributions and their poor representation in fully protected areas.

We hope that this study has contributed to the science of systematic conservation planning by developing an innovative method for detecting duplicates in large data sets. We also identified the types of information needed in metadata associated with species records to produce reliable results. Additionally, the occurrence of endemic and threatened species and gaps on species sampling to guide future biodiversity inventories in the Basin are also identified in Chapter 2. Furthermore, in Chapter 3, we presented the richness and assemblage patterns modelled for major biological groups to understand how they are distributed in the Basin. We described how we extracted the best information available from the database produced in Chapter 2 and discussed the strengths and weaknesses of this approach. In the process of area prioritization, in Chapter 4, we evaluated for the first time the simultaneous use of species and assemblage-level data and the complementary perspectives they offer. We also compared area prioritizations based on different sets of biodiversity features: species, assemblages, threatened and endemic species, and threatened ecosystems. In Chapter 5, we presented area prioritization considering the future impacts of climate change on species distributions. In this final chapter, we have discussed the pros and cons of these approaches and how we addressed the limitations of data and analyses.

6.3. Addressing and informing uncertainties

Uncertainties are things that are not known, or known only imprecisely [220], which may reflect incomplete knowledge of many risks associated with decisions [221]. Uncertainties are not completely avoidable, but they might be minimized or quantified so that they can be accounted for in decisions [220]. In this study, uncertainties permeate all phases of the research, from collecting field data to the data analysis and area prioritization involving prediction of future climate. Uncertainties were not quantitatively addressed in this study, but some procedures were adopted to reduce uncertainties and so reach decisions that were reliable.

An intensive investment in removing duplicate species records minimized errors in the database that would have been a distraction in the planning process. As we have seen, more than 85% of the initial data records were discarded, eliminating more than 47% of the species names that were contaminating the dataset. However, errors may still be present. Species misidentifications during field work cannot be detected if voucher numbers are not associated with the observational data. For deposited material, errors are mainly related to taxonomic interpretation. Scientific names represent a taxonomic hypothesis [222]. Hence there is instability and uncertainty associated with each named specimen. A consensus in some taxonomic determinations might not be reached, even among specialists. The steps adopted to choose which records to keep for analysis (see Chapter 2), while therefore not removing uncertainty, at least reduced it.

Another type of error associated with species data is related to assigned geographic coordinates. This kind of error was found in more than 5,000 records in this study. Removing or correcting this type of error could have avoided incorrect associations of species with environmental data and, consequently, increased error in SDMs and GDMs [212].

Species respond differently to different environmental variables [223]. Modelling such a large number of species at one time using the same predictors might therefore introduce errors. In this sense, some correlative species models might not represent true environmental niches as well as species-specific models would (see species variation responses to environmental variables in Figure 3.3, for example). Additionally, models based on species recorded in only a few locations might not establish the true niche-based signal, as discussed in Chapter 3. Consequently, estimates of suitability for species and patterns of richness come with a certain amount of error, which is only exacerbated in predicted distributions under a future climate, as discussed in Chapter 5. For the pair-wise dissimilarity models used in GDM, these problems are of less concern since correlations are made assuming species assemblages instead of distinct species. However, the inclusion of a large number of sites with few species per site in the GDM analysis contributed to the scatter observed in the full models (see Figure 3.6) and to relatively small deviance explained. Aggregated data from multiple sources is likely to involve substantial spatial and environmental biases, therefore masking robust signal variation in both species and assemblage models as discussed in Chapter 3.

The uncertainties described above will certainly have affected the area prioritizations to some extent. However, assigning higher values to the best models available on species and assemblage distributions will have reduced the influence of inadequate information while also maximizing the use of the scarce information on biodiversity for conservation planning.

6.4. Further investigations

Continuing studies in the Paranã River Basin to improving knowledge for biodiversity conservation should combine enhanced use of the information that is already available with the results of additional investments in data gathering. Besides transparency, making the dataset freely available might incentivize researchers to contribute to investigations that could accelerate conservation actions within the region. Considering that, some important issues not covered by this study that need closer investigation are related to the modelling process. The use of different species distribution models (SDMs), global circulation models (GCMs), and gas emission scenarios (GESs) might offer distinct solutions to estimate species distributions for current and future climates, and might improve evaluation of shifting biodiversity patterns in response to climate change [see 71]. Additionally, consensus maps of species and assemblage distributions and a map of the variance components obtained across multiple models would be useful to address uncertainties in area prioritization and give more confidence in identifying the most important areas for conservation within the Basin.

Given the high level of uncertainty associated with SDMs based on data-poor species (see Chapter 3), future investigations in species modelling and area prioritization should explore the use of only reliable models (for species recorded in 20 or more localities that were tested for accuracy). Additionally, different weighting schemes should be tested considering, at least, equal weights as a benchmark to systematically evaluate the effects of the weighting process on area prioritization. Comparison of alternative weighting systems should consider the proportions of distributions of highly valued features in the top fractions of priority areas.

Additional investments in collecting species data are necessary to fill gaps in biodiversity information. This study has already identified gaps in species occurrence across the Basin (Chapter 2). To address these shortcomings, one of the strategies already agreed with professors from the University of Brasília and researchers from Brazilian Corporation of Agricultural Research (EMBRAPA) – all collaborators of previous inventories – is to develop Ecology and Zoology field courses within the Basin. Besides their educational value, these courses would add information on species biology and knowledge of species occurrences in areas that are presently poorly sampled.

Additional investments in fine-resolution spatial data on vegetation, land cover, geology, and soils will also be important as a basis for conservation planning. The coarse-resolution maps used in this study are insufficient to meet the needs of more robust planning. For vegetation and land cover, we already have the tools and data necessary for fine-resolution mapping. For soils and geology, though, it will be necessary to establish partnerships with other researchers and institutions.

More data on species occurrences from different taxonomic groups will provide the basis for more reliable models of species distributions and assemblage, especially for those groups - other than birds, mammals, and plants – for which models are presently not available. Additionally, the assemblage-level analysis must be extended to consider assemblages under future scenarios of climate change.

Further biodiversity inventories should be accompanied by monitoring of species movements and composition of assemblages in relation to climate change to a better understanding of the impacts of climate change on biological diversity of the Paranã River Basin. Understanding species and assemblage responses to a changing climate would provide insights to improve conservation planning in the Basin and in the whole Cerrado biome. Monitoring ecological responses to climate change is non-existent or incipient in Brazil. Establishment of climate stations across the montane gradient, from the bottom of the Basin to its highlands, followed by monitoring biodiversity at those stations would, for instance, provide a comprehensive understanding of species and assemblage responses to changes in climate, the velocity and direction of changes, and system vulnerability [197, 199, 201, 224]. Such information is essential to identify and propose mitigation measures to promote species long-term persistence or adaptation.

Another topic to be further explored is related to policy and tools for conservation. Brazil has one of the most progressive policies in conservation in the world [225], but why these policies are not adequately translated into conservation actions is still poorly understood. The prioritization of economic development over sustainability is, as usual, one of the leading causes of the problem. However, resistance to government decisions taken towards conservation is observed in all sectors of society, irrespective of the possible impacts on productive sectors. Among the policy instruments that need to be better trialled and evaluated are: motivational, educational and informational instruments (promotion of positive attitudes towards biodiversity conservation); voluntary instruments (incentives to participate in conservation programs); property rights instruments (compensation for forgone economic uses); financial mechanisms (conditional grants and tax concessions); and regulation (restricting resource use) [226]. Engaging and informing the local community in a participatory way has proved a useful tool to involve people in the creation of protected areas of sustainable use in other parts of the Cerrado biome, like the Nascentes Geraizeiras Sustainable Development Reserve in the north of Minas Gerais state [227]. However, this is not an inexpensive or short-term process. More than 12 years were needed to involve and convince the beneficiary communities and government agencies of the importance of the reserve in Minas Gerais for supporting the joint goals of local livelihoods, biodiversity, and ecosystem services.

Considering that conservation planning cannot produce static products [14], the Brazilian priority areas for conservation [29] must evolve to accommodate new information on species occurrence, new analysis tools, and even loss and gains of areas of interest for conservation. Thus, a revision of the Brazilian priority areas for the Basin should be scheduled for the near future to consider the new information provided by this study. Even with its limitations, the results of this study are currently the best basis on which to plan biodiversity conservation of the Paranã River Basin. The framework developed here and presented in the format of data-chapters can be revised for further investigations in the Basin, but also adapted for other regions in Brazil or around the world.

Supporting Information

Supporting Information 1 – Digging deep for biodiversity information in data-deficient areas: the case of the Paranã River Basin within a global biodiversity hotspot

---

Contents

SI 1.1. The Paranã River Basin biodiversity database

SI 1.2. Deduplication process: a complementary description of data-attributes standardization procedures

SI 1.1. The Paranã River Basin biodiversity database

This Supporting Information presents a database with all biological information gathered for the Paranã River Basin summarized in Table 2.1. Due to space limitation, it is available only as a web repository. To access this database, please refer to the link: http://eresearch.jcu.edu.au/tdh/datasets/parana_river_basin/species_occurrence/species_occurrence.xlsx.

SI 1.2. Deduplication process: a complementary description of data-attributes standardization procedures adopted

Prior to identifying and removing species duplicates in our biodiversity database gathered from multiple sources, we had to define a protocol of codes to operate database debugging in RecordLinkage package [78, 79]. These codes were used to identify duplicates and a minimum number of parameters (data descriptors in the metadata) associated with each species record. When we identified the best comparison parameters, i.e., the best data descriptors to be used for detecting duplicates, they were standardized in metadata. After numerous attempts, we identified at least eleven parameters that should be present and needed to be standardized for effective identification of data duplication and errors as described in the Methods section of Chapter 2: species binomial, geographic coordinates, collection date, collector's name, species specialist identifier, date of species identification, geographic location, municipality of occurrence, collection code, name of the collection, source of information.

Since protocols to identify duplicate records in the database were not perfect, we had to run the database deduplication several times to identify and fix omission errors. Standardizations consisted of correcting coordinate positions, errors in species scientific names, conflicts between species' determination, i.e., same species and/or individuals with different scientific names in various collections, and synonymies. Additionally, we had to standardize names of cities (where specimens were collected), names of collectors, date when specimens were collected, names of species identifiers, date when specimens were identified, and names of collections where specimens were deposited.

Conversions between latitude and longitude were the most common errors found in species records (more than 5,000). We plotted species records and observed their locations to identify problems associated with coordinates. The most difficult problems were duplicates of the same specimens that were identified in different collections with different, but still valid, scientific names. In some cases, these duplicates were identified by the same specialists, in others by different specialists.

We considered as valid names only those recorded for the study area or the biome. However, we found an extreme case of one specimen with a voucher referring to five duplicates deposited in different herbaria around the world. In each herbarium, the specimen duplicate received a different valid species name, all given by the same specialist. From the five different valid names, all found in the Cerrado biome, two were known for the study area. For this case, we considered only the two species names recognized for the region and we assumed the last identification as correct during the process of deduplication. When the same specimen had received different valid names by different specialists, we considered the identification given by the most renowned expert.

Considering most recent identification provided by the same specialist as the best choice or the species determination by a more experienced expert as correct assumed that species identification is improving over time among specialists.

These are only some of the examples illustrating the myriad problems and time spent to extract the best possible information on species occurrence during database preparation. The approach described above, however, allowed consistency in resolving duplicates and avoided discarding a very large number of reliable species records.

After duplicate records had been identified, they still needed to be removed manually to guarantee reliability of the remaining data. Once these protocols were established, they were applied to our database, reducing the total number of records from ~140,000 to ~20,000 and reducing the number of species from ~10, 000 (or ~12,000 including records identified only to a level higher than species) to ~5,000.

Data duplication is, therefore, one of the biggest sources of errors in databases for conservation planning when data from multiple sources are used. Data duplication can artificially inflate the number of species in a region or mask species occurrence because of synonymies. Duplication can also misrepresent species’ ranges due to errors in coordinates. On the other hand, utilising information available in online databases is becoming very common for modelling species and community distribution and for planning for biodiversity conservation. We have shown that the indiscriminate use of information from multiple sources could involve substantial errors. A careful inspection of the data is necessary to avoid errors that contribute uncertainties to spatial analyses and conservation planning.

Supporting Information 2 – Biodiversity patterns within the Paranã River Basin: what we can learn from species and assemblage-level distributions?

Contents

SI 2 – Table 1. List of environmental variables, their ecological rationale, models in which variables can be potentially used, and sources of information. Bold characters on "model used" column indicate the use of that variable in the respective final model.

SI 2 – Table 2. Pearson's correlation coefficient between variables used to model species and assemblage distributions within the Paranã River Basin. Variable labels follow the same codes used on Table 1, except for Lat = Latitude and Long = Longitude.

SI 2 – Additional Figure 1. Species richness maps for the full models (stacked models from the Maxent logistic output of species models tested and non-tested for model accuracy) and elevation classes within the Paranã River Basin.

SI 2 – Additional Figure 2. Species richness maps based on models tested with a 10-fold cross-validation (stacked models from the Maxent logistic output of species recorded in 20 or more localities) and elevation classes within the Paranã River Basin.

SI 2 – Table 1. List of environmental variables, their ecological rationale, models in which variables can be potentially used, and sources of information.

Bold characters on “model used” column indicate the use of that variable in the respective final model.

