TOWARDS IMPROVING BIODIVERSITY ASSESSMENTS FOR MAMMALS AND BIRDS WITHIN THE NEOTROPICS

By

Beth E. Gerstner

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Fisheries and Wildlife – Doctor of Philosophy
Ecology, Evolutionary Biology, and Behavior – Dual Major

2023

ABSTRACT

Global change is greatly affecting biodiversity and the vital ecosystem functions and services it provides. Biodiversity hotspots, areas with high levels of diversity that are under strong anthropogenic pressures, are especially prevalent in the tropics, making them conservation priorities. However, these areas often lack comprehensive data on species spatial distributions and functional traits, hindering our ability to perform biodiversity assessments and assess species and ecosystem vulnerability. Species distribution models (SDMs) can help delineate species ranges and inform conservation decisions. Yet, SDMs often solely focus on local environmental factors and overlook filtering processes occurring at broader-scales. Further, conservation efforts have mainly emphasized taxonomic diversity (TD), assuming equal species importance, but a more comprehensive approach should consider the diversity of roles species play in ecosystems –functional diversity (FD). To develop robust FD metrics across broad geographic regions, we need improved trait data that fills gaps and harmonizes existing datasets. Overcoming these challenges demands improved methods, data availability, and embracing multiple dimensions of biodiversity.

To address the need to consider FD alongside TD, and the issue of data paucity in the Neotropics, I compiled an open-access functional trait dataset, containing ecological, life-history, morphological, and geographical traits for frugivorous birds and mammals within Neotropical moist forests called *Frugivoria* (Chapter 1). This open and accessible dataset and workflow enables practitioners and researchers to investigate patterns of FD for taxa performing the essential ecosystem function of seed dispersal.

To enhance the capacity of SDMs to more fully capture the distributions of species and to better incorporate environmental filtering processes that influence species occurrence, I

generated scale-dependent geodiversity variables (Chapter 2) and incorporated them into SDMs for Colombian mammals. Models incorporating geodiversity variables generally performed better than those constructed without. This finding demonstrates the utility of geodiversity variables for generating robust geographic estimates useful for species risk assessments.

In Chapter 3, I aimed to develop a more comprehensive perspective on diversity that extends beyond traditional TD and considers the roles of species in sustaining ecosystem function. To achieve this, I used the *Frugivoria* dataset to explore spatial alignments and mismatches between TD and FD, the extent to which protected areas (PAs) encompassed TD and FD, and assessed the distribution of Functionally Unique, Specialized, and Endangered (FUSE) species. The results identified strong spatial mismatches and few areas of alignment between dimensions of diversity for mammals and birds, with only ~30% protected. Further, many areas with high proportions of FUSE species remain unprotected. FD was also found to be higher in community-managed lands. This finding emphasizes the importance of including community-managed lands in strategies aimed at achieving biodiversity targets.

The concluding chapter discusses potential ways to use *Frugivoria* data to inform conservation practice and policy. It further discusses how to apply the framework for testing geodiversity variables to improve SDMs to other regions and taxa. Lastly, it highlights priorities for conservation when considering the spatial tradeoffs of TD, FD, and FUSE species distributions. By integrating these different elements, this dissertation improves our ability to understand the distribution of bird and mammal biodiversity and will therefore help generate better-informed species risk assessments and set more inclusive spatial conservation priorities within the Neotropics.

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to the following undergraduate students who provided invaluable assistance with data entry for the first chapter of my work: Arpita Nayak, Minali Bhatt, Erik Ralston, Olivia Melville, Taylor Stehouwer, Aaron Guggenheimer, Ashley Ezzo, Chakata Hart, Caroline Blommel, Alison McCormick, Kate Kelly, Nellie Chalem, and Emma Mushaka. Your dedication and time significantly enriched the quality of this dataset and allowed me to carry out the bulk of my dissertation.

There are so many others who also contributed to this dissertation. I want to extend my special thanks to Patrick Bills for his invaluable support in streamlining and implementing workflows across multiple chapters, a task that greatly contributed to the efficiency and rigor of this project. I also want to acknowledge the guidance and advice provided by Laura Twardochleb, Jamie Kass, and Casey Youngflesh at various stages of this work. Your insights were invaluable in shaping the direction and scope of this research. I am particularly grateful to my committee member Mary Blair and collaborator Elkin Noguera-Urbano for their expert guidance and support in navigating the complexities of utilizing BioModelos data for my second chapter.

I am deeply grateful for the financial support that made this research possible. Funding for this work was provided by NASA FINESST Grant #80NSSC19K1332, the National Science Foundation (CC*Compute: #2200792), and NASA grant #80NSSC18K0406. I would also like to acknowledge the university for their financial support through the Kellogg Biological Station Academic Programs Fellowship, the University Enrichment Fellowship, and the MSU EEB Digital Fellowship. This work was supported in part through computational resources and services provided by the Institute for Cyber-Enabled Research at Michigan State University. Their assistance was instrumental in carrying out the computational aspects of this research.

I would also like to extend my sincere gratitude to all those who have played a supporting role in bringing this dissertation to fruition. First and foremost, I want to express my deep appreciation to my friends and family for their unwavering support and encouragement throughout this journey. Your belief in me has been a driving force behind my accomplishments.

I am truly grateful to my dedicated advisor, committee members Kyla Dahlin and Robert Montgomery, MSU Spatial and Community Ecology (SpaCE) Lab mates, and colleagues for their insightful guidance and constructive feedback during my academic journey. Your expertise has significantly shaped the trajectory of my research. I specifically want to acknowledge my advisor, Phoebe Zarnetske, for her continued mentorship, guidance, and encouragement. Her knowledge and commitment to my academic growth have been invaluable and I will always be grateful for her taking me under her wing all those years ago.

A special mention goes to my husband, Andrew, whose constant belief in my abilities and infinite patience during this process have been invaluable. Your unwavering support has kept me motivated even during the most challenging times.

Finally, I can't forget to thank my loyal canine companion, Ande. Your presence provided much-needed moments of levity and comfort throughout my years in graduate school.

To everyone who has contributed to my academic journey, whether through advice, encouragement, or simply being there, your impact is immeasurable. Thank you from the bottom of my heart.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AOH	Area of Habitat
AOO	Area of Occupancy
API	Application Programming Interface
AUC	Area Under the Curve
CART	Categorical and Regression Trees
CBI	Continuous Boyce Index
DD	Data Deficient
EDI	Environmental Data Initiative
EOO	Extent of Occurrence
FD	Functional Diversity
Fdis	Functional Dispersion
FII	Forest Integrity Index
Fric	Functional Richness
FUSE	Functional Unique, Specialized, and Endangered
GBF	Global Biodiversity Framework
GBIF	Global Biodiversity Information Facility
IUCN	International Union for the Conservation of Nature
IPBES	Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services
LQ	Linear Quadratic
MAD	Mean Absolute Deviation
MaxEnt	Maximum Entropy

MTP	Minimum Training Presence
OECMs	Other Effective Conservation Measures
PA	Protected Area
PAM	Presence/Absence Matrix
PCoA	Principle Coordinates Analysis
PMM	Predictive Mean Matching
RF	Random Forest
SDM	Species Distribution Model
SQ	Root Mean Square Roughness
SRTM	Shuttle Radar Topography Mission
TD	Taxonomic Diversity
WCS	Wildlife Conservation Society

INTRODUCTION

Evaluating broad-scale biodiversity patterns is often a first step towards identifying priority areas for conservation efforts (Margules & Pressey, 2000). During the 2022 Convention on Biodiversity in Montreal, nations committed to the mission of safeguarding 30% of terrestrial and marine landscapes under the aegis of the United Nations' Global Biodiversity Framework, leading an essential question to take center stage: 'Which 30% should we protect?'. This question is especially important to consider for regions such as the tropics, which host a large share of global biodiversity while under strong anthropogenic pressures such as habitat loss and climate change (Myers et al., 2000; Bellard et al., 2014; IPCC, 2022). Among these tropical ecosystems, forests cover a mere 5% of Earth's surface yet encapsulate almost half of the world's species (Brandon, 2014). Despite this high species richness, knowledge of the spatial dynamics of this and other dimensions of biodiversity has been and remains incomplete. This is due limitations in the datasets and tools allowing us to quantify and measure this diversity (Collen et al., 2008), making it difficult to set spatial conservation priorities.

This dearth of information on biodiversity in the tropics is especially apparent in the availability of species geographic ranges and knowledge of the traits that allow them to persist in those environments (Collen et al. 2008). These well-known gaps (sometimes termed “shortfalls”; Hortal et al., 2015) are essential to fill for improving fundamental understanding of species’ biogeography and to inform more robust conservation assessments and decisions. With emerging tools and methodologies, such as global repositories for biodiversity data (e.g., Global Biodiversity Information Facility [GBIF, 2023]; Vertnet [Guralnick & Constable, 2010] etc.), software to make modeling species distributions and calculation of biodiversity change metrics easier for conservation practitioners (Kass et al., 2018; Merow et al. 2022; Galante et al., 2023), and the increasing availability of species trait datasets (e.g., through digitization of museum specimens and online repositories; Nelson & Ellis, 2019) – these gaps are beginning to narrow.

The understanding of species distributions across the vast spectrum of Earth's species suffers from what is known as the Wallacean shortfall (Hortal et al. 2015). This shortfall indicates that even among formally described species, comprehension of their global, regional, or local geographic ranges remains incomplete (Whittaker et al., 2005; Brito, 2010). This lack of information on species ranges is particularly salient for species inhabiting tropical regions due to factors such as limited sampling stemming from true species rarity, geopolitical instability, challenges in accessing sampling sites, and funding constraints (Collen et al., 2008; Engemann et al., 2015). This challenge becomes especially important given the utility of species range estimates in spatial prioritization of conservation efforts. The International Union for the Conservation of Nature (IUCN), a preeminent global authority on species extinction risk, considers metrics related to species distributions as a factor in assessing species risk (e.g., extent of occurrence [EOO], and area of occupancy [AOO]) (IUCN, 2022). Thus, ensuring reliable estimates of species ranges is of paramount importance in setting spatial conservation priorities.

Species distribution models (SDMs) employ statistical methods to relate species occurrence data to multiple environmental variables and generate predictions about species-environment relationships (Peterson, 2011; Franklin, 2013). They are the most widely used modeling tool for understanding the effects of global change on biodiversity (Ferrier, 2002; Guisan et al., 2013), enabling the delineation of species' ranges in support of future sampling efforts, management, and conservation prioritization (Franklin, 2010). Though often used as a single species approach via the process of overlaying predictions from multiple species, it is possible to generate predictions of species richness and endemism (Zurell et al., 2020). Both of these applications provide detailed spatial insights that are integral for making informed decisions regarding conservation prioritization and reserve design (Taylor et al., 2017). Nonetheless, the effectiveness of SDMs in spatial conservation prioritization hinges on addressing methodological challenges that influence model performance and the resulting range estimates. These challenges encompass issues such as biases in the data collected from species occurrences (Syfert et al., 2013; Boria et al., 2014), the careful delineation of appropriate study extent (Anderson & Raza, 2010; Vander Wall & Beck, 2012), model evaluation (Anderson & Gonzalez, 2011; Radosavljevic & Anderson, 2014), consideration of how models should be applied to new areas or time periods (Guevara et al., 2018; Charney et al., 2021; Helmstetter et al., 2021), and selection of relevant environmental variables for analysis (Austin & Van Niel, 2011; Guevara et al. 2018). Neglecting these challenges could lead to less reliable estimates of species distributions, and ultimately misguided allocation of limited conservation resources. However, given its status as a burgeoning field of study, novel methods, techniques, and standards are continually being developed to effectively apply SDMs (Araújo et al., 2019; Merow et al., 2019; Zurell et al., 2020).

Traditionally, SDMs only incorporate environmental variables at the local scale where occurrences are recorded and therefore only reflect local environmental conditions. However, the occurrence of a species is intricately linked to the species' realized niche, which is a result of environmental filtering operating across multiple scales beyond the local occurrence point (e.g., dispersal limitations, habitat configuration, climatic variations, and biotic interactions) (Kraft et al., 2015; Pearson et al., 2018). A promising way to capture these broad scale filtering processes is through the utilization of "geodiversity", which refers to the variation of abiotic features and processes present on Earth's surface (Parks & Mulligan, 2010; Lawler et al., 2015; Zarnetske et al., 2019). It is known that regions characterized by high geodiversity likely exhibit increased levels of biodiversity due to the heterogeneity and variety of habitats in these areas (Lawler et al., 2015; Bailey et al., 2017; Tukiainen, 2019)—potentially shaping biodiversity patterns and the distributions of species (Dufour et al., 2006). Geodiverse regions are expected to serve as refuges for species as the climate changes, leading conservationists to prioritize these areas as a strategic approach for safeguarding biodiversity in a dynamic climate (Reynard & Brilha, 2018; Schrodt et al., 2019; Crofts, 2022). Existing research has predominantly concentrated on quantifying the relationship between geodiversity and biodiversity, particularly in terms of species richness (Stein et al., 2014; Bailey et al., 2017; Zarnetske et al., 2019). However, the exact relationship between geodiversity and the distributions of individual species remains relatively unexplored (Bailey et al. 2017). Geodiversity variables can capture these filtering processes and therefore can further explain species-environment relationships. By testing the utility of geodiversity variables for improving SDMs, we can close this gap in knowledge and potentially improve geographic estimates for species in regions of high heterogeneity.

Focusing on species distributions offers direct practicality for spatial prioritization for both single species and scaling up to species richness; however, this is just one dimension of biodiversity. Prevailing biodiversity and conservation assessments have focused on the sheer number of species in a given location (Belmaker & Jetz, 2013; Engemann et al., 2015; Durán et al., 2016), often sidelining other facets of diversity that could yield more insightful indications of irreplaceability and maintenance of ecosystem health (Cadotte et al., 2011; Pollock et al., 2017), for example, phylogenetic diversity (i.e., evolutionary relatedness; Faith et al., 2004) and functional diversity (i.e., diversity of species roles with importance for maintaining ecosystem functioning; Petchey & Gaston, 2006).

Analysis of different dimensions of biodiversity adds a nuanced understanding of diversity in a given region, however no one dimension should be used as a proxy of another. In terms of functional diversity, it has been suggested by some studies that phylogenetic diversity is a sufficient proxy because of the relationship between shared evolutionary history and traits (Winter et al., 2013; Mazel et al., 2018). However, this is problematic because phylogenetic diversity does not explicitly reflect species' ecological strategies or their distinct ecological roles and captures functional diversity unreliably (Mazel et al., 2018; Cooke et al., 2020). Similarly, taxonomic diversity is not always a sufficient proxy of functional diversity. While the presence of a greater number of species often implies a wider range of traits, it's important to recognize that two communities with an equal number of species can still exhibit different levels of functional diversity (Borges et al., 2021). This variation can be attributed to a concept known as functional redundancy, where species within a community are more similar to each other in terms of their functional traits than would be expected by chance (Mouchet et al., 2010). This phenomenon can arise due to different environmental filters influencing the relationship between species and their functional traits (Mouchet et al., 2010). Because species' responses to environmental changes are largely influenced by their specific traits, there is a growing recognition of the significance of considering species' functional traits as a critical component of biodiversity (Petchey & Gaston, 2006; Devictor et al., 2010; Brum et al., 2017; González-Maya et al., 2017; Gomez et al., 2021). This perspective underscores the idea that assessing functional diversity, which accounts for the unique traits and roles of species within a community, provides a more comprehensive understanding of how ecosystems respond to environmental changes and maintain their functioning. It's not just the number of species that matters but also the specific functional traits they possess and how those traits interact with the environment.

The global community, represented by collaborations such as the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES; IPBES, 2022), is increasingly acknowledging the significance of alternative biodiversity metrics, including functional and phylogenetic diversity. Multiple studies have demonstrated that different dimensions of diversity can have very different spatial patterns (Brum et al. 2017; Devictor et al. 2010; Pollock et al. 2017; Borges et al. 2021). As such, relying solely on one dimension of diversity to guide conservation strategies should be avoided, as this one-dimensional approach may not provide a comprehensive view of diversity (Brooks et al., 2006; Flynn et al., 2009; Devictor et al., 2010). Given the limitations in conservation resources—not least of which are funding and time—further research is required to explore potential trade-offs between prioritization approaches concerning species with important roles for ecosystem functioning (Devictor et al. 2010; Mazel et al. 2018). Spatial discrepancies between various dimensions of diversity carry significant implications for the management and design of protected areas (Cumming & Child, 2009).

Despite the increased attention to multiple dimensions of diversity, there are scarce, if any, practical instances of functional diversity being explicitly incorporated into spatial conservation planning within the literature (but see Herrera, 2017). Further, IUCN documentation on identifying Key Biodiversity Areas (KBAs)–areas identified as sites contributing to global persistence of biodiversity–acknowledges the importance of taxonomic diversity as a criterion for identifying these areas but does not explicitly consider other dimensions like phylogenetic and functional diversity (IUCN, 2016). The significance of functional diversity within conservation is underscored by the fact that species' responses to environmental changes are influenced by their traits. Thus, the inclusion of functional diversity is pivotal for gauging sensitivity to future changes (Newbold et al., 2014). This is particularly important for biodiversity hotspots, as they are often data poor and sensitive to changes (Bellard et al., 2014). The limited application of functional diversity in conservation stems largely from the demanding data prerequisites for its computation (Cadotte & Tucker, 2018). The calculation necessitates complete datasets of species traits and how they relate to ecosystem function. These data are often absent for numerous taxa, a gap known as the Raunkiæran shortfall (Hortal et al., 2015), or are scattered across diverse sources, rendering their aggregation and quantification challenging (Cadotte & Tucker, 2018). These challenges necessitate increased availability of trait data for use in studies of functional diversity, particularly for biodiversity hotspots such as in Neotropical forests.

This dissertation aims to enhance biodiversity assessments by contributing new open and accessible data for species traits, by improving SDMs through incorporating geodiversity predictors to capture environmental heterogeneity, and by evaluating the spatial mismatches and alignments among taxonomic and functional diversity to understand potential biodiversity trade-offs relevant for the conservation of mammals and birds. Chapter 1 contributes an open-access functional trait dataset (*Frugivoria*) for frugivorous birds and mammals within Neotropical moist forests. This chapter fills major data gaps and provides an open workflow, enabling investigations into patterns of functional diversity for taxa performing the essential ecosystem service of seed dispersal. Chapter 2 provides the first comprehensive assessment of incorporating scale-dependent geodiversity variables into SDMs through a case study with Colombian mammals. Chapter 3 leverages *Frugivoria* to quantify spatial alignments and mismatches between taxonomic and functional diversity for frugivorous birds and mammals within the Tropical Andes and evaluates the effectiveness of current protected areas and community-managed areas in capturing this diversity. By integrating these different components, this dissertation contributes data and quantitative assessments to improve understanding of multidimensional biodiversity patterns, and to help foster well-informed and efficacious conservation strategies tailored to capturing this diversity within the Neotropics.

REFERENCES

Anderson, R.P. & Gonzalez, I. (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. Ecological Modelling, 222, 2796–2811.

Anderson, R.P. & Raza, A. (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus Nephelomys) in Venezuela. Journal of Biogeography, 37, 1378–1393.

Araújo, M.B., Anderson, R.P., Barbosa, A.M., Beale, C.M., Dormann, C.F., Early, R., Garcia, R.A., Guisan, A., Maiorano, L., Naimi, B., O'Hara, R.B., Zimmermann, N.E. & Rahbek, C. (2019) Standards for distribution models in biodiversity assessments. Science Advances, 5, eaat4858.

Austin, M.P. & Van Niel, K.P. (2011) Improving species distribution models for climate change studies: variable selection and scale. Journal of Biogeography, 38, 1–8.

Bailey, J.J., Boyd, D.S., Hjort, J., Lavers, C.P. & Field, R. (2017) Modelling native and alien vascular plant species richness: At which scales is geodiversity most relevant? Global Ecology and Biogeography, 26, 763–776.

Bellard, C., Leclerc, C., Leroy, B., Bakkenes, M., Veloz, S., Thuiller, W. & Courchamp, F. (2014) Vulnerability of biodiversity hotspots to global change. Global Ecology and Biogeography, 23, 1376–1386.

Belmaker, J. & Jetz, W. (2013) Spatial scaling of functional structure in bird and mammal assemblages. The American Naturalist, 181, 464–478.

Borges, F.J.A., Fortunato, D. de S. & Loyola, R. (2021) Critical areas for retaining multiple dimensions of bird diversity in the Cerrado. Journal for Nature Conservation, 64, 126079.

Boria, R.A., Olson, L.E., Goodman, S.M. & Anderson, R.P. (2014) Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. Ecological Modelling, 275, 73–77.

Brandon, K. (2014) Ecosystem Services from Tropical Forests: Review of Current Science, Social Science Research Network, Rochester, NY.

Brito, D. (2010) Overcoming the Linnean shortfall: Data deficiency and biological survey priorities. Basic and Applied Ecology, 11, 709–713.

Brooks, T.M., Mittermeier, R.A., Fonseca, G.A.B. da, Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D. & Rodrigues, A.S.L. (2006) Global Biodiversity Conservation Priorities. Science, 313, 58–61.

Brum, F.T., Graham, C.H., Costa, G.C., Hedges, S.B., Penone, C., Radeloff, V.C., Rondinini, C., Loyola, R. & Davidson, A.D. (2017) Global priorities for conservation across multiple dimensions of mammalian diversity. Proceedings of the National Academy of Sciences, 114, 7641–7646.

Cadotte, M.W., Carscadden, K. & Mirotchnick, N. (2011) Beyond species: functional diversity and the maintenance of ecological processes and services. Journal of Applied Ecology, 48, 1079–1087.

Cadotte, M.W. & Tucker, C.M. (2018) Difficult decisions: Strategies for conservation prioritization when taxonomic, phylogenetic and functional diversity are not spatially congruent. Biological Conservation, 225, 128–133.

Charney, N.D., Record, S., Gerstner, B.E., Merow, C., Zarnetske, P.L. & Enquist, B.J. (2021) A Test of Species Distribution Model Transferability Across Environmental and Geographic Space for 108 Western North American Tree Species. Frontiers in Ecology and Evolution, 9.

Collen, B., Ram, M., Zamin, T. & McRae, L. (2008) The Tropical Biodiversity Data Gap: Addressing Disparity in Global Monitoring. Tropical Conservation Science, 1, 75–88.

Cooke, R.S.C., Eigenbrod, F. & Bates, A.E. (2020) Ecological distinctiveness of birds and mammals at the global scale. Global Ecology and Conservation, 22, e00970.

Crofts, R. (2022) Progress and future challenges for geoconservation in protected and conserved areas. Parks Stewardship Forum, 38.

Cumming, G.S. & Child, M.F. (2009) Contrasting spatial patterns of taxonomic and functional richness offer insights into potential loss of ecosystem services. Philosophical Transactions of the Royal Society B: Biological Sciences, 364, 1683–1692.

Devictor, V., Mouillot, D., Meynard, C., Jiguet, F., Thuiller, W. & Mouquet, N. (2010) Spatial mismatch and congruence between taxonomic, phylogenetic and functional diversity: the need for integrative conservation strategies in a changing world. Ecology Letters, 13, 1030–1040.

Dufour, A., Gadallah, F., Wagner, H.H., Guisan, A. & Buttler, A. (2006) Plant species richness and environmental heterogeneity in a mountain landscape: effects of variability and spatial configuration. Ecography, 29, 573–584.

Durán, A.P., Inger, R., Cantú-Salazar, L. & Gaston, K.J. (2016) Species richness representation within protected areas is associated with multiple interacting spatial features. Diversity and Distributions, 22, 300–308.

Engemann, K., Enquist, B.J., Sandel, B., Boyle, B., Jørgensen, P.M., Morueta-Holme, N., Peet, R.K., Violle, C. & Svenning, J.-C. (2015) Limited sampling hampers "big data" estimation of species richness in a tropical biodiversity hotspot. Ecology and Evolution, 5, 807–820.

Faith, D.P., Reid, C. a. M. & Hunter, J. (2004) Integrating Phylogenetic Diversity, Complementarity, and Endemism for Conservation Assessment. Conservation Biology, 18, 255–261.

Ferrier, S. (2002) Mapping Spatial Pattern in Biodiversity for Regional Conservation Planning: Where to from Here? Systematic Biology, 51, 331–363.

Flynn, D.F.B., Gogol-Prokurat, M., Nogeire, T., Molinari, N., Richers, B.T., Lin, B.B., Simpson, N., Mayfield, M.M. & DeClerck, F. (2009) Loss of functional diversity under land use intensification across multiple taxa. Ecology Letters, 12, 22–33.

Franklin, J. (2010) Moving beyond static species distribution models in support of conservation biogeography. Diversity and Distributions, 16, 321–330.

Franklin, J. (2013) Species distribution models in conservation biogeography: developments and challenges. Diversity & Distributions, 19, 1217–1223.

Galante, P.J., Chang Triguero, S., Paz, A., Aiello-Lammens, M., Gerstner, B.E., Johnson, B.A., Kass, J.M., Merow, C., Noguera-Urbano, E.A., Pinilla-Buitrago, G.E. & Blair, M.E. (2023) changeRangeR: An R package for reproducible biodiversity change metrics from species distribution estimates. Conservation Science and Practice, 5, e12863.

GBIF. (2023) The Global Biodiversity Information Facility. What is GBIF?. Available from https://www.gbif.org/what-is-gbif.

Gómez, C., Tenorio, E.A. & Cadena, C.D. (2021). Change in avian functional fingerprints of a Neotropical montane forest over 100 years as an indicator of ecosystem integrity. Conservation Biology, 35, 1552–1563

González-Maya, J.F., Martínez-Meyer, E., Medellín, R. & Ceballos, G. (2017) Distribution of mammal functional diversity in the Neotropical realm: Influence of land-use and extinction risk. PLOS ONE, 12, e0175931.

Guevara, L., Gerstner, B.E., Kass, J.M. & Anderson, R.P. (2018) Toward ecologically realistic predictions of species distributions: A cross-time example from tropical montane cloud forests. Global Change Biology, 24, 1511–1522.

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H.P. & Buckley, Y.M. (2013) Predicting species distributions for conservation decisions. Ecology Letters, 16, 1424–1435.

Guralnick, R. & Constable, H. (2010) VertNet: Creating a Data-sharing Community. BioScience, 60, 258–259.

Helmstetter, N.A., Conway, C.J., Stevens, B.S. & Goldberg, A.R. (2021) Balancing transferability and complexity of species distribution models for rare species conservation. Diversity and Distributions, 27, 95–108.

Herrera, J.P. (2017) Prioritizing protected areas in Madagascar for lemur diversity using a multidimensional perspective. Biological Conservation, 207, 1–8.

Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. & Ladle, R.J. (2015) Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity. Annual Review of Ecology, Evolution, and Systematics, 46, 523–549.

IPBES (2022). Thematic Assessment Report on the Sustainable Use of Wild Species of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Fromentin, J. M., Emery, M. R., Donaldson, J., Danner, M. C., Hallosserie, A., and Kieling, D. (eds.). IPBES secretariat, Bonn, Germany.

IPCC, 2022: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press. Cambridge University Press, Cambridge, UK and New York, NY, USA, 3056 pp., doi:10.1017/9781009325844.

IUCN. (2022) The IUCN Red List of Threatened Species. Version 2022-2. https://www.iucnredlist.org.

IUCN (2016). A Global Standard for the Identification of Key Biodiversity Areas, Version 1.0 . First edition. Gland, Switzerland: IUCN.

Kass, J.M., Vilela, B., Aiello-Lammens, M.E., Muscarella, R., Merow, C. & Anderson, R.P. (2018) Wallace: A flexible platform for reproducible modeling of species niches and distributions built for community expansion. Methods in Ecology and Evolution, 9, 1151–1156.

Kraft, N.J.B., Adler, P.B., Godoy, O., James, E.C., Fuller, S. & Levine, J.M. (2015) Community assembly, coexistence and the environmental filtering metaphor. Functional Ecology, 29, 592–599.

Lawler, J.J., Ackerly, D.D., Albano, C.M., Anderson, M.G., Dobrowski, S.Z., Gill, J.L., Heller, N.E., Pressey, R.L., Sanderson, E.W. & Weiss, S.B. (2015) The theory behind, and the challenges of, conserving nature's stage in a time of rapid change. Conservation Biology, 29, 618–629.

Margules, C. & Pressey, R. (2000) Systematic Conservation Planning. Nature, 405, 243–253.

Mazel, F., Pennell, M.W., Cadotte, M.W., Diaz, S., Dalla Riva, G.V., Grenyer, R., Leprieur, F., Mooers, A.O., Mouillot, D., Tucker, C.M. & Pearse, W.D. (2018) Prioritizing phylogenetic diversity captures functional diversity unreliably. Nature Communications, 9, 1–9.

Merow, C., Maitner, B.S., Owens, H.L., Kass, J.M., Enquist, B.J., Jetz, W. & Guralnick, R. (2019) Species' range model metadata standards: RMMS. Global Ecology and Biogeography, 28, 1912–1924.

Merow, C., Galante, P.J., Kass, J.M., Aiello-Lammens, M.E., Babich Morrow, C., Gerstner, B.E., Grisales Betancur, V., Moore, A.C., Noguera-Urbano, E.A., Pinilla-Buitrago, G.E., Velásquez-Tibatá, J., Anderson, R.P. & Blair, M.E. (2022) Operationalizing expert knowledge in species' range estimates using diverse data types. Frontiers of Biogeography, 0.

Mouchet, M.A., Villéger, S., Mason, N.W.H. & Mouillot, D. (2010) Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. Functional Ecology, 24, 867–876.

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

Nelson, G. & Ellis, S. (2019) The history and impact of digitization and digital data mobilization on biodiversity research. Philosophical Transactions of the Royal Society B: Biological Sciences, 374, 20170391.

Newbold, T., Scharlemann, J.P.W., Butchart, S.H.M., Şekercioğlu, Ç.H., Joppa, L., Alkemade, R. & Purves, D.W. (2014) Functional traits, land-use change and the structure of present and future bird communities in tropical forests. Global Ecology and Biogeography, 23, 1073–1084.

Parks, K.E. & Mulligan, M. (2010) On the relationship between a resource based measure of geodiversity and broad scale biodiversity patterns. Biodiversity and Conservation, 19, 2751–2766.

Pearson, D.E., Ortega, Y.K., Eren, Ö. & Hierro, J.L. (2018) Community Assembly Theory as a Framework for Biological Invasions. Trends in Ecology & Evolution, 33, 313–325.

Petchey, O.L. & Gaston, K.J. (2006) Functional diversity: back to basics and looking forward. Ecology Letters, 9, 741–758.

Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of evidence: Ecological niche conservatism. Journal of Biogeography, 38, 817–827.

Pollock, L.J., Thuiller, W. & Jetz, W. (2017) Large conservation gains possible for global biodiversity facets. Nature, 546, 141–144.

Radosavljevic, A. & Anderson, R.P. (2014) Making better Maxent models of species distributions: complexity, overfitting and evaluation. Journal of Biogeography, 41, 629–643.

Brilha, J. & Reynard, E. (2018) Geoheritage and Geoconservation: The Challenges. Geoheritage, pp. 433–438.

Schrodt, F., Bailey, J.J., Kissling, W.D., Rijsdijk, K.F., Seijmonsbergen, A.C., Ree, D. van, Hjort, J., Lawley, R.S., Williams, C.N., Anderson, M.G., Beier, P., Beukering, P. van, Boyd, D.S., Brilha, J., Carcavilla, L., Dahlin, K.M., Gill, J.C., Gordon, J.E., Gray, M., Grundy, M., Hunter, M.L., Lawler, J.J., Monge-Ganuzas, M., Royse, K.R., Stewart, I., Record, S., Turner, W., Zarnetske, P.L. & Field, R. (2019) Opinion: To advance sustainable stewardship, we must document not only biodiversity but geodiversity. Proceedings of the National Academy of Sciences, 116, 16155–16158.

Stein, A., Gerstner, K. & Kreft, H. (2014) Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. Ecology Letters, 17, 866–880.

Syfert, M.M., Smith, M.J. & Coomes, D.A. (2013) The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. PLOS ONE, 8, e55158.

Taylor, C., Cadenhead, N., Lindenmayer, D.B. & Wintle, B.A. (2017) Improving the Design of a Conservation Reserve for a Critically Endangered Species. PLOS ONE, 12, e0169629.

Tukiainen, H. (2019) Multi-scale relationship between geodiversity and biodiversity across high-latitude environments: implications for nature conservation. Nordia Geographical Publications, 48, 1–54.

Vander Wall, S.B. & Beck, M.J. (2012) A Comparison of Frugivory and Scatter-Hoarding Seed-Dispersal Syndromes. The Botanical Review, 78, 10–31.

Whittaker, R.J., Araújo, M.B., Jepson, P., Ladle, R.J., Watson, J.E.M. & Willis, K.J. (2005) Conservation Biogeography: assessment and prospect. Diversity and Distributions, 11, 3–23.

Winter, M., Devictor, V. & Schweiger, O. (2013) Phylogenetic diversity and nature conservation: where are we? Trends in Ecology & Evolution, 28, 199–204.

Zarnetske, P.L., Read, Q.D., Record, S., Gaddis, K.D., Pau, S., Hobi, M.L., Malone, S.L., Costanza, J., Dahlin, K.M., Latimer, A.M., Wilson, A.M., Grady, J.M., Ollinger, S.V. & Finley, A.O. (2019) Towards connecting biodiversity and geodiversity across scales with satellite remote sensing. Global Ecology and Biogeography, 28, 548–556.

Zurell, D., Franklin, J., König, C., Bouchet, P.J., Dormann, C.F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., Guisan, A., Lahoz-Monfort, J.J., Leitão, P.J., Park, D.S., Peterson, A.T., Rapacciuolo, G., Schmatz, D.R., Schröder, B., Serra-Diaz, J.M., Thuiller, W., Yates, K.L., Zimmermann, N.E. & Merow, C. (2020) A standard protocol for reporting species distribution models. Ecography, 43, 1261–1277.

CHAPTER 1:

FRUGIVORIA: A TRAIT DATABASE FOR BIRDS AND MAMMALS EXHIBITING FRUGIVORY ACROSS CONTIGUOUS NEOTROPICAL MOIST FORESTS

Published in Global Ecology & Biogeography (2023): https://doi.org/10.1111/geb.13716

Co-authors include: Phoebe L. Zarnetske and Patrick Bills

1.1 Abstract

Biodiversity in many areas is rapidly declining because of global change. As such, there is an urgent need for new tools and strategies to help identify, monitor, and conserve biodiversity hotspots. This is especially true for frugivores, species consuming fruit, because of their important role in seed dispersal and maintenance of forest structure and health. One way to identify these areas is by quantifying functional diversity, which measures the unique roles of species within a community and is valuable for conservation because of its relationship with ecosystem functioning. Unfortunately, the functional trait information required for these studies can be sparse for certain taxa and specific traits and difficult to harmonize across disparate data sources, especially in biodiversity hotspots. To help fill this need, we compiled Frugivoria, a trait database containing ecological, life-history, morphological and geographical traits for mammals and birds exhibiting frugivory. Frugivoria encompasses species in contiguous moist montane forests and adjacent moist lowland forests of Central and South America—the latter specifically focusing on the Andean states. Compared with existing trait databases, Frugivoria harmonizes existing trait databases, adds new traits, extends traits originally only available for mammals to birds also and fills gaps in trait categories from other databases. Furthermore, we create a cross-taxa subset of shared traits to aid in analysis of mammals and birds. In total, Frugivoria adds 8662 new trait values for mammals and 14,999 for birds and includes a total of 45,216 trait

entries with only 11.37% being imputed. Frugivoria also contains an open workflow that harmonizes trait and taxonomic data from disparate sources and enables users to analyze traits in space. As such, this open-access database, which aligns with FAIR data principles, fills a major knowledge gap, enabling more comprehensive trait-based studies of species in this ecologically important region.

1.2 Introduction

In a time of rapid global change and significant declines in biodiversity, there is an urgent need to identify, monitor and conserve biodiversity hotspots—regions of high endemic biodiversity with significant anthropogenic pressures (Myers et al., 2000). When identifying biodiversity hotspots, the level of biodiversity is often defined as endemic species richness (Myers et al., 2000). However, focusing on the number of endemic species assumes that each species plays an equal role in the environment and does not represent a holistic view of biodiversity (Devictor et al., 2010; Pollock et al., 2017). Functional traits of species are increasingly being considered in conservation with a focus on conserving functional diversity (i.e., the unique roles and functions of species in an ecosystem; Cadotte et al., 2011; Devictor et al., 2010; González-Maya et al., 2017; Gómez et al., 2021). Conserving areas with higher functional diversity—where species exhibit many functional roles—can help maintain ecosystem functioning (Cadotte et al., 2011; Cadotte & Tucker, 2018). Therefore, a functional trait approach to conservation is promising because species' responses to environmental change largely depend on their traits and incorporating functional diversity can help assess sensitivity to future changes (Newbold et al., 2014). When combined, functional diversity, taxonomic diversity, and phylogenetic diversity provide an even more comprehensive assessment of an area's biodiversity, which is especially important for biodiversity hotspots such as in the Neotropics (Devictor et al., 2010; Pollock et al., 2017).

Neotropical forests contain numerous biodiversity hotspots yet lack sufficient data on species occurrences and traits (Collen et al., 2008), making it difficult to generate robust species habitat maps and forecasts of biodiversity change and species vulnerability. Despite tropical forests covering less than 5% of Earth's surface (Brandon, 2014), they contain almost half of the world's biodiversity (Dinerstein et al., 2017) and many of their species are considered data deficient (DD) by the International Union for the Conservation of Nature (IUCN; IUCN, 2022a). These DD species do not have enough information on population status, spatial distribution or threats, or combinations therein, to make an official threat status designation (IUCN, 2022a). For example, as of 2021, 15.6% of mammals, 0.15% of birds, 14.3% of reptiles, and 16.8% of amphibians are DD within the tropical forests of Central and South America; the percentages for this biodiversity hotspot are higher than the global forest percentages for both mammals and amphibians (mammal: 14.4%, bird: 0.47%, reptiles: 14.4%, amphibians: 16.5%) (IUCN, 2021). This data paucity is especially pressing in montane cloud forests—regions high in elevation that are frequented by cloud cover—because of their unique environment for many endemic species and vulnerability to climate change (Foster, 2001; Ponce-Reyes et al., 2012). Biodiversity in these regions is especially vulnerable to climate change because cloud forests exist in a narrow elevational band and are naturally fragmented (Foster, 2001). This pattern of fragmentation limits dispersal, and further, the lagging tree line shift (i.e., the upper elevational range of cloud forests) will make it difficult for species to shift their distributions to track changing climates (Fricke et al., 2022; Rehm & Feeley, 2015). Importantly, the distribution and community composition of forests in the tropics are largely determined by seed dispersal processes that are mediated by frugivores, fauna consuming fruit (Sales et al., 2021).

Seed dispersal is considered a key biotic interaction for maintaining biodiversity-ecosystem function, especially in tropical regions where almost 90% of woody plants rely on frugivores to disperse their seeds (Howe & Smallwood, 1982; Sales et al., 2021). For example, there have been multiple studies on the crucial role of frugivorous mammals and birds in dispersing seeds and how changes in their abundance and distribution influence seed dispersal services and the ability of plants to track changes in climate (Fricke et al., 2022; Sales et al., 2021). In addition to dispersing seeds for many plants, frugivores also contribute to nutrient cycling and are prey for carnivores such as felids and raptors (Farwig & Berens, 2012; Tóbon et al., 2004; Wilkie et al., 2011). Due to the strong interaction between frugivores and many fruiting plants, the trait composition of frugivore assemblages (e.g., gape size, body size) can likewise affect plant community composition (i.e., prevalence of plants with small or large seeds), and ecological resilience to perturbations (Sales et al., 2021). Comprehensive trait databases will help quantify the outsized roles that frugivorous species play in maintaining biodiversity and ecosystem functions for conservation and macroecological studies.

Over the last 20 years, there has been significant growth in the use of vouchered data from specimens housed in museum collections (Nelson & Ellis, 2019). Large quantities of data have been stored in these physical repositories for much longer; however, access to these global institutions has not always been feasible (Nelson & Ellis, 2019). Recently, increased efforts and funding have made it possible to digitize these collections for scientific research purposes (Miralles et al., 2020; Nelson & Ellis, 2019). In particular, national and international funding over the last two decades has facilitated more open access data through institutional websites and databases (e.g., GBIF; GBIF, 2023, VertNet; Guralnick & Constable, 2010; Nelson & Ellis, 2019). This mobilization and digitization of once inaccessible data has opened a vast resource for studies in conservation, ecology, and systematics, allowing researchers to perform biodiversity analyses without having to conduct costly field research or work in museums. This digitization has increased the openness and accessibility of trait data through online repositories and downloadable datasets (e.g., EltonTraits; Wilman et al., 2014, PanTHERIA; Jones et al., 2009; VertNet; Guralnick & Constable, 2010; COMBINE; Soria et al., 2021). However, existing databases only provide a subset of taxa and their traits. The lack of comprehensive trait databases is due to multiple reasons. First, compiling detailed ecological trait data for species rich clades is time-consuming. Second, actively maintaining databases is costly and time-consuming; changing taxonomies and newly discovered species require constant updating. Third, data gaps in existing databases can result from bias towards regions that are more easily accessed (Engemann et al., 2015), leading to the exclusion of more remote and rare species. As a result, tropical regions in particular have significant data gaps (Collen et al., 2008; Ferrier, 2002). Filling these data gaps is especially important in biodiversity hotspot regions which are often data poor and more sensitive to changes (Bellard et al., 2014).

To help fill the aforementioned data and knowledge gaps and aid in the use of species traits for use in conservation analyses (Cooke et al., 2019; González-Maya et al., 2017; González-Suárez & Revilla, 2013), we created a trait database, ‘Frugivoria’. Frugivoria is a comprehensive trait database of species exhibiting frugivory in the Neotropics of Central America and the Andean states of South America within habitat designated ‘Forest—Subtropical/Tropical Moist Montane’ and/or ‘Forest—Subtropical/Tropical Moist Lowland’ by the IUCN. These data encompass Neotropical contiguous tropical moist montane forests and their adjacent lowland forests. Traits in Frugivoria encompass any morphological, phenological, physiological, and behavioral characteristics of a species (Kissling et al., 2018), and we further extend this definition to key geographical traits relevant for species extinction risk assessments. Frugivoria harmonizes bird and mammal trait data and scientific names from existing databases and online species accounts (e.g., PanTHERIA (Jones et al., 2009), EltonTraits (Wilman et al., 2014), AnAge (Tacutu et al., 2018), Encyclopedia of Life (Parr et al., 2014), Cornell's Birds of the World (Billerman et al., 2022), University of Michigan's Animal Diversity Web (Myers et al., 2023), IUCN (IUCN, 2022a)) and adds 23,661 new records for 1732 species including information related to diet breadth, habitat breadth, habitat specialization, and adds new species and their associated traits as well as IUCN threat statuses. Furthermore, once only available to mammals, Frugivoria updates and expands select traits from the PanTHERIA dataset to birds. These traits include derived breadth traits (diet and habitat breadth) and range-based geographical traits (e.g., mean annual temperature, mean annual precipitation, human impacts across the range for multiple years), using more recent and relevant sources (e.g., CHELSA bioclimatic variables, WCS Human Footprint data) (Karger et al., 2017, 2018; Sanderson et al., 2022). The inclusion of these newly derived traits expands the utility of cross-taxa trait comparisons for mammals and birds.

Creating a reproducible workflow that aligns with FAIR (Findable, Accessible, Interoperable and Reusable) data principles (Wilkinson et al., 2016) is crucial for enabling transparent and efficient data sharing and reuse. I used these principles to generate Frugivoria, and therefore provide code for all aspects of database building, which can facilitate the creation of similar datasets for other taxa or different diet preferences. Furthermore, we provide code to explore and summarize the trait data and link associated traits with species occurrences. Many of the species within Frugivoria exist in other databases, particularly for mammals (e.g., EltonTraits and PanTHERIA), yet existing databases are not entirely overlapping, collectively they omit many frugivorous Neotropical species and traits, and certain geographical traits are outdated. In addition, there is no existing comprehensive trait database for frugivorous birds in this region. Together, Frugivoria and its open workflow in R fill essential data gaps for biodiversity conservation.

1.3 Methods

1.3.1 Frugivoria overview

Frugivoria encompasses frugivorous species of the classes 'Mammalia' and 'Aves' in contiguous moist montane forests and adjacent moist lowland forests of Central and South America—the latter specifically focusing on the Andean states (i.e., Mexico, Guatemala, Costa Rica, Panama, El Salvador, Belize, Nicaragua, Ecuador, Colombia, Peru, Bolivia, Argentina, Venezuela and Chile).

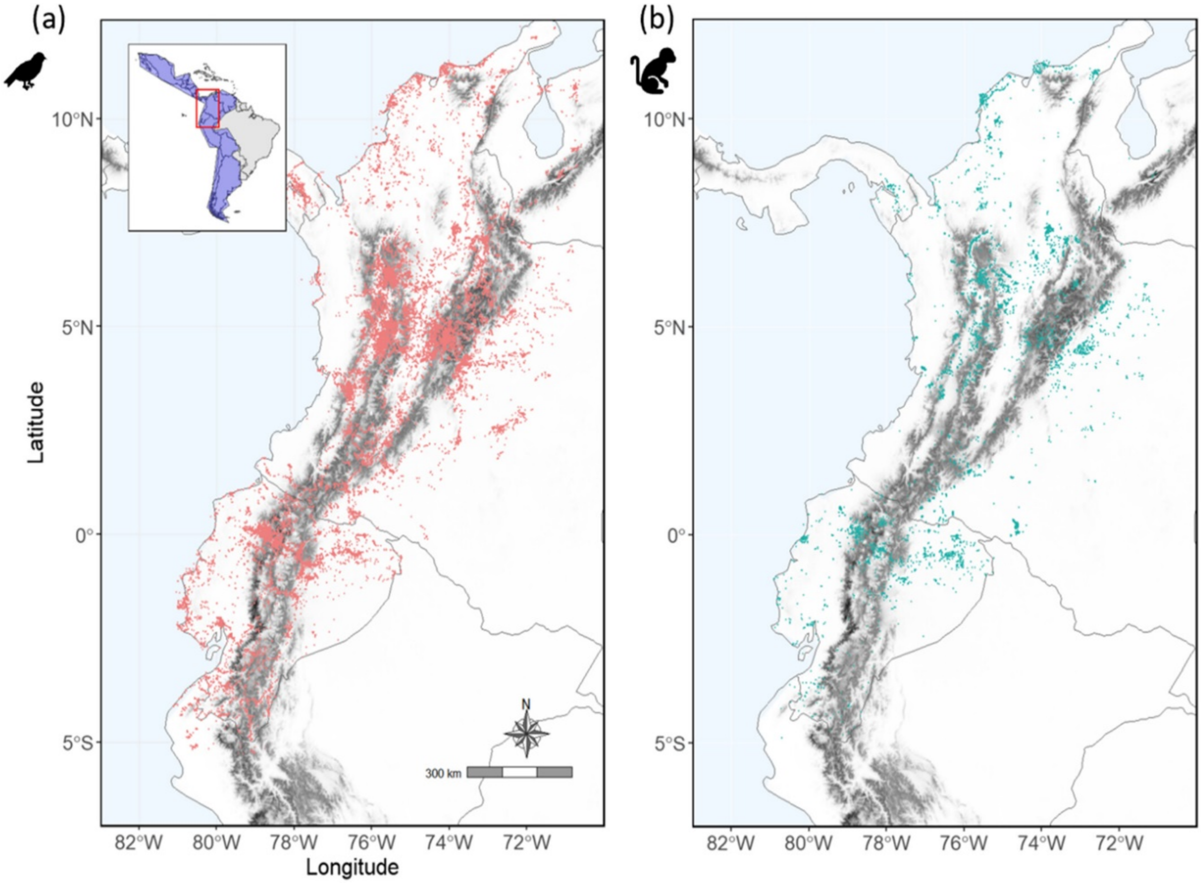
The species list for this region was obtained through the IUCN API (IUCN, 2015) and subset to species with IUCN habitat designations of 'Forest—Subtropical/Tropical Moist Montane' and/or 'Forest—Subtropical/Tropical Moist Lowland' (Figure 1.1).

I chose to use the IUCN as the basis for the database species list because their database includes the most current taxonomic information, has an open access inventory of all known species present per country that have been officially assessed and is the most appropriate species list for conservation applications.

Figure 1.1: Inset map—Distribution of all Frugivoria data (purple) with a red box indicating the zoomed in region presented here. Panels show maps of occurrence records of frugivorous bird (panel a) and mammal (panel b) species in the greater Colombia and Ecuador region. Records were obtained through GBIF (GBIF, 2022a; GBIF, 2022b; GBIF, 2022c; GBIF, 2022d) and pulled for differing time frames for viewing purposes (birds: 2021–2022; mammals: 2016–2022). Darker

Figure 1.1 (cont'd)

background areas show higher probability of cloud forest presence (moist montane forest; Wilson & Jetz, 2016).



1.3.2 Existing databases

Databases composing a significant portion of Frugivoria include the IUCN (IUCN, 2022a), EltonTraits (Wilman et al., 2014) and PanTHERIA (Jones et al., 2009; Figure 2a), with supplemental information coming from the AnAge database (Tacutu et al., 2018). The IUCN Red List of Threatened Species, which formally assesses the conservation status of species, provides a large online database of formal assessments for many animals, including information on taxonomy, habitat, life history, and threats, and serves as a powerful tool for biodiversity conservation and policy change (IUCN, 2022a). EltonTraits is a global database of activity patterns and feeding habits of mammals and birds (Wilman et al., 2014). PanTHERIA consists of morphological, life-history, and geographical traits of mammals globally (Jones et al., 2009). In some instances, data permitting, longevity information for birds and mammals were obtained through the AnAge database when primary sources were unavailable (Tacutu et al., 2018). For a comprehensive list of all traits included in Frugivoria, their column names, definitions and sources, see Table 1.1.

Since the publication of all these data sources, new studies and genetic analyses have led to taxonomic revisions for many genera, which were standardized to IUCN taxonomic classifications in Frugivoria during the harmonization process (described below in Frugivoria Workflow part 2: Harmonizing).

Table 1.1: Frugivoria traits, their associated column name in the database, their definitions, units, and sources. Traits lacking a source are those specific to the Frugivoria dataset, or minimum and maximum values of the original trait definition. Some traits are only available for mammals (*), only available for birds (**). Traits included in cross-taxa subsets are denoted with (†).

GROUPING FE	TRAIT	COLUMN NAME	DESCRIPTION	UNIT	CITATION
LIFE HISTORY	Longevity	longevity†	Average lifespan; if only range is available this value indicates the maximum longevity.	years	
		min_longevity	Minimum longevity found if range was available or differing values from multiple sources.	years	
		max_longevity	Maximum longevity found if range was available or differing values from multiple sources.	years	

Table 1.1 (cont’d)

	max_longevity_p*	Maximum adult age measured.	months	Jones et al. 2009
Generation time	generation_time†	The average age of parents of the current cohort (i.e., newborn individuals); if only a range is available, this value indicates the maximum generation time.	years	IUCN 2022
	min_generation_time	The minimum age of parents of the current cohort (i.e., newborn individuals) if range was available or there were differing values from multiple sources.	years	IUCN 2022
	max_generation_time	The maximum age of parents of the current cohort (i.e., newborn individuals).	years	IUCN 2022
Age at eye opening	age_at_eye_opening_d_p*	Age at which both eyes are fully open after birth; measure of central tendency.	days	Jones et al. 2009
Age at first birth	age_at_first_birth_d_p*	Age at which females give birth to their first litter (eutherians), or their young attach to teats (metatherians) or hatch out (monotremes); measure of central tendency.	days	Jones et al. 2009
Gestation length	gestation_len_d_p*	Length of time of non-inactive fetal growth; all measures of central tendency.	days	Jones et al. 2009
Interbirth interval	inter_birth_interval_d*	The length of time between successive births of the same female(s) after a successful or unspecified litter; all measures of central tendency.	days	Jones et al. 2009
Litter size	litter_size_p*	Number of offspring born per litter per female, either counted before birth, at birth or after birth; all measures of central tendency.		Jones et al. 2009

Table 1.1 (cont'd)

	Litters per year	litters_per_year_p*	Number of litters per female per year; all measures of central tendency.		Jones et al. 2009
	Sexual maturity age	sexual_maturity_age_d_p*	Age when individuals are first physically capable of reproducing, defined as either physically sexually mature, age at first mating or unspecified (males and females), age at first estrus or age at first pregnancy (females only), age at spermatogenesis or age at testes descent (males only), using captive, wild, provisioned, or unspecified populations; all measures of central tendency.	days	Jones et al. 2009
	Weaning age	weaning_age_d_p*	Age when primary nutritional dependency on the mother ends and independent foraging begins to make a major contribution to the offspring's energy requirements; all measures of central tendency.	days	Jones et al. 2009
ECOLOGY	% Composition diet	diet_inv_e, diet_vend_e, diet_vect_e, diet_vfish_e, diet_vunk_e, diet_scav_e diet_fruit_e, diet_nect_e, diet_seed_e, diet_plant_e	% Prevalence of foraging for: invertebrates, endotherms, ectotherms, fish,vertebrates-general or unknown, scavenging, fruit, nectar, seeds, plants.	%	Wilman et al. 2014
	Diet category	diet_cat_e†, diet_cat†	Overall diet category (assigned using same method as in Wilman et al. 2014). Assigned to the dominant diet categories based on the percent composition of diet for each species: plant and seeds, fruits and nectar, invertebrates, and vertebrates, fish, and carrion; omnivore: score of <= 50% in each category (Wilman et al. 2014).		Wilman et al. 2014

Table 1.1 (cont'd)

		trophic_level_p*	Trophic level of each species; species were defined as (1) herbivore (not vertebrate and/or invertebrate), (2) omnivore (vertebrate and/or invertebrate plus any of the other categories) and (3) carnivore (vertebrate and/or invertebrate only); all measures of central tendency.		Jones et al 2009
	Diet breadth	diet_breadth_p*	Number of dietary categories eaten by each species; categories were defined as vertebrate, invertebrate, fruit, flowers/nectar/pollen, leaves/branches/bark, seeds, grass and roots/tubers; all measures of central tendency.		Jones et al. 2009
z		diet_breadth†	Diet breadth was calculated using the Shannon Diversity index and is based on EltonTraits % composition of diet columns. This metric takes into account the number of different types of food consumed, as well as the relative abundance of each type.		Wilman et al. 2014
	Forest stratum	for_strat_value*†	Assigned forest strata value; marine, ground level, including aquatic foraging, scansorial, arboreal, or aerial.		Wilman et al. 2014
			Prevalence of: Foraging below on ground, understory, mid-height, canopy, and in flight.		