#	Code	Description	Variable Group	Ecological Type Action	Physiolo-gical proximity	Rationale	Model used	Source
1	AMT	Annual mean temperature	Tempera-ture	Direct to plants and animals	Proximal to plants and animals	Species richness and assemblage patterns are related directly or indirectly to environmental temperature, which influences on individual metabolism of ectothermic organisms and plants. According to the Metabolic Theory of Ecology, the high temperatures lead an increase of the metabolic rate through of acceleration of biological processes (at cellular level) on ectothermic organisms, which leads an increase in the mutation rate and, consequently, in speciation rate. Within the Paranã River Basin (PRB), higher mean temperature values are related to the bottom of the Basin, reaching 27.1 ^{0}C on the north region. Temperature average drops to 19.5 ^{0}C on the western and south portion of the mountain range of the Basin.	**Maxent** and **GDM**	http://www.worldclim.org/download
2	MDR	Mean diurnal range(mean of monthly (max temp - min temp))	Tempera-ture	Indirect to plants and animals	Distal to plants and animals	It represents temperature fluctuation during the day along the months. Within the PRB, areas in the north portion of the bottom of the Basin have a mean diurnal range higher (15.7 ^{0}C) than other areas of the bottom, revealing a temperature behaviour of arid environments. The minimum values are found in the mountain regions (9.5 ^{0}C) of the western portion of the Basin.	Maxent and GDM	http://www.worldclim.org/download

#	Code	Description	Variable Group	Ecological Type Action	Physiolo-gical proximity	Rationale	Model used	Source
3	IT	Isothermality ((mean diurnal range/temperature annual range)*100)	Tempera-ture	Indirect to plants and animals	Distal to plants and animals	It represents seasonal variation in temperature. Within the PRB areas in the north of the bottom of the Basin have isothermality higher than areas in the highlands of the western and northeastern portions.	Maxent and **GDM**	http://www.worldclim.org/download
4	TS	Temperature seasonality (coefficient of variation of mean monthly temperatures)	Tempera-ture	Indirect to plants and animals	Distal to plants and animals	It represents seasonal variation in temperature. In the PBR, the effects of temperature seasonality are more evident in the southern and on the northern portion of the Basin.	**Maxent** and **GDM**	http://www.worldclim.org/download

#	Code	Description	Variable Group	Ecological Type Action	Physiolo-gical proximity	Rationale	Model used	Source
5	MTWM	Maximum temperature of warmest month	Tempera-ture	Direct to plants and animals	Proximal to plants and animals	Maximum temperature conditions are generally recorded during the day (after midday), and it is related to physiological conditions of plants and animals. It influences growth and development. The maximum is indicative of the warmest summer conditions that may require specialised drought and heat adaptive responses or morphology. Within the PRB, higher values of maximum temperatures are present in the northern portion of the Basin's bottom. There, maximum temperature can reach 35.5 ^{0}C. In the highlands, maximum temperature during the warmest month is ~10 ^{0}C lower than in the bottom, reaching 26.8 ^{0}C.	Maxent and GDM	http://www.worldclim.org/download
6	MTCM	Minimum temperature of coldest month	Tempera-ture	Direct to plants and animals	Proximal to plants and animals	Minimum temperature conditions are generally recorded overnight (near dawn) and mainly relate to physiological conditions on plants and animals supporting growth and development or constraints. The minimum is indicative of the depth of winter cold conditions that may require specialised adaptive responses. In the PRB, the minimum temperature is more pronounced in the highlands of the western and the southern portion of the Basin (10.3 ^{0}C) while higher values are present in the northeastern (18.7 ^{0}C).	Maxent and **GDM**	http://www.worldclim.org/download

#	Code	Description	Variable Group	Ecological Type Action	Physiolo-gical proximity	Rationale	Model used	Source
7	TAR	Temperature annual range (max temperature of warmest month - min temperature of coldest month)	Tempera-ture	Indirect to plants and animals	Distal to plants and animals	It is a different way to represent seasonal variation in temperature. Larger range are related to the areas in the northern part of the bottom of the Basin where temperature can vary in 22.9 ^{0}C, and shorter range to the areas in highlands, where temperature varies on 13.6 ^{0}C). Environments with high amplitude in temperature range are more stressful for organisms than areas with a shorter amplitude of variation. This variation can affect organisms physiologically, selecting those more adapted to that variation.	Maxent and GDM	http://www.worldclim.org/download
8	MTWQ	Mean temperature of wettest quarter	Tempera-ture	Direct to plants and animals	Proximal to plants and animals	It is an indicative of summer condition when respiration and metabolism are most active in plants and animals and favoured by water availability. Within the PRB, higher temperature values are found in the bottom of the Basin, mainly on the north portion (27.2 ^{0}C) while, in highlands, temperatures are lower (19.5 ^{0}C).	Maxent and **GDM**	http://www.worldclim.org/download
9	MTDQ	Mean temperature of driest quarter	Tempera-ture	Direct to plants and animals	Proximal to plants and animals	It is an indicative of winter condition when process of respiration and metabolism are less active in plants and animals, additionally depressed by drought period. Temperature pattern in the winter period follows the same pattern of summer within the PRB. Higher temperature values are found in the bottom of the Basin, mainly in the north portion (26.7 ^{0}C) while, in highlands, temperatures are lower (18.7 ^{0}C).	Maxent and **GDM**	http://www.worldclim.org/download

#	Code	Description	Variable Group	Ecological Type Action	Physiolo-gical proximity	Rationale	Model used	Source
10	MTWaQ	Mean temperature of warmest quarter	Tempera-ture	Direct to plants and animals	Proximal to plants and animals	Warmest temperature occurs during spring condition when process of respiration and metabolism are most active in plants and animals. Within the PRB, higher values are found in the bottom of the Basin, mainly in the northeastern region (27.7 ^{0}C), while lower temperatures are present in highlands, mainly in the western mountain range (20.5 ^{0}C)	Maxent and **GDM**	http://www.worldclim.org/download
11	MTCQ	Mean temperature of coldest quarter	Tempera-ture	Direct to plants and animals	Proximal to plants and animals	It is an indicative of winter condition when process of respiration and metabolism are less active in plants and animals, additionally depressed by drought during the winter period. Within the PRB, higher values are found on the bottom of the Basin, mainly in the northeastern region (26.3 ^{0}C), while lower temperatures are present on highlands, mainly in the western mountain range (18.3 ^{0}C)	Maxent and GDM	http://www.worldclim.org/download
12	AP	Annual precipitation	Water	Direct to plants and animals	Proximal to plants and animals	It is related to soil water availability for plants and atmospheric moisture. Maximum values indicate potential flood conditions correlated with high cloud cover and low levels of radiation. Minimum values indicate potential drought conditions correlated with low cloud cover and high levels of radiation. Highland areas, mainly in western and on northeastern portions of the Basin receive high average volume of water from rain (2161 mm), when compared to the bottom in regions occupied by dry forests (868 mm).	**Maxent** and GDM	http://www.worldclim.org/download

#	Code	Description	Variable Group	Ecological Type Action	Physiolo-gical proximity	Rationale	Model used	Source
13	PWM	Precipitation of wettest month	Water	Direct to plants and animals	Proximal to plants and animals	Maximum values indicate potential flood conditions correlated with high cloud cover and low levels of radiation in the summer period. Within the PRB areas of high concentrated values of precipitation are related to the mountain regions, mainly in the western and northeastern portions of the Basin, where it can rain 405 mm a month. In the bottom, precipitation values reach 150 mm.	Maxent and **GDM**	http://www.worldclim.org/download
14	PDM	Precipitation of driest month	Water	Direct to plants and animals	Proximal to plants and animals	Maximum values represent less stressful condition in terms of water availability during the period. Within the PRB, precipitation in the driest month is more pronounced in mountain regions of the western border and in the southeastern border where rains can reach 12 mm. Drought are pronounced in an axis from the centre to the northwest of the bottom of the Basin, along the course of the Paranã River, with no rains.	Maxent and **GDM**	http://www.worldclim.org/download
15	PS	Precipitation seasonality	Water	Indirect to plants and animals	Distal to plants and animals	It is a measure of seasonal variation in rain. Areas of lower values of seasonality indicate more water availability through the year than areas of higher values. Higher precipitation seasonality suggests drought periods. Additionally, higher seasonality contributes to select organisms adapted to periods of drought. Within the PRB, seasonality is greater in the bottom of the Basin than in the mountain regions of the centre western and southeastern.	**Maxent** and **GDM**	http://www.worldclim.org/download

#	Code	Description	Variable Group	Ecological Type Action	Physiolo-gical proximity	Rationale	Model used	Source
16	PWQ	Precipitation of wettest quarter	Water	Direct to plants and animals	Proximal to plants and animals	Maximum values indicate potential flood conditions correlated with high cloud cover and low levels of radiation. Within the PRB, precipitation in wettest quarter coincides with the summer period. It is more pronounced in mountain regions of the centre western, northeastern and southeastern borders, which values can reach 1103 mm. Less rain volume are received at the bottom of the Basin, mainly on an axis from southwestern to the centre eastern and in the north region of the Basin, where it rains 425 mm during this period.	Maxent and GDM	http://www.worldclim.org/download
17	PDQ	Precipitation of driest quarter	Water	Direct to plants and animals	Proximal to plants and animals	It indicates a trend during the winter period. Maximum values represent less stressful condition in terms of water availability during the winter period. Within the PRB, precipitation during this period is more pronounced in mountain regions of the centre western border and in the southeastern border where rains can reach 45 mm. Drought are pronounced at the bottom of the Basin, mainly in an axis from the centre to the northwestern, following the course of the Paranã River, with no rains.	**Maxent** and GDM	http://www.worldclim.org/download

#	Code	Description	Variable Group	Ecological Type Action	Physiolo-gical proximity	Rationale	Model used	Source
18	PWQ	Precipitation of warmest quarter	Water	Direct to plants and animals	Proximal to plants and animals	It indicates a trend during the spring period. High temperature within the PRB is not correlated with high precipitation, which occurs during the summer period. During the warmest quarter, organisms are subject to desiccation due to lower availability of water at the bottom of the Basin (77 mm). Rains during the warmest period are mainly located in the mountains range of the south portion (400 mm) and the centre western (270 mm) portion of the Basin. Even in low amount, these rains are important during this period for dry forest formations when most of the tree plants are flowering/fruiting.	Maxent and **GDM**	http://www.worldclim.org/download
19	PCQ	Precipitation of coldest quarter	Water	Direct to plants and animals	Proximal to plants and animals	It indicates a trend during the winter period. Maximum values represent less stressful condition in terms of water availability during the winter. Within the PRB precipitation is mainly located in the mountains of the centre west, southwestern, northeastern and southeastern, which values reach 207 mm during this period while in the bottom of the Basin there are no rains.	Maxent and **GDM**	http://www.worldclim.org/download

#	Code	Description	Variable Group	Ecological Type Action	Physiological proximity	Rationale	Model used	Source
20	ARI	Aridity index (mean annual precipitation /mean annual potential evapotranspiration)	Water / Temperature	Indirect to plants and animals	Distal to plants and animals	It is used to quantify precipitation availability over atmospheric water demand. It is an indicative of atmospheric moisture conditions (relative water availability or water stress) as the ratio between rainfall and evaporation. It shows moisture availability for potential growth of reference vegetation excluding the impact of soil mediating water runoff events. Maximum value is dominated by precipitation and minimum is dominated by evaporation. Within the PRB, maximum values are related to the mountain region where precipitation values are bigger than on the bottom of the Basin. Minimum values are found on the lower part of the Basin where air moisture can frequently drop to less than 10% during the dry season.	Maxent and GDM	http://www.cgiar-csi.org/data/global-aridity-and-pet-database

#	Code	Description	Variable Group	Ecological Type Action	Physiological proximity	Rationale	Model used	Source
21	PET	Potential evapotranspiration	Water / Temperature	Indirect to plants and animals	Distal to plants and animals	It influences soil water availability and atmospheric moisture. Maximum values of evapotranspiration indicate potential heat or drought conditions correlated with temperature and radiation. Minimum values indicate potential flood or cold conditions correlated with temperature and radiation. Within the PRB, there is a gradient between higher and lower values following the structure of the vegetation. From higher to lower we have: dry forests, dense savannas, savannas, open savannas, and grasslands. Thus, higher values are found mainly in regions of dry forest on the eastern side of the Basin, while lower values are located in mountainous regions that surround the Basin.	Maxent and GDM	http://www.cgiar-csi.org/data/global-aridity-and-pet-database
22	DEM	Digital elevation model	Terrain	Indirect to plants and animals	Distal to plants and animals	It comprises information from different variables, mainly associated with changes in temperature and precipitation. DEM can be associated with changes in land cover as well since intact areas are generally relegated to the slope areas while flat areas are preferably used for agricultural purposes. On the PRB, however, occupation is preferably done on flat areas independently of the elevational positional.	Maxent and GDM	http://srtm.csi.cgiar.org
23	CTI	Compound topographic index	Substrate	Indirect to plants and animals	Distal to plants and animals	Soil retention humidity based on topography. Generally, on PRB areas on mountains will retain less water than areas on the bottom of the Basin	Maxent and GDM	Extracted from DEM