Table 1.1 (cont'd)

		for_strat_ground,**† for_strat_understory**†, for_strat_midheight**†, for_strat_canopy**†, forstrait_aerial**†			Wilman et al. 2014
		terrestriality_p*	Degree of terrestriality of each species; species were defined as fossorial and/or ground dwelling only and above ground dwelling; all measures of central tendency.		Jones et al. 2009
t	Activity patterns	activity_nocturnal_e† activity_crepuscular_e* activity_diurnal_e*	Presence of foraging activity at night, at dawn and dusk, and during the day.		Wilman et al. 2014
		activity_cycle_p*	Activity cycle for wild populations: nocturnal only, nocturnal/crepuscular, cathemeral, crepuscular or diurnal/crepuscular and diurnal only. Measures of central tendency.		Jones et al. 2009
	Habitat specialization	habitat_specialization†	Descriptions of species habitat mention cloud forest as an important or sole component.		
	Habitat category	habitat† habitat_suitability_lowland habitat_suitability_montane habitat_season_lowland habitat_season_montane habitat_major_importance_lowland habitat_major_importance_montane	Degree of suitability, seasonality, and importance of habitat type to species as designated by IUCN		IUCN 2022
	Habitat breadth	habitat_breadth_p*	Number of habitat layers used by each species. Categories were defined as above ground dwelling, aquatic, fossorial and ground dwelling; all measures of central tendency.		Jones et al. 2009

Table 1.1 (cont'd)

		habitat_breadth†	Number of habitat layers inhabited by each species according to the IUCN. The total number of different habitats inhabited by species within this dataset is 42. See the "habitats_all_species.csv" for the full list of possible habitat types for each species.		
	Population statistics	population_density_n.km2_p*	Number of individuals per square kilometer; all measures of central tendency.		Jones et al. 2009
		population_grp_size_p*	Number of individuals, adults or definition unspecified in a group that spends the majority of their time in a 24 hour cycle together; all measures of central tendency.		Jones et al. 2009
		social_grp_size_p*	Number of individuals, adults or definition unspecified in a group that spends the majority of their time in a 24 hour cycle together where there is some indication that these individuals form a social cohesive unit; all measures of central tendency.		Jones et al. 2009
MORPHOLOGY	Body mass	body_mass_g†	Average mass of animal.	grams	Wilman et al. 2014
		adult_body_mass_g_p*	Mass of adult; measures of central tendency.	grams	Jones et al. 2009
		basal_met_rate_mass_g_p*	Mass of individual(s) from which the basal metabolic rate was taken (see metabolic trait below).	grams	Jones et al. 2009
		neonate_body_mass_g_p*	Mass of live or freshly killed specimens of infants at either a near term embryonic stage, birth, immediately after birth or up to an age of seven days after birth; all measures of central tendency.	grams	Jones et al. 2009
		weaning_body_mass_g_p*	Mass of live or freshly-killed specimens of weanlings; all measures of central tendency.	grams	Jones et al. 2009

Table 1.1 (cont'd)

	Body size	body_size_mm†	Length of animal (mammals: nose to base of tail; birds: beak to end of tail). This value indicates the mean body size unless a range of values are available. In the latter case, this value indicates the maximum body size.	mm	Ansell, 1965
		min_body_size_mm	Minimum body length (mammals: nose to base of tail; birds: beak to end of tail) if range was available or there were differing values from multiple sources.	mm	
		max_body_size_mm	Maximum body length (mammals: nose to base of tail; birds: beak to end of tail) if range was available or there were differing values from multiple sources.	mm	
		adult_forearm_len_mm_p*	Total length from elbow to wrist of adult; measures of central tendency.	mm	Jones et al. 2009
		adult_head_body_len_mm_p*	Total length from tip of nose to anus or base of tail; measures of central tendency.	mm	Jones et al. 2009
		weaning_head_body_len_mm_p*	Total length from tip of nose to anus or base of tail of live or freshly-killed specimens of weanlings.	mm	Jones et al. 2009
	Sexual dimorphism	sexual_dimorphism†	Body pattern differs between males and females.		
	Teat number	teat_number_p*	Total number of teats present; all measures of central tendency.		Jones et al. 2009
METABOLIC	Basal metabolic rate	basal_met_rate_mLO2hr_p*	Basal metabolic rate of adult; measures of central tendency.	mL.O2/hr	Jones et al. 2009

Table 1.1 (cont'd)

GEOGRAPHIC	Home range size	home_range_size†	Area traversed by an individual in its normal activities of food gathering, mating, and caring for young; if only range is available, this value indicates the maximum home range size.	Km2	Burt, 1943
		min_home_range	Minimum area traversed by an individual in its normal activities of food gathering, mating, and caring for young; if range was available or there were differing values from multiple sources.	Km2	
		max_home_range	Maximum area traversed by an individual in its normal activities of food gathering, mating, and caring for young; if range was available or there were differing values from multiple sources.	Km2	
		home_range_km2_p*	Size of the area within which everyday activities of individuals or groups (of any type) are typically restricted; all measures of central tendency.	Km2	Jones et al. 2009
		home_range_indiv_km2_p*	Size of the area within which everyday activities of individuals are typically restricted; all measures of central tendency (see Jones et al. 2009 for details).	Km2	Jones et al. 2009
	Range size	observed_range_size†	Range size calculated from IUCN range maps and projected using a global equal-area projection; derived only from parts of the range with the presence code "Extant". This is the range size of areas the species is known to occur. See Workflow section 4: Range Size.	Km2	IUCN 2022
		inferred_range_size†	Range size calculated from IUCN range maps and projected using a global equal-area projection; derived from parts of the range using presence codes "Extant", "Probably Extant" ,and "Possibly Extant". This range size incorporates inferred areas where the species may exist and represents potentially suitable habitat rather than just the known range of the species. See Workflow section 4: Range Size.	Km2	IUCN 2022
		GR_area_km2_p*	Geographic range size of species based on digital maps from Sechrest (2003). Range areas were calculated using ArcGIS; projected using a global equal-area projection.	Km2	Jones et al. 2009

Table 1.1 (cont'd)

Range extent	GR_max_lat_dd_p*, GR_min_lat_dd_p*, GR_mid_range_lat_dd_p*, GR_max_long_dd_p*, GR_min_long_dd_p*, GR_mid_range_long_dd_p*	Extent of each species range (as of 2003) calculated using a global geographic projection (decimal degrees).	Decimal degrees	Jones et al. 2009
Human population density	hu_pop_den_min_n.km2_p_OD*, hu_pop_den_mean_n.km2_p_OD*, hu_pop_den_5p_n.km2_p_OD*	Minimum, mean and 5th percentile human population density (persons per km2) over the species range (as of 2003) using the Gridded Population of the World (GPW) (CIESIN and CIAT 2005).	n/km2	Jones et al. 2009
	hu_pop_den_change_p_OD	Mean rate of increase in human population density over the species range (as of 2003) using the Gridded Population of the World (GPW) (CIESIN and CIAT 2005) for 1990 and 1995 as: (1995–1990)/1990.		Jones et al. 2009
Human footprint	mean_human_fp_range_2010†, mean_human_fp_range_2020†	Average human footprint based on the WCS Human Footprint dataset for 2010 and 2020; human influence is mapped using the weighted sum of individual maps of population density, infrastructure such as roads, railways, factories, accessibility, use of electricity; https://wcshumanfootprint.org/.		Sanderson et al. 2022
	percent_change_hf_2010_2020†	Percent change in average human footprint calculated between the years 2010 - 2020	%	
Average precipitation	precip_mean_mm_p_OD*	Mean monthly precipitation (mm) of species range as of 2003.	mm	Jones et al. 2009
Average precipitation	mean_CHELSA_bio12_1981.2010_V.2.1†	Average annual mean precipitation from the CHELSA bioclimatic dataset calculated over the IUCN range map for each species. Derived only from parts of the range with the presence code "Extant" or "Possibly Extant".	mm	Karger et al. 2017a; Karger et al. 2017b
Average temperature	temp_mean_01degC_p_OD*	Mean monthly temperature (0.1°C) of species range as of as of 2003.	°C	Jones et al. 2009

Table 1.1 (cont’d)

	Average temperature	mean_CHELSA_bio1_1981.2010_V.2.1†	Average annual mean temperature from the CHELSA bioclimatic dataset calculated over the IUCN range map for each species. Derived only from parts of the range with the presence code “Extant” or "Possibly Extant".	°C	Karger et al. 2017a; Karger et al. 2017b
	AET mean	AET_mean_mm_p_OD*	Mean monthly AET (Actual Evapotranspiration Rate) of species range (as of 2003) from 1920 to 1980 (mm). Calculated using the Global Resource Information Database of UNEP.	mm	Jones et al. 2009
‘	PET mean	PET_mean_mm_p_OD*	Mean monthly PET (Potential Evapotranspiration Rate) over species range (as of 2003), from 1920 to 1980 (mm) calculated using the Global Resource Information Database of UNEP.	mm	Jones et al. 2009

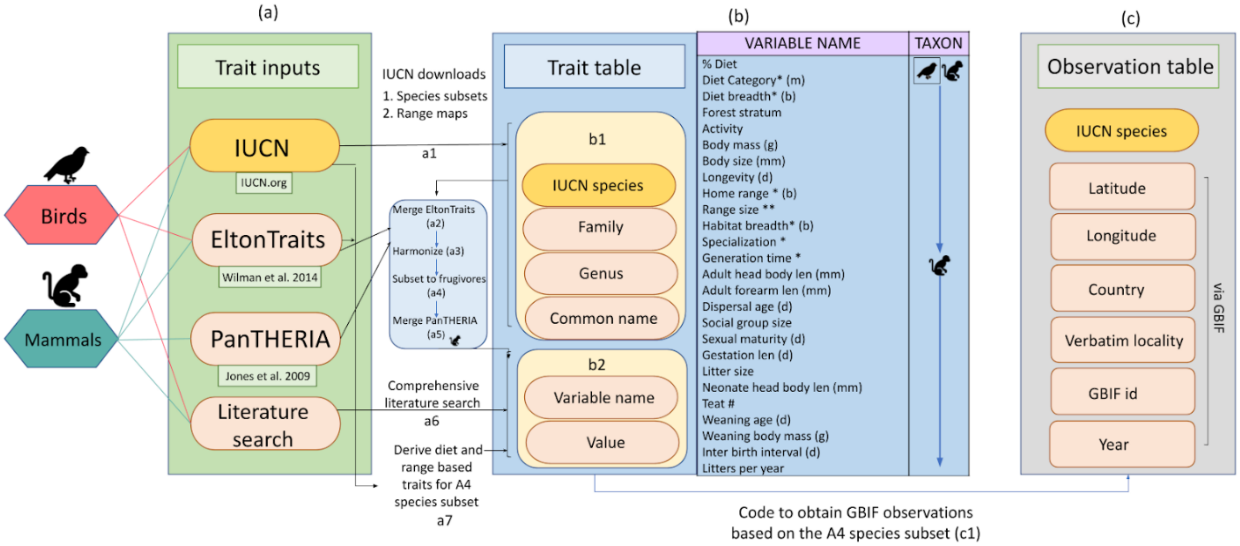
1.3.3 Frugivoria Workflow

I outlined a workflow (Figure 2) for creating and replicating the Frugivoria database and to help with the organization of future databases with similar aims to this one. This workflow delineates each step of the database building process, including trait data sources, code for extracting and harmonizing the data, and highlighting the end products.

Figure 1.2: Workflow diagram of Frugivoria database generation for frugivorous birds and mammals in Central and South American moist forest. An IUCN based species list was first downloaded for regions of interest (a1: 1) along with IUCN range maps for birds and mammals of the world (a1: 2). The species list was then subset to mammals and birds in “Subtropical/Tropical Moist Montane” and/or “Subtropical/Tropical Moist Lowland” habitat (a1; L0: 1_IUCN_species_list_subset.R). This IUCN species list was then merged with EltonTraits (Wilman *et al.*, 2014) and species names were harmonized between databases (a2 & a3; L0: 2_external_trait_database_merge.R). Once these databases were merged, they were subset to only those species eating at least 10% fruit (a4; L0: 3_frugivore_subset.R). This final subset was then merged with the PanTHERIA dataset (Jones et al., 2009; a5; L0: 4_mammal_merge_pantheria.R). Traits found in the literature and other credible species accounts (a6), PanTHERIA, and EltonTraits were combined for mammals, whereas only EltonTraits and traits from the literature and species accounts were combined for birds. I then derived traits such as diet breadth from the EltonTraits dataset, habitat breadth from the IUCN habitat designations, and geographic traits based on the IUCN range maps (a1:2) for the a4

Figure 1.2 (cont'd)

species subset (a7). Code to obtain observations per species is also provided and allows for spatial trait analyses (Panel c; c1; L2: downloading_gbif_records.R).



Each step in the data workflow to construct Frugivoria was scripted with the R language (R Core Team, 2022) to maximize reproducibility. The data in Frugivoria are stored as comma separated data tables and R scripts, to convert raw data (Level 0) to harmonized data (Levels 1, 2) (Figure 1.2). The steps to generate Frugivoria are numbered below with reference to lettered sections within Figure 1.2 and specific scripts in the workflow.

Subsetting—I downloaded static trait datasets for mammals and birds, EltonTraits (Wilman et al., 2014) and PanTHERIA (Jones et al., 2009). I used the R package 'rredlist' (version 0.6.0) and the function 'rl_sp_country', which downloads IUCN Red List information over an API, to obtain IUCN species lists for each country of interest as defined above (IUCN, 2021). These lists were then subset to the classes 'Aves' and 'Mammalia' and further filtered to include only those that include an IUCN habitat designation of 'Forest—Subtropical/Tropical Moist Montane' and/or 'Forest—Subtropical/Tropical Moist Montane' (Figure 1.2a.1; L0:

1_IUCN_species_list_subset.R). I then merged this IUCN based species list with EltonTraits for mammals and birds (Figure 1.2a.2; L0: 2_external_trait_database_merge.R). After harmonizing and subsetting the dataset to include frugivorous species only (Figure 1.2a.3 & a4; L0: 3_frugivore_subset.R), PanTHERIA was merged with the mammal database (Figure 1.2a.5; L0:4_mammal_merge_pantheria.R). I then filled in new traits using information from existing literature and credible online sources (Figure 1.2a.6), and provided code to obtain occurrences from GBIF for all species within the final database (Figure 1.2c.1; L2_downloading_gbif_records.R).

Harmonizing—Disparities in naming conventions between the IUCN, EltonTraits (as mentioned previously) and PanTHERIA led to some species not aligning correctly between databases and perpetuated the need for harmonization (Figure 1.2a.2; L0_ 2_external_trait_database_merge.R). To harmonize scientific names between databases, we first merged the final IUCN species list by scientific name with EltonTraits. I resolved conflicts in species name merges between databases using known synonyms of the IUCN species names that matched species names in the EltonTraits database. In general, any species name disparities were set aside for manual checking of synonyms using the IUCN website, existing literature and Avibase (for birds; Lepage et al., 2014), and later appended to the dataset. I generated a lookup table, hosted in the Environmental Data Initiative (EDI) repository, showing the corresponding species name in each database for all mismatched species (mammals: n = 390; birds: n = 873; lookup_table_all_mammals.csv; lookup_table_all_birds.csv), which should help facilitate the construction of databases for other species (not only frugivorous species) in this region. In Frugivoria, there were name disparities for 171 mammal and 195 bird species in the EltonTraits database. New species discovered since the publication of these databases (mammals: n = 42, birds: n = 2) and those species that had been split (taxonomically reclassified; mammals: n = 132, birds: n = 182) were assigned values of sister taxa, based on known phylogenies, or at the genus level, which was simple in EltonTraits since in some cases entire genera had matching trait values. For mammals, we then merged our final database with PanTHERIA (Jones et al., 2009; L0: 4_mammal_merge_pantheria.R) and resolved taxonomic naming issues in the same way as above, resolving the same number of species names.

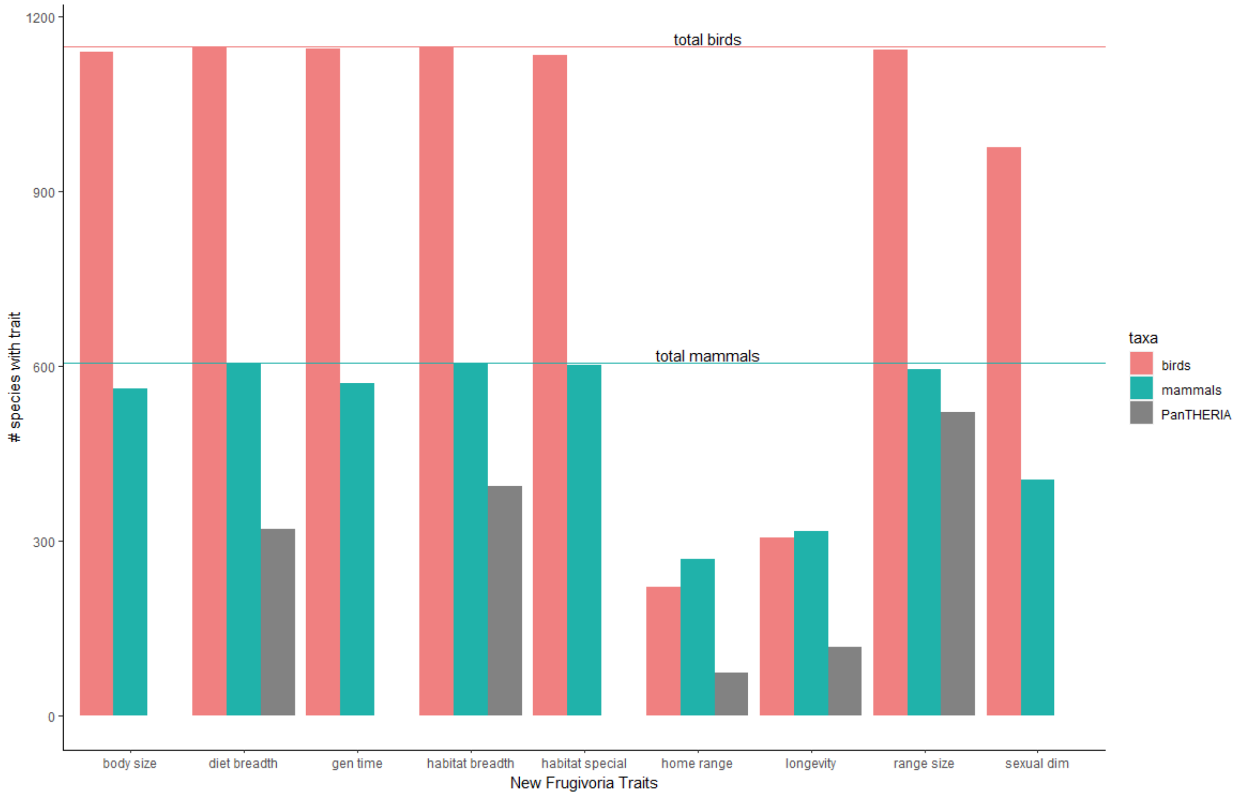
I have included a 'taxonomic_disparity' column where a value of '1' indicates differences in taxonomic nomenclature between databases for some species. For each database merged in Frugivoria, I have retained the original database species name with a suffix '_e' (EltonTraits) and '_p' (PanTHERIA) as a unique column for ease of interpretation. These suffixes are also used for any trait originating from those respective databases.

Frugivore Selection—Once the IUCN species list and EltonTraits dataset were harmonized and successfully merged for 946 mammals and 2818 birds, I used EltonTraits 'diet_fruit' to select and retain species that had a frugivorous diet at or above 10%, leaving a total of 160 mammal genera (586 species and 16 subspecies) and 329 bird genera (1148 species) (L0: 3_frugivore_subset.R).

Trait data—I then merged our frugivorous species subset and the PanTHERIA dataset for mammals (L0: 4_mammal_merge_pantheria.R). New traits and missing traits values that did not occur in existing openly accessible trait databases were then obtained by exhaustively identifying missing species and traits from the IUCN database, peer-reviewed literature, online reference material such as species accounts Encyclopedia of Life (Parr et al., 2014); University of Michigan's Animal Diversity Web (Myers et al., 2023); Cornell Lab of Ornithology's Birds of the World (Billerman et al., 2022), and reference books (Burton & Robert, 1974; Emmons &

Feer, 1997; Eisenberg & Redford, 1999; Thorington et al., 2012; Fleagle, 2013 etc.; for a full list of materials see 'Frugivoria: Sources …' in the EDI hosted dataset). In total, Frugivoria contains 45,216 trait inputs with some traits having minimum and maximum values; 24,494 trait values for mammals (8662 newly added), and 20,722 for birds (14,999 newly added) (Figure 1.3).

Figure 1.3: Newly added traits in Frugivoria for birds (red) and mammals (teal). Horizontal lines indicate the total number of taxa for each taxonomic group within the database (birds: $n = 1148$, mammals: $n = 602$). Certain traits for mammals overlapped with the PanTHERIA dataset (i.e., home range, longevity, and range size), however the methodology used for the trait calculations differs between Frugivoria and PanTHERIA. I have highlighted PanTHERIA traits in gray for comparison. Overall, for mammals, EltonTraits contributed 14.7% of traits, PanTHERIA 49.8%, and 35.36% of traits are newly added traits. For birds, EltonTraits contributed 27.62% and newly added traits comprise 72.38% of the dataset.



If multiple values were found for a given species trait, I gave preference to primary literature and the most recent data source. However, if I found multiple values from recent sources, I incorporated this information into a range of values with each source being cited in the source column for that trait but separated by a ‘|’ symbol. For these traits, specific details about each source may be found in the trait's notes column. Relevant traits have a minimum and maximum column for instances where ranges of values were present or different values were available from alternate yet appropriate sources. Traits with no available information for a given species were assigned to the genus level where possible and this was noted in the ‘*_level’ column for the applicable trait. Including prior imputations in the original EltonTraits dataset and our further imputations to congenerics for reclassified and newly added species in both EltonTraits and PanTHERIA datasets, 5142 traits (11.37%) were imputed to either genus or family, with 9 phylogenetically imputed traits for mammals (previously imputed from EltonTraits). For the newly added traits, 6.92% were assigned to the level of genus or family (birds: 2.16% genus, 0.75% family; mammals: 9.02% genus, 4.83% family) if species-specific information was not available. To minimize observer error within new trait data assignments for species, all new trait data records in Frugivoria were independently entered by two technicians. Any discrepancies between the two technicians were noted and reviewed by the lead author and resolved by replacing the former value with that from the most accurate and reputable source.

Due to the disparate nature of methodology of data sources used to build the Frugivoria dataset, there are some overlapping traits between databases. Because of this, I retain them all here in their original form for ease of interpretation. For example, mass is measured as averages in EltonTraits (Wilman et al., 2014), and in PanTHERIA they are measures of central tendency (often using the median; Jones et al., 2009), thus the origin of these values is not the same and they have non-comparable units. This is true for traits shared between EltonTraits and PanTHERIA such as activity patterns and mass, as well as PanTHERIA and the data I have collected and derived for this dataset, such as body size, diet breadth, home range size, range size and longevity. For some PanTHERIA traits, I added an additional suffix '_OD', for range-based geographical traits I suspect may be outdated (e.g., mean population size, mean annual precipitation and temperature), but may be useful for some cross-time analyses or comparisons. For these traits, I have derived more recent versions and applied them across taxa, whereas the original PanTHERIA traits were only available for mammals. Similarly, diet and habitat breadth, both traits available in the PanTHERIA dataset, were recalculated and applied across taxa.

Range-based metrics were calculated based on species range maps (L0_spatial_traits.R; Figure 1.2a.7). I obtained range maps from the IUCN Spatial Dataset (Figure 1.2a.1; IUCN, 2022a). I then computed the range sizes in km2 using the 'st_area' function built into the 'sf' package in R. I used the geography transformation of the 'shape' column within the IUCN Spatial Dataset (IUCN, 2022a). I calculated these range sizes in two ways: (1) using presence code 'Extant' and (2) using presence codes 'Extant', 'Probably Extant' and 'Possibly Extant'. The latter incorporates inferred areas where the species may exist and therefore represents potentially suitable habitat rather than just the known range of the species. I chose to include all 'origin' (e.g., native, re-introduced, introduced, vagrant) and 'seasonality' (e.g., resident, breeding season, non-breeding season, passage) codes as these all constitute a part of the range where the species would be performing the function of seed dispersal. I then appended the resulting range sizes to the full database for those species with range information available (n = 1734; missing ranges for n = 16 species). It is important to note that IUCN range maps have the potential to either overestimate or underestimate the true range of species (Gaston & Fuller, 2009; Ramesh et al., 2017) and therefore these range size estimates should be interpreted with discretion. Also, these ranges differ from those found in the PanTHERIA dataset for mammals, with nearly two-thirds of species having smaller ranges in our dataset calculated using the most recent data from the IUCN. These differences in range size may be due to different methodologies in the estimation of species ranges (IUCN, 2022a; Sechrest, 2003), differences in nomenclature between datasets (e.g., a species has been recently split and is species level for the IUCN range value and the sister species value in PanTHERIA) and the effects of ongoing habitat loss and changes in climate since the publication of sources used for calculating range size in the PanTHERIA dataset (Sechrest, 2003).

I also used the IUCN range maps to derive average climate and human impacts over species ranges, which can be used to understand the species' climatic tolerances as well as quantify potential anthropogenic impacts across the range. Specifically, I calculated the average values of two CHELSA bioclimatic variables (Karger et al., 2017, 2018), mean annual air temperature (Bio 1) and mean annual precipitation (Bio 12) over the observed and inferred portions of the range (e.g., parts of the range designated ‘extant’, ‘probably extant’ and ‘possibly extant’). To quantify anthropogenic impacts across the range, I used the WCS human footprint datasets for both 2010 and 2020 (Sanderson et al., 2022). This metric quantifies human impact by considering population density, infrastructure such as roads, railways, factories, accessibility, and the use of electricity. I then calculated the percent change between years as an additional metric to indicate the degree to which these human impacts have escalated over time.

In an effort to increase the number of cross-taxa traits, I recalculated diet and habitat breadth for mammals based on different data than that of PanTHERIA and extended these traits to birds (L0_7_breadth_traits.R; Figure 1.2a.7). For diet breadth, a measure of diet diversity, I used the Shannon Diversity Index (Shannon & Weaver, 1949). Similar to Santini et al. (2018), I calculated the Shannon Diversity Index using the 10 food categories from EltonTraits (Wilman et al., 2014) representing the proportional makeup of the species' diet. The maximum potential diversity value was calculated using the natural log of the number of possible diet categories. Here, the maximum diet diversity value is 2.3 and indicates a strong generalist species that consumes each food category evenly, whereas a value of zero indicates a monotonous diet specializing in a single source of food.

I also calculated habitat breadth by summing the number of suitable habitat types for each species based on level 2 of the IUCN Habitat Classification scheme (IUCN, 2022a). This was done using the 'rredlist' package in R and extracting the habitat types used by each species using the 'rl_habitat' function (IUCN, 2015). A full list of habitats suitable for each species is included in the published EDI dataset.

1.3.4 Occurrence Data

I provide code to extract GBIF (GBIF, 2023) records for species contained within this database (L2_downloading_gbif_records.R). This code is modified from that provided by the GBIF blog, which overcomes the issue of pulling large numbers of records through GBIF for multiple species at once (Waller & Grosjean, 2019). A free GBIF account is required to implement this code.

1.3.5 Frugivoria datasets

In addition to providing the full Frugivoria dataset, I also provide a simplified subset of the full Frugivoria dataset (e.g., Frugivoria_mammal_database_simple.csv, Frugivoria_bird_database_simple.csv). This subset represents shared cross-taxa traits, simplifying comparisons among birds and mammals (code contained in L0_final_database_edits.R). These traits are generally well filled; however, I also include home range and longevity traits despite their missing values (filled for birds—longevity: 26.7%, home range: 19.3%; mammals—longevity: 52.1%, home range: 44.2%; Figure 3) because of their ecological relevance and their scarcity in the literature, which may make them of heightened interest to users. This simplified subset excludes repetitive traits (e.g., ranges of values, percent composition of diet, habitat suitability), and retains traits that encapsulate and synthesize this information (e.g., trait averages, diet category, diet breadth, habitat breadth).

1.4 Results & Discussion

Frugivoria and its workflow facilitate studies on frugivorous mammal and bird traits encompassing ecology, life-history, morphology and geographical occurrences for a region of great ecological importance. Although mammals and birds are among the taxonomic groups with higher degrees of sampling and research, there remain significant gaps in our knowledge of their traits in this region. The existing trait databases for birds and mammals only include a subset of high-level traits (EltonTraits, Wilman et al., 2014; Phylacine, Faurby et al., 2018) or are missing values for many species in biodiversity hotspots (PanTHERIA; Jones et al., 2009)—gaps which Frugivoria helps fill.

In total, I have added 8709 new traits for mammals (35.36% of all mammal traits) and 14,999 for birds (72.38% of all bird traits) (Figures 3 and 4) and of those, only 6.92% of traits were imputed. I added traits for 44 new species, and updated the taxonomy for 314 species. Of the new morphological, ecological and life-history traits that did not require explicit calculation (i.e., home range, longevity, generation time, sexual dimorphism and body size), 1285 come directly from the literature, which is 21.6% of new traits in those categories. The remaining 79.4% of these traits (4653) were collated across disparate online sources and datasets including species accounts requiring explicit interpretation from Encyclopedia of Life (Parr et al., 2014), Cornell's Birds of the World (Billerman et al., 2022), University of Michigan's Animal Diversity Web (Myers et al., 2023), IUCN (IUCN, 2022a) and AnAge database (Tacutu et al., 2018).

I also derive new traits such as diet breadth and habitat breadth and calculate geographical range-based traits such as observed and inferred range size, climate-based traits, and different aspects of human impact across the range, which all have the potential to be used to estimate extinction risk. Both the breadth traits and new geographical range-based traits update and expand the applicability of these traits across taxa, as these were traits once only available for mammals. Furthermore, I increased the completeness of traits in PanTHERIA (e.g., body mass, body size, range size, home range size and longevity for mammals)—in some cases more than doubling the traits available for mammals in PanTHERIA (Figure 1.3; home range and longevity). Not only does Frugivoria help fill existing data gaps and generate comparable cross-taxa traits for birds and mammals, but it also harmonizes existing databases into a single unified source for mammals and birds, making studies of these taxa much less cumbersome and time-consuming (Etard et al., 2020). This database (and its reproducible workflow for other taxa and regions of the world) is particularly beneficial for studying functional diversity in the Neotropics, where the sheer number of species and the complexity of their interactions can make it difficult to identify patterns of ecological importance.

The increased trait resolution (i.e., low levels of imputation and greater levels of filled traits for species) and spatial and taxonomic coverage in Frugivoria provides vital information to address fundamental and applied aspects of conservation biology. Specifically, Frugivoria comprises a unified and comprehensive source that can be used to understand community assembly (species coexistence; Zamudio et al., 2016), spatial patterns of biodiversity, trait distributions (Figure 1.4), and can help assess the vulnerability of species to environmental changes (Pacifici et al., 2015) in moist montane and lowland regions of Central and South America. Some examples of more focused investigations that are now possible with Frugivoria include understanding how frugivore traits help mediate ecological processes in moist montane systems (Lim et al., 2020; Sekercioğlu et al., 2004); quantifying how traits vary across montane regions (Dehling et al., 2014; Santillán et al., 2019; Figure 4); and understanding how certain traits relate to species extinction risk (Bland et al., 2015; González-del-Pliego et al., 2019; Ripple et al., 2017). The latter topic is particularly relevant given the high endemicity in this region (Gradstein et al., 2008; Myers et al., 2000), the projected rapid changes in climate (IPCC, 2022) and land use (Armenteras et al., 2011; González-Maya et al., 2017; Powers & Jetz, 2019) and the anticipated shifts in ecological communities over the next century (Williams et al., 2007). Montane regions and cloud forests are especially sensitive to environmental changes (Elsen & Tingley, 2015; Foster, 2001; Ponce-Reyes et al., 2012; Toledo-Aceves et al., 2011); thus, species traits can aid in understanding and predicting shifts in biodiversity and associated ecosystem functions and services.

The IUCN categorizes species into risk categories to assess extinction risk and help prioritize species and habitat conservation (IUCN, 2022b). Despite the increase in and utility of trait-based approaches in assessing extinction risk and vulnerability (Foden et al., 2013; Kosman et al., 2019; Pacifici et al., 2015), setting spatial priorities for parks and reserves (Kukkala & Moilanen, 2017), and mapping diversity patterns (Cadotte & Tucker, 2018; Devictor et al., 2010), the IUCN does not explicitly incorporate trait-based approaches into the official IUCN Red List assessment process (IUCN, 2022b). Instead, the IUCN uses a population trend-based approach (changes in abundance or current and potential changes in distributions; Etard et al., 2020), relying on information often unavailable for many species. Of the species contained in Frugivoria, 12.9% are classified as threatened by the IUCN, with 10.1% of lowland species and 16.4% of montane species having threat categories of either vulnerable, endangered, or critically endangered. The higher proportion of threatened montane species is not unexpected, since montane ecosystems are particularly vulnerable to threats such as climate change and fragmentation, and species in these regions often have restricted dispersal capabilities.

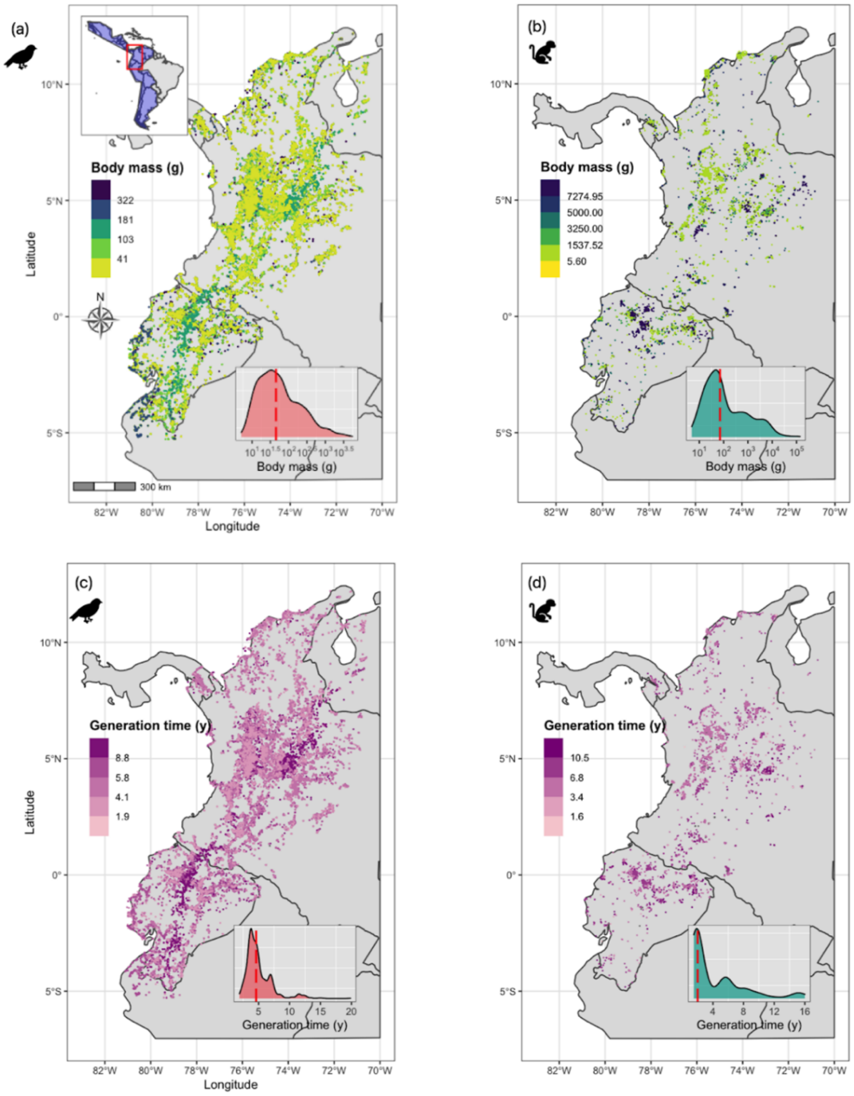
Of mammals specifically, 14.6% are classified as DD, because they lack one or both population and distributional (range size; EOO; IUCN, 2022a) information to make an official extinction risk assessment. Species such as this provide an opportunity to test trait-based approaches for assessing conservation statuses.

Trait-based approaches to conservation offer an alternative approach to trend-based approaches because trait-based approaches rely on species' sensitivity to particular threats (Etard et al., 2020). If the response of species to a threat consistently aligns with certain traits (e.g., both narrow diet breadth, which indicates a specialized diet, and large-bodied and small-ranged species being potential indicators of extinction risk; Boyles & Storm, 2007; Harris & Pimm, 2008; Ripple et al., 2017), traits can be used as part of a proactive approach to generalize patterns or set rules for assigning extinction risk for species that do not have the sufficient population or geographical data (Cardillo & Meijaard, 2012) or can be used in addition to trend-based approaches. For instance, Bland et al. (2015) estimated the extinction risk of DD terrestrial mammals using predictive models based on life history, geographical ranges, and information on threats, increasing the estimate of globally threatened terrestrial mammals by 5%. Frugivoria has the potential to be used for this purpose for frugivorous species in the montane Neotropics, as traits such as habitat specialization (strong reliance on cloud forest habitat), diet breadth

Figure 1.4: Distribution of mass and generation time for birds (a & c) and mammals (b & d) across the Northern Andes and Isthmus of Panama. Occurrence records were obtained through GBIF (GBIF, 2022a; GBIF, 2022b; GBIF, 2022c; GBIF, 2022d). Lighter colors indicate lower values. Birds and mammals differed in their minimums and maximums for each trait (birds a: 5.1 - 5,525.0 g, c: 1.9 - 19.8 y; mammals b: 4.8 - 140,000.6 g, d: 0.94 - 18.3 y). Natural breaks were used to bin these data for plotting, with a few larger values included in the uppermost bin. The

Figure 1.4 (cont'd)

density curves in the lower right corner indicate the distribution of the mapped trait in the Frugivoria dataset. For mass (panels a & b), these densities were log transformed for visualization purposes. Red dashed lines indicate the median value for each trait (birds a: 45.3 g, c: 4.6 y; mammals b: 71.0 g, d: 4.2 y).



(based on the Shannon Index and demonstrating the degree of diet specialization), and range size can indicate levels of risk to anthropogenic pressures such as habitat degradation and climate change (Bland et al., 2015). For example, range size can be used to infer sensitivity to environmental change because narrow-ranged species, which often have very little data, are more sensitive to anthropogenic disturbances and tend to have higher extinction risk than those with broad ranges (Collen et al., 2016; Ripple et al., 2017). Furthermore, in Frugivoria, I include a geographical trait—the average human footprint—explicitly quantifying anthropogenic forces acting across the species range. As such, it would be straightforward to use Frugivoria to generate lists of potentially at-risk species in this region using these traits listed above. This dataset can be used in applied conservation to set targets for maintaining areas of high functional diversity and spatially prioritizing regions containing species with unique ecosystem roles.

The species currently in Frugivoria cover an important and highly diverse region and habitat type, yet it does not provide a complete picture of Central and South American frugivore biodiversity. Importantly, it excludes explicit incorporation of countries lacking contiguous mountain ranges, for example in Southeastern Brazil, the IUCN designated habitats 'subtropical/tropical dry forest' (e.g., seasonally dry inter-Andean valleys) and 'tropical high altitude' (e.g., páramo; regions above the timberline). Frugivoria has been compiled in an open and reproducible way that facilitates its future expansion to high-altitude montane and lowland regions of Central and South America—an essential undertaking for gaining a more complete picture of trait diversity for frugivorous species in the Neotropics. Though the species contained in Frugivoria are only a subset of the world's vast biodiversity, I hope the data and workflow facilitate further study in the Neotropics and beyond, as filling in more data gaps in hotspots like these continues to be a research priority.

1.5 Data Availability Statement

The Frugivoria database and associated metadata are openly available as .csv files through the Environmental Data Initiative.

(EDI; https://doi.org/10.6073/pasta/168e95f04d4726d31d868bfe22d749a5) (Gerstner et al. 2023). I also provide the R scripts used for the database building workflow and analysis for this manuscript on EDI and GitHub (https://github.com/bioXgeo/neotropical_frugivores.git).

Published manuscript citation:

Gerstner, B. E., Bills, P. & Zarnetske, P. L. (2023). Frugivoria: A trait database for birds and mammals exhibiting frugivory across contiguous Neotropical moist forests. Global Ecology and Biogeography, 32, 1466–1484.

REFERENCES

Ansell, W. F. H. (1965). Standardization of field data on mammals. Zoologica Africana, 1, 97–113.

Armenteras, D., Rodríguez, N., Retana, J., & Morales, M. (2011). Understanding deforestation in montane and lowland forests of the Colombian Andes. Regional Environmental Change, 11, 693–705.

Bellard, C., Leclerc, C., Leroy, B., Bakkenes, M., Veloz, S., Thuiller, W., & Courchamp, F. (2014). Vulnerability of biodiversity hotspots to global change. Global Ecology and Biogeography, 23, 1376–1386.

Billerman, S. M.,. Keeney, B. K , Rodewald, P. G., & Schulenberg, T. S. (Eds.). (2022). Birds of the world. Cornell Laboratory of Ornithology. https://birdsoftheworld.org/bow/home

Bland, L. M., Collen, B., Orme, C. D. L., & Bielby, J. (2015). Predicting the conservation status of data-deficient species. Conservation Biology, 29, 250–259.

Boyles, J. G., & Storm, J. J. (2007). The perils of picky eating: Dietary breadth is related to extinction risk in insectivorous bats. PLoS One, 2, e672.

Brandon, K. (2014) Ecosystem Services from Tropical Forests: Review of Current Science, Social Science Research Network, Rochester, NY.

Burt, W. H. (1943). Territoriality and home range concepts as applied to mammals. Journal of Mammalogy, 24, 346–352.

Burton, M., & Robert, B. (1974). The Funk & Wagnalls wildlife encyclopedia. Funk and Wagnalls. New York, New York.

Cadotte, M. W., Carscadden, K., & Mirotchnick, N. (2011). Beyond species: Functional diversity and the maintenance of ecological processes and services. Journal of Applied Ecology, 48, 1079–1087.

Cadotte, M. W., & Tucker, C. M. (2018). Difficult decisions: Strategies for conservation prioritization when taxonomic, phylogenetic and functional diversity are not spatially congruent. Biological Conservation, 225, 128–133.

Cardillo, M., & Meijaard, E. (2012). Are comparative studies of extinction risk useful for conservation? Trends in Ecology & Evolution, 27, 167–171.

Center for International Earth Science Information Network – CIESIN – Columbia University, and Centro Internacional de Agricultura Tropical – CIAT. (2005). Gridded population of the world, version 3 (GPWv3): Population density grid. NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4XK8CG2

Collen, B., Dulvy, N. K., Gaston, K. J., Gärdenfors, U., Keith, D. A., Punt, A. E., Regan, H. M., Böhm, M., Hedges, S., Seddon, M., Butchart, S. H. M., Hilton-Taylor, C., Hoffmann, M., Bachman, S. P., & Akçakaya, H. R. (2016). Clarifying misconceptions of extinction risk assessment with the IUCN red list. Biology Letters, 12, 20150843.

Collen, B., Ram, M., Zamin, T., & McRae, L. (2008). The tropical biodiversity data gap: Addressing disparity in global monitoring. Tropical Conservation Science, 1, 75–88.

Cooke, R. S. C., Bates, A. E., & Eigenbrod, F. (2019). Global trade-offs of functional redundancy and functional dispersion for birds and mammals. Global Ecology and Biogeography, 28, 484–495.

Dehling, D. M., Töpfer, T., Schaefer, H. M., Jordano, P., Böhning-Gaese, K., & Schleuning, M. (2014). Functional relationships beyond species richness patterns: Trait matching in plant–bird mutualisms across scales. Global Ecology and Biogeography, 23, 1085–1093.

Devictor, V., Mouillot, D., Meynard, C., Jiguet, F., Thuiller, W., & Mouquet, N. (2010). Spatial mismatch and congruence between taxonomic, phylogenetic and functional diversity: The need for integrative conservation strategies in a changing world. Ecology Letters, 13, 1030–1040.

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., Hansen, M., Locke, H., Ellis, E. C., Jones, B., Barber, C. V., Hayes, R., Kormos, C., Martin, V., Crist, E., … Saleem, M. (2017). An ecoregion-based approach to protecting half the terrestrial realm. Bioscience, 67, 534–545.

Eisenberg, J. F., & Redford, K. H. (1999). Mammals of the neotropics (Vol. 3). University of Chicago Press.

Elsen, P. R., & Tingley, M. W. (2015). Global mountain topography and the fate of montane species under climate change. Nature Climate Change, 5, 772–776.

Emmons, L., & Feer, F. (1997). Neotropical rainforest mammals: A field guide. University of Chicago Press.

Engemann, K., Enquist, B. J., Sandel, B., Boyle, B., Jørgensen, P. M., Morueta-Holme, N., Peet, R. K., Violle, C., & Svenning, J.-C. (2015). Limited sampling hampers “big data” estimation of species richness in a tropical biodiversity hotspot. Ecology and Evolution, 5, 807–820.

Etard, A., Morrill, S., & Newbold, T. (2020). Global gaps in trait data for terrestrial vertebrates. Global Ecology and Biogeography, 29, 2143–2158.

Farwig, N., & Berens, D. G. (2012). Imagine a world without seed dispersers: A review of threats, consequences and future directions. Basic and Applied Ecology, 13, 109–115.

Faurby, S., Davis, M., Pedersen, R. O., Schowanek, S. D., Antonelli, A., & Svenning, J.-C. (2018). PHYLACINE 1.2: The phylogenetic atlas of mammal macroecology. Ecology, 99, 2626.

Ferrier, S. (2002). Mapping spatial pattern in biodiversity for regional conservation planning: Where to from here? Systematic Biology, 51, 331–363.

Fleagle, J. G. (2013). Primate adaptation and evolution. Academic Press.

Foden, W. B., Butchart, S. H. M., Stuart, S. N., Vié, J.-C., Akçakaya, H. R., Angulo, A., DeVantier, L. M., Gutsche, A., Turak, E., Cao, L., Donner, S. D., Katariya, V., Bernard, R., Holland, R. A., Hughes, A. F., O'Hanlon, S. E., Garnett, S. T., Şekercioğlu, Ç. H., & Mace, G. M. (2013). Identifying the World's Most climate change vulnerable species: A systematic trait-based assessment of all birds, amphibians and corals. PLoS ONE, 8, e65427.

Foster, P. (2001). The potential negative impacts of global climate change on tropical montane cloud forests. Earth-Science Reviews, 55, 73–106.

Fricke, E. C., Ordonez, A., Rogers, H. S., & Svenning, J.-C. (2022). The effects of defaunation on plants' capacity to track climate change. Science, 375, 210–214.

Gaston, K. J., & Fuller, R. A. (2009). The sizes of species' geographic ranges. Journal of Applied Ecology, 46, 1–9.

GBIF: The Global Biodiversity Information Facility. (2023). What is GBIF? https://www.gbif.org/what-is-gbif

GBIF.org (26 August 2022.a) GBIF Occurrence Download https://doi.org/10.15468/dl.wmksrj

GBIF.org (26 August 2022.b) GBIF Occurrence Download https://doi.org/10.15468/dl.k2szsu

GBIF.org (26 August 2022.c) GBIF Occurrence Download https://doi.org/10.15468/dl.aeyeem

GBIF.org (26 August 2022.d) GBIF Occurrence Download https://doi.org/10.15468/dl.fqmhp2

Gómez, C., Tenorio, E. A., & Cadena, C. D. (2021). Change in avian functional fingerprints of a neotropical montane forest over 100 years as an indicator of ecosystem integrity. Conservation Biology, 35, 1552–1563.

González-del-Pliego, P., Freckleton, R. P., Edwards, D. P., Koo, M. S., Scheffers, B. R., Pyron, R. A., & Jetz, W. (2019). Phylogenetic and trait-based prediction of extinction risk for data-deficient amphibians. Current Biology, 29, 1557–1563.e3.

González-Maya, J. F., Martínez-Meyer, E., Medellín, R., & Ceballos, G. (2017). Distribution of mammal functional diversity in the neotropical realm: Influence of land-use and extinction risk. PLoS One, 12, e0175931.

González-Suárez, M., & Revilla, E. (2013). Variability in life-history and ecological traits is a buffer against extinction in mammals. Ecology Letters, 16, 242–251.
S. R. Gradstein, J. Homeier, & D. Gansert (Eds.). (2008). The tropical mountain forest: Patterns and processes in a biodiversity hotspot. Göttingen University Press.

Guralnick, R., & Constable, H. (2010). VertNet: Creating a data-sharing community. Bioscience, 60, 258–259.

Harris, G., & Pimm, S. L. (2008). Range size and extinction risk in Forest birds. Conservation Biology, 22, 163–171.

Howe, H. F., & Smallwood, J. (1982). Ecology of seed dispersal. Annual Review of Ecology and Systematics, 13, 201–228.

IPCC. (2022). In H.-O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, & B. Rama (Eds.), Climate change 2022: Impacts, adaptation, and vulnerability. Contribution of working group II to the sixth assessment report of the intergovernmental panel on climate change (3056). Cambridge University Press. https://doi.org/10.1017/9781009325844

IUCN. (2015). The IUCN red list of threatened species. Version 2022-2. https://www.iucnredlist.org

IUCN. (2021). The IUCN Red List of Threatened Species. Version 2021. https://www.iucnredlist.org

IUCN. (2022a). The IUCN red list of threatened species. Version 2022-2. https://www.iucnredlist.org

IUCN. (2022b). IUCN Standards and Petitions Committee. Guidelines for Using the IUCN Red List Categories and Criteria. Version 15.1. Prepared by the Standards and Petitions Committee. https://www.iucnredlist.org/documents/RedListGuidelines.pdf

Jones, K. E., Bielby, J., Cardillo, M., Fritz, S. A., O'Dell, J., Orme, C. D. L., Safi, K., Sechrest, W., Boakes, E. H., Carbone, C., Connolly, C., Cutts, M. J., Foster, J. K., Grenyer, R., Habib, M., Plaster, C. A., Price, S. A., Rigby, E. A., Rist, J., … Purvis, A. (2009). PanTHERIA: A species-level database of life history, ecology, and geography of extant and recently extinct mammals. Ecology, 90, 2648.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. Scientific Data, 4, 170122.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2018). Data from: Climatologies at high resolution for the earth's land surface areas. Dryad Digital Repository https://doi.org/10.5061/dryad.kd1d4

Kissling, W. D., Walls, R., Bowser, A., Jones, M. O., Kattge, J., Agosti, D., Amengual, J., Basset, A., van Bodegom, P. M., Cornelissen, J. H. C., Denny, E. G., Deudero, S., Egloff, W., Elmendorf, S. C., Alonso García, E., Jones, K. D., Jones, O. R., Lavorel, S., Lear, D., … Guralnick, R. P. (2018). Towards global data products of essential biodiversity variables on species traits. Nature Ecology & Evolution, 2, 1531–1540.

Kosman, E., Burgio, K. R., Presley, S. J., Willig, M. R., & Scheiner, S. M. (2019). Conservation prioritization based on trait-based metrics illustrated with global parrot distributions. Diversity and Distributions, 25, 1156–1165.

Kukkala, A. S., & Moilanen, A. (2017). Ecosystem services and connectivity in spatial conservation prioritization. Landscape Ecology, 32, 5–14.

Lepage, D., Vaidya, G., & Guralnick, R. (2014). Avibase – A database system for managing and organizing taxonomic concepts. ZooKeys, 420, 117–135.

Lim, J. Y., Svenning, J.-C., Göldel, B., Faurby, S., & Kissling, W. D. (2020). Frugivore-fruit size relationships between palms and mammals reveal past and future defaunation impacts. Nature Communications, 11, 4904.

Miralles, A., Bruy, T., Wolcott, K., Scherz, M. D., Begerow, D., Beszteri, B., Bonkowski, M., Felden, J., Gemeinholzer, B., Glaw, F., Gloeckner, F. O., Hawlitschek, O., Kostadinov, I., Nattkemper, T. W., Printzen, C., Renz, J., Rybalka, N., Stadler, M., Weibulat, T., … Vences, M. (2020). Repositories for taxonomic data: Where we are and what is missing. Systematic Biology, 69, 1231–1253.

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

Myers, P., Espinosa, R., Parr, C. S., Jones, T., Hammond, G. S., & Dewey, T. A. (2023). The Animal Diversity Web (online). https://animaldiversity.org

Nelson, G., & Ellis, S. (2019). The history and impact of digitization and digital data mobilization on biodiversity research. Philosophical Transactions of the Royal Society B: Biological Sciences, 374, 20170391.

Newbold, T., Scharlemann, J. P. W., Butchart, S. H. M., Şekercioğlu, Ç. H., Joppa, L., Alkemade, R., & Purves, D. W. (2014). Functional traits, land-use change and the structure of present and future bird communities in tropical forests. Global Ecology and Biogeography, 23, 1073–1084.

Pacifici, M., Foden, W. B., Visconti, P., Watson, J. E. M., Butchart, S. H. M., Kovacs, K. M., Scheffers, B. R., Hole, D. G., Martin, T. G., Akçakaya, H. R., Corlett, R. T., Huntley, B., Bickford, D., Carr, J. A., Hoffmann, A. A., Midgley, G. F., Pearce-Kelly, P., Pearson, R. G., Williams, S. E., … Rondinini, C. (2015). Assessing species vulnerability to climate change. Nature Climate Change, 5, 215–224.