#	Code	Description	Variable Group	Ecological Type Action	Physiolo-gical proximity	Rationale	Model used	Source
24	Rough	Surface roughness	Terrain	Indirect to plants and animals	Distal to plants and animals	Coefficient of variation in elevation potentially related to local terrain heterogeneity and associated substrate conditions. Roughness is related to environmental complexity. On PRB, higher values of roughness are related to limestone outcrops and the mountain regions where soils depth is low.	**Maxent** and **GDM**	Extracted from DEM
25	DTS	Distance to streams	Terrain	Direct to plants and animals	Proximal to plants and animals	It is related with water availability on landscape. In different proportions, animals and plants are limited by water directly accessibility what interfere with their distribution.	Maxent and **GDM**	Extracted from DEM
26	GeoCl	Geology class	Substrate	Indirect to plants and animals	Distal to plants and animals	It is a categorical variable describing different geology classes. Different classes are expected to support different types of vegetation, i.e. different plant populations and communities and, consequently, different animals populations and communities.	Maxent	CPRM – ISBN 85-7499-099-4
27	GeoAge	Geology age (Geological age in Millions of years BP)	Substrate	Indirect to plants and animals	Distal to plants and animals	Rock age is indirectly related to substrate weathering, soil formation and nutrient status, which will have an influence on plant communities, and fauna associated.	Maxent and GDM	CPRM – ISBN 85-7499-099-4

#	Code	Description	Variable Group	Ecological Type Action	Physiolo-gical proximity	Rationale	Model used	Source
28	SCl	Soil class	Substrate	Direct to plants Indirect to animals	Proximal to plants Distal to Animals	It is a categorical variable describing different soils classes. Different classes are expected to support different types of vegetation, i.e. different plant populations and communities and, consequently, different animals populations and communities.	Maxent	http://sosgisbr.com/2012/06/04/mapa-de-solos-do-brasil-2011-embrapa/
29	SB	Soil base saturation	Substrate	Direct to plants Indirect to animals	Proximal to plants Distal to Animals	Saturation of Base refers to the amount of base (generally calcium and magnesium) available in the soil for plants. Base saturation was attributed based on soils description on literature [228-231]. Within PRB, soils with 50% or more in saturation of bases are occupied by dry forests while savannas formation occupies soils with less than 50%. Base saturation is a proxy for dry forest formation within the PRB.	**Maxent** and **GDM**	Extracted from SCl
30	SD	Soil depth	Substrate	Direct to plants Indirect to animals	Proximal to plants Distal to Animals	Soils depth, from top to rock material, was attributed based on soils description on literature [228-231]. Soil depth can affect root exploration volume for water and nutrients as well as water-holding capacity. Higher depth soils allow root development better than shallow soils. Generally, on PRB, arboreal savannas and dry forests are located on depth soils, while grasslands and open savannas are situated on shallow soils.	**Maxent** and GDM	Extracted from SCl

#	Code	Description	Variable Group	Ecological Type Action	Physiolo-gical proximity	Rationale	Model used	Source
31	SP	Soil permeability	Substrate	Direct to plants Indirect to animals	Proximal to plants Distal to Animals	Soils permeability to water (index) was attributed based on soils description on literature [228-231]. On the PRB, generally, soils with low permeability to water are seasonally inundated what restricts the root development. Usually, shallow inundate soils are occupied with wet grasslands. Median to depth soils can have some impediments to water percolation on the horizon given by underground calcareous lenses, occupied by dry forests, or by laterite crusts, occupied by different types of savannas (from open to dense).	**Maxent** and **GDM**	Extracted from SCI
32	SI	Soil index (SI = SB*SD*SP)	Substrate	Direct to plants Indirect to animals	Proximal to plants Distal to Animals	A soil index to comprise soils characteristics information	Maxent and GDM	Extracted from SCI
33	Veg	Vegetation class	Substrate	Direct to plants Direct to animals	Proximal to plants Proximal to animals	It is a categorical variable describing vegetation types within the PRB.	**Maxent**	http://mapas.mma.gov.br/mapas/aplic/probio/datadownload.htm?/cerrado

#	Code	Description	Variable Group	Ecological Type Action	Physiolo-gical proximity	Rationale	Model used	Source
34	VCo	Vegetation coverage	Substrate	Direct to plants Direct to animals	Proximal to plants Proximal to animals	Indicates coverage's percentage of a given vegetation class. For each vegetation class was attributed a value of coverage based on vegetation description on literature [59, 129, 232] and our knowledge of the vegetation of the Basin. Higher is the coverage; less light will be reaching the ground, which minimizes the effect of drought.	Maxent and **GDM**	Extracted from Veg
35	VH	Vegetation height	Substrate	Direct to plants Direct to animals	Proximal to plants Proximal to animals	Indicates total height of a given vegetation class. For each vegetation class was attributed a value of height based on vegetation description of literature [59, 129, 232] and our knowledge of the vegetation of the Basin. Higher is height; more structured is the environment to be explored by animals such birds, amphibians, reptiles, mammals and insects.	Maxent and **GDM**	Extracted from Veg
36	VS	Vegetation strata	Substrate	Direct to plants Direct to animals	Proximal to plants Proximal to animals	Indicates number of strata of a given vegetation class. For each vegetation class was attributed a value of number of strata based on the description of literature [59, 129, 232] and our knowledge of the vegetation within the Basin. More stratified is the vegetation more structured is the environment to be explored by animals such birds, amphibians, reptiles, mammals and insects.	Maxent and GDM	Extracted from Veg

#	Code	Description	Variable Group	Ecological Type Action	Physiolo-gical proximity	Rationale	Model used	Source
37	VHC	Vegetation height and cover index (VHC = VH*VCo)	Substrate	Direct to plants Direct to animals	Proximal to plants Proximal to animals	It is a vegetation diversity index combining total height and percentage of coverage of each class of vegetation.	Maxent and GDM	Extracted from Veg
38	VC	Vegetation complexity index (VC = VCo*VH*VS)	Substrate	Direct to plants Direct to animals	Proximal to plants Proximal to animals	It is a vegetation diversity index combining total height, percentage of cover and number of strata of each class of vegetation.	Maxent and **GDM**	Extracted from Veg

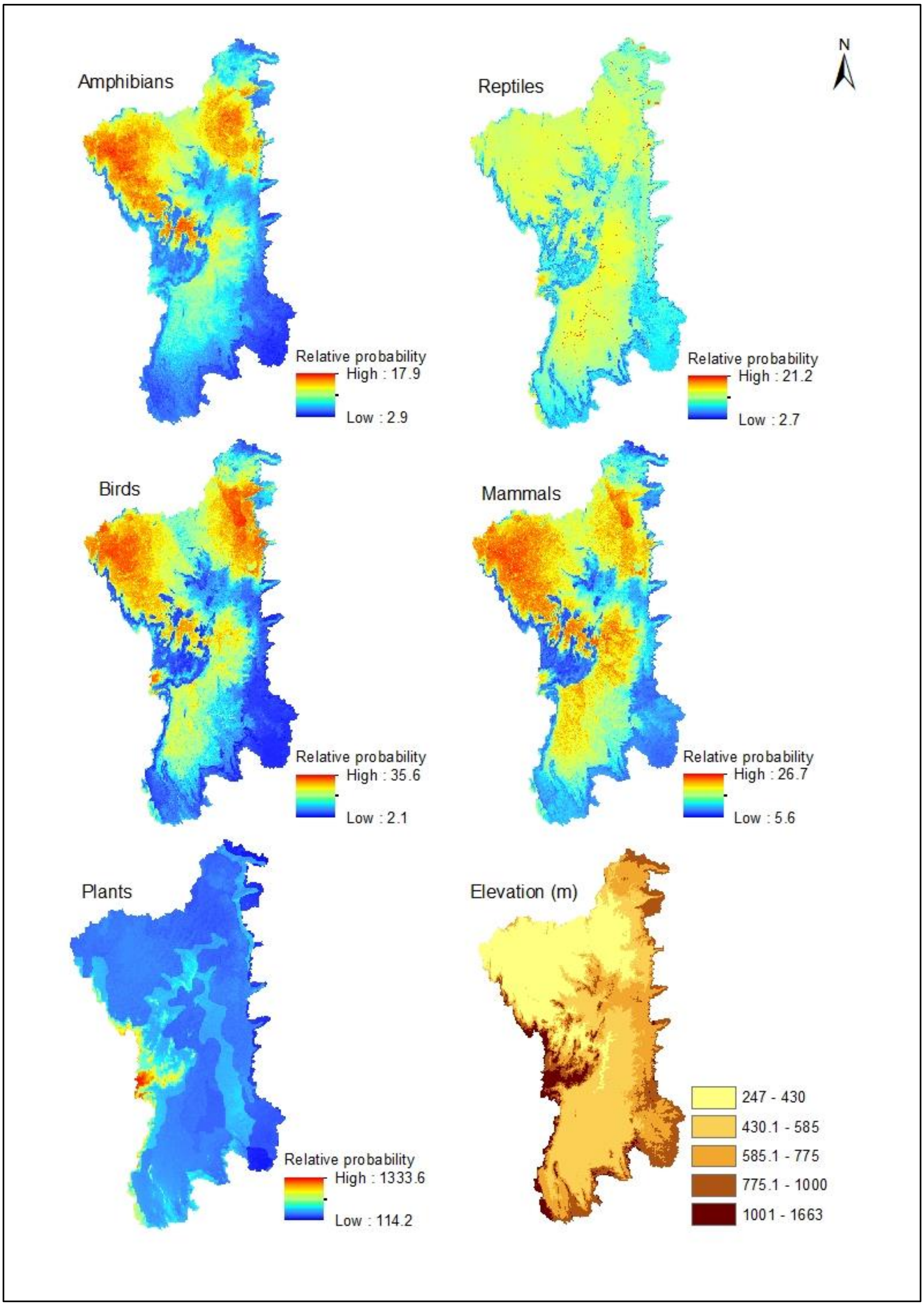
SI 2 – Table 2. Pearson's correlation coefficient between variables used to model species and assemblage distributions within the Paranã River Basin.

Variable labels follow the same codes used on Table 1, except for Lat = Latitude and Long = Longitude.