Parr, C. S., Wilson, N., Leary, P., Schulz, K. S., Lans, K., Walley, L., Hammock, J. A., Goddard, A., Rice, J., Studer, M., Holmes, J. T. G., & Corrigan, R. J., Jr. (2014). The encyclopedia of life v2: Providing global access to knowledge about life on earth. Biodiversity Data Journal, 2, e1079. https://doi.org/10.3897/BDJ.2.e1079

Pollock, L. J., Thuiller, W., & Jetz, W. (2017). Large conservation gains possible for global biodiversity facets. Nature, 546, 141–144.

Ponce-reyes, R., Reynoso-rosales, V., Watson, J. E. M., Vanderwal, J., Fuller, R. A., Pressey, R. L., & Possingham, H. P. (2012). Vulnerability of cloud forest reserves in Mexico to climate change. Nature Climate Change, 2, 448–452.

Powers, R. P., & Jetz, W. (2019). Global habitat loss and extinction risk of terrestrial vertebrates under future land-use-change scenarios. Nature Climate Change, 9, 323–329.

R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Ramesh, V., Gopalakrishna, T., Barve, S., & Melnick, D. J. (2017). IUCN greatly underestimates threat levels of endemic birds in the Western Ghats. Biological Conservation, 210, 205–221.

Rehm, E. M., & Feeley, K. J. (2015). The inability of tropical cloud forest species to invade grasslands above treeline during climate change: Potential explanations and consequences. Ecography, 38, 1167–1175.

Ripple, W. J., Wolf, C., Newsome, T. M., Hoffmann, M., Wirsing, A. J., & McCauley, D. J. (2017). Extinction risk is most acute for the world's largest and smallest vertebrates. Proceedings of the National Academy of Sciences of the United States of America, 114, 10678–10683.

Sales, L. P., Kissling, W. D., Galetti, M., Naimi, B., & Pires, M. M. (2021). Climate change reshapes the eco-evolutionary dynamics of a neotropical seed dispersal system. Global Ecology and Biogeography, 30, 1129–1138.

Sanderson, E. W., Fisher, K., Robinson, N., Sampson, D., Duncan, A., & Royte, L. (2022). The march of the human footprint. EcoEvoRxiv.

Santillán, V., Quitián, M., Tinoco, B. A., Zárate, E., Schleuning, M., Böhning-Gaese, K., & Neuschulz, E. L. (2019). Different responses of taxonomic and functional bird diversity to forest fragmentation across an elevational gradient. Oecologia, 189, 863–873.

Santini, L., González-Suárez, M., Russo, D., Gonzalez-Voyer, A., Hardenberg, A., & Ancillotto, L. (2018). One strategy does not fit all: determinants of urban adaptation in mammals. Ecology Letters, 22(2), 365–376. https://doi.org/10.1111/ele.13199.

Sechrest, W. (2003). Global diversity, endemism and conservation of mammals. University of Virginia.

Sekercioğlu, C. H., Daily, G. C., & Ehrlich, P. R. (2004). Ecosystem consequences of bird declines. Proceedings of the National Academy of Sciences of the United States of America, 101, 18042–18047.

Shannon, C. E., & Weaver, W. (1949). The mathematical theory of communication. University of Illinois Press.

Soria, C. D., Pacifici, M., Marco, M. D., Stephen, S. M., & Rondinini, C. (2021). COMBINE: A coalesced mammal database of intrinsic and extrinsic traits. Ecology, 102, e03344.

Tacutu, R., Thornton, D., Johnson, E., Budovsky, A., Barardo, D., Craig, T., Diana, E., Lehmann, G., Toren, D., Wang, J., Fraifeld, V. E., & de Magalhães, J. P. (2018). Human ageing genomic resources: New and updated databases. Nucleic Acids Research, 46, D1083–D1090.

Thorington, R. W., Koprowski, J. L., Steele, M. A., & Whatton, J. F. (2012). Squirrels of the world. Johns Hopkins University Press.

Tóbon, C., Sevink, J., & Verstraten, J. M. (2004). Solute fluxes in throughfall and stemflow in four Forest ecosystems in Northwest Amazonia. Biogeochemistry, 70, 1–25.

Toledo-Aceves, T., Meave, J. A., González-Espinosa, M., & Ramírez-Marcial, N. (2011). Tropical montane cloud forests: Current threats and opportunities for their conservation and sustainable management in Mexico. Journal of Environmental Management, 92, 974–981.

Waller, J., & Grosjean, M. (2019). Downloading occurrences from a long list of species in R and python. GBIF Data Blog. https://data-blog.gbif.org/post/downloading-long-species-lists-on-gbif/

Wilkie, D. S., Bennett, E. L., Peres, C. A., & Cunningham, A. A. (2011). The empty forest revisited. Annals of the New York Academy of Sciences, 1223, 120–128.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. Scientific Data, 3, 160018.

Williams, J. W., Jackson, S. T., & Kutzbach, J. E. (2007). Projected distributions of novel and disappearing climates by 2100 AD. Proceedings of the National Academy of Sciences of the United States of America, 104, 5738–5742.

Wilman, H., Belmaker, J., Simpson, J., de la Rosa, C., Rivadeneira, M. M., & Jetz, W. (2014). EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals. Ecology, 95, 2027.

Wilson, A. M., & Jetz, W. (2016). Remotely sensed high-resolution global cloud dynamics for predicting ecosystem and biodiversity distributions. PLoS Biology, 14, e1002415.

Zamudio, K. R., Bell, R. C., & Mason, N. A. (2016). Phenotypes in phylogeography: Species' traits, environmental variation, and vertebrate diversification. Proceedings of the National Academy of Sciences of the United States of America, 113, 8041–8048.

CHAPTER 2:

THE INFLUENCE OF SCALE-DEPENDENT GEODIVERSITY ON SPECIES DISTRIBUTION MODELS IN A BIODIVERSITY HOTSPOT

Resubmitted to Philosophical Transactions A special issue on "Geodiversity science for society."

Co-authors include: Phoebe L. Zarnetske, Mary E. Blair, Patrick Bills, Cristian A.C. Rodriguez

2.1 Abstract

Improving models of species' distributions is essential for conservation, especially in light of global change. Species distribution models (SDMs) often rely on mean environmental conditions, yet species distributions are also a function of environmental heterogeneity and filtering acting at multiple spatial scales. Geodiversity, which is the variation in Earth's features and processes, has the potential to improve SDMs and conservation assessments, but geodiversity variables have not been sufficiently tested in SDMs. I tested a range of geodiversity variables computed at varying scales using climate and elevation data. I compared the predictive performance of MaxEnt SDMs generated using CHELSA bioclimatic variables to those also including geodiversity variables for 31 mammalian species in Columbia. Results show that the inclusion of scale dependent geodiversity variables in SDMs improves model performance both statistically and spatially, especially for species in areas of high heterogeneity. Some variables consistently exhibited an increasing or decreasing trend in variable importance with spatial grain, showing slight scale-dependence and indicating that some geodiversity variables are more relevant at particular scales for some species. Incorporating geodiversity variables into SDMs, and doing so at the appropriate spatial scales to best reflect the species-habitat relationship, enhances the ability to model species-environment relationships, thereby contributing to the conservation and management of biodiversity.

2.2 Introduction

In light of the unprecedented global changes threatening biodiversity, there is an increasing need for effective tools and strategies to aid in the spatial prioritization of conservation efforts. One proposed strategy is to focus on “geodiversity”, which has a range of definitions (Parks & Mulligan, 2010; Lawler et al., 2015), but for which I define here as the diversity of abiotic features and processes of Earth’s critical zone (including the lithosphere, atmosphere, hydrosphere, and cryosphere) and is thereby inclusive of climate (Zarnetske et al., 2019). Geodiverse areas are expected to harbor higher levels of biodiversity because they provide more niche opportunities than areas with lower geodiversity (Lawler et al., 2015; Bailey et al., 2017; Tukiainen, 2019). This relationship is thought to influence patterns of biodiversity and species distributions due to the varied landscape and associated abiotic and biotic conditions which can increase the size of available niche space (Dufour et al., 2006). Geodiverse areas, which harbor a diversity of abiotic and biotic conditions, are likely to serve as refugia for species, and conservationists have proposed focusing on them to protect biodiversity in a changing climate (Brost & Beier, 2012; Reynard & Brilha, 2018; Schrodt et al., 2019; Crofts, 2022). Existing research has primarily focused on quantifying the relationship between geodiversity and biodiversity, particularly species richness (Stein et al., 2014; Bailey et al., 2017; Zarnetske et al., 2019), yet the relationship between geodiversity, measured as abiotic spatial heterogeneity within a site, and individual species distributions remains largely unexplored. Given that a majority of conservation decisions still focus on individual species (Brooks et al., 2006; Brum et al., 2017) and that species distribution models are widely regarded as a useful and often key approach for assessing extinction risk and setting spatial conservation priorities

(Lawler et al., 2015; IUCN, 2022), there is a need to understand how geodiversity variables may influence species distributions and therefore their utility within SDMs.

Understanding the complex interplay between measures of geodiversity and biodiversity as well as their spatial scaling relationships can be essential for developing effective conservation strategies, particularly in regions with high levels of topographic complexity (Zarnetske et al., 2019). Geodiversity plays a crucial role in determining the physical boundaries of species' ranges by influencing the physiological constraints imposed by species' tolerances towards environmental conditions. Distributional limits can be further influenced by structural barriers to dispersal that might arise from topographic complexity, and the arrangement of habitat patches which can either facilitate or impede biotic interactions among species, as noted by Urban et al. (2013). Further, the effects of environmental heterogeneity on species distributions will vary depending on the scale at which a species responds to the environment (Graham et al., 2019). Further, this scale often differs among species or their associated functional groups (i.e., ecological groupings of species sharing traits and life strategies; Blaum et al., 2011).

The occurrence of a species is intricately linked to its realized niche, which emerges from environmental filtering operating across multiple scales beyond the local occurrence point (Kraft et al., 2015; Pearson et al., 2018). This filtering process, which broadly determines the occurrence of species, involves a range of factors, including dispersal limitations, habitat configuration, climatic variations, and biotic interactions. For example, the presence of a river, a local dispersal barrier, or a competitor, might deter a small primate population from moving to suitable habitat less than 20 meters away, while variation in climate might be more gradual and prevent the primate species’ range from extending across vastly different temperature or precipitation zones. These combined filtering processes play a fundamental role in shaping species distributions and offer valuable insights into the intricate dynamics governing species occurrence (Guisan & Thuiller, 2005; Soberón & Nakamura, 2009). The most common approach to understanding and predicting species-environment relationships is species distribution modeling (SDM) (Guisan & Thuiller, 2005). Yet typically, SDMs only incorporate environmental variables such as bioclimatic variables (variables that summarize annual, seasonal, and monthly trends in temperature and precipitation), at the local pixel scale. In a typical SDM, single pixel environmental values are intersected with a species' occurrence point. Reliance on this local scale relationship discounts the broader contextual environmental information of areas surrounding occurrence points. To better incorporate environmental filtering and associated broader scales of environmental conditions surrounding a species' occurrence point, SDMs could also include environmental heterogeneity in areas surrounding the occurrence points. Without including this broader environmental heterogeneity, SDMs are limited to reflecting the finer-scale local species-habitat relationships and are therefore less complete explanations and predictions of species distributions.

Species-environment relationships can also be highly scale-dependent, with the strength and direction of the interactions between biotic and abiotic factors varying across different spatial scales (Stein et al., 2014; Bailey et al., 2017, 2018; Zarnetske et al., 2019). There are numerous scale-dependent relationships between species and their environmental drivers (Elith & Leathwick, 2009; Fournier et al., 2017). For example, the distribution of species is determined by a combination of factors, including climate, which has likely influenced occurrence at broad spatial scales, as noted by Blach-Overgaard et al. (2010), and habitat factors, such as availability and fragmentation, at more local spatial scales (Virkkala et al., 2005; Luoto et al., 2007).

Therefore, it is essential to investigate scale-dependency in species-geodiversity relationships. Incorporating geodiversity in terms of spatial heterogeneity or variability into species distribution models and assessing scale dependency has potential to improve our understanding of the factors that govern species distributions and may help refine resulting distribution maps. This has important implications for conservation as distribution maps are often used as a tool for assessing extinction risk (e.g., area of occupancy [AOO] and extent of occurrence [EOO] for the International Union of the Conservation of Nature [IUCN]; IUCN, 2022), and for determining potential areas for future sampling and priority areas for conservation.

I tested the utility of incorporating geodiversity variables computed at varying spatial scales across a rasterized landscape into SDMs. These geodiversity variables capture the spatial heterogeneity within a defined neighborhood around species occurrences and might offer insights into the underlying processes that either facilitate or hinder species presence. My approach addresses the need to incorporate environmental filtering at broader scales surrounding species occurrence points, and scale-dependency in species-environment relationships. As geodiversity variables can reflect the availability of microclimates or landscape variability, they hold promise for improving SDMs and provide a more comprehensive understanding of species-environment relationships (Stein et al., 2014; Zarnetske et al., 2019; Blair et al., 2022b).

While it has been established that environmental heterogeneity can influence species distributions and diversity patterns at multiple spatial scales, it is also possible that species traits might be mediating these patterns. For instance, each species possesses unique functional traits (any traits that allow species to survive and reproduce in a given environment; Violle et al., 2007) and evolutionary histories, resulting in different sensitivities to and preferences for specific environmental conditions (Pollock et al., 2012). Most research aiming to understand the influence of functional traits on species distributions, however, has focused on plants (Pollock et al., 2012; Maharjan et al., 2021; Wang & Wan, 2021; Tang et al., 2022) or aquatic animals (Wittmann et al., 2016; Bosch-Belmar et al., 2021), limiting our generalized understanding of these dynamics more broadly. Ultimately, understanding the complex relationship between geodiversity and species functional traits, such as body mass (e.g., relationship with trophic level, dispersal ability, and home range size) and diet preference (e.g., relationship with trophic level and habitat use) (Ruggiero & Kitzberger, 2004), can offer valuable insights into the underlying ecological processes that influence species distributions. While there is limited consensus about appropriate scales and important predictors for species belonging to specific functional groups (i.e., groups of species sharing similar ecological characteristics and roles in the environment), understanding the scaling relationships between geodiversity and species traits can help to identify potential predictors and scales that are relevant for specific groups of species. To address this need, I assessed how traits influence the species-environment relationships with scale-dependent geodiversity variables.

Recent advances in satellite remote sensing and climate reanalysis products, like MERRAclim (Vega et al., 2017) and CHELSA bioclimatic variables (Karger et al., 2017), as well as methods to measure spatial heterogeneity offer opportunities to improve the performance of SDMs and the conservation assessments derived from their outputs. For example, gradient surface metrics (e.g., average roughness, root mean square height, surface kurtosis, etc.) can capture spatial heterogeneity at varying spatial scales for any raster dataset (e.g., through the '*geodiv*' R package; Smith et al., 2021) and these rasters can be incorporated into SDMs. These measures of geodiversity now enable us to capture factors important for species distributions at finer resolutions as demonstrated by some studies (Bailey et al., 2018; Blair et al., 2022b).

Consequently, these metrics have important implications for understanding both species distributions and the overall patterns of biodiversity (Parks & Mulligan, 2010; Bailey et al., 2018). Using climate reanalysis and remotely sensed products in combination with gradient surface metrics may improve the performance of SDMs.

Here I examined the influence of scale-dependent geodiversity variables on the performance of SDMs and evaluate the ability of these variables to explain species-environment relationships for mammals in the Northern Andes – a region characterized by high topographic and climatic heterogeneity – primarily in Colombia, one of the world's most biodiverse countries. I compare the performance of MaxEnt SDMs generated using CHELSA bioclimatic variables only, to those additionally including geodiversity variables quantified at multiple scales. I aimed to determine: 1) whether scale-dependent geodiversity improves understanding of species-environment relationships and SDM performance, 2) if there are scales at which geodiversity consistently improves model performance or species in different functional groups (i.e., species exhibiting similar body mass and dietary preferences); and 3) whether the species-geodiversity relationship differs by biogeographic region.

I expected that:

1. Incorporating geodiversity variables computed at varying spatial scales surrounding species occurrence points in SDMs will improve model predictions as well as our understanding of species-environment relationships, in line with the principles of environmental filtering theory (Kraft et al., 2015; Pearson et al., 2018). This theory suggests that species distributions are shaped by a filtering process involving multiple abiotic and biotic factors (e.g., dispersal barriers, habitat configuration, climatic variation, competitors, etc.; Guisan & Thuiller, 2005; Soberón & Nakamura, 2009). By including

geodiversity variables in SDMs, I aim to capture the spatial heterogeneity associated with many of these filtering processes and gain a deeper understanding of the complex dynamics that govern species occurrence. Further, the relationship between geodiversity and species-environment relationships is likely to exhibit scale-dependency (Stein et al., 2014; Bailey et al., 2017; Zarnetske et al., 2019).

2. The scales at which geodiversity best explains species distributions will differ among functional groups. Considering that functional traits are closely tied to how species perceive and interact with their environment, I anticipate that the effects of geodiversity will vary depending on species' specific functional characteristics, such as body mass and feeding type (Violle et al., 2007). Smaller mammals may show stronger associations with fine-scale geodiversity, while larger mammals may respond more to geodiversity at coarser scales, which reflects their dispersal capabilities (McNab, 1963). In terms of feeding habits, fruit/nectar specialists and folivores may be more sensitive to fine scale geodiversity variations as their home ranges are typically smaller, which for folivores is due to the energetic costs of a leaf-based diet (Milton & May, 1976; Tucker et al., 2014). In contrast, omnivorous or frugivorous mammals, which likely have to “hunt” for food, may exhibit a more flexible response to geodiversity at both fine and coarse scales, as they can adapt to a wider range of available resources and their home ranges are typically larger than those of folivores (McNab, 1963; Milton & May, 1976).
3. The relationship between species and geodiversity will vary across different biogeographic regions given the differing levels of heterogeneity across the Northern Andes (Hernandez et al., 2006). The unique environmental conditions, habitat types, and ecological dynamics of each region are likely to shape the species-geodiversity

relationship differently and I expect geodiversity to be more important for species in ecoregions with high topographic complexity.

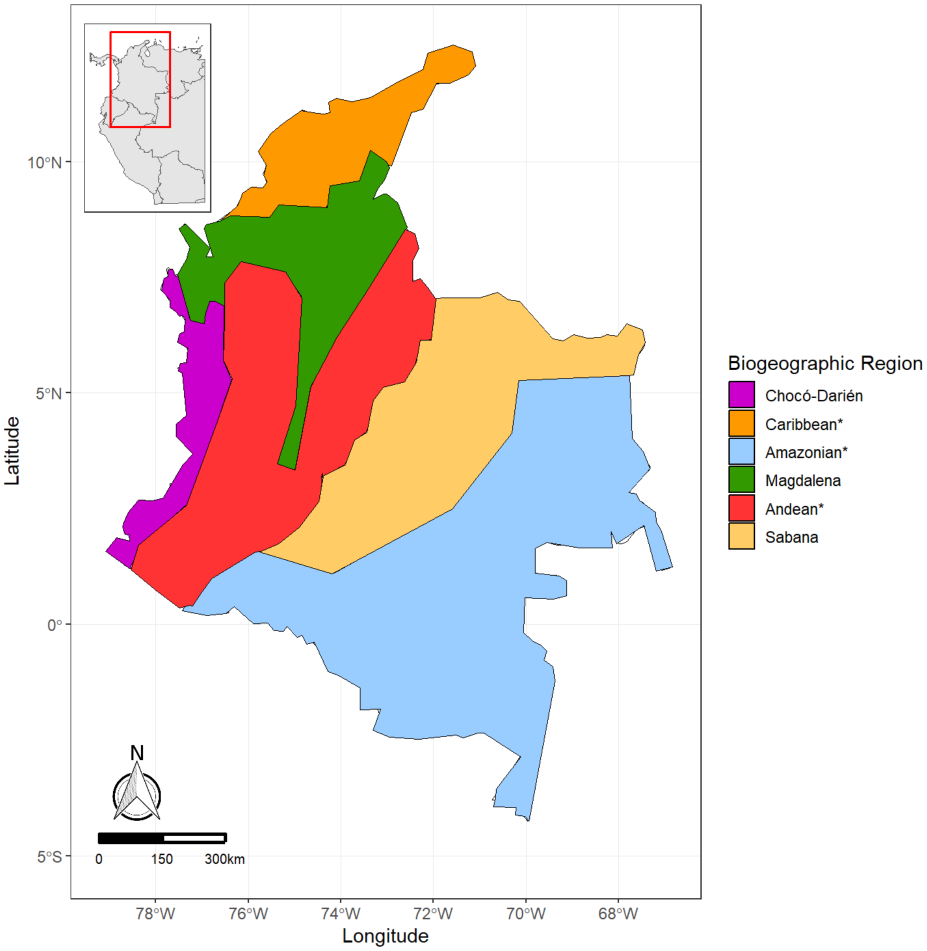
2.3 Methods

2.3.1 Study Region

Colombia is in the northwest corner of South America. With only 0.77% of the planet's land cover and approximately 10% of the world's biota, Colombia is recognized as one of the world's megadiverse countries. This diversity stems from its unique geographical location, providing it with increased sunlight exposure year-round compared to the southern regions of South America as well as its diverse geomorphology, which leads to multiple ecosystem types (Hernandez et al., 2006). Additionally, it serves as a crucial bridge between South and Central America, facilitating the exchange and intermingling of diverse species (Hernandez et al., 2006). Because of all these characteristics, Colombia offers a unique opportunity to study the influence of geodiversity on the distribution of species.

Here I focus on five primary biogeographic regions which represent distinct ecological zones with varied topography and climatic conditions where most of the study species are distributed (Figure 2.1). The Andean region, located in the central and western part of the country, features the Andes Mountain range with elevations above 1,000 meters above sea level (González-Orozco, 2021; referred to as Páramo). In contrast, the Chocó-Darién region encompasses the Pacific hyper-humid coastal and alluvial plains, while the Sabana region in the east experiences seasonal flooding and includes the vast Llanos Orientales plain extending into Venezuela (Hamilton et al., 2004). The Amazonian region covers the southeastern part of Colombia, mainly comprising the Amazon rainforest (González-Orozco, 2021; referred to as Imerí), and the Magdalena region represents a transitional zone between the central, eastern, and western Pacific Andean regions (González-Orozco, 2021). These diverse habitats support a high level of biodiversity and endemism (Olson & Dinerstein, 1998; Myers et al., 2000), making Colombia a priority for biodiversity conservation and an ideal study site for evaluating the role of geodiversity in shaping species distributions.

Figure 2.1: Major biogeographic regions within Colombia based on regions defined by González-Orozco et al. (2021). Fine-scale details have been simplified for clarity, while still depicting the main biogeographic regions. Region names denoted with (*) have had their names modified from the original publication to ensure easier recognition and understanding.



2.3.2 Study Species and Occurrence Records:

Our study includes a diverse set of mammal species (encompassing 17 genera) with ranges spanning most of the biogeographic regions mentioned above (Table 1). I obtained validated occurrence data for 29 Colombian species through BioModelos (Velásquez-Tibatá et al., 2019), an innovative digital tool that facilitates communication and collaboration among biodiversity experts in the development of species distribution models. In addition, I also obtained expert maps from BioModelos that were generated using the same set of occurrence records. These maps represent the most up-to-date version of species distribution ranges in Colombia. These 29 species included all primates with over 15 occurrence records (following post-spatial thinning, as described in the Modeling section), as well as the Andean bear (*Tremarctos ornatus*) (Akçakaya et al, 2019). To complement this dataset, I referenced recent publications (Gerstner et al., 2018; Medrano-Vizcaíno & Gutiérrez-Salazar, 2020; Ramírez-Chaves et al., 2022) to obtain occurrence data for two additional species, namely the olinguito (*Bassaricyon neblina*) and the western mountain coati (*Nasuella olivacea*), as their expert maps from BioModelos were still awaiting validation. By utilizing this subset of mammal species, which encompasses a diverse range of environmental roles and requirements, and by incorporating validated occurrence data and expert-made range maps, this study offers a comprehensive assessment of geodiversity as a tool to enhance SDMs.

2.3.3 Climate and topography data:

Of the 19 bioclimatic variables from CHELSA (Karger et al., 2017), which summarize annual, seasonal, and monthly trends of temperature and precipitation data, I selected four: bio5 (maximum temperature of the warmest month), bio6 (minimum temperature of the coldest month), bio13 (precipitation of the wettest month), and bio14 (precipitation of the driest month), which represent temperature and precipitation extremes that may be limiting to tropical species, particularly those in montane regions (Guevara et al., 2018). Additionally, I included the MODIS-derived mean annual cloud cover product (Wilson & Jetz, 2016), which has been demonstrated to enhance species distribution models for species in the Northern Andes (Wilson & Jetz, 2016). I also included the Shuttle Radar Topography Mission (SRTM; Farr et al., 2007) digital elevation model (SRTM30) to capture fine-scale variations in terrain known to influence species distributions (Coblentz & Riitters, 2004; Oke & Thompson, 2015; Leong et al., 2018). All variables were used at a spatial resolution of 30 arcseconds (~1 km^2).

2.3.4 Geodiversity data:

I used the 'geodiv' package (version 1.0.5; Smith et al., 2021) in R (version 4.2.3; Olson & Dinerstein, 1998) to calculate the root mean square roughness (SQ) of the areas surrounding each pixel for the same variables defined above (variables denoted with *_sq), and those neighborhood calculations became the value of the focal pixel. These neighborhood calculations were conducted over varying distances, which I will henceforth refer to as spatial grains, however it is important to note that the resolution of each geodiversity variable remained 30 arcseconds. Spatial grains of these neighborhood calculations ranged from 3 km, which characterizes the spatial scale at which most species in this study experience their environment, to 33 km, which is large enough to likely encompass (at least seasonally) the home range of the species with the largest dispersal capacity in this study, the Andean bear (*Tremarctos ornatus*; Castellanos, 2011).

2.3.5 Analysis:

Modeling – Specifically, I used MaxEnt, a machine learning approach, to generate species distribution models, as it is a widely used and effective approach, particularly with presence-only data (Urbina-Cardona et al., 2019). Following a similar methodology from Bailey et al. (2018), I opted for a machine learning approach, given the intricate and relatively unknown relationships between species distributions and geodiversity variables in this study. Compared to other modeling methods, MaxEnt has numerous advantages, including its ability to handle complex predictor-species relationships, and its insensitivity to collinearity among variables (Elith et al., 2011; Dormann et al., 2013; Blair et al., 2022a) owing to a regularization parameter that minimizes the influence of correlated variables by shrinking regression coefficients (Elith et al., 2011).

To set up and pre-process data before running species distribution models, I used the R package ‘wallace’ (version 2022.09.09.1; Kass et al., 2018), which is a GUI based ecological modeling software that allows for the building, evaluating, and visualizing of species distribution models in a guided and stepwise fashion. I used the base code for Wallace and their stepwise workflow for most of the data pre-processing pipeline. However, to increase computational efficiency and mitigate sampling bias, I spatially thinned occurrence records prior to using Wallace (usually a step within Wallace). To remove potential sampling biases and artefactual spatial autocorrelation, I used the 'SpThin' package (version 0.2.0; Aiello-Lammens et al., 2015) to thin occurrence records at a 10 km distance. This distance was deemed to be appropriate given the steep elevational gradients and overall heterogeneity of the region (Anderson & Raza, 2010; Boria et al., 2014; Gerstner et al., 2018). Next, as part of the Wallace pipeline, I created species-specific study regions for each species by generating 1-degree point buffers around all occurrence records to create a single unified polygon. These species-specific regions were used as the environmental background for randomly sampling 10,000 background points. Finally, I built and evaluated models using the R package ‘ENMeval’ (version 2.0.4; Kass et al., 2021).