Variable	Lat	Long	AMT	MDR	IT	TS	MT WM	MT CM	TAR	MT WQ	MT DQ	MT WaQ	MT CQ	AP	PWM	PDM	PS	PWQ	PDQ	PWQ	PCQ	ARI	PET	DEM	CTI	Roug	DTS	GC	GA	SCl	SB	SD	SP	SI	Veg	VCo	VH	VS	VHC	VC
Lat	1.00	0.00	0.66	0.26	0.04	0.26	0.79	0.53	0.25	0.56	0.71	0.67	0.74	0.22	-0.16	-0.68	-0.12	0.07	-0.59	-0.55	-0.12	-0.07	0.11	-0.51	0.10	-0.14	-0.02	-0.52	0.06	-0.17	-0.03	-0.02	0.00	-0.06	-0.42	0.18	0.21	-0.11	0.16	0.18
Long		1.00	-0.31	0.63	0.59	0.63	-0.05	-0.49	0.50	-0.28	-0.40	-0.23	-0.35	-0.71	-0.44	-0.40	0.61	-0.61	-0.56	0.23	-0.38	-0.71	-0.21	0.23	0.21	-0.27	0.03	-0.10	-0.57	-0.18	0.01	-0.26	0.11	-0.11	0.36	-0.10	-0.07	-0.12	-0.04	-0.07
AMT			1.00	-0.04	-0.25	-0.04	0.90	0.85	0.02	0.98	0.97	0.99	0.98	0.29	0.07	-0.50	0.07	0.26	-0.42	-0.62	-0.19	0.01	0.33	-0.82	0.07	-0.14	-0.12	-0.48	0.27	0.00	0.02	0.07	-0.17	-0.01	-0.51	0.37	0.35	0.16	0.30	0.33
MDR				1.00	0.11	1.00	0.37	-0.52	0.97	0.05	-0.19	0.04	-0.12	-0.61	-0.59	-0.37	0.38	-0.60	-0.46	0.31	-0.32	-0.79	-0.13	-0.21	0.18	-0.28	0.00	-0.23	-0.31	-0.03	-0.14	-0.10	0.04	-0.23	0.08	-0.11	-0.08	-0.10	-0.03	-0.08
IT					1.00	0.11	-0.21	-0.11	-0.11	-0.31	-0.20	-0.23	-0.19	-0.15	-0.04	-0.20	0.18	-0.13	-0.26	0.02	-0.07	-0.13	-0.20	0.43	0.17	-0.20	0.05	0.01	-0.46	-0.16	-0.04	-0.29	0.08	-0.12	0.27	-0.17	-0.13	-0.23	-0.13	-0.13
TS						1.00	0.37	-0.52	0.97	0.05	-0.19	0.04	-0.12	-0.61	-0.59	-0.37	0.38	-0.60	-0.46	0.31	-0.32	-0.79	-0.13	-0.21	0.18	-0.28	0.00	-0.23	-0.31	-0.03	-0.14	-0.10	0.04	-0.23	0.08	-0.11	-0.08	-0.10	-0.03	-0.08
MTWM							1.00	0.58	0.42	0.90	0.83	0.93	0.87	0.06	-0.18	-0.64	0.18	0.02	-0.60	-0.49	-0.29	-0.28	0.24	-0.83	0.13	-0.22	-0.10	-0.55	0.14	-0.03	-0.03	0.03	-0.12	-0.09	-0.47	0.29	0.29	0.07	0.26	0.27
MTCM								1.00	-0.49	0.75	0.92	0.80	0.90	0.57	0.35	-0.33	-0.14	0.52	-0.22	-0.73	0.00	0.41	0.31	-0.52	0.00	-0.01	-0.09	-0.32	0.32	-0.04	0.08	0.06	-0.13	0.10	-0.46	0.34	0.31	0.11	0.25	0.30
TAR									1.00	0.13	-0.14	0.10	-0.08	-0.57	-0.58	-0.32	0.34	-0.56	-0.40	0.30	-0.31	-0.76	-0.08	-0.31	0.14	-0.23	-0.01	-0.24	-0.21	0.01	-0.13	-0.04	0.02	-0.20	0.02	-0.06	-0.04	-0.04	0.00	-0.04
MTWQ										1.00	0.91	0.98	0.92	0.17	0.01	-0.46	0.15	0.17	-0.40	-0.51	-0.24	-0.10	0.34	-0.86	0.08	-0.17	-0.14	-0.46	0.25	0.04	0.00	0.09	-0.19	-0.03	-0.48	0.37	0.35	0.20	0.31	0.33
MTDQ											1.00	0.95	0.99	0.44	0.18	-0.46	-0.07	0.39	-0.36	-0.70	-0.09	0.18	0.32	-0.73	0.04	-0.09	-0.10	-0.45	0.31	-0.03	0.02	0.06	-0.16	0.01	-0.53	0.35	0.33	0.11	0.27	0.31
MTWaQ												1.00	0.97	0.22	0.02	-0.54	0.13	0.21	-0.47	-0.60	-0.23	-0.07	0.32	-0.83	0.09	-0.17	-0.12	-0.50	0.24	-0.01	0.02	0.06	-0.17	-0.02	-0.49	0.37	0.35	0.15	0.30	0.33
MTCQ													1.00	0.39	0.12	-0.51	-0.03	0.33	-0.41	-0.69	-0.13	0.12	0.31	-0.75	0.06	-0.11	-0.10	-0.48	0.28	-0.04	0.03	0.06	-0.15	0.01	-0.53	0.35	0.34	0.11	0.28	0.32
AP														1.00	0.77	0.24	-0.51	0.94	0.38	-0.38	0.49	0.93	0.19	-0.10	-0.18	0.26	0.01	0.01	0.41	0.10	0.06	0.14	-0.02	0.16	-0.33	0.09	0.04	-0.04	0.00	0.01
PWM															1.00	0.31	-0.02	0.90	0.39	-0.18	0.38	0.79	0.21	0.04	-0.16	0.24	0.01	0.14	0.28	0.13	0.04	0.12	0.05	0.16	-0.17	0.06	-0.04	-0.05	-0.05	-0.10
PDM																1.00	-0.42	0.25	0.90	0.43	0.48	0.43	-0.07	0.38	-0.21	0.28	0.05	0.51	0.15	0.11	-0.02	0.04	-0.10	0.01	0.21	-0.16	-0.16	0.14	-0.13	-0.13
PS																	1.00	-0.24	-0.56	0.03	-0.50	-0.54	0.10	-0.16	0.14	-0.20	-0.04	-0.14	-0.22	0.05	0.03	-0.02	0.08	0.01	0.09	0.06	0.02	-0.03	0.04	-0.03
PWQ																		1.00	0.36	-0.33	0.43	0.89	0.24	-0.11	-0.18	0.26	0.01	0.05	0.40	0.16	0.08	0.17	-0.02	0.19	-0.28	0.09	0.03	-0.03	0.00	-0.02
PDQ																			1.00	0.35	0.54	0.56	0.01	0.32	-0.23	0.32	0.05	0.46	0.28	0.14	0.01	0.13	-0.07	0.08	0.12	-0.13	-0.15	0.12	-0.12	-0.12
PWQ																				1.00	0.06	-0.25	-0.18	0.35	-0.01	0.02	0.04	0.33	-0.19	0.14	-0.09	-0.01	0.05	-0.09	0.32	-0.19	-0.21	0.00	-0.16	-0.21
PCQ																					1.00	0.54	-0.01	0.22	-0.13	0.19	0.07	0.15	0.11	0.02	-0.03	0.03	-0.02	0.01	-0.02	-0.09	-0.11	-0.02	-0.10	-0.10
ARI																						1.00	0.12	0.17	-0.23	0.34	0.04	0.20	0.39	0.10	0.11	0.14	0.01	0.22	-0.17	0.02	-0.02	-0.03	-0.05	-0.04
PET																							1.00	-0.36	0.03	-0.09	0.02	-0.13	0.22	0.20	0.08	0.11	-0.03	0.11	-0.19	0.20	0.19	0.13	0.18	0.18
DEM																								1.00	-0.02	0.08	0.07	0.40	-0.25	-0.08	-0.02	-0.14	0.18	0.00	0.42	-0.32	-0.30	-0.19	-0.28	-0.28
CTI																									1.00	-0.66	-0.06	-0.22	-0.23	0.01	-0.14	-0.06	0.09	-0.13	0.00	0.01	0.00	-0.07	0.00	-0.01
Roug																										1.00	0.06	0.34	0.30	-0.03	0.23	0.07	-0.08	0.22	0.02	-0.02	-0.01	0.06	-0.01	0.01
DTS																											1.00	0.02	-0.06	0.01	0.01	0.02	0.03	0.02	0.04	-0.04	-0.03	-0.03	-0.03	-0.03
GC																												1.00	0.21	0.03	0.11	-0.03	-0.04	0.09	0.30	-0.21	-0.18	0.03	-0.16	-0.15
GA																													1.00	0.12	0.18	0.17	-0.13	0.22	-0.25	0.06	0.04	0.06	0.01	0.04
SCl																														1.00	-0.03	0.17	0.07	0.05	-0.06	0.07	0.01	0.04	0.01	-0.02
SB																															1.00	0.14	0.10	0.83	0.12	0.18	0.24	0.21	0.22	0.28
SD																																1.00	0.52	0.56	-0.15	-0.05	-0.09	-0.11	-0.08	-0.13
SP																																	1.00	0.47	0.02	0.01	-0.05	-0.17	-0.06	-0.12
SI																																		1.00	0.01	0.10	0.10	0.03	0.08	0.09
Veg																																			1.00	-0.11	0.02	0.18	0.02	0.10
VCo																																				1.00	0.90	0.72	0.83	0.74
VH																																					1.00	0.77	0.96	0.90
VS																																						1.00	0.75	0.85
VHC																																							1.00	0.88
VC																																								1.00

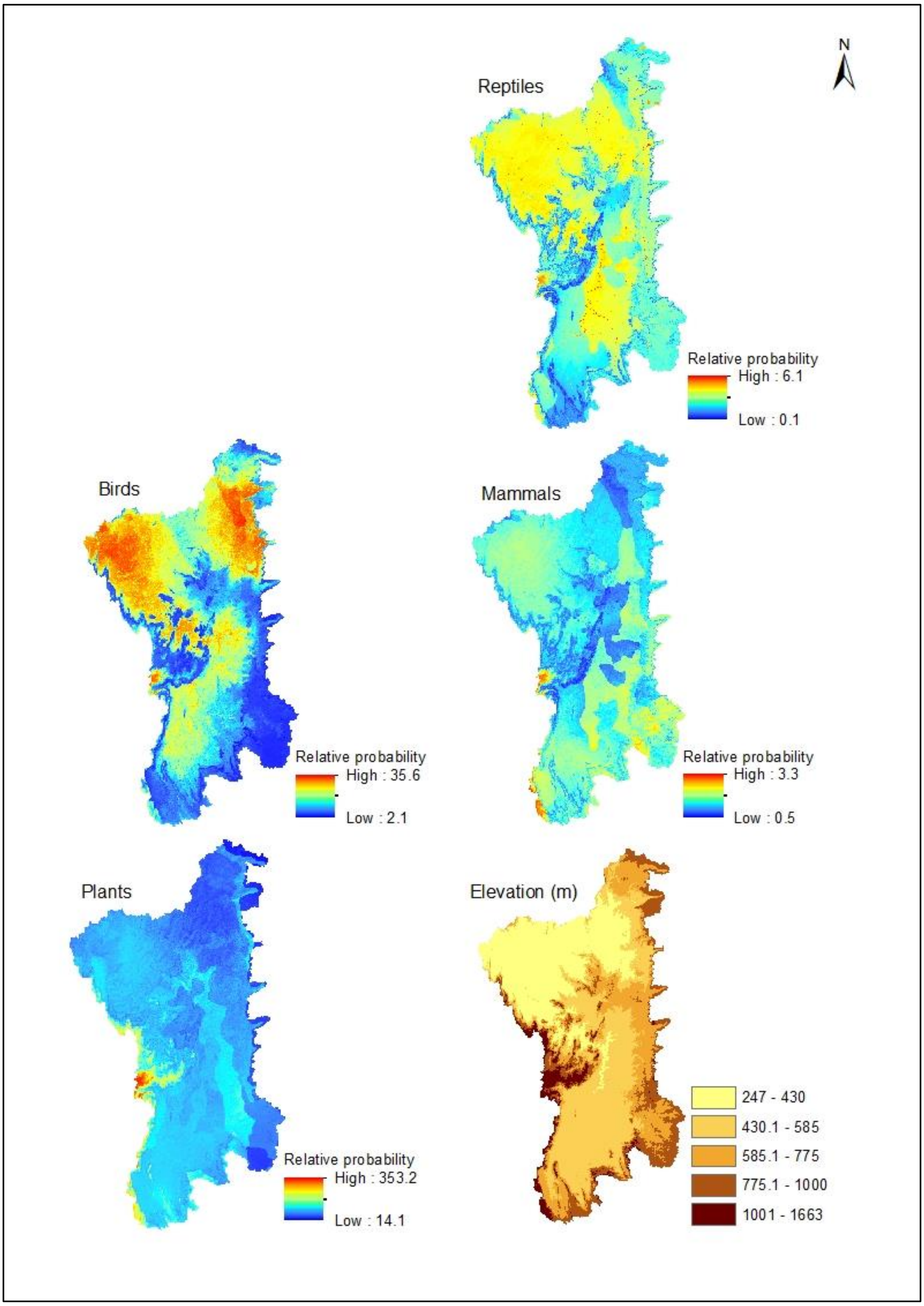
SI 2 – Additional Figures

As an exploratory analysis of the behaviour of spatial richness distribution using continuous distribution in stacked SDMs we used the same data and methods used to build the stacked SDMs presented in the main text (see Chapter3). The only exception is related to the use of a threshold to stacking predicted species distributions. In the following figures, the spatial species richness maps for the full models and tested models (see methods section in Chapter 3 for description) were derived from the Maxent logistic output.



SI 2 – Additional Figure 1. Species richness maps for the full models (stacked models from the Maxent logistic output of species models tested and non-tested for model accuracy) and elevation classes within the Paranã River Basin.

Warmer (reddish) and colder (bluish) colors indicate higher and lower summation of relative probability values of species occurrence of amphibians, reptiles, birds, mammals and plants.



SI 2 – Additional Figure 2. Species richness maps based on models tested with a 10-fold cross-validation (stacked models from the Maxent logistic output of species recorded in 20 or more localities) and elevation classes within the Paranã River Basin.

Note that amphibians was not included in the tested models given that small numbers of records precluded SDMs. Warmer (reddish) and colder (bluish) colors indicate higher and lower summation of relative probability values of species occurrence of amphibians, reptiles, birds, mammals and plants.

Supporting Information 3 – Conservation in a biodiversity hotspot: a multi-solution approach to conservation prioritization

---

Contents

SI 3. A brief description of the Zonation software and the additive benefit function cell removal rule

SI 3. A brief description of the Zonation software and the additive benefit function cell removal rule

The Zonation algorithm is a reserve selection software for large-scale high-resolution spatial planning using GIS grid data, such as predicted spatial features distributions (species distributions, species assemblages and ecosystems types, for example). It is primarily intended for binary type problems (select or not, protected or not), instead of direct targeting of multiple alternative conservation actions[172]. Generally speaking, the Zonation produces a hierarchical grid cell prioritization (i.e. the most valuable 2% fraction of the landscape is within the most valuable 5%, that is within the most valuable 10% and so on) that seeks to maximize retention of weighted features in each cell. The hierarchical prioritization in Zonation is a reverse process, i.e. it starts from the full landscape and iteratively removes cells one by one from the landscape, using the strategy of minimization of marginal loss as the criterion to decide which cell is removed next [172]. The hierarchy is generated via iterated removal of that cell whose loss causes the smallest decrease in the conservation value of the remaining reserve network [177]. The presence in a particular grid cell of each spatial layer (feature) included in the analysis, adds up to the total value of that cell. Assuming equal distributions and weight for each feature, the principle is: the more features are present in a particular cell, the higher is the value of retention of that cell. Each round, the cell with less value is removed and the value of each cell is recalculated considering the remaining distribution of the features.

The cell-remove rule in Zonation determines which cell leads to smallest marginal loss of biodiversity value. The additive benefit function cell-removal rule in Zonation [177, 179] is a function that gives more weight to locations (grid cells) with high features values taking into account all features proportions in a given cell [172]. Thus, the additive benefit function results in a reserve network concentrate towards features-rich locations, where cells have high aggregate value over features, i.e. which has a higher performance on average over all features retained. However, a lower minimum proportion of the original distributions for the worst-off features is retained [172]. In other words, features-rich areas are favoured over features-poor areas that might be, for example, significant for the existence of one or few species. In this sense, additive benefit function allows a trade-off between features representations, since the cells in sites where distributions of several features overlap receive a higher value than the cells where only one feature occurs.

References

References

1. Pimm SL, Russell GJ, Gittleman JL, Brooks TM. The future of biodiversity. Science. 1995;269:347-50.
2. CDB. Global biodiversity outlook 2. Montreal: Secretariat of the Convention on Biological Diversity; 2006. 81 + vii p.
3. Barnosky AD, Matzke N, Tomiya S, Wogan GOU, Swartz B, Quental TB, et al. Has the Earth's sixth mass extinction already arrived? Nature. 2011;471(7336):51-7.
4. Vane-Wright RI, Humphries CJ, Williams PH. What to protect? Systematics and the agony of choice. Biol Conserv. 1991;55(3):235-54. doi: http://dx.doi.org/10.1016/0006-3207(91)90030-D.
5. Multilateral Convention on biological diversity (with annexes). Concluded at Rio de Janeiro on 5 June 1992, 30619 (1992).
6. Butchart SHM, Walpole M, Collen B, van Strien A, Scharlemann JPW, Almond REA, et al. Global biodiversity: indicators of recent declines. Science. 2010:science.1187512. doi: 10.1126/science.1187512.
7. CBD/UNEP. COP 10 Decision X/2. Strategic plan for biodiversity 2011/2020: SCBD - Secretariat of the Convention on Biological Diversity.; 2010 [cited 2014 26th of November of 2014]. Available from: https://www.cbd.int/decision/cop/?id=12268.
8. Drummond JA, Franco JLdA, Ninis AB. Brazilian federal conservation units: a historical overview of their creation and of their current status. Environment and History. 2009;15:463-91. doi: 10.3197/096734009X12532652872036.
9. Margules CR, Pressey RL, Williams PH. Representing biodiversity: data and procedures for identifying priority areas for conservation. Journal of Biosciences. 2002;27(4):309-26. PubMed PMID: ISI:000176944900003.
10. Margules CR, Pressey RL. Systematic conservation planning. Nature. 2000;405(6783):243-53.
11. Cowling RM, Pressey RL. Introduction to systematic conservation planning in the Cape Floristic Region. Biol Conserv. 2003;112(1-2):1-13. doi: 10.1016/S0006-3207(02)00418-4.
12. Miller KR, Lanou SM. National biodiversity planning: guidelines based on early experiences around the world. Washington D. C; Nairobi; Gland, Switzerland: World Resources Institute, United Nations Environment Programme and The World Conservation Union; 1995. 161 p.
13. Pressey RL, Cabeza M, Watts ME, Cowling RM, Wilson KA. Conservation planning in a changing world. Trends in Ecology & Evolution. 2007;22(11):583-92. doi: http://dx.doi.org/10.1016/j.tree.2007.10.001.
14. Pressey RL, Bottrill MC. Approaches to landscape and seascape scale conservation planning: convergence, contrasts and challenges. Oryx. 2009;43(4):464-75. doi: 10.1017/S0030605309990500.
15. Parmesan C, Yohe G. A globally coherent fingerprint of climate change impacts across natural systems. Nature. 2003;421(6918):37-42.
16. Root TL, Price JT, Hall KR, Schneider SH, Rosenzweig C, Pounds JA. Fingerprints of global warming on wild animals and plants. Nature. 2003;421(6918):57-60.

17. Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, et al. Extinction risk from climate change. Nature. 2004;427(6970):145-8.

18. Williams P, Hannah L, Andelman S, Midgley G, Araújo M, Hughes G, et al. Planning for climate change: identifying minimum-dispersal corridors for the Cape Proteaceae. Conserv Biol. 2005;19(4):1063-74. doi: 10.1111/j.1523-1739.2005.00080.x.

19. Thomas CD, Bodsworth EJ, Wilson RJ, Simmons AD, Davies ZG, Musche M, et al. Ecological and evolutionary processes at expanding range margins. Nature. 2001;411(6837):577-81.

20. Walther G-R, Post E, Convey P, Menzel A, Parmesan C, Beebee TJC, et al. Ecological responses to recent climate change. Nature. 2002;416(6879):389-95.

21. Heller NE, Zavaleta ES. Biodiversity management in the face of climate change: a review of 22 years of recommendations. Biol Conserv. 2009;142(1):14-32.

22. Lawler JJ. Climate change adaptation strategies for resource management and conservation planning. Annals of the New York Academy of Sciences. 2009;1162(The Year in Ecology and Conservation Biology 2009):79-98.

23. Williams SE, Bolitho EE, Fox S. Climate change in Australian tropical rainforests: an impending environmental catastrophe. P Roy Soc Lond B Bio. 2003;270(1527):1887-92. doi: 10.1098/rspb.2003.2464.

24. Shoo LP, Williams SE, Hero J-M. Detecting climate change induced range shifts: where and how should we be looking? Austral Ecology. 2006;31(1):22-9. doi: 10.1111/j.1442-9993.2006.01539.x.

25. Hannah L, Midgley G, Andelman S, Araújo M, Hughes G, Martinez-Meyer5 E, et al. Protected area needs in a changing climate. Front Ecol Environ. 2007;5(3):131-8.

26. Fuller RA, McDonald-Madden E, Wilson KA, Carwardine J, Grantham HS, Watson JEM, et al. Replacing underperforming protected areas achieves better conservation outcomes. Nature. 2010;466(7304):365-7. doi: 10.1038/nature09180.

27. Peterson GD, Cumming GS, Carpenter SR. Scenario planning: a tool for conservation in an uncertain world. Conserv Biol. 2003;17(2):358-66.

28. MMA. Avaliação e identificação de áreas e ações prioritárias para a conservação, utilização sustentável e repartição de benefícios da biodiversidade nos biomas brasileiros. Brasília, DF: Ministério do Meio Ambiente - MMA; Centro de Informação e Documentação Luiz Eduardo Magalhães - CID Ambiental; 2002.

29. MMA. Priority areas for the conservation, sustainable use, and benefits sharing of Brazilian biological diversity. Update: MMA Administrative Ruling n^{o} 9, of 23 January 2007. Brasília - DF: Ministry of the Environment - MMA. National Secretariat of Biodiversity and Forests. Department of Biodiversity Conservation. Environmental CID; 2007.

30. Mittermeier RA, Gil PR, Hoffman M, Pilgrim J, Brooks T, Mittermeier CG, et al. Hotspots revisited: earth's biologically richest and most threatened terrestrial ecoregions. Chicago: The University of Chicago Press; 2005. 392 p.

31. Scariot A, Sevilha AC. Biodiversidade, estrutura e conservação de florestas estacionais deciduais no Cerrado. In: Scariot A, Felfili JM, Souza-Silva JC, editors. Cerrado: ecologia, biodiversidade e conservação. Brasília: Ministério do Meio Ambiente; 2005. p. 121-39.

32. BirdLife-International. Endemic bird area factsheet: deciduous forests of Minas Gerais and Goiás. Downloaded from http://www.birdlife.org on 28/05/2013. 2013.

33. Werneck FP, Colli GR. The lizard assemblage from seasonally dry tropical forest enclaves in the Cerrado biome, Brazil, and its association with the Pleistocenic Arc. Journal of Biogeography. 2006;33(11):1983-92. doi: 10.1111/j.1365-2699.2006.01553.x.

34. Olmos F, Martuscelli P, Silva e Silva R. Distribution and dry-season ecology of Pfrimer's Conure *Pyrrhura pfrimeri*, with a reappraisal of Brazilian *Pyrrhura* "*leucotis*". Ornitologia Neotropical. 1997;(8):121-32.

35. Silva JMC, Bates JM. Biogeographic patterns and conservation in the South American Cerrado: a tropical savanna hotspot. Bioscience. 2002;52:225-33. doi: 10.1641/0006-3568(2002)052[0225:bpacit]2.0.co;2.

36. Trajano E, Bichuette ME. Área cárstica de São Domingos, alto Tocantins, nordeste de Goiás: a maior diversidade de peixes subterrâneos no Brasil. O Carste. 2003;15(4):114-25.

37. Bichuette ME, Trajano E. Epigean and subterranean ichthyofauna from the São Domingos karst area, upper Tocantins River Basin, central Brazil. J Fish Biol. 2003;63(5):1100-21. doi: 10.1046/j.1095-8649.2003.00227.x.

38. Camargo NF, Aguiar LMS. Análise de dados secundários sobre a fauna de vertebrados da unidade fitogeográfica do Vale do Paranã. Documentos. 182. Planaltina-DF: Embrapa Cerrados; 2007. p. 56.

39. Scariot A, Sevilha AC. Diversidade, estrutura e manejo de florestas deciduais e as estratégias para a conservação. In: Cavalcanti TB, Walter BMT, editors. Tópicos atuais em botânica: palestras convidadas do 51º Congresso Nacional de Botânica. Brasília: Embrapa Recursos Genéticos e Biotecnologia; 2000. p. 183-8.

40. Mata RAd, Roque F, Tidon R. Drosophilids (Insecta, Diptera) of the Paranã Valley: eight new records for the Cerrado biome. Biota Neotropica. 2008;8:55-60.

41. Mata RAd, McGeoch M, Tidon R. Drosophilids (Insecta, Diptera) as tools for conservation biology. Natureza & Conservação. 2010;8(1):60-5. doi: 10.4322/natcon.00801009.

42. Sevilha AC, Colli GR, Constantino R, Marini MA, Marinho-Filho J, Tidon R, et al. Inventário da biodiversidade do vale e serra do Paranã e do sul do Tocantins - relatório final apresentado ao Ministério do Meio Ambiente - MMA. Brasília: Embrapa Recursos Genéticos e Biotecnologia - Embrapa-Cenargen; Universidade de Brasília - UnB; Pesquisa e Conservação do Cerrado - Pequi, 2007.

43. Sevilha AC, Scariot A, Noronha SE. Estado atual da representatividade de unidades de conservação em florestas estacionais deciduais no Brasil. In: Silva AF, editor. LV Congresso Nacional de Botânica, XXVI Encontro Regional de Botânicos BA, MG, ES, Simpósio sobre Unidades de Conservação nos Biomas Brasileiros; Julho de 2004; Viçosa: Sociedade Botânica do Brasil, Universidade Federal de Viçosa; 2004. p. 1-60.

44. Janzen DH. Tropical dry forest: the most endangered major tropical ecosystem. In: Wilson EO, editor. Biodiversity. Washington, DC, USA.: National Academy Press; 1988. p. 130-7.

45. Gardner G. Viagem ao interior do Brasil, principalmente nas províncias do Norte e nos distritos do ouro e do diamante durante os anos de 1836-1841. São Paulo: Editora da Universidade de São Paulo e Livraria e Itatiaia Editora Ltda.; 1975.

46. Spix JB, Martius CFP. Viagem pelo Brasil: 1817-1820. 3rd edition ed. São Paulo: Melhoramentos; 1976. 270 p.

47. Barbosa AG. As estratégias de conservação da biodiversidade na Chapada dos Veadeiros: conflitos e oportunidades. Brasília: Universidade de Brasília; 2008.

48. Barbosa AS, Schimiz PI. Ocupação indígena do Cerrado: esboço de uma história. In: Sano SM, Almeida SP, editors. Cerrado: ambiente e flora. Brasília - DF: Embrapa - CPAC; 1988. p. 3-42.

49. Luíz GC. Estudo do impacto ambiental a partir da análise espaço-temporal - caso da região Vão do Paranã - GO [Dissertação de Mestrado.]. Goiânia: Universidade Federal de Goiás; 1998.

50. Espírito-Santo MM, Sevilha AC, Anaya FC, Barbosa R, Fernandes GW, Sanchez-Azofeifa GA, et al. Sustainability of tropical dry forests: two case studies in southeastern and central Brazil. For Ecol Manage. 2009;258(6):922-30.

51. Tiburcio BA, Valente ALEF. O comércio justo e solidário é alternativa para segmentos populacionais empobrecidos? Estudo de caso em Território Kalunga (GO). Revista de Economia e Sociologia Rural. 2007;45(2):497-519.

52. Fioravanti MCS, Sereno JRB, Neiva ACGR, Abud LJ, Lobo, J.R., Francescantônio D, Cardoso WS, et al., editors. Reintrodução do gado curraleiro na comunidade quilombola Kalunga de Cavalcante, Goiás, Brasil: resultados preliminares. II International Symposium on Tropical Savannas; 2008; Brasília: Embrapa.

53. Scariot A, Sevilha AC, Colli GR, Cavalcanti TB, Roma JC, Dalponte J. Subsídios à conservação da biodiversidade na bacia do rio Paranã. Relatório parcial de atividades apresentado à Conservação Internacional do Brasil. Brasília: Embrapa Recursos Genéticos e Biotecnologia - Embrapa-Cenargen; Universidade de Brasília - Deptos. Zoologia e Ecologia; Pesquisa e Conservação do Cerrado - Pequi, 2004.

54. IBGE. Zoneamento geoambiental e agroecológico do estado de Goiás: região nordeste. Moreira HL, editor. Rio de Janeiro, Brasil: IBGE/Divisão de Geociências do Centro-Oeste; 1995. 178 p.