To train and test the models, I utilized two distinct methods. For species with 25 or fewer records, I implemented the 'jackknife' approach, which involves leaving each occurrence record out of the model once to use for testing, as a special case of *k*-1 cross-validation (Shcheglovitova

& Anderson, 2013). Model statistics were then averaged across all iterations. For species with more than 25 records, I used standard *k*-1 cross-validation. To ensure consistency, I parameterized all models with the same regularization multiplier and feature class of 'LQ1', which strikes a balance between capturing the complexity of the response to environmental conditions and avoiding excessive complexity. While I acknowledge the importance of species-specific tuning to obtain optimal species distribution models (Anderson & Gonzalez, 2011), tuning would render comparisons across model sets impractical since each set could potentially be parameterized differently for the same species. If I had performed species-specific tuning, differences between models would not be attributed to the inclusion of geodiversity variables, but rather to differences in regularization and feature class selections.

Model Sets – Analyses were performed for two model sets:

1. Local pixel climate and topography predictors: This set included six variables (described in the Climate and Topography data section) representing local pixel climate and topographic conditions across the study area (30 arcseconds; ~1km).
2. Local pixel climate and topography predictors (6) + geodiversity (SQ of neighborhood) versions of the same predictors (set 1): In this set, geodiversity variables were incorporated by calculating root mean square height (SQ) versions of the local pixel climate and topography predictors (same as set 1). The variability around each local pixel was calculated at different spatial grains, specifically at 3 km, 9 km, 15 km, 21 km, 27 km, and 33 km and assigned to each local pixel. Importantly, the spatial resolution of each geodiversity variable remained 30 arcseconds.

Each species had a total of seven model runs: one local level run, and six runs with geodiversity predictors additionally incorporated at each spatial grain. Like Schnase et al. (2021), I performed three replicates of each run per species and averaged all modeling outputs to minimize any random variation in performance statistics and permutation importance values.

Model Evaluation – The Continuous Boyce Index (CBI; Hirzel et al., 2006) and Area Under the Receiver Operating Characteristic Curve (AUC) are commonly used to evaluate the performance of species distribution models. That being said, AUC has been criticized for its insensitivity to rare species with low occurrence records, leading to inflated scores in such cases (Lobo et al., 2008; Peterson et al., 2008). CBI measures the agreement between model predictions and a random distribution of observed presences across prediction gradients, ranging from -1 (perfect disagreement) to 1 (perfect agreement), with values above 0 indicating better-than-random performance (Hirzel et al., 2006). It is designed specifically for presence-only data, is not influenced by prevalence, and does not rely on a presence/absence threshold and therefore I chose to use it for this study. To investigate the influence of different spatial grains on model performance, I averaged the performance across all species for each spatial grain. I assessed the significance of performance changes across spatial grains using the Mann-Whitney U test.

Grouping analyses – Additionally, I categorized species into groups (trait-based and biogeographic) to assess whether the species-geodiversity relationships varied by traits and biogeographic region. For traits, I used quantiles of mass and feeding type (assigned based on % prevalence in diet; Wilman et al., 2014), aiming to identify scales at which model performance was higher (Table 2.1). For diet, frugivores were defined as species consuming ≥ 60 % fruit, while the fruit/nectar specialists were species whose diet consisted of ≥ 60% fruit and nectar combined. Further, I grouped species by biogeographic region. Subgroups were created for the Amazonian region due to its breadth and differences in species distributions in that area. Amazonian-1, are restricted distributions near the foothills of the Cordillera Oriental.

Amazonian-2 are large distributions with a significant portion of the range at the foothills of the Cordillera Oriental. Amazonian-mix, are large distributions combining Amazonian, Andean, Sabana and Magdalena, and Amazonian are distributions primarily in the Amazon.

Table 2.1: Functional groups for study species organized by biogeographic region. Species were grouped by diet and quartiles of body mass (Wilman et al. 2014; Helgen et al. 2013) as well as biogeographic region (González-Orozco, 2021). For diet, frugivores were defined as species consuming ≥ 60 % fruit, while fruitnect were species whose diet consisted of ≥ 60% fruit and nectar combined. Subgroups were created for the Amazonian region due to its breadth and differences in species distributions in that area. Amazonian-1, are restricted distributions near the foothills of the Cordillera Oriental. Amazonian-2 are large distributions with a significant portion of the range at the foothills of the Cordillera Oriental. Amazonian -mix, are large distributions combining Amazonian, Andean, Sabana and Magdalena, and Amazonian are distributions primarily in the Amazon.

Species	Diet group	Body mass (g)	Body mass quartile	Biogeographic group
Alouatta palliata	folivore	7274.95	Q4	Chocó-Darién
Ateles fusciceps	frugivore	9100	Q4	Chocó-Darién
Cebus capucinus	omnivore	2733.32	Q3	Chocó-Darién

Table 2.1 (cont'd)

Aotus zonalis	omnivore	889	Q2	Chocó-Darién
Saguinus geoffroyi	fruitnect	486.5	Q1	Chocó-Darién
Cebus albifrons	omnivore	2629	Q3	Amazonian
Pithecia hirsuta	frugivore	387	Q1	Amazonian
Cacajao melanocephalus	frugivore	3100	Q4	Amazonian
Cheracebus lucifer	omnivore	3000	Q3	Amazonian
Leontocebus fuscus	frugivore	6299.99	Q4	Amazonian
Pithecia milleri	omnivore	2240.99	Q3	Amazonian-1
Plecturocebus caquetensis	omnivore	1537.52	Q3	Amazonian-1
Plecturocebus ornatus	frugivore	1170.5	Q2	Sabana
Cheracebus lugens	omnivore	1500	Q2	Amazonian-2
Plecturocebus discolor	omnivore	915	Q2	Amazonian-2
Ateles belzebuth	frugivore	5000	Q4	Amazonian-2
Cebuella pygmaea	nectarivore	125	Q1	Amazonian-2
Lagothrix lagotricha	omnivore	1011.32	Q2	Amazonian-mix

Table 2.1 (cont'd)

Saimiri cassiquiarensis	omnivore	743.24	Q1	Amazonian-mix
Sapajus apella	omnivore	2500	Q3	Amazonian-mix
Aotus brumbacki	omnivore	875	Q2	Amazonian-mix
Alouatta seniculus	folivore	6145.54	Q4	Amazonian-mix
Saguinus leucopus	fruitnect	440	Q1	Magdalena
Cebus versicolor	omnivore	2629	Q3	Magdalena
Ateles hybridus	frugivore	6394.85	Q4	Magdalena
Aotus griseimembra	omnivore	872.99	Q1	Magdalena
Saguinus oedipus	fruitnect	430	Q1	Magdalena
Tremarctos ornatus	frugivore	140000.63	Q4	Andean
Nasuella olivacea	fruitnect	1339.99	Q2	Andean
Aotus lemurinus	omnivore	872.99	Q1	Andean
Bassaricyon neblina	frugivore	872	Q1	Andean

Post-processing of SDMs and model comparisons – To generate binary suitability maps for each species I thresholded both the model without geodiversity (henceforth termed 'non-geodiversity models') and the optimal geodiversity model (i.e., the model at the spatial grain with the highest CBI for a species) based on either the minimum training presence (MTP) or the 10% omission rate, depending on the number of occurrences (MTP for less than or equal to 25, and 10% omission for greater than 25). Next, I used known information about species ranges and structural barriers as provided by the International Union for Conservation of Nature (IUCN; IUCN 2022) as well as obvious structural boundaries within the expert maps, to create range boundary polygons beyond which I excluded areas where the species was unlikely to disperse. These post-processed models were then visually inspected, and comparisons were made between expert maps (available in BioModelos and based on MaxEnt models & expert opinion or land cover types), non-geodiversity models, and optimal geodiversity models. I evaluated gain and loss in predicted areas, omission rates, and Schoener's D, a measure of spatial overlap, for each model set to understand differences in all predictions.

2.4 Results

In this study, the incorporation of geodiversity variables improved the average predictive performance of the SDMs. On average, the CBI of the non-geodiversity models was 0.80 and the CBI of geodiversity models was 0.93. Specifically, I observed an average increase of 17.2% in the Continuous Boyce Index (CBI) across the optimal models for all species when geodiversity variables were included. When compared with non-geodiversity models at 1 km (reflecting local pixel values), all other models improved in performance across all evaluated spatial grains (Figure 2.2; Mann-Whitney U tests, $p < 0.05$). However, I identified an interesting exception for the western mountain coati *(Nasuella olivacea*), where model performance was found to be higher in the model without geodiversity variables compared to the "optimal" geodiversity model (Table 2.2). Additionally, when comparing the average model performance across all spatial grains, I found marginal superiority for finer grains, particularly 3 km and 9 km. However, there was no difference in average model performance among these spatial grains.

Figure 2.2: The average Continuous Boyce Index (CBI), represented by a diamond, reflects the mean value, while the upper and lower whiskers depict the range of observations within 1.5

times the interquartile range above the upper hinge or below the lower hinge. This provides an overview of the variations in model performance across different spatial grains and highlights the impact of incorporating geodiversity variables on the CBI. At every spatial where geodiversity was tested (3-33 km), all models with geodiversity variables increased in CBI when compared with the local 1 km pixel non-geodiversity models (Mann-Whitney U tests, $p < 0.05$).

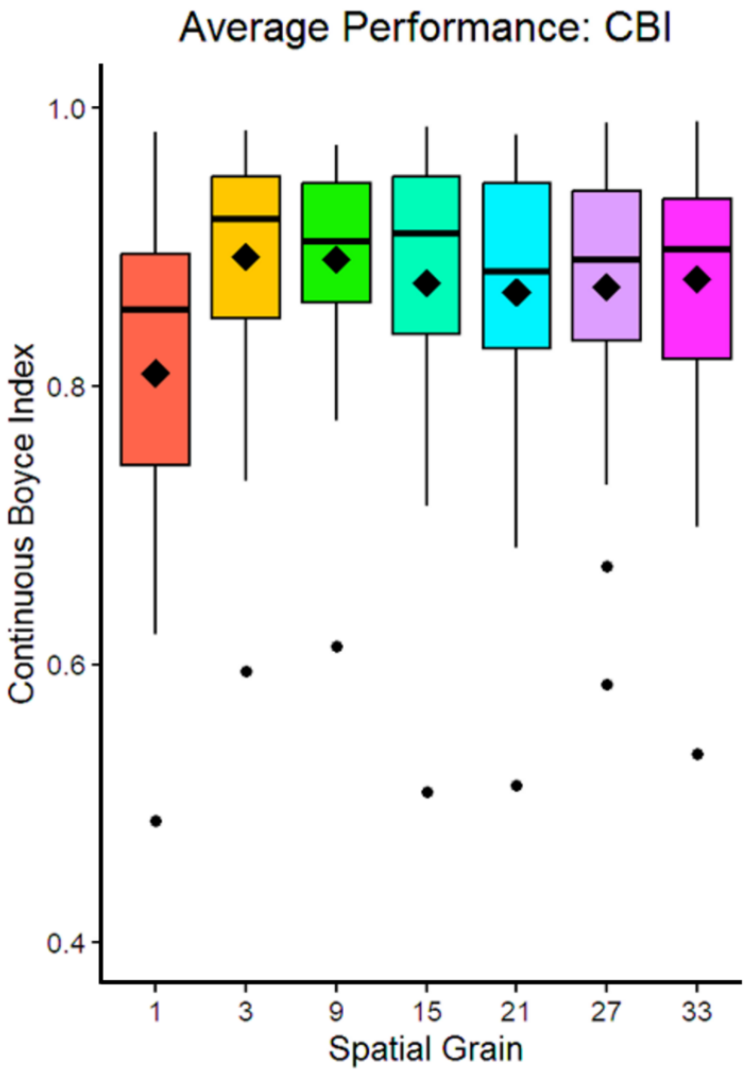
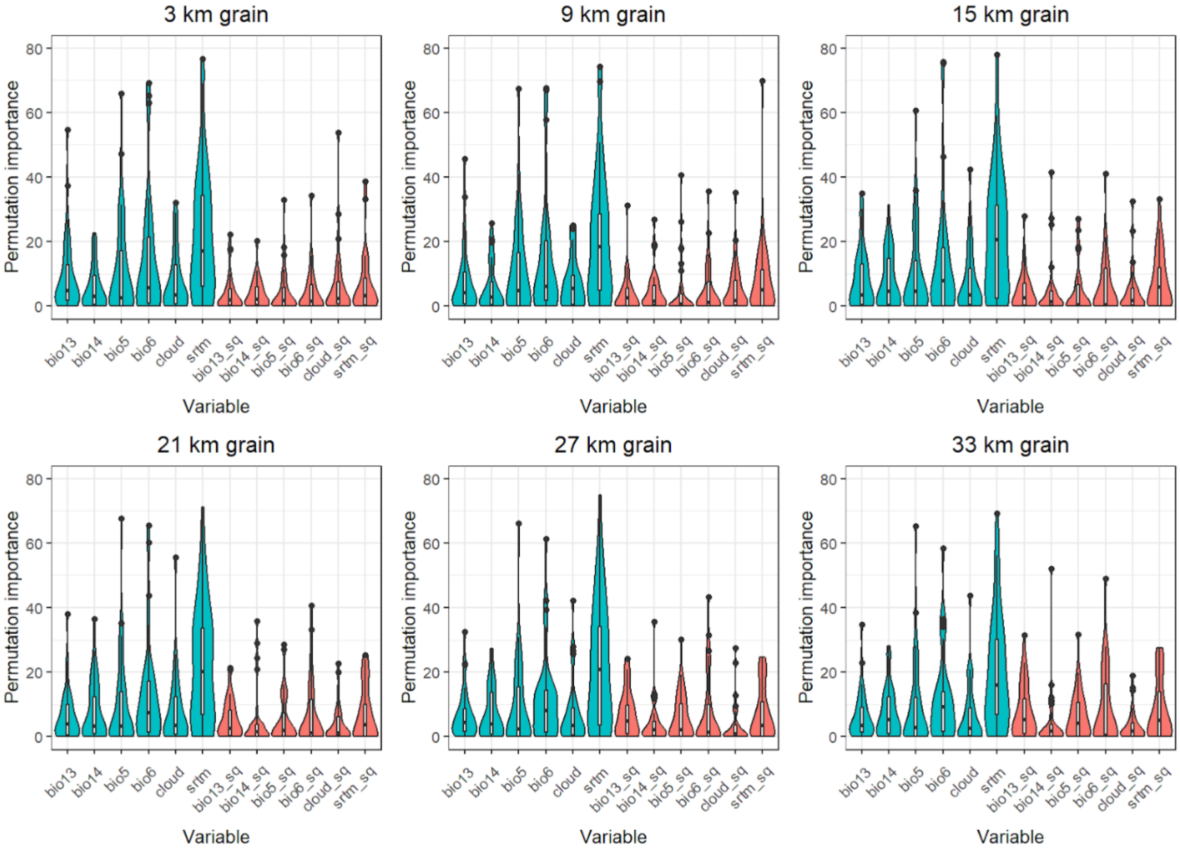


Figure 2.3: Permutation importance values (i.e., impact or contribution of individual environmental variables in a MaxEnt) across geodiversity variables calculated at different spatial grains. Blue bars indicate non-geodiversity variables and red indicate geodiversity variables. The shape of each bar represents the density distribution of the permutation importance values for each predictor across all species.



When assessing the permutation importance of variables, clear differences were observed between non-geodiversity and geodiversity variables. Non-geodiversity variables had higher average permutation importance (11.57%) compared to geodiversity variables (5.57%) across all species distribution models (Figure 2.3). Among the geodiversity variables, several variables stood out with higher average permutation importance (> 5%), including srtm_sq (7.48%), bio6_sq (6.19%), bio13_sq (5.39%), and bio5_sq (5.15%) (Table 2B.1). Notably, the geodiversity variable srtm_sq exhibited consistently higher average permutation importance than the non-geodiversity variable bio14, indicating that there may be instances where geodiversity variables are more informative than non-geodiversity variables.

I found some evidence for scale-dependency in the importance of the explanatory variables. Non-geodiversity variables generally showed a decreasing trend in importance as the spatial grain of geodiversity variables increased, except for bio14 and srtm, indicating their diminishing influence as geodiversity was incorporated at coarser scales (Table 2B.1). Further, certain geodiversity variables had a modest yet noticeable increase in importance with increasing spatial scale, such as bio5_sq (1.84%), bio6_sq (3.73%), bio13_sq (3.75%), and srtm_sq (1.55%), whereas the importance of cloud_sq decreased (3.1%) as the spatial scale increased (Table 2B.1). The frequency at which certain geodiversity variables were incorporated into models also varied with spatial scale. Bio5_sq and bio13_sq were more frequently included in models at coarser scales, while bio6_sq and cloud_sq were more frequently incorporated at finer scales. In general, the variables srtm and srtm_sq were frequently included in the top models across scales, indicating their robust influence in capturing species-environment relationships. Overall, geodiversity variables ranked within the top three variables in terms of permutation importance for the optimal geodiversity models of 23 species (Table 2.2), with an average permutation importance of 19.7%.

The responses of individual species to geodiversity variables at different spatial scales were highly variable, highlighting the complexity of species-environment relationships. Notably, models of species such as the common woolly monkey (*Lagothrix lagotricha*) and the Andean bear (*Tremarctos ornatus*) had substantial increases in the importance of geodiversity variables with scale. For instance, in *L. lagotricha* models, the permutation importance of bio6_sq, increased from 0% at 3 km to 19.97% at 33 km—the spatial grain that resulted in the highest model performance for this species. Similarly, as the spatial scale increased from 3 km to 33 km for *T. ornatus* models, the permutation importance of srtm_sq increased from 3.49% to 8.4%, and for bio6_sq, increased from 4.5% to 21.35%. Interestingly, for *T. ornatus*, the model with the highest performance was at 3 km spatial grain. While scale dependence was evident for certain variables, the magnitude and direction of the effects varied considerably by species.

Table 2.2: Percent increase in Continuous Boyce Index (CBI) model performance with inclusion of geodiversity variables. Percent increase in model performance achieved by incorporating geodiversity variables compared to models without geodiversity variables. The optimal geodiversity grain where model performance was highest for each species and whether this grain is idiosyncratic when considering optimal grains for specific traits (Figure 2.4) is also noted. Species denoted with (*) indicate those for which geodiversity variables were ranked within the top three in terms of permutation importance.

Species	**CBI: non-geodiversity**	**CBI: geodiversity**	**CBI: % increase**	**Optimal geodiversity grain (km^2)**	**Optimal grain idiosyncratic? (yes/no)**
*Alouatta palliata**	0.79	0.89	13.31	3	no
Alouatta seniculus	0.98	0.99	0.82	33	no
*Aotus brumbacki**	0.74	0.90	20.96	15	no
*Aotus griseimembra**	0.94	0.98	4.46	27	yes
*Aotus lemurinus**	0.77	0.91	18.60	15	no
*Aotus zonalis**	0.75	0.88	18.08	27	yes
*Ateles belzebuth**	0.88	0.96	8.52	15	no
*Ateles fusciceps**	0.87	0.93	7.14	15	no
Ateles hybridus	0.96	0.97	1.67	9	no

Table 2.2 (cont'd)

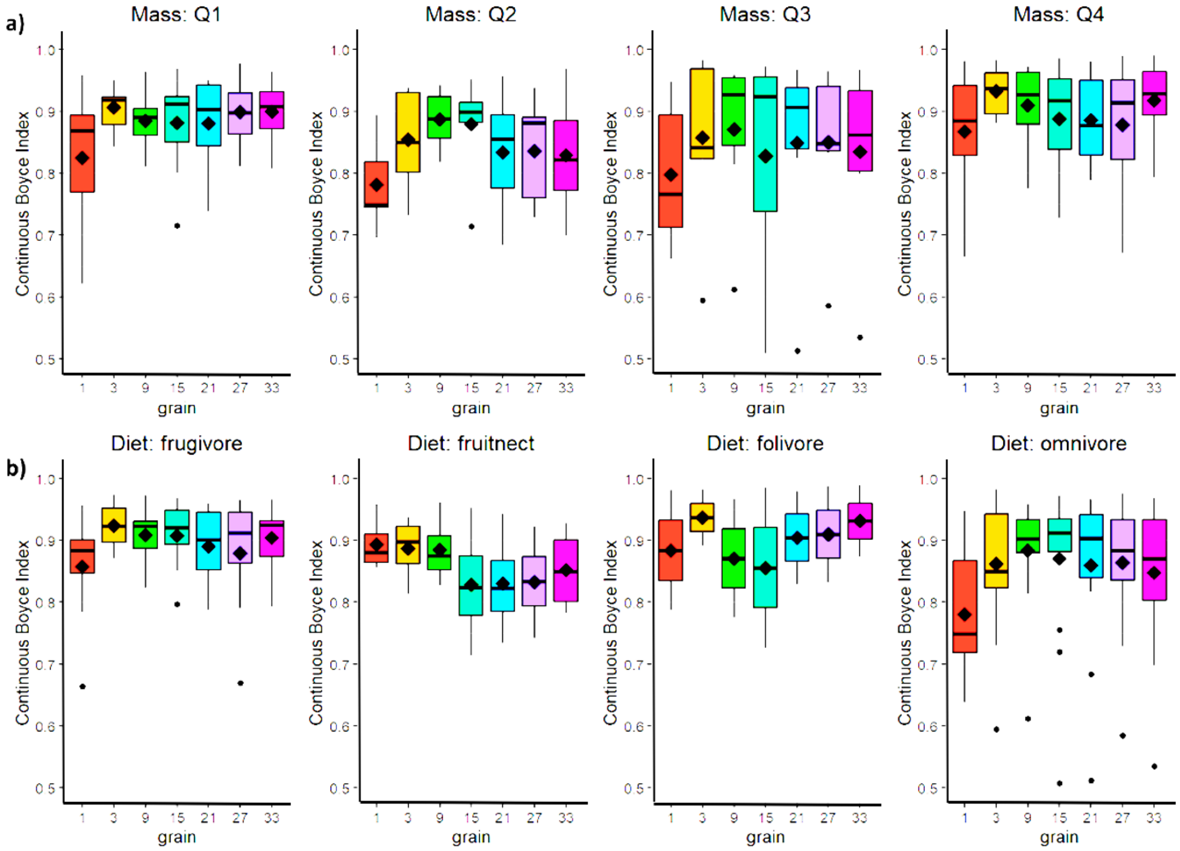
Bassaricyon neblina	0.89	0.89	0.04	15	no
*Cacajao melanocephalus**	0.66	0.93	39.31	33	no
*Cebuella pygmaea**	0.62	0.93	50.05	33	-
*Cebus albifrons**	0.95	0.97	2.78	3	yes
*Cebus capucinus**	0.91	0.97	7.24	15	no
*Cebus versicolor**	0.88	0.96	9.20	3	no
Cheracebus lucifer	0.68	0.88	28.71	9	no
*Cheracebus lugens**	0.75	0.94	24.97	3	yes
*Lagothrix lagotricha**	0.85	0.97	13.43	33	yes
*Leontocebus fuscus**	0.49	0.90	84.12	3	no
Nasuella olivacea	0.89	0.83	-7.45	9	no
Pithecia hirsuta	0.88	0.97	9.58	15	no
*Pithecia milleri**	0.77	0.83	8.66	27	yes
*Plecturocebus caquetensis**	0.74	0.82	11.00	3	no
*Plecturocebus discolor**	0.70	0.89	27.43	9	no
Plecturocebus ornatus	0.78	0.94	19.62	3	no
Saguinus geoffroyi	0.87	0.89	2.84	33	no
Saguinus leucopus	0.96	0.96	0.42	9	no
*Saguinus oedipus**	0.86	0.92	7.32	3	no
*Saimiri cassiquiarensis**	0.64	0.94	47.05	21	yes
*Sapajus apella**	0.66	0.98	48.84	3	no
*Tremarctos ornatus**	0.93	0.97	5.15	3	no

The responses of individual species to geodiversity variables at different spatial scales were highly variable, highlighting the complexity of species-environment relationships. Notably, models of species such as the common woolly monkey (*Lagothrix lagotricha*) and the Andean bear (*Tremarctos ornatus*) had substantial increases in the importance of geodiversity variables with scale. For instance, in *L. lagotricha* models, the permutation importance of bio6_sq, increased from 0% at 3 km to 19.97% at 33 km—the spatial grain that resulted in the highest model performance for this species. Similarly, as the spatial scale increased from 3 km to 33 km for *T. ornatus* models, the permutation importance of srtm_sq increased from 3.49% to 8.4%, and for bio6_sq, increased from 4.5% to 21.35%. Interestingly, for *T. ornatus*, the model with the highest performance was at 3 km spatial grain. While scale dependence was evident for certain variables, the magnitude and direction of the effects varied considerably by species.

I conducted additional analyses to evaluate the model performance of species belonging to specific functional groups, providing valuable insights into their relationships with the environment. These functional groups were defined based on quartiles of mass and diet preference. The results revealed that spatial grain had varying impacts on model performance within these functional groups. Specifically, when grouping species by mass, I observed that differences in model performance across spatial grains were relatively subtle. Quantile 1 and 4 species exhibited slight increases in average performance at both fine (3 km) and coarse (33 km) spatial grains, while Quantiles 2 and 3 showed higher performance at a finer scale of 9 km (Figure 2.4a). In contrast, when considering feeding types, I observed more pronounced differences in model performance across spatial grains. Folivores demonstrated the highest average performance at both fine (3 km) and coarse (33 km) spatial grains, while frugivores had highest average performance at low to intermediate scales (3 - 15 km) with another increase at 33 km, and fruit/nectar specialists displayed the highest performance at fine scales (3 - 9 km) (Figure 2.4b). Omnivores exhibited the highest average performance at low (9 km) to intermediate (15 km) scales (Figure 2.4b). However, similar to the analysis conducted on all species, it is important to emphasize that the optimal models for individual species within these functional groups sometimes exhibited idiosyncratic patterns (i.e., scale of optimal model performance for a species not aligning with highest performing grain sizes for at least one of the species' associated traits; 22.6% of species; Table 2.2), highlighting the species-specific responses to geodiversity variables and the considerations of spatial scale.

During the evaluation of the SDMs, I conducted a spatial assessment and compared them with expert-generated maps. Overall, the models incorporating geodiversity variables performed well and predicted distributions that aligned with species ecology. To assess model performance, I examined spatial gain and loss, Schoener's D, and the omission rate for expert, non-geodiversity, and optimal geodiversity models (models with the highest CBI for each species) (Table 2A.1). On average, expert models exhibited a higher omission rate (20%) compared to both the non-geodiversity models (14.1%) and geodiversity models (13.84%). The geodiversity models, on average, had slightly fewer omissions compared to the non-geodiversity models. Both the non-geodiversity and geodiversity models demonstrated substantial gains and losses compared to the expert models. Specifically, the geodiversity models showed slightly fewer gains (7.79%) but more losses (7.45%) than the non-geodiversity models (gains: 8.35%, losses: 5.28%). This indicates that, in general, the geodiversity models predicted less suitable areas than the expert and non-geodiversity models.