55. Jarvis A, Reuter HI, Nelson A, Guevara E. Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT) 2008 [23/05/2012]. Available from: http://srtm.csi.cgiar.org.

56. IBGE, cartographer Mapa de unidades de relevo do Brasil. Brasília: Ministério do Planejamento, Orçamento e Gestão. Instituto de Geografia e Estatística. Diretoria de Geociências; 2006.

57. ANA, cartographer Corpos hídricos superficiais. Brasília: Superintendência de Planejamento de Recursos Hídricos - SPR. Coordenação do Conjuntura e Gestão da Informação; 2016.

58. Sevilha AC, Pressey RL, Wiederhecker HC, VanDerWal J, Williams SE, Colli GR, et al. Digging deep for biodiversity information in data deficient areas: the case of the Paranã River Basin within a global biodiversity hotspot. Chapter 2 - in preparation.

59. Ribeiro JF, Walter BMT. Fitofisionomias do bioma Cerrado. In: Sano SM, Almeida SP, Ribeiro JF, editors. Cerrado: ecologia e flora. 1. Brasília, DF: Embrapa Cerrados; 2008.

60. MMA, cartographer Mapeamento da cobertura vegetal do bioma Cerrado (1:5000000). Brasília: Ministério do Meio Ambiente. (http://www.mma.gov.br/biomas/cerrado/mapa-de-cobertura-vegetal); 2007.

61. Diniz-Filho JAF, Bini LM, Pinto MP, Terribile LC, de Oliveira G, Vieira CM, et al. Conservation planning: a macroecological approach using the endemic terrestrial

vertebrates of the Brazilian Cerrado. Oryx. 2008;42(04):567-77. doi: 10.1017/S0030605308001129.

62. Valdujo PH, Silvano DL, Colli G, Martins M. Anuran species composition and distribution patterns in Brazilian Cerrado, a Neotropical hotspot. South American Journal of Herpetology. 2012;7(2):63-78. doi: 10.2994/057.007.0209.

63. Diniz-Filho JAF, Bini LM, Pinto MP, Rangel TFLVB, Carvalho P, Bastos RP. Anuran species richness, complementarity and conservation conflicts in Brazilian Cerrado. Acta Oecologica. 2006;29(1):9-15. doi: http://dx.doi.org/10.1016/j.actao.2005.07.004.

64. Nogueira C, Ribeiro S, Costa GC, Colli GR. Vicariance and endemism in a Neotropical savanna hotspot: distribution patterns of Cerrado squamate reptiles. Journal of Biogeography. 2011;38(10):1907-22. doi: 10.1111/j.1365-2699.2011.02538.x.

65. Costa GC, Nogueira C, Machado RB, Colli GR. Squamate richness in the Brazilian Cerrado and its environmental–climatic associations. Diversity and Distributions. 2007;13(6):714-24. doi: 10.1111/j.1472-4642.2007.00369.x.

66. Lima-Ribeiro MS, Felizola Diniz-Filho JA, Barberi M. Climate stability and the current patterns of terrestrial vertebrate species richness on the Brazilian Cerrado. Quaternary International. 2010;222(1–2):230-6. doi: http://dx.doi.org/10.1016/j.quaint.2009.10.003.

67. Faleiro FV, Machado RB, Loyola RD. Defining spatial conservation priorities in the face of land-use and climate change. Biol Conserv. 2013;158(0):248-57. doi: http://dx.doi.org/10.1016/j.biocon.2012.09.020.

68. Marini MÂ, Barbet-Massin M, Lopes LE, Jiguet F. Predicted climate-driven bird distribution changes and forecasted conservation conflicts in a Neotropical savanna. Conserv Biol. 2009;23(6):1558-67. doi: 10.1111/j.1523-1739.2009.01258.x.

69. Simon LM, Oliveira Gd, Barreto BdS, Nabout JC, Rangel TFLVB, Diniz-Filho JAF. Effects of global climate changes on geographical distribution patterns of economically important plant species in cerrado. Revista Árvore. 2013;37:267-74.

70. Collevatti R, Lima-Ribeiro MS, Diniz-Filho JA, Oliveira G, Dobrovolski R, Terribile LC. Stability of Brazilian seasonally dry forests under climate change: inferences for long-term conservation. American Journal of Plant Sciences. 2013;4(4):792-805. doi: 10.4236/ajps.2013.44098.

71. Diniz-Filho JAF, Bini LM, Rangel TF, Loyola RD, Hof C, Nogués-Bravo D, et al. Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. Ecography. 2009;32(6):897-906. doi: 10.1111/j.1600-0587.2009.06196.x.

72. Ladle RJ, Whittaker RJ. Prospects and challenges. Conservation Biogeography: John Wiley & Sons, Ltd; 2011. p. 245-58.

73. Pressey RL, Mills M, Weeks R, Day JC. The plan of the day: managing the dynamic transition from regional conservation designs to local conservation actions. Biol Conserv. 2013;166(0):155-69. doi: http://dx.doi.org/10.1016/j.biocon.2013.06.025.

74. Bonvicino CR, Lindberg SM, Faria MB, Brezerra AMR. The eastern boundary of the Brazilian Cerrado: a hotspot region. Zoological Studies. 2012;51(7):1207-18.

75. Werneck FP, Costa GC, Colli GR, Prado DE, Sites Jr JW. Revisiting the historical distribution of seasonally dry tropical forests: new insights based on palaeodistribution modelling and palynological evidence. Global Ecol Biogeogr. 2010:no-no. doi: 10.1111/j.1466-8238.2010.00596.x.

76. Sobrevila C, Bath P. Evaluación ecológica rápida: um manual para usuários de América Latina y el Caribe. Arlington, VA.: The Nature Conservancy; 1992. 232 p.

77. MMA/IBAMA, cartographer Monitoramento do desmatamento nos biomas brasileiros por satélite. (1:250.000). Brasília: Ministério do Meio Ambiente/Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis. *in* http://www.mma.gov.br/portalbio (Accessed 25/04/2013); 2011.

78. Borg A, Sariyar M. Record Linkage in R. 0.4-1 ed2012.

79. Sariyar M, Borg A. The RecordLinkage package: detecting errors in data. The R Journal. 2010;2(2):61-7.

80. Haber M. Algorithm AS 207: fitting a general log-linear model. Journal of the Royal Statistical Society Series C (Applied Statistics). 1984;33(3):358-62. doi: 10.2307/2347724.

81. Do CB, Batzoglou S. What is the expectation maximization algorithm? Nat Biotech. 2008;26(8):897-9. doi: http://www.nature.com/nbt/journal/v26/n8/suppinfo/nbt1406_S1.html.

82. IUCN. IUCN red list of threatened species. Version 2012.2 2012. Available from: http://www.iucnredlist.org/technical-documents/spatial-data.

83. MMA. Lista nacional das espécies da fauna brasileira ameaçadas de extinção. Instrução Normativa no. 003, de 26 de Maio de 2003. Brasília: Ministério do Meio Ambiente; 2003.

84. MMA. Lista oficial das espécies da flora brasileira ameaçadas de extinção. Anexos I e II. Instruçao Normativa no. 006, de 23 de Setembro de 2008. Brasília: Ministério do Meio Ambiente; 2008.

85. Marinho-Filho J, Rodrigues FHG, Juarez KM. The Cerrado mammals: diversity, ecology, and natural history. In: Oliveira PS, Marquis RJ, editors. The Cerrados of Brazil. Columbia, NY: Columbia University Press; 2002. p. 266-84.

86. Lista de espécies da flora do Brasil 2013. *in* http://floradobrasil.jbrj.gov.br/ (Accessed 25/06/2013) [Internet]. Jardim Botânico do Rio de Janeiro. Rio de Janeiro. 2013 [cited 25/06/2013].

87. MMA. Biodiversidade do Cerrado e Pantanal: áreas e ações prioritárias para conservação. Brasília: Ministério do Meio Ambiente; 2007.

88. MMA/IBAMA. Monitoramento do desmatamento nos biomas brasileiros por satélite. Acordo de cooperação técnica MMA/IBAMA. Monitoramento do Bioma Cerrado 2008-2009. Brasília: Centro de Informação, Documentação Ambiental e Editoração Luís Eduardo Magalhães – CID Ambiental; Ministério do Meio Ambiente – MMA; 2011. p. 55.

89. BDGEx. Banco de dados geográficos do Exército. Manual do usuário. Versão 2.3: Ministério da Defesa. Exército Brasileiro; 2012 [cited 2012 May]. Available from: http://www.geoportal.eb.mil.br/index.php.

90. Colwell RK. EstimateS: statistical estimation of species richness and shared species from samples. Version 9. User's guide and application. Published at: http://viceroy.eeb.uconn.edu/estimates. 2013.

91. Chao A. Species estimation and applications. In: Balakrishnan N, Read CB, Vidakovic B, editors. Encyclopedia of Statistical Sciences. 12. 2nd ed. New York: Wiley; 2005. p. 7907-16.

92. Pullin AS, Knight TM. Assessing conservation management's evidence base: a survey of management-plan compilers in the United Kingdom and Australia. Conserv Biol. 2005;19(6):1989-96. doi: 10.1111/j.1523-1739.2005.00287.x.

93. Pohl JE. Viagem no interior do Brasil; empreendida nos anos de 1817 a 1821 e publicada por ordem de Sua Majestade o Imperador da Áustria Francisco Primeiro. Rio de Janeiro: Instituto Nacional do Livro; 1951.

94. de Saint-Hilaire A. Viagem à província de Goiás: Livraria Itatiaia Editora; 1975.

95. Machado RB, Aguiar LMS, Castro AAJF, Nogueira CC, Neto MBR. Caracterização da fauna e flora do Cerrado. In: Faleiro FGdFN, Lopes A, editors. Savanas: desafios e estratégias para o equilíbrio entre sociedade, agronegócio e recursos naturais Planaltina: Embrapa Cerrados2008.

96. Kremen C, Colwell RK, Erwin TL, Murphy DD, Noss RF, Sanjayan MA. Terrestrial arthropod assemblages: their use in conservation planning. Conserv Biol. 1993;7(4):796-808. doi: 10.1046/j.1523-1739.1993.740796.x.

97. Whittaker RJ, Araújo MB, Paul J, Ladle RJ, Watson JEM, Willis KJ. Conservation biogeography: assessment and prospect. Diversity and Distributions. 2005;11(1):3-23. doi: 10.1111/j.1366-9516.2005.00143.x.

98. Walther BA, Morand S. Comparative performance of species richness estimation methods. Parasitology. 1998;116 (Pt 4):395-405. Epub 1998/05/20. PubMed PMID: 9585941.

99. Hortal J, Borges PAV, Gaspar C. Evaluating the performance of species richness estimators: sensitivity to sample grain size. J Anim Ecol. 2006;75(1):274-87. doi: 10.1111/j.1365-2656.2006.01048.x.

100. Margules C, Sarkar S. Systematic conservation planning. Usher M, Adam P, Peet R, Dobson A, editors. Cambridge: Cambridge University Press; 2007.

101. Ferrier S, Manion G, Elith J, Richardson K. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Diversity and Distributions. 2007;13(3):252-64. doi: 10.1111/j.1472-4642.2007.00341.x.

102. Hone P. The financial value of cultural, heritage and scientific collections: a public management necessity. Australian Accounting Review. 1997;7(13):38-43. doi: 10.1111/j.1835-2561.1997.tb00026.x.

103. Burley HM, Laffan SW, Williams KJ. Spatial non-stationarity and anisotropy of compositional turnover in eastern Australian Myrtaceae species. International Journal of Geographical Information Science. 2012;26(11):2065-81. doi: 10.1080/13658816.2012.692371.

104. Franklin J. Mapping species distribution. Spatial inference and prediction. 5th ed. Usher M, Saunders D, Peet R, Dobson A, editors. Cambridge: Cambridge University Press; 2011. 320 p.

105. Hernandez PA, Graham CH, Master LL, Albert DL. The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography. 2006;29(5):773-85. doi: 10.1111/j.0906-7590.2006.04700.x.

106. Uniyal S. Who will collect and name them? Biodivers Conserv. 2014;23(2):511-2. doi: 10.1007/s10531-013-0601-6.

107. Guisan A, Zimmermann NE. Predictive habitat distribution models in ecology. Ecological Modelling. 2000;135(2–3):147-86. doi: http://dx.doi.org/10.1016/S0304-3800(00)00354-9.

108. Yackulic CB, Chandler R, Zipkin EF, Royle JA, Nichols JD, Campbell Grant EH, et al. Presence-only modelling using Maxent: when can we trust the inferences? Methods in Ecology and Evolution. 2012;4(3):236-43. doi: 10.1111/2041-210x.12004.

109. Blas MB, Luis C, Fabio SA. The impact of modelling choices in the predictive performance of richness maps derived from species-distribution models: guidelines to build better diversity models. Methods in Ecology and Evolution. 2013. doi: 10.1111/2041-210x.12022.

110. Elith J, H. Graham C, P. Anderson R, Dudík M, Ferrier S, Guisan A, et al. Novel methods improve prediction of species' distributions from occurrence data. Ecography. 2006;29(2):129-51. doi: 10.1111/j.2006.0906-7590.04596.x.