Figure 2.4: Boxplots of model performance for functional groups based on mass and diet preference. The functional groups were defined using quartiles of mass and diet information (Wilman et al. 2014, Helgen et al. 2013). The analysis reveals varying impacts of spatial grain on model performance within these groups. When considering mass, subtle differences in performance were observed across spatial grains, with Quantile 1 and 4 species showing slight average increases in performance at fine and coarse scales, while Quantile 2 exhibited higher performance at low and intermediate scales and Quantile 3 had higher performance at finer scales. In terms of feeding types, more pronounced differences in model performance were found. Folivores demonstrated the highest average performance at both fine and coarse scales, and frugivores had highest average performance at fine to intermediate scales. Fruit/nectar specialists had the highest average performance at fine scales. Omnivores exhibited the highest performance at low to intermediate scales. This figure excludes one nectivorous species.

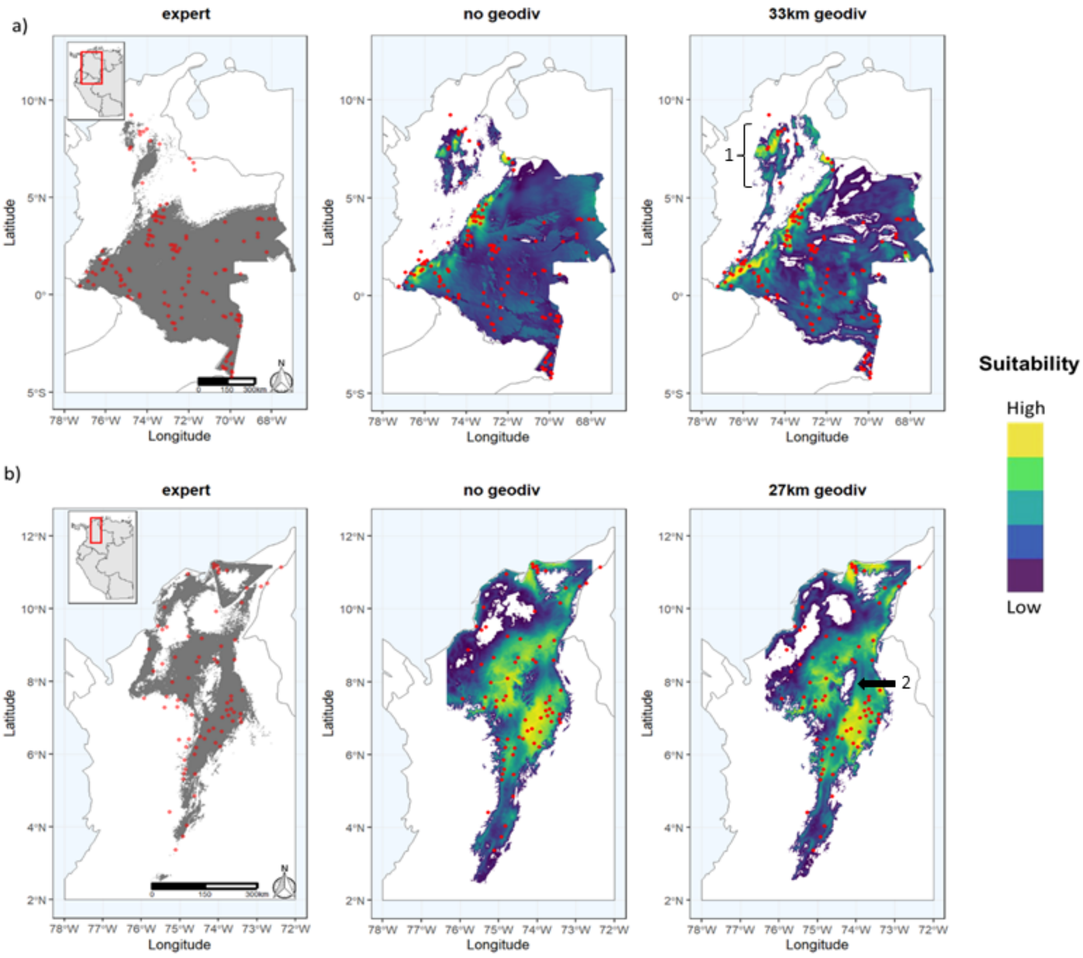


Furthermore, the assessment of Schoener's D values, representing the overlap between the geodiversity models and the expert models, revealed that, on average, the geodiversity models exhibited lower values (65.7%) compared to the non-geodiversity models (70.2%), indicating less overlap with the expert models (Table 2A.1). I also evaluated variations observed among species from different biogeographic regions. For instance, species in the Andean region generally showed more gains in suitable habitat in the geodiversity models, resulting in slightly higher Schoener's D values and lower omission rates compared to the non-geodiversity models (Table 2A.1). Similarly, species in the Magdalena and Amazonian-1 regions, also areas of high topographic heterogeneity, demonstrated a closer alignment between the geodiversity models and expert models. In contrast, the non-geodiversity models are better aligned with the expert models in the Amazonian-2 and Amazonian-mix regions, and marginally better aligned in the Chocó-Darién region, regions of lower topographic heterogeneity.

It is worth noting that many of the geodiversity models in these regions still produced ecologically reasonable predictions, despite differences from the expert-generated maps. For example, even though the geodiversity model for the Common woolly monkey (*Lagothrix lagotricha*) in general had a slightly higher omission rate compared to the expert model, it predicted increased suitable area and had lower omission in the northern part of the range when compared to both the non-geodiversity model and expert model (Figure 2.5a). Further, there was lower suitability in the Colombian Llanos (Sabana region; Figure 2.1), which are shown as not being suitable in the expert model. For the Gray-handed night monkey, *Aotus griseimembra*, both the non-geodiversity and geodiversity models better captured the occurrence records than the expert model, however the non-geodiversity model predicted suitability in high elevation areas whereas the geodiversity model does not, which is more closely aligned to the species' ecology as a lowland primate (Figure 2.5b). In contrast, there were instances where the geodiversity models exhibited limitations in capturing the full distribution range of certain species. For six species (*Cheracebus lugens, Pithecia hirsuta, Plecturocebus caquetensis, Saimiri cassiquiarensis, Cebus albifrons, Cebuella pygmae*) in the Amazonian regions, the optimal geodiversity models appeared to be constrained to the distribution of rivers in the

Amazon. This constraint was most evident from the average difference in Schoener's D of 0.17, indicating challenges in fully representing the complete distribution range of these species (Table 2A.1).

Figure 2.5: Comparisons of expert maps and thresholded models made without and with geodiversity variables for two species, the Common woolly monkey (*Lagothrix lagotricha*) and Gray-handed night monkey (*Aotus griseimembra*). Lighter colors indicate higher suitability and occurrence records for each species are denoted by red circles. Panel a) represents the expert map and thresholded models for species *L. lagotricha*, where there is less suitability in northeastern Colombia in the geodiversity model than the model without geodiversity and aligns better with the expert map. Predictions in the northernmost part of the species range in the geodiversity model (label 1) better capture the occurrence records than both the expert map and the non-geodiversity model. Panel b) represents the expert maps and models for *A. griseimembra*. Both the non-geodiversity and geodiversity models capture the occurrence records better than the expert model; however, the non-geodiversity model predicts suitability in high elevation areas whereas the geodiversity model does not (label 2), the latter being more closely aligned to the species' ecology as a lowland primate.



2.5 Discussion

This study provides valuable insights into the influence of geodiversity on species distribution models (SDMs) in the Northern Andes, encompassing both general patterns and species-specific responses. By incorporating geodiversity variables, I observed a significant impovement in SDM performance both statistically and spatially, which aligned with my expectations. While non-geodiversity variables predominantly shaped species distributions, certain geodiversity variables, such as topographic roughness and temperature and precipitation variations, exhibited notable influences. The response to geodiversity also exhibited species-specific variation, underscoring the individualistic nature of species-environment interactions and the challenge of predicting optimal performance grains based on shared traits. Furthermore, the influence of geodiversity varied across biogeographic regions, with topographic heterogeneity playing a pivotal role, while the efficacy of geodiversity predictors for enhancing model performance diminished in regions characterized by low heterogeneity.

2.5.1 Statistical model performance and scale-dependency (Expectation 1)

Consistent with my expectations, incorporating geodiversity variables yielded significant improvements in the statistical performance of SDMs, as indicated by an average increase of 17.2% in the Continuous Boyce Index (CBI), a measure of the predictive performance of the model, despite geodiversity variables having on average lower permutation importance than local level variables. These improvements were consistently observed across various spatial grains, which can be attributed to the complementary information provided by geodiversity variables, which capture variability of the physical environment. This suggests that geodiversity variables may capture crucial ecological information that goes beyond traditional predictors, providing valuable insights into species-environment relationships and improving the predictive power of the models. These findings highlight the potential of geodiversity variables in refining SDMs and enhancing our understanding of species distributions.

I observed clear differences between non-geodiversity and geodiversity variables in terms of permutation importance of variables. Non-geodiversity variables generally had higher average permutation importance (11.57%) compared to geodiversity variables (5.57%) across all species distribution models. This suggests that factors other than geodiversity, such as local level climate or topography, play a more prominent role in shaping species distributions. Among the geodiversity variables examined, topographic roughness (srtm_sq) exhibited the highest average permutation importance (7.48%), indicating its stronger influence on species-environment relationships. Additionally, geodiversity variables related to temperature and precipitation, namely min temperature of coldest month (bio6_sq), precipitation of wettest month (bio13_sq), and max temperature of warmest month (bio5_sq), had higher levels of permutation importance (> 5%), suggesting that variation in topographic roughness and certain climate extremes can also play a role in shaping these species’ distributions. Specifically, topographic roughness may indicate important dispersal limitations for species, while temperature and precipitation geodiversity variables reflect the spatial variation of important ecological drivers influencing species' physiological tolerances and resource availability.

Incorporating geodiversity variables in SDMs provides valuable complementary information and captures variability of the physical environment. However, it is important to recognize that the responses of individual species to geodiversity variables can be idiosyncratic. One notable example is the western mountain coati (*Nasuella olivacea*), for which the model without geodiversity variables statistically outperformed the "optimal" geodiversity model (Table 2.2). This suggests that factors other than geodiversity variables may play a more influential role in shaping the distribution patterns of this particular species. However, despite the lower CBI in the geodiversity model for this species, the spatial performance of the model remained ecologically reasonable and actually omitted fewer occurrence records than both the expert and non-geodiversity models (Table 2A.1). Therefore, although geodiversity may not be the dominant driver for this species (only 8% permutation importance), it still contributes valuable information that improves the model's ability to predict suitability.

I found some evidence for scale dependence in the importance of the geodiversity variables. Non-geodiversity variables generally decreased in importance as the spatial grain increased. At these coarser scales, some geodiversity variables become more influential in shaping species distributions, possibly reflecting the importance of broader landscape patterns and environmental gradients. The variables with the greatest increase in permutation importance with spatial grain were climate variables minimum temperature of the coldest month (bio6_sq) and precipitation of the wettest month (bio13_sq) (Table 2B.1), which is in line with other research showing the role of climate increases at broader scales (Pearson & Dawson, 2003; Blach-Overgaard et al., 2010). However, despite the average increase in permutation importance for some of these explanatory variables, the frequency at which geodiversity variables were incorporated into models also varied with spatial scale (Table 2B.1). Maximum temperature of the warmest month (bio5_sq) and precipitation of the wettest month (bio13_sq) were more frequently included in models at coarser scales, while minimum temperature of the coldest month (bio6_sq) and mean annual cloud cover (cloud_sq) were more frequently incorporated at finer scales (Table 2B.1). These findings suggest that some geodiversity variables may be more informative at specific scales, reflecting the scale- and species-dependent nature of geodiversity in shaping species distributions. However, similar to Bailey et al. (2018), the elevation variables (srtm and srtm_sq) consistently demonstrated high permutation importance and were frequently incorporated into models across scales (Table 2B.1). These variables, representing elevation and topographic roughness, respectively, likely play crucial roles in shaping species distributions across scales.

2.5.2 Functional groups and the influence of geodiversity (Expectation 2)

For certain species' traits, the optimal spatial grains of geodiversity aligned with my expectations whereas for others, they differed from expectations (Figure 2.4; Table 2.2). Specifically, I anticipated that larger-bodied species would have optimal models at larger spatial grains, and folivores would exhibit higher model performance at finer spatial grains. However, the results indicate that both fine and coarse spatial grains contribute to better model performance for these groups (Figure 2.4). Further, for omnivores I expected higher performance at fine and coarse grains, but for most species, model performance was highest at low to intermediate grains. Frugivores and fruit/nectar specialists did follow expected patterns, with frugivores having optimal grains across fine and coarse scales and fruit/nectar specialists having higher performance at fine grains.

In the case of the Andean bear (*Tremarctos ornatus*), characterized by its large body size (Figure 2.4a; Q4) and primarily frugivorous diet (Figure 2.4b; Frugivore), the optimal geodiversity model was at a spatial grain of 3 km, corresponding to the spatial grains associated with the highest average performance for these traits (Figure 2.4), which aligned with my original expectations for frugivorous species. Despite geodiversity variables contributing only 12.8% to the optimal model, this model exhibited closer alignment with the expert model than the non-geodiversity model. The importance of fine-scale variation may be particularly relevant for the Andean bear due to its specific habitat requirements and ecological adaptations. Being a large-bodied mammal, the Andean bear relies on extensive home ranges to meet its resource needs. Despite its body size, fine-scale variations in habitat conditions, including terrain roughness and microclimate gradients, play a crucial role in providing suitable foraging opportunities, shelter, and access to resources such as food and water (Figure 2.4b; Frugivore) (García-Rangel, 2012). The species is known to inhabit diverse montane ecosystems with rugged mountainous terrain, where fine-scale variations in terrain roughness and microclimate conditions may influence the availability of suitable den sites, access to preferred food sources, and the bear's ability to navigate through challenging landscapes (García-Rangel, 2012), likely leading to an optimal model with geodiversity variables reflecting the spatial grain of this variability.

The results also suggest that even if species are closely related in terms of their shared traits, they respond differently to geodiversity and their response also varies by spatial scale. This finding highlights the unique nature of species' interactions with their environment and suggests that shared traits do not necessarily determine species' responses to geodiversity variables and their scales of influence. For instance, based on the trait-grouping results in Figure 2.4, I expected the Common woolly monkey *(Lagothrix lagotricha*), an omnivorous species in the Q2 mass quartile, to have an optimal grain at low to intermediate scales. However, I found that the optimal spatial grain was 33 km, suggesting that omnivores like this species have a variable response to geodiversity across both fine and coarse scales, which supports my original expectation for this trait group. Geodiversity played a significant role for this species, with a notable permutation importance of 57.87%. The Common woolly monkey is primarily found in lowland primary terra firma forests, occasionally utilizing secondary and disturbed habitats, and they seasonally enter flooded forests to feed on fruits (Stevenson et al., 1994). Woolly monkeys have a diverse diet consisting of fruits, arthropods, leaves, seeds in unripe fruits, flowers, and other minor items. The composition of their diet varies throughout the year, depending on fruit abundance, which tends to be higher in the rainy season when precipitation is higher. During periods of fruit scarcity, they rely more on leaves, unripe fruits, and flowers (Stevenson et al., 1994). These dietary preferences and seasonal movements may influence the optimal spatial grain of the geodiversity model, where the permutation importance of spatial variation in bio13 (precipitation of the wettest month; bio13_sq) was actually higher (6.75%) than the non-geodiversity version of that variable (5.49%). The broader-scale patterns of fruit availability and distribution within the lowland forest landscape might be better captured at a spatial grain of 33 km, allowing for more robust predictions of suitable habitats for the species (Figure 2.5a).

2.5.3 Biogeographic regions and influence of geodiversity in SDMs (Expectation 3)

I found support for my expectation that species-geodiversity relationships differed by biogeographic region, likely due to biogeographic differences in habitat heterogeneity. Diverse and varied landscapes provide more opportunities for geodiversity variables to capture important ecological patterns (Lawler et al., 2015). Regions with high topographic geodiversity, such as the Andean, Magdalena, and Amazonian-1 regions, likely exhibit greater heterogeneity in terms of topography and climate. This heterogeneity provides a range of microhabitats and ecological niches, allowing species to occupy diverse habitats within these regions. Conversely, regions in the Amazonian, and Chocó-Darién may have different characteristics, such as less pronounced heterogeneity or a higher proportion of homogeneous habitats. Specifically for the Amazonian, Amazonian-2, and Amazonian-mix habitats certain species had distributions constrained to rivers. It is possible that this issue stems from overfitting to noise in areas with generally low habitat heterogeneity (Merow et al., 2013). Rivers, being prominent features in the landscape, may introduce a significant amount of variability that is unrelated to the ecological requirements of these species (excluding *Cebuella pygmaea*). This can lead to models that overly associate species presence with riverine habitats, incorrectly constraining their distributions along waterways and omitting many occurrence records (species: *Cheracebus lugens, Pithecia hirsuta, Plecturocebus caquetensis, Saimiri cassiquiarensis, Cebus albifrons, Cebuella pygmae*; Table 2.1). This may have led to discrepancies between the geodiversity models for these species and the expert models, which consider a broader range of ecological factors and account for species' ecological requirements beyond just the presence of rivers. Due to this, caution should be exercised to avoid overfitting to noise or artifacts in the data, especially in areas with low habitat heterogeneity where there is less benefit to using these kinds of explanatory variables.

2.5.4 Conclusions:

The inclusion of geodiversity variables in species distribution models (SDMs) in this study offers valuable insights into the role of spatially-varying environmental heterogeneity on species distributions. Model performance across different spatial grains indicated that geodiversity had the strongest influence on most species at finer scales, particularly at 3 km and 9 km. Two thirds of all species had optimal geodiversity models at spatial grains of 3 - 15 km, with only one third of species having optimal spatial grains of 27 km and above (Table 2.2). Incorporating geodiversity variables at fine to intermediate scales may be sufficient to increase model performance for many species and may better represent species-environment relationships and environmental filtering at these scales. To effectively implement this approach, careful selection of geodiversity variables is crucial and it may be prudent to test geodiversity variables at multiple scales given that a “one size fits all” approach does not work for all species. This study highlights the importance of incorporating topographic roughness (srtm_sq) and climate-related variables, such as bio6_sq and bio13_sq, which consistently demonstrated high importance for improving model performance (Table 2B.1). These variables capture key topographic and climatic factors that shape species distributions in this region. However, other variables may also be promising and perhaps more appropriate for certain species (e.g., species found in the Amazon basin) including annual averages of climate and precipitation, and variables related to vegetation including vertical canopy structure (i.e., from Global Ecosystem Dynamics Investigation; GEDI) and even soil types (Fischer et al. 2008).

In general, geodiversity models made for species in regions with high heterogeneity had higher or equal levels of spatial overlap with expert models than non-geodiversity models. However, it is important to acknowledge that expert models often incorporate broader ecological knowledge beyond the specific variables considered in geodiversity models. Expert models may encompass historical or anecdotal evidence, species-specific nuances, and additional ecological factors like known species interactions that are not routinely captured in species distribution models (Choy et al., 2009; Velásquez-Tibatá et al., 2019; Skroblin et al., 2021) and not explicitly represented in geodiversity variables. This broader ecological context in expert models can lead to different extents of suitable area compared to our models, making interpretation of differences challenging.

To ensure a comprehensive understanding of species-environment relationships and effectively refine species distribution models (SDMs) for conservation purposes, it is essential to embrace an integrated approach that incorporates geodiversity alongside expert knowledge and field observations (Urbina-Cardona et al., 2019; Merow et al., 2022). By combining these complementary methods, we can harness the strengths of both approaches, leading to more robust and reliable predictions (Velásquez-Tibatá et al., 2019; Merow et al., 2022) which are essential given the utility of SDMs for conservation such as target species prioritization, guiding future sampling efforts, and as inputs into biodiversity assessments (Franklin, 2013). Hence, a collaborative and comprehensive strategy that integrates geodiversity with expert insights presents a promising avenue for advancing conservation strategies and safeguarding biodiversity for generations to come.

2.6 Data Availability Statement

All code used in this study, including data processing, analysis, and visualizations, is publicly available on GitHub at https://github.com/bioXgeo/neotropical_geodiv.

REFERENCES

Aiello-Lammens, M.E., Boria, R.A., Radosavljevic, A., Vilela, B. & Anderson, R.P. (2015) spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. Ecography, 38, 541–545.

Akçakaya, H.R., Rodrigues, A.S.L., Keith, D.A., Milner-Gulland, E.J., Sanderson, E.W., Hedges, S., Mallon, D.P., Grace, M.K., Long, B., Meijaard, E. & Stephenson, P.J. Assessing ecological function in the context of species recovery. Conservation Biology, n/a.

Anderson, R.P. & Gonzalez, I. (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. Ecological Modelling, 222, 2796–2811.

Anderson, R.P. & Raza, A. (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus Nephelomys) in Venezuela. Journal of Biogeography, 37, 1378–1393.

Bailey, J.J., Boyd, D.S. & Field, R. (2018) Models of upland species' distributions are improved by accounting for geodiversity. Landscape Ecology, 33, 2071–2087.

Bailey, J.J., Boyd, D.S., Hjort, J., Lavers, C.P. & Field, R. (2017) Modelling native and alien vascular plant species richness: At which scales is geodiversity most relevant? Global Ecology and Biogeography, 26, 763–776.

Blach-Overgaard, A., Svenning, J.-C., Dransfield, J., Greve, M. & Balslev, H. (2010) Determinants of palm species distributions across Africa: the relative roles of climate, non-climatic environmental factors, and spatial constraints. Ecography, 33, 380–391.

Blair, M.E., Le, M.D. & Xu, M. (2022a) Species distribution modeling to inform transboundary species conservation and management under climate change: promise and pitfalls. Frontiers of Biogeography, 14.

Blair, M.E., Nguyen, T.A., Le, M.D., Liu, Z., Meng, T., Horning, N., Sterling, E.J., Thach, H.M., Xu, M. & Galante, P.J. (2022b) Karst as an abiotic driver of François' langur distribution, with predictions for biological communities on karst under climate change. Frontiers of Biogeography, 14.

Blaum, N., Mosner, E., Schwager, M. & Jeltsch, F. (2011) How functional is functional? Ecological groupings in terrestrial animal ecology: towards an animal functional type approach. Biodiversity and Conservation, 20, 2333–2345.

Boria, R.A., Olson, L.E., Goodman, S.M. & Anderson, R.P. (2014) Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. Ecological Modelling, 275, 73–77.

Bosch-Belmar, M., Giommi, C., Milisenda, G., Abbruzzo, A. & Sarà, G. (2021) Integrating functional traits into correlative species distribution models to investigate the vulnerability of marine human activities to climate change. Science of The Total Environment, 799, 149351.

Brooks, T.M., Mittermeier, R.A., Fonseca, G.A.B. da, Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D. & Rodrigues, A.S.L. (2006) Global Biodiversity Conservation Priorities. Science, 313, 58–61.

Brost, B.M. & Beier, P. (2012) Use of land facets to design linkages for climate change. Ecological Applications, 22, 87–103.

Brum, F.T., Graham, C.H., Costa, G.C., Hedges, S.B., Penone, C., Radeloff, V.C., Rondinini, C., Loyola, R. & Davidson, A.D. (2017) Global priorities for conservation across multiple dimensions of mammalian diversity. Proceedings of the National Academy of Sciences, 114, 7641–7646.

Castellanos, A. (2011) Andean bear home ranges in the Intag region, Ecuador. Ursus, 22, 65–73.

Choy, S.L., O'Leary, R. & Mengersen, K. (2009) Elicitation by design in ecology: using expert opinion to inform priors for Bayesian statistical models. Ecology, 90, 265–277.

Coblentz, D.D. & Riitters, K.H. (2004) Topographic controls on the regional-scale biodiversity of the south-western USA. Journal of Biogeography, 31, 1125–1138.

Crofts, R. (2022) Progress and future challenges for geoconservation in protected and conserved areas. Parks Stewardship Forum, 38.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D. & Lautenbach, S. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography, 36, 27–46.

Dufour, A., Gadallah, F., Wagner, H.H., Guisan, A. & Buttler, A. (2006) Plant species richness and environmental heterogeneity in a mountain landscape: effects of variability and spatial configuration. Ecography, 29, 573–584.

Elith, J. & Leathwick, J.R. (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. Annual Review of Ecology Evolution and Systematics, pp. 677–697. Annual Reviews, Palo Alto.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists: Statistical explanation of MaxEnt. Diversity and Distributions, 17, 43–57.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D. & Alsdorf, D. (2007) The Shuttle Radar Topography Mission. Reviews of Geophysics, 45.

Fischer, G., F. Nachtergaele, S. Prieler, H.T. van Velthuizen, L. Verelst, D. Wiberg, (2008) Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008). IIASA, Laxenburg, Austria and FAO, Rome, Italy.

Franklin, J. (2013) Species distribution models in conservation biogeography: developments and challenges. Diversity & Distributions, 19, 1217–1223.

Fournier, A., Barbet-Massin, M., Rome, Q. & Courchamp, F. (2017) Predicting species distribution combining multi-scale drivers. Global Ecology and Conservation, 12, 215–226.

García-Rangel, S. (2012) Andean bear Tremarctos ornatus natural history and conservation. Mammal Review, 42, 85–119.

Gerstner, B.E., Kass, J.M., Kays, R., Helgen, K.M. & Anderson, R.P. (2018) Revised distributional estimates for the recently discovered olinguito (Bassaricyon neblina), with comments on natural and taxonomic history. Journal of Mammalogy, 99, 321–332.

González-Orozco, C.E. (2021) Biogeographical regionalisation of Colombia: a revised area taxonomy. Phytotaxa, 484, 247–260.

Graham, L.J., Spake, R., Gillings, S., Watts, K. & Eigenbrod, F. (2019) Incorporating fine-scale environmental heterogeneity into broad-extent models. Methods in Ecology and Evolution, 10, 767–778.

Grantham, H.S., Duncan, A., Evans, T.D., Jones, K.R., Beyer, H.L., Schuster, R., Walston, J., Ray, J.C., Robinson, J.G., Callow, M., Clements, T., Costa, H.M., DeGemmis, A., Elsen, P.R., Ervin, J., Franco, P., Goldman, E., Goetz, S., Hansen, A., Hofsvang, E., Jantz, P., Jupiter, S., Kang, A., Langhammer, P., Laurance, W.F., Lieberman, S., Linkie, M., Malhi, Y., Maxwell, S., Mendez, M., Mittermeier, R., Murray, N.J., Possingham, H., Radachowsky, J., Saatchi, S., Samper, C., Silverman, J., Shapiro, A., Strassburg, B., Stevens, T., Stokes, E., Taylor, R., Tear, T., Tizard, R., Venter, O., Visconti, P., Wang, S. & Watson, J.E.M. (2020) Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity. Nature Communications, 11, 5978.

Guevara, L., Gerstner, B.E., Kass, J.M. & Anderson, R.P. (2018) Toward ecologically realistic predictions of species distributions: A cross-time example from tropical montane cloud forests. Global Change Biology, 24, 1511–1522.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. Ecology Letters, 8, 993–1009.