111. Araújo MB, Guisan A. Five (or so) challenges for species distribution modelling. Journal of Biogeography. 2006;33(10):1677-88. doi: 10.1111/j.1365-2699.2006.01584.x.

112. Williams KJ, Belbin L, Austin MP, Stein JL, Ferrier S. Which environmental variables should I use in my biodiversity model? International Journal of Geographical Information Science. 2012;26(11):2009-47. doi: 10.1080/13658816.2012.698015.

113. Guisan A, Lehmann A, Ferrier S, Austin M, Overton JMC, Aspinall R, et al. Making better biogeographical predictions of species'distributions. Journal of Applied Ecology. 2006;43:386-92. doi: 10.1111/j.1365-2664.2006.01164.x.

114. Pearson RG, Raxworthy CJ, Nakamura M, Townsend Peterson A. ORIGINAL ARTICLE: Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography. 2007;34(1):102-17. doi: 10.1111/j.1365-2699.2006.01594.x.

115. Ferrier S, Guisan A. Spatial modelling of biodiversity at the community level. Journal of Applied Ecology. 2006;43(3):393-404. doi: 10.1111/j.1365-2664.2006.01149.x.

116. Fitzpatrick MC, Sanders NJ, Ferrier S, Longino JT, Weiser MD, Dunn R. Forecasting the future of biodiversity: a test of single- and multi-species models for ants in North America. Ecography. 2011;34(5):836-47. doi: 10.1111/j.1600-0587.2011.06653.x.

117. Ferrier S. Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? Systematic Biology. 2002;51(2):331-63. doi: 10.2307/3070795.

118. Phillips SJ, Dudík M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography. 2008;31(2):161-75. doi: 10.1111/j.0906-7590.2008.5203.x.

119. Phillips SJ, Dudík M, Schapire RE. A maximum entropy approach to species distribution modeling. Proceedings of the twenty-first international conference on machine learning; Banff, Alberta, Canada. 1015412: ACM; 2004. p. 83.

120. Ferrier S, Drieslma M, Manion G, Watson G. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. II. Community-level modelling. Biodivers Conserv. 2002;11:2309-38.

121. Ferrier S, Powell GVN, Richardson KS, Manion G, Overton JM, Allnutt TF, et al. Mapping more of terrestrial biodiversity for global conservation assessment. Bioscience. 2004;54(12):1101-9. doi: 10.1641/0006-3568(2004)054[1101:mmotbf]2.0.co;2.

122. Austin M. Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecological Modelling. 2007;200(1–2):1-19. doi: http://dx.doi.org/10.1016/j.ecolmodel.2006.07.005.

123. Austin MP, Van Niel KP. Improving species distribution models for climate change studies: variable selection and scale. Journal of Biogeography. 2011;38(1):1-8. doi: 10.1111/j.1365-2699.2010.02416.x.

124. Austin MP. Vegetation and environment: discontinuities and continuities. Vegetation Ecology: John Wiley & Sons, Ltd; 2013. p. 71-106.

125. McCain CM, Grytnes J-A. Elevational gradients in species richness. Encyclopedia of Life Sciences (ELS). Chichester: John Wiley & Sons, Ltd:; 2010. p. 1-10.

126. Field R. Latitudinal diversity gradients. eLS: John Wiley & Sons, Ltd; 2001. p. 1-8.

127. Fox BJ. Niche parameters and species richness. Ecology. 1981;62(6):1415-25. doi: 10.2307/1941497.

128. Eiten G. Vegetação do cerrado. In: Novaes Pinto M, editor. Cerrado: caracterização, ocupação e perspectivas. 2 ed. Brasília: Editora Universidade de Brasília; 1993. p. 17-74.

129. Eiten G. The cerrado vegetation of Brazil. The Botanical Review. 1972;38(2):201-341.

130. MacArthur RH. Geographical Ecology. Patterns in distribution of species. Princeton: Princeton University Press; 1972. p. 269.

131. Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A, et al. Effects of sample size on the performance of species distribution models. Diversity and Distributions. 2008;14:763-73. doi: 10.1111/j.1472-4642.2008.00482.x.

132. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecological Modelling. 2006;190:231-59. doi: 10.1016/j.ecolmodel.2005.03.026.

133. Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ. A statistical explanation of Maxent for ecologists. Diversity and Distributions. 2011;17(1):43-57. doi: 10.1111/j.1472-4642.2010.00725.x.

134. Syfert MM, Smith MJ, Coomes DA. The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. PLoS ONE. 2013;8(2):e55158. doi: 10.1371/journal.pone.0055158.

135. Manel S, Williams HC, Ormerod SJ. Evaluating presence–absence models in ecology: the need to account for prevalence. Journal of Applied Ecology. 2001;38(5):921-31. doi: 10.1046/j.1365-2664.2001.00647.x.

136. Anderson RP, Gonzalez Jr I. Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. Ecological Modelling. 2011;222(15):2796-811. doi: http://dx.doi.org/10.1016/j.ecolmodel.2011.04.011.

137. Pearson RG, Thuiller W, Araújo MB, Martinez-Meyer E, Brotons L, McClean C, et al. Model-based uncertainty in species range prediction. Journal of Biogeography. 2006;33(10):1704-11. doi: 10.1111/j.1365-2699.2006.01460.x.

138. Swets J. Measuring the accuracy of diagnostic systems. Science. 1988;240(4857):1285-93. doi: 10.1126/science.3287615.

139. Phillips JT. Species' distribution modeling for conservation educators and practitioners. Exercise. Lessons in Conservation [Internet]. 2009 30/05/2015; (3):[107-35 pp.]. Available from: http://ncep.amnh.org/linc.

140. VanDerWal J, Falconi L, Januchowski S, Shoo L, Storlie C. Species distribution modelling tools: tools for processing data associated with species distribution modelling exercises. 1.1-13 ed2013.

141. Legendre P. Spatial autocorrelation: trouble or new paradigm? Ecology. 1993;74(6):1659-73. doi: 10.2307/1939924.

142. Manion G. .NET GD Modeller. Canberra: Office of the Environment and Heritage, New South Wales Government, Australia; 2012.

143. Belbin L, Marshall C, Faith D. Representing relationships by automatic assignment of colour. Australian Computer Journal. 1983;15(4): 160-3 7 refs. PubMed PMID: Journal.

144. Rodrigues MT, Acosta-Galvis A, Lavilla E, Solís F, Ibáñez R, Jaramillo C, et al. *Elachistocleis ovalis*. IUCN 2013 IUCN Red List of Threatened Species Version 20132 <wwwiucnredlistorg> Downloaded on 07 February 2014 [Internet]. 2010.

145. Schneider CJ, Smith TB, Larison B, Moritz C. A test of alternative models of diversification in tropical rainforests: ecological gradients vs. rainforest refugia. Proceedings of the National Academy of Sciences. 1999;96(24):13869-73. doi: 10.1073/pnas.96.24.13869.

146. Faith DP, Ferrier S. Linking species-compositional dissimilarities and environmental data for biodiversity assessment. Available at: http://australianmuseumnetau/uploads/documents/28679/d%20p%20faith,%20s%20ferrier,%20australian%20museum%20online%20(2002)pdf. 2002.

147. Hawkins BA. Ecology's oldest pattern? Endeavour. 2001;25(3):133. doi: http://dx.doi.org/10.1016/S0160-9327(00)01369-7.

148. Cassemiro FAS, Diniz-Filho JAF. Diversity gradients and the metabolic theory of ecology2010.

149. Gaston KJ. Global patterns in biodiversity. Nature. 2000;405(6783):220-7.

150. Hill JL, Hill RA. Why are tropical rain forests so species rich? Classifying, reviewing and evaluating theories. Progress in Physical Geography. 2001;25(3):326-54. doi: 10.1177/030913330102500302.

151. Pianka ER. Latitudinal gradients in species diversity - a review of concepts. Am Nat. 1966;100(910):33-46.

152. Pacheco JF, Olmos F. As aves do Tocantins 1: região Sudeste. Revista Brasileira de Ornitologia. 2006;14(2):85-100.

153. Silva JMC. Endemic bird species and conservation in the Cerrado Region, South America. Biodiversity & Conservation. 1997;6(3):435-50. doi: 10.1023/a:1018368809116.

154. Simon MF, Proença C. Phytogeographic patterns of *Mimosa* (Mimosoideae, Leguminosae) in the Cerrado biome of Brazil: an indicator genus of high-altitude centers of endemism? Biol Conserv. 2000;96(3):279-96. doi: http://dx.doi.org/10.1016/S0006-3207(00)00085-9.

155. Guillera-Arroita G, Lahoz-Monfort JJ, Elith J, Gordon A, Kujala H, Lentini PE, et al. Is my species distribution model fit for purpose? Matching data and models to applications. Global Ecol Biogeogr. 2015;24(3):276-92. doi: 10.1111/geb.12268.

156. Kramer-Schadt S, Niedballa J, Pilgrim JD, Schröder B, Lindenborn J, Reinfelder V, et al. The importance of correcting for sampling bias in Maxent species distribution models. Diversity and Distributions. 2013;19(11):1366-79. doi: 10.1111/ddi.12096.

157. Phillips SJ, Dudik M, Elith J, Graham CH, Lehmann A, Leathwick J, et al. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecological Applications. 2009;1(19):181-97.

158. Elith J, Leathwick J. Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. Diversity and Distributions. 2007;13(3):265-75. doi: 10.1111/j.1472-4642.2007.00340.x.

159. VanDerWal J, Shoo LP, Graham C, Williams SE. Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? Ecological Modelling. 2009;220(4):589-94. doi: http://dx.doi.org/10.1016/j.ecolmodel.2008.11.010.

160. Calabrese JM, Certain G, Kraan C, Dormann CF. Stacking species distribution models and adjusting bias by linking them to macroecological models. Global Ecol Biogeogr. 2014;23(1):99-112. doi: 10.1111/geb.12102.

161. Ryti RT. Effect of the focal taxon on the selection of nature reserves. Ecological Applications. 1992;2(4):404-10. doi: 10.2307/1941875.

162. Decreto No. 8447, de 06 de Maio de 2015. Dispõe sobre o Plano de Desenvolvimento Agropecuário do Matopiba e a criação de seu Comitê Gestor. Presidência da República, Casa Civil, Subchefia de Assuntos Jurídicos, (2015).

163. Watson JEM, Grantham HS, Wilson KA, Possingham HP. Systematic conservation planning: past, present and future. Conservation Biogeography: John Wiley & Sons, Ltd; 2011. p. 136-60.

164. Pressey RL, Cowling RM, Rouget M. Formulating conservation targets for biodiversity pattern and process in the Cape Floristic Region, South Africa. Biol Conserv. 2003;112(1-2):99-127. PubMed PMID: ISI:000182913800007.

165. Lehtomäki J, Moilanen A. Methods and workflow for spatial conservation prioritization using Zonation. Environmental Modelling & Software. 2013;47(0):128-37. doi: http://dx.doi.org/10.1016/j.envsoft.2013.05.001.

166. Green A, Smith SE, Lipsett-Moore G, Groves C, Peterson N, Sheppard S, et al. Designing a resilient network of marine protected areas for Kimbe Bay, Papua New Guinea. Oryx. 2009;43(04):488-98. doi: 10.1017/S0030605309990342.

167. Bini LM, Diniz-Filho JAF, Rangel TFLVB, Bastos RP, Pinto MP. Challenging Wallacean and Linnean shortfalls: knowledge gradients and conservation planning in a biodiversity hotspot. Diversity and Distributions. 2006;12(5):475-82. doi: 10.1111/j.1366-9516.2006.00286.x.

168. Sevilha AC, Pressey RL, VanDerWal J, Williams SE, Wiederhecker HC, Anderson BJ, et al. Conservation in a biodiversity hotspot: a multi-solution aproach to conservation prioritization. Chapter 4 - in preparation.

169. Lombard AT, Cowling RM, Pressey RL, Rebelo AG. Effectiveness of land classes as surrogates for species in conservation planning for the Cape Floristic Region. Biol Conserv. 2003;112(1-2):45-62.

170. Levin N, Watson JEM, Joseph LN, Grantham HS, Hadar L, Apel N, et al. A framework for systematic conservation planning and management of Mediterranean landscapes. Biol Conserv. 2013;158(0):371-83. doi: http://dx.doi.org/10.1016/j.biocon.2012.08.032.

171. Carwardine J, Klein CJ, Wilson KA, Pressey RL, Possingham HP. Hitting the target and missing the point: target-based conservation planning in context. Conservation Letters. 2009;2(1):4-11.

172. Moilanen A, Meller L, Leppänen J, Pouzols FM, Kujala H, Arponen A. Zonation spatial conservation planning framework and software V3.1 - User manual. 2013. 288 p.

173. Arponen A, Moilanen A, Ferrier S. A successful community-level strategy for conservation prioritization. Journal of Applied Ecology. 2008;45(5):1436-45. doi: 10.1111/j.1365-2664.2008.01513.x.

174. Leathwick JR, Moilanen A, Ferrier S, Julian K. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. Biol Conserv. 2010;143(4):984-91. doi: http://dx.doi.org/10.1016/j.biocon.2010.01.012.

175. Sevilha AC, Pressey RL, VanDerWal J, Wiederhecker HC, Williams KJ, Ferrier S, et al. Biodiversity patterns within Paranã River Basin: what we can learn from species and community-level distributions. Chapter 3 - in preparation.

176. Franco AMA, Anderson BJ, Roy DB, Gillings S, Fox R, Moilanen A, et al. Surrogacy and persistence in reserve selection: landscape prioritization for multiple taxa in Britain. Journal of Applied Ecology. 2009;46(1):82-91. doi: 10.1111/j.1365-2664.2008.01598.x.

177. Moilanen A. Landscape Zonation, benefit functions and target-based planning: unifying reserve selection strategies. Biol Conserv. 2007;134(4):571-9. doi: http://dx.doi.org/10.1016/j.biocon.2006.09.008.

178. Moilanen A, Franco AMA, Early RI, Fox R, Wintle B, Thomas CD. Prioritizing multiple-use landscapes for conservation: methods for large multi-species planning problems. Proceedings of the Royal Society B: Biological Sciences. 2005;272(1575):1885-91. doi: 10.1098/rspb.2005.3164.

179. Arponen A, Heikkinen RK, Thomas CD, Moilanen A. The value of biodiversity in reserve selection: representation, species weighting, and benefit functions. Conserv Biol. 2005;19(6):2009-14. doi: 10.1111/j.1523-1739.2005.00218.x.

180. Lei nº 12.651, de 25 de Maio de 2012, (2012).

181. Ås S. Invasion of matrix species in small habitat patches. Conservation Ecology [Internet]. 1999 08/07/2016; 3(1):[12 p.]. Available from: http://www.ecologyandsociety.org/vol3/iss1/art1/inline.html.

182. Smith TB, Wayne RK, Girman DJ, Bruford MW. A role for ecotones in generating rainforest biodiversity. Science. 1997;276(5320):1855-7. doi: 10.1126/science.276.5320.1855.

183. Rodrigues ASL, Gaston KJ. How large do reserve networks need to be? Ecol Lett. 2001;4:602-9.

184. Willi Y, Van Buskirk J, Hoffmann AA. Limits to the adaptive potential of small populations. Annual Review of Ecology Evolution and Systematics. Annual Review of Ecology Evolution and Systematics. 37. Palo Alto: Annual Reviews; 2006. p. 433-58.

185. Shoo L, Hoffmann A, Garnett S, Pressey R, Williams Y, Taylor M, et al. Making decisions to conserve species under climate change. Clim Change. 2013;119(2):239-46. doi: 10.1007/s10584-013-0699-2.

186. Millar CI, Stephenson NL, Stephens SL. Climate change and forests of the future: managing in the face of uncertainty. Ecological Applications. 2007;17(8):2145-51. doi: 10.1890/06-1715.1.

187. Hannah L, Midgley GF, Millar D. Climate change-integrated conservation strategies. Global Ecol Biogeogr. 2002;11(6):485-95. PubMed PMID: ISI:000179978200006.

188. Hagerma SM, Chan KM. Climate change and biodiversity conservation: impacts, adaptation strategies and future research directions. F1000 Biol Reports. 2009;1(16). doi: 10.3410/B1-16.

189. Noss RF. Beyond Kyoto: Forest management in a time of rapid climate change. Conserv Biol. 2001;15(3):578-90. doi: 10.1046/j.1523-1739.2001.015003578.x.

190. Iverson LR, Schwartz MW, Prasad AM. Potential colonization of newly available tree-species habitat under climate change: an analysis for five eastern US species. Landscape Ecol. 2004;19:787-99.

191. Serra-Diaz JM, Franklin J, Ninyerola M, Davis FW, Syphard AD, Regan HM, et al. Bioclimatic velocity: the pace of species exposure to climate change. Diversity and Distributions. 2014;20(2):169-80. doi: 10.1111/ddi.12131.

192. Peterson AT. Future projections for Mexican faunas under global climate change scenarios. Nature. 2002;416:626-9.

193. Carroll C, Dunk JR, Moilanen A. Optimizing resiliency of reserve networks to climate change: multispecies conservation planning in the Pacific Northwest, USA. Global Change Biol. 2010;16(3):891-904. doi: 10.1111/j.1365-2486.2009.01965.x.

194. Cai X, Wang D, Zhu T, Ringler C. Assessing the regional variability of GCM simulations. Geophysical Research Letters. 2009;36(2):L02706. doi: 10.1029/2008gl036443.

195. Murphy JM, Sexton DMH, Barnett DN, Jones GS, Webb MJ, Collins M, et al. Quantification of modelling uncertainties in a large ensemble of climate change simulations. Nature. 2004;430(7001):768-72. doi: 10.1038/nature02771.

196. IPCC. Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. [Core Writing Team, Pachauri, R.K and Reisinger, A. (eds.)]. Geneva, Switzerland: IPCC; 2007. p. 104 pp.

197. Garcia RA, Cabeza M, Rahbek C, Araújo MB. Multiple dimensions of climate change and their implications for biodiversity. Science. 2014;344(6183). doi: 10.1126/science.1247579.

198. Colwell RK, Brehm G, Cardelús CL, Gilman AC, Longino JT. Global warming, elevational range shifts, and lowland biotic attrition in the wet tropics. Science. 2008;322(5899):258-61. doi: 10.1126/science.1162547.

199. VanDerWal J, Murphy HT, Kutt AS, Perkins GC, Bateman BL, Perry JJ, et al. Focus on poleward shifts in species' distribution underestimates the fingerprint of climate change. Nature Clim Change. 2013;3(3):239-43. doi: 10.1038/nclimate1688.

200. Chen I-C, Hill JK, Ohlemüller R, Roy DB, Thomas CD. Rapid range shifts of species associated with high levels of climate warming. Science. 2011;333(6045):1024-6. doi: 10.1126/science.1206432.

201. Loarie SR, Duffy PB, Hamilton H, Asner GP, Field CB, Ackerly DD. The velocity of climate change. Nature. 2009;462(7276):1052-5. doi: 10.1038/nature08649.

202. Damschen EI, Haddad NM, Orrock JL, Tewksbury JJ, Levey DJ. Corridors increase plant species richness at large scales. Science. 2006;313(5791):1284-6. doi: 10.1126/science.1130098.

203. Damschen EI, Brudvig LA, Haddad NM, Levey DJ, Orrock JL, Tewksbury JJ. The movement ecology and dynamics of plant communities in fragmented landscapes. Proceedings of the National Academy of Sciences. 2008;105(49):19078-83. doi: 10.1073/pnas.0802037105.

204. Pennington RT, Lavin M, Oliveira-Filho A. Woody plant diversity, evolution, and ecology in the tropics: perspectives from seasonally dry tropical forests. Annual Review of Ecology, Evolution, and Systematics. 2009;40(1):437-57. doi: 10.1146/annurev.ecolsys.110308.120327.

205. Garwood NC. Seed germination in a seasonal tropical forest in Panama: a community study. Ecological Monographs. 1983;53(2):159-81. doi: 10.2307/1942493.

206. Lima VVFd, Vieira DLM, Sevilha AC, Salomão AN. Germinação de espécies arbóreas de floresta estacional decidual do vale do rio Paranã em Goiás após três tipos de armazenamento por até 15 meses. Biota Neotropica. 2008;8:0-.

207. Thuiller W, Araujo MB, Pearson RG, Whittaker RJ, Brotons L, Lavorel S. Biodiversity conservation: uncertainty in predictions of extinction risk. Nature. 2004;430(6995).

208. Williams SE, Hero JM. Multiple determinants of Australian tropical frog biodiversity. Biol Conserv. 2001;98(1):1-10.

209. Hopper SD. OCBIL theory: towards an integrated understanding of the evolution, ecology and conservation of biodiversity on old, climatically buffered, infertile landscapes. Plant Soil. 2009. doi: 10.1007/s11104-009-0068-0.

210. Thuiller W. Patterns and uncertainties of species' range shifts under climate change. Global Change Biol. 2004;10(12):2020-7. doi: 10.1111/j.1365-2486.2004.00859.x.

211. Guisan A, Thuiller W. Predicting species distribution: offering more than simple habitat models. Ecol Lett. 2005;8(9):993-1009. doi: 10.1111/j.1461-0248.2005.00792.x.

212. Buisson L, Thuiller W, Casajus N, Lek S, Grenouillet G. Uncertainty in ensemble forecasting of species distribution. Global Change Biol. 2010;16(4):1145-57. doi: 10.1111/j.1365-2486.2009.02000.x.

213. Sistema Nacional de Unidades de Conservação - SNUC, (2000).

214. Watts ME, Ball IR, Stewart RS, Klein CJ, Wilson K, Steinback C, et al. Marxan with Zones: Software for optimal conservation based land- and sea-use zoning. Environmental Modelling & Software. 2009;24(12):1513-21.

215. Pressey RL, M. Watts, M. Ridges & T. Barrett. C-Plan conservation planning software. User Manual: NSW Department of Environment and Conservation; 2005.

216. Delavenne J, Metcalfe K, Smith RJ, Vaz S, Martin CS, Dupuis L, et al. Systematic conservation planning in the eastern English Channel: comparing the Marxan and Zonation decision-support tools. ICES Journal of Marine Science: Journal du Conseil. 2012;69(1):75-83. doi: 10.1093/icesjms/fsr180.

217. Allnutt TF, McClanahan TR, Andréfouët S, Baker M, Lagabrielle E, McClennen C, et al. Comparison of marine spatial planning methods in Madagascar demonstrates value of

alternative approaches. PLoS ONE. 2012;7(2):e28969. doi: 10.1371/journal.pone.0028969. PubMed PMID: PMC3281012.

218. Rogelj J, Meinshausen M, Knutti R. Global warming under old and new scenarios using IPCC climate sensitivity range estimates. Nature Clim Change. 2012;2(4):248-53. doi: 10.1038/nclimate1385.

219. MMA. Brazilian forest at a glance - Data from 2007 to 2012. Ministry of Environment. Brazilian Forest Service. Brasília2013.

220. McManus HL, Hastings DE. A framework for understanding uncertainty and its mitigation and exploitation in complex systems. Fifteenth Annual International Symposium of the International Council on Systems Engineering (INCOSE); July, 2005; Rochester, NY: INCOSE; 2005. p. 20.

221. Paté-Cornell ME. Uncertainties in risk analysis: Six levels of treatment. Reliability Engineering & System Safety. 1996;54(2-3):95-111.

222. Rapini A. Modernizando a taxonomia. Biota Neotropica. 2004;4:1-4.

223. Austin MP. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling. 2002;157(2–3):101-18. doi: http://dx.doi.org/10.1016/S0304-3800(02)00205-3.

224. Preston BL, Stafford-Smith M. Framing vulnerability and adaptive capacity assessment: Discussion paper: CSIRO Climate Adaptation Flagship Working paper No. 2.; 2009 [cited 2010 20/05/2010]. Available from: http://www.csiro.au/org/ClimateAdaptationFlagship.html

225. Hance J. Controversial changes to Brazilian forest law passes first barrier: Mongabay.com; 2010 [cited 2010 30/08/2010]. Available from: http://news.mongabay.com/2010/0708-hance_brazil_amend.html.

226. Young MD, Gunningham N, Elix J, Lamber J, Howard B, Grabosky P, et al. Reimbursing the future: an evaluation of motivational, voluntary, price-based, property-right, and regulatory incentives for the conservation of biodiversity. Canberra: Department of the Environment, Sport and Territories; 1996. p. 205; 199.

227. Decreto de 13 de Outubro de 2014. Cria a Reserva de Desenvolvimento Sustentável Nascentes Geraizeiras, localizada nos municípios de Montezuma, Rio Pardo de Minas e Vargem Grande do Rio Pardo, Estado de Minas Gerais, (2014).

228. EMBRAPA, cartographer O novo mapa de solos do Brasil. Legenda atualizada. Escala 1:5000000. (http://sosgisbr.com/2012/06/04/mapa-de-solos-do-brasil-2011-embrapa/). Brasilia: Embrapa. Ministério da Agricultura e do Abastecimento; 2011.

229. EMBRAPA-SNLCS. Levantamento de reconhecimento de média intensidade dos solos e avaliação de aptidão agrícola das margens direita do Rio Paranã, Estado de Goiás. Empresa Brasileira de Pesquisa Agropecuária. Serviço Nacional de Levantamento e Conservação dos Solos. Rio de Janeiro. Boletim de Pesquisa. 1983;(23):503.

230. Santos HGd, Jacomine PKT, Anjos LHCd, Oliveira VÁd, Lumbreras JF, Coelho MR, et al. Sistema Brasileiro de Classificação dos Solos. 3rd ed. Brasília, DF: Embrapa; 2013. 353 p.

231. Correia JR, Burle ML, Calderano SB, Spera ST, Gomes IA, Santos RDd, et al. Caracterização de ambientes na Chapada dos Veadeiros / Vale do Rio Paranã: uma contribuição para a classificação brasileira de solos. Embrapa Cerrados. Embrapa Solos. Documentos. 2001;(47):79.

232. Eiten G. Formas fisionômicas do Cerrado. Revta Brasil Bot. 1979;2:139-48.

The end

“... *Nestes últimos anos nada de novo há no rugir das tempestades.*
Não estamos alegres, é certo, mas também por que razão haveríamos de ficar tristes?
O mar da história é agitado.
As ameaças e as guerras havemos de atravessá-las, rompê-las ao meio, cortando-as como uma quilha corta as ondas” (Vladimir Mayakovsky, 1927).