Hamilton, S.K., Sippel, S.J. & Melack, J.M. (2004) Seasonal inundation patterns in two large savanna floodplains of South America: the Llanos de Moxos (Bolivia) and the Llanos del Orinoco (Venezuela and Colombia). Hydrological Processes, 18, 2103–2116.

Hernandez, P.A., Graham, C.H., Master, L.L. & Albert, D.L. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography, 29, 773–785.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C. & Guisan, A. (2006) Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling, 199, 142–152.

Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, H.P. & Kessler, M. (2017) Climatologies at high resolution for the earth's land surface areas. Scientific Data, 4, 170122.

IUCN (2022). The IUCN Red List of Threatened Species. Version 2022-2. https://www.iucnredlist.org.

Kass, J.M., Muscarella, R., Galante, P.J., Bohl, C.L., Pinilla-Buitrago, G.E., Boria, R.A., Soley-Guardia, M. & Anderson, R.P. (2021) ENMeval 2.0: Redesigned for customizable and reproducible modeling of species' niches and distributions. Methods in Ecology and Evolution, 12, 1602–1608.

Kass, J.M., Vilela, B., Aiello-Lammens, M.E., Muscarella, R., Merow, C. & Anderson, R.P. (2018) Wallace: A flexible platform for reproducible modeling of species niches and distributions built for community expansion. Methods in Ecology and Evolution, 9, 1151–1156.

Kraft, N.J.B., Adler, P.B., Godoy, O., James, E.C., Fuller, S. & Levine, J.M. (2015) Community assembly, coexistence and the environmental filtering metaphor. Functional Ecology, 29, 592–599.

Lawler, J.J., Ackerly, D.D., Albano, C.M., Anderson, M.G., Dobrowski, S.Z., Gill, J.L., Heller, N.E., Pressey, R.L., Sanderson, E.W. & Weiss, S.B. (2015) The theory behind, and the challenges of, conserving nature's stage in a time of rapid change. Conservation Biology, 29, 618–629.

Leong, R.C., Friess, D.A., Crase, B., Lee, W.K. & Webb, E.L. (2018) High-resolution pattern of mangrove species distribution is controlled by surface elevation. Estuarine, Coastal and Shelf Science, 202, 185–192.

Lobo, J.M., Jiménez-Valverde, A. & Real, R. (2008) AUC: a misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography, 17, 145–151.

Luoto, M., Virkkala, R. & Heikkinen, R.K. (2007) The role of land cover in bioclimatic models depends on spatial resolution. Global Ecology and Biogeography, 16, 34–42.

Maharjan, S.K., Sterck, F.J., Dhakal, B.P., Makri, M. & Poorter, L. (2021) Functional traits shape tree species distribution in the Himalayas. Journal of Ecology, 109, 3818–3834.

McNab, B.K. (1963) Bioenergetics and the Determination of Home Range Size. The American Naturalist, 97, 133–140.

Medrano-Vizcaíno, P. & Gutiérrez-Salazar, P. (2020) Current and future suitable habitat areas for Nasuella olivacea (Gray, 1865) in Colombia and Ecuador and analysis of its distribution across different land uses. Biodiversity Data Journal, 8, e49164.

Merow, C., Galante, P.J., Kass, J.M., Aiello-Lammens, M.E., Babich Morrow, C., Gerstner, B.E., Grisales Betancur, V., Moore, A.C., Noguera-Urbano, E.A., Pinilla-Buitrago, G.E., Velásquez-Tibatá, J., Anderson, R.P. & Blair, M.E. (2022) Operationalizing expert knowledge in species' range estimates using diverse data types. Frontiers of Biogeography, 14.

Merow, C., Smith, M.J. & Silander, J.A. (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography, 36, 1058–1069.

Milton, K. & May, M.L. (1976) Body weight, diet and home range area in primates. Nature, 259, 459–462.

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

Oke, O.A. & Thompson, K.A. (2015) Distribution models for mountain plant species: The value of elevation. Ecological Modelling, 301, 72–77.

Olson, D.M. & Dinerstein, E. (1998) The Global 200: A Representation Approach to Conserving the Earth's Most Biologically Valuable Ecoregions. Conservation Biology, 12, 502–515.

Parks, K.E. & Mulligan, M. (2010) On the relationship between a resource based measure of geodiversity and broad scale biodiversity patterns. Biodiversity and Conservation, 19, 2751–2766.

Pearson, D.E., Ortega, Y.K., Eren, Ö. & Hierro, J.L. (2018) Community Assembly Theory as a Framework for Biological Invasions. Trends in Ecology & Evolution, 33, 313–325.

Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? Global Ecology and Biogeography, 12, 361–371.

Peterson, A.T., Papeş, M. & Soberón, J. (2008) Rethinking receiver operating characteristic analysis applications in ecological niche modeling. Ecological Modelling, 213, 63–72.

Pollock, L.J., Morris, W.K. & Vesk, P.A. (2012) The role of functional traits in species distributions revealed through a hierarchical model. Ecography, 35, 716–725.

Ramírez-Chaves, H.E., Ossa-López, P.A., Velásquez-Guarín, D., Colmenares-Pinzón, J., Noguera-Urbano, E.A., Mejía-Fontecha, I.Y., Acosta-Castañeda, A., Chacón-Pacheco, J., Rivera Páez, F.A. & Suárez-Castro, A.F. (2022) New genetic information and geographic distribution of charismatic carnivores: the olingos (Procyonidae: Bassaricyon) in Colombia. Mammalian Biology.

Reynard, E. & Brilha, J. eds. (2017) Geoheritage: Assessment, Protection, and Management, 1st edition. Elsevier, Amsterdam, Netherlands ; Cambridge, Massachusetts.

Ruggiero, A. & Kitzberger, T. (2004) Environmental correlates of mammal species richness in South America: effects of spatial structure, taxonomy and geographic range. Ecography, 27, 401–417.

Schnase, J.L., Carroll, M.L., Gill, R.L., Tamkin, G.S., Li, J., Strong, S.L., Maxwell, T.P., Aronne, M.E. & Spradlin, C.S. (2021) Toward a Monte Carlo approach to selecting climate variables in MaxEnt. PLOS ONE, 16, e0237208.

Schrodt, F., Bailey, J.J., Kissling, W.D., Rijsdijk, K.F., Seijmonsbergen, A.C., Ree, D. van, Hjort, J., Lawley, R.S., Williams, C.N., Anderson, M.G., Beier, P., Beukering, P. van, Boyd, D.S., Brilha, J., Carcavilla, L., Dahlin, K.M., Gill, J.C., Gordon, J.E., Gray, M., Grundy, M., Hunter, M.L., Lawler, J.J., Monge-Ganuzas, M., Royse, K.R., Stewart, I., Record, S., Turner, W., Zarnetske, P.L. & Field, R. (2019) Opinion: To advance sustainable stewardship, we must document not only biodiversity but geodiversity. Proceedings of the National Academy of Sciences, 116, 16155–16158.

Shcheglovitova, M. & Anderson, R.P. (2013) Estimating optimal complexity for ecological niche models: A jackknife approach for species with small sample sizes. Ecological Modelling, 269, 9–17.

Skroblin, A., Carboon, T., Bidu, G., Chapman, N., Miller, M., Taylor, K., Taylor, W., Game, E.T. & Wintle, B.A. (2021) Including indigenous knowledge in species distribution modeling for increased ecological insights. Conservation Biology, 35, 587–597.

Smith, A.C., Dahlin, K.M., Record, S., Costanza, J.K., Wilson, A.M. & Zarnetske, P.L. (2021) The geodiv r package: Tools for calculating gradient surface metrics. Methods in Ecology and Evolution, 12, 2094–2100.

Soberón, J. & Nakamura, M. (2009) Niches and distributional areas: Concepts, methods, and assumptions. Proceedings of the National Academy of Sciences, 106, 19644–19650.

Stein, A., Gerstner, K. & Kreft, H. (2014) Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. Ecology Letters, 17, 866–880.

Stevenson, P.R., Quinones, M.J. & Ahumada, J.A. (1994) Ecological strategies of woolly monkeys (Lagothrix lagotricha) at Tinigua National Park, Colombia. American Journal of Primatology, 32, 123–140.

Tang, L., Morris, W.K., Zhang, M., Shi, F. & Vesk, P.A. (2022) Exploring how functional traits modulate species distributions along topographic gradients in Baxian Mountain, North China. Scientific Reports, 12, 994.

Tucker, M.A., Ord, T.J. & Rogers, T.L. (2014) Evolutionary predictors of mammalian home range size: body mass, diet and the environment. Global Ecology and Biogeography, 23, 1105–1114.

Tukiainen, H. (2019) Multi-scale relationship between geodiversity and biodiversity across high-latitude environments: implications for nature conservation. Nordia Geographical Publications, 48, 1–54.

Urban, M.C., Zarnetske, P.L. & Skelly, D.K. (2013) Moving forward: dispersal and species interactions determine biotic responses to climate change. Annals of the New York Academy of Sciences, 1297, 44–60.

Urbina-Cardona, N., Blair, M.E., Londoño, M.C., Loyola, R., Velásquez-Tibatá, J. & Morales-Devia, H. (2019) Species Distribution Modeling in Latin America: A 25-Year Retrospective Review. Tropical Conservation Science, 12, 1940082919854058.

Vega, R., Vázquez-Domínguez, E., White, T.A., Valenzuela-Galván, D. & Searle, J.B. (2017) Population genomics applications for conservation: the case of the tropical dry forest dweller Peromyscus melanophrys. Conservation Genetics, 18, 313–326.

Velásquez-Tibatá, J., Olaya-Rodríguez, M.H., López-Lozano, D., Gutiérrez, C., González, I. & Londoño-Murcia, M.C. (2019) BioModelos: A collaborative online system to map species distributions. PLOS ONE, 14, e0214522.

Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007) Let the concept of trait be functional! Oikos, 116, 882–892.

Virkkala, R., Luoto, M., Heikkinen, R.K. & Leikola, N. (2005) Distribution patterns of boreal marshland birds: modelling the relationships to land cover and climate. Journal of Biogeography, 32, 1957–1970.

Wang, C.-J. & Wan, J.-Z. (2021) Functional trait perspective on suitable habitat distribution of invasive plant species at a global scale. Perspectives in Ecology and Conservation, 19, 475–486.

Wilman, H., Belmaker, J., Simpson, J., de la Rosa, C., Rivadeneira, M.M. & Jetz, W. (2014) EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals. Ecology, 95, 2027–2027.

Wilson, A.M. & Jetz, W. (2016) Remotely Sensed High-Resolution Global Cloud Dynamics for Predicting Ecosystem and Biodiversity Distributions. PLOS Biology, 14, e1002415.

Wittmann, M.E., Barnes, M.A., Jerde, C.L., Jones, L.A. & Lodge, D.M. (2016) Confronting species distribution model predictions with species functional traits. Ecology and Evolution, 6, 873–879.

Zarnetske, P.L., Read, Q.D., Record, S., Gaddis, K.D., Pau, S., Hobi, M.L., Malone, S.L., Costanza, J., Dahlin, K.M., Latimer, A.M., Wilson, A.M., Grady, J.M., Ollinger, S.V. &

Finley, A.O. (2019) Towards connecting biodiversity and geodiversity across scales with satellite remote sensing. Global Ecology and Biogeography, 28, 548–556

APPENDIX 2A: SPATIAL EVALUATION METRICS OF SPECIES DISTRIBUTION MODEL PERFORMANCE

Table 2A.1: Spatial evaluation of species distribution model performance. This table provides an assessment of the performance of species distribution models through a spatial comparison with expert-generated maps. The evaluation aimed to identify similarities and variations between the models and expert maps, with a focus on the effectiveness of incorporating geodiversity variables. The models incorporating geodiversity variables demonstrated overall good performance, generating distributions that were consistent with species ecology. The evaluation of model performance included the analysis of percent spatial gain and loss, Schoener's D (noted "D"), and the omission rate, which indicates the percentage of records left out of the model prediction. These metrics were calculated for the expert models, models without geodiversity variables, and the optimal geodiversity models.

SPECIES	GAIN GEODIVERSITY	LOSS GEODIVERSITY	GAIN NON-GEODIVERSITY	LOSS NON-GEODIVERSITY	D GEODIVERSITY	D NON-GEODIVERSITY	OMISSION GEODIVERSITY	OMISSION NO GEODIVERSITY	OMISSION EXPERT	BIOGEOGRAPHIC REGION
Alouatta palliata	8.01	0.18	7.35	0.25	0.23	0.24	11.11	11.11	16.67	Choco-Darien
Aotus zonalis	4.88	2.61	4.21	4.61	0.89	0.89	5.88	5.88	23.53	Choco-Darien
Ateles fusciceps	4.28	11.93	4.89	7.76	0.61	0.75	9.09	9.09	9.09	Choco-Darien
Cebus capucinus	6.16	3.10	5.91	3.82	0.79	0.80	12.33	12.33	24.66	Choco-Darien
Saguinus geoffroyi	4.76	10.97	4.47	14.47	0.70	0.61	12.50	12.50	0.00	Choco-Darien
Average	**5.62**	**5.76**	**5.37**	**6.18**	**0.65**	**0.66**	**10.18**	**10.18**	**14.79**	
Cacajao melanocephalus	0.67	10.76	0.85	0.86	0.71	0.98	6.67	6.67	6.67	Amazonian
Cebus albifrons	1.73	12.50	2.30	6.98	0.64	0.80	69.49	69.49	66.95	Amazonian
Cheracebus lucifer	0.80	20.57	0.56	18.50	0.33	0.40	0.00	0.00	6.67	Amazonian
Leontocebus fuscus	1.69	5.37	1.73	3.87	0.82	0.87	7.41	7.41	11.11	Amazonian
Pithecia hirsuta	1.53	23.97	1.35	22.63	0.37	0.41	16.67	16.67	4.17	Amazonian
Average	**1.28**	**14.63**	**1.36**	**10.57**	**0.58**	**0.69**	**20.05**	**20.05**	**19.11**	
Pithecia milleri	2.75	1.89	2.90	1.11	0.93	0.93	5.56	5.56	11.11	Amazonian-1
Plecturocebus caquetensis	16.84	13.80	23.63	11.66	0.54	0.48	50.00	50.00	50.00	Amazonian-1
Average	**7.15**	**7.36**	**9.66**	**6.33**	**0.76**	**0.74**	**21.30**	**21.30**	**20.37**	
Ateles belzebuth	6.23	5.11	6.30	3.52	0.85	0.86	14.55	14.55	25.45	Amazonian-2

Table 2A.1 (cont'd)

Cebuella pygmaea	1.35	8.68	1.47	1.76	0.57	0.91	22.22	22.22	27.78	Amazonian-2
Cheracebus lugens	3.33	15.15	4.63	3.93	0.62	0.89	12.50	12.50	4.17	Amazonian-2
Plecturocebus discolor	9.49	1.65	12.42	0.00	0.88	0.85	50.00	50.00	50.00	Amazonian-2
Average	**6.37**	**8.23**	**7.35**	**2.83**	**0.70**	**0.83**	**19.89**	**19.89**	**19.85**	
Alouatta seniculus	0.64	16.67	0.58	14.75	0.63	0.67	6.02	6.02	1.88	Amazonian-mix
Aotus brumbacki	6.36	6.61	7.49	3.83	0.88	0.88	17.65	17.65	17.65	Amazonian-mix
Lagothrix lagotricha	9.52	3.51	10.25	2.44	0.74	0.74	8.73	8.73	7.94	Amazonian-mix
Saimiri cassiquiarensis	8.29	15.26	9.03	5.36	0.54	0.76	11.32	11.32	3.77	Amazonian-mix
Sapajus apella	0.86	7.90	0.92	5.34	0.83	0.88	9.02	10.53	3.76	Amazonian-mix
Average	**5.14**	**9.99**	**5.66**	**6.34**	**0.72**	**0.78**	**10.55**	**10.85**	**7.00**	
Ateles hybridus	22.66	1.21	22.81	0.98	0.48	0.48	12.82	12.82	35.90	Magdalena
Cebus versicolor	13.02	12.41	14.42	5.99	0.74	0.75	0.00	0.00	18.18	Magdalena
Saguinus leucopus	34.45	0.92	33.56	1.29	0.33	0.34	7.26	7.26	58.87	Magdalena
Saguinus oedipus	19.85	0.01	19.93	0.00	0.56	0.56	4.62	4.62	27.69	Magdalena
Aotus griseimembra	16.27	1.32	18.56	0.67	0.59	0.56	8.11	8.11	31.08	Magdalena
Average	**21.25**	**3.17**	**21.86**	**1.79**	**0.54**	**0.54**	**6.56**	**6.56**	**34.34**	
Aotus lemurinus	20.56	3.87	20.79	3.46	0.62	0.62	0.00	0.00	12.99	Andean
Bassaricyon neblina	1.48	5.73	1.27	6.18	0.73	0.71	0.00	0.00	18.75	Andean
Nasuella olivacea	1.73	0.98	1.20	1.50	0.80	0.81	17.86	25.00	21.43	Andean
Tremarctos ornatus	9.36	0.05	10.69	0.00	0.61	0.58	11.54	10.58	22.12	Andean
Average	**8.28**	**2.66**	**8.49**	**2.79**	**0.69**	**0.68**	**7.35**	**8.89**	**18.82**	
Plecturocebus ornatus	1.86	6.38	2.43	6.23	0.79	0.80	8.33	8.33	0.00	Sabana
Overall average	7.79	7.45	8.35	5.28	0.66	0.70	13.85	14.09	20.00	

APPENDIX 2B: PERMUTATION IMPORTANCE ACROSS SPATIAL GRAIN

Table 2B.1: Average permutation importance for each variable across spatial grains. The numbers in parentheses represent the frequency of each variable being among the top seven contributing variables in a species distribution model.

	Spatial grain (km)					
variable	**3**	**9**	**15**	**21**	**27**	**33**
bio13	9.7 (27)	7.76 (23)	8.13 (22)	6.42 (22)	6.77 (23)	6.45 (22)
bio13_sq	3.95 (15)	3.78 (14)	4.69 (19)	5.36 (22)	6.82 (23)	7.71 (22)
bio14	5.61 (16)	6.14 (17)	7.84 (18)	7.96 (21)	7.03 (21)	7.07 (21)
bio14_sq	3.88 (16)	4.65 (18)	4.93 (16)	4.86 (17)	4.46 (16)	4.71 (14)
bio5	11.36 (20)	11.46 (23)	10.35 (23)	9.91 (21)	9.34 (17)	9.63 (18)
bio5_sq	4.45 (15)	4.78 (11)	4.28 (13)	5.24 (11)	5.88 (17)	6.29 (19)
bio6	14.19 (19)	13.94 (23)	13.42 (23)	12.8 (23)	11.93 (24)	11.8 (23)
bio6_sq	4.38 (17)	4.95 (17)	5.61 (14)	6.68 (16)	7.39 (14)	8.13 (14)
cloud	7.77 (22)	7.31 (21)	7.65 (20)	7.97 (22)	7.62 (21)	6.79 (18)
cloud_sq	6.61 (21)	5.67 (19)	4.49 (18)	3.74 (13)	3.76 (12)	3.51 (14)
srtm	21.56 (25)	20.91 (25)	20.99 (25)	22.06 (25)	22.04 (25)	19.8 (27)
srtm_sq	6.55 (18)	8.66 (20)	7.62 (20)	6.99 (18)	6.97 (18)	8.1 (18)

CHAPTER 3:

EVALUATING THE EFFECTIVENESS OF PROTECTED AREAS AND COMMUNITY-MANAGED LANDS IN CAPTURING MULTIPLE DIMENSIONS OF FRUGIVOROUS BIODIVERSITY IN THE TROPICAL ANDES

For submission to Biological Conservation

Co-authors include: Phoebe L. Zarnetske

3.1 Abstract

The Tropical Andes is known for its remarkable biodiversity; however, this region faces significant anthropogenic pressures such as habitat loss and climate change, making it a conservation priority. Conservation efforts often concentrate on protecting areas with high taxonomic diversity (TD), overlooking the essential roles played by species that contribute to fotherecosystem functioning– functional diversity (FD). To ensure comprehensive conservation priorities, a broader perspective considering various dimensions of diversity is necessary. Further, there needs to be a better understanding of how community managed lands contribute to the protection of biodiversity as they are essential for reaching spatial conservation targets. Here, I utilized the Frugivoria trait database to evaluate the functional diversity of frugivorous birds and mammals, which perform the vital role of seed dispersal, an essential ecosystem service necessary to maintain forest structure and ecosystem health. Across the Tropical Andes region, I quantified the spatial alignment and mismatch between taxonomic and functional diversity for areas with the highest diversity values in the region. Our findings revealed many spatial misalignments between the highest levels of FD and TD, emphasizing the limitations of relying solely on TD for conservation and the potential biodiversity trade-offs of doing so. Nevertheless, some alignment between taxonomic and functional diversity emerged for mammals and birds, identifying potential areas for multidimensional biodiversity conservation. Further, I found that

some stricter protected areas (PAs) encapsulate different diversity dimensions better than less-strict areas and this differed among taxa. Similarly, I examined the distribution of Functionally Unique, Specialized, and Endangered (FUSE) species and their protection status across strict and less strict PAs. Some stricter protected zones had better coverage of FUSE species distributions, though many areas of high FUSE species richness remain unprotected. I also found that community managed lands had higher levels of FD higher than other protected areas within the same protected area category. This has strong implications for the utility of other effective conservation measures (OECMs) such as Indigenous Lands and community managed areas in protecting areas of high functional diversity. Our results highlight the need for a more holistic conservation approach that considers multiple dimensions of diversity in the context of varying degrees of protection and PA management. As global conservation goals target the protection of 30% of the Earth's land by 2030, our study underscores the importance of considering multiple dimensions of diversity and emphasizes the potential of OECMs in ensuring effective and sustainable conservation strategies within the Tropical Andes.

3.2 Introduction

The Tropical Andes, despite covering less than 0.5% of the Earth's land surface, is a remarkable biodiversity hotspot, housing an impressive 10% of all known species and with the highest number of endemic plants and vertebrates in the world (Myers et al., 2000). However, this biodiversity is threatened by global change, primarily driven by deforestation and anthropogenic climate change (IPCC, 2022). In response to the pressing need for global biodiversity conservation, the UN Biodiversity Conference (COP15) adopted the "Kunming-Montreal Global Biodiversity Framework" (GBF) in 2022. The GBF outlines ambitious targets, including effective conservation and management of at least 30% of the world's lands, inland waters, coastal areas, and oceans, with a focus on areas of particular importance for biodiversity and ecosystem functioning and services and emphasizes the importance of indigenous and traditional territories for doing so. However, while the Tropical Andes boasts exceptional biodiversity, most studies investigating the spatial distribution of biodiversity have only focused on metrics such as species richness, species endemism, and vulnerability. These approaches tend to prioritize taxonomic diversity, the number of species present, and may overlook other vital dimensions of biodiversity, such as the diversity of traits exhibited by species within an ecosystem that help maintain ecosystem functioning (i.e., functional diversity; Petchey and Gaston, 2006) and evolutionary relatedness (phylogenetic diversity; Faith et al., 2004). Further, there is an incomplete understanding of how lands governed by Indigenous peoples and local communities, which cover almost 37% of all remaining natural lands across Earth, contribute the protection of these different dimensions of diversity, which has strong implications for the ability to reach biodiversity conservation targets (Garnett et al. 2018) such as the GBF. Quantifying multiple dimensions of diversity is essential for more comprehensive assessments of biodiversity (Borges et al., 2021; Brum et al., 2017; Devictor et al., 2010; Pollock et al., 2017). In particular, spatial analysis of multiple dimensions of diversity can determine mismatches and alignments among the dimensions and inform management and prioritization efforts aimed at conserving areas with high multidimensional diversity. Functional diversity, encompassing the variety of ecological functions and traits exhibited by species within an ecosystem, plays a pivotal role in maintaining ecosystem functioning and services (Cadotte et al., 2011; Cadotte and Tucker, 2018; Mouchet et al., 2010). It has been suggested by some studies that prioritizing phylogenetic diversity for conservation, which is easier to measure, is a sufficient proxy for conserving functional diversity because of the relationship between shared evolutionary history and traits

(Winter et al., 2013; Mazel et al., 2018). However, this is problematic because phylogenetic diversity does not explicitly reflect species' ecological strategies or their distinct ecological roles and captures functional diversity unreliably (Mazel et al., 2018; Cooke et al., 2020). Similarly, protecting areas solely based on taxonomic diversity can also lead to spatial mismatches, potentially neglecting the conservation of regions that harbor important functional groups or species with specialized functional roles. For example, spatial mismatches between taxonomic and functional diversity in French protected areas (PAs) revealed that bird taxonomic diversity was overrepresented, while functional diversity was underrepresented (Devictor et al. 2010). Similarly, Pollock et al. (2017), in a global study of multiple dimensions of biodiversity, revealed that approximately 12% of the functional tree of life for mammals and birds was missing from PAs, suggesting an incomplete representation of the diversity of ecological functions and roles in these regions. Overlooking functional diversity in conservation efforts can have significant consequences for ecosystem resilience and functioning, as functionally diverse communities are better equipped to adapt to changing environmental conditions, such as climate change and habitat disturbances (Cadotte et al. 2011).

Understanding the spatial mismatches among dimensions of diversity is especially important for groups of species that perform essential ecosystem services such as seed dispersal. Frugivores, animals that consume fruits, represent a diverse group of species with seed dispersal roles. The interactions between frugivores and the fruiting species they consume create complex relationships that shape and sustain diversity and aid in the regeneration of tropical forest plant communities (Howe and Smallwood, 1982; Sales et al., 2021). Multiple studies have highlighted the essential contributions of frugivorous mammals and birds in seed dispersal, whose abundance and distribution changes can impact seed dispersal services and plants' abilities to adapt to climate changes (Fricke et al., 2022; Sales et al., 2021). However, it is essential to recognize that not all species contribute equally to functional diversity due to their unique traits and roles within the ecosystem. For instance, a study focusing on the Atlantic Forest used an extensive dataset of frugivore-plant interactions and found that endangered specialist frugivores played an outsized role in maintaining tropical forest seed dispersal. Their extinction led to a rapid and disproportionate loss of tree species that relied nearly exclusively on those at-risk frugivores for seed dispersal, significantly affecting the forest's regeneration processes and plant diversity (Lamperty and Brosi, 2022). These findings underscore the importance of threatened frugivores in shaping and sustaining the diversity and regeneration of tropical forest plant communities.

To address the challenges of overlooking functionally important species like threatened frugivores, the new concept of Functionally Unique, Specialized, and Endangered (FUSE) species has emerged as a promising strategy (Pimiento et al., 2020; Griffin et al. 2020). The concept of FUSE species is like that of EDGE species (Issac et al. 2007), which is a way of triaging conservation efforts based on phylogenetic diversity and focuses on conserving threatened species that are evolutionarily distinct. However, FUSE species are assigned conservation value by evaluating their contribution to functional diversity and are identified based on their position in multidimensional trait space (Griffin et al. 2020). While EDGE species might reflect evolutionarily distinct lineages that may often exhibit unique traits, many traits are not phylogenetically conserved (Mazel et al. 2018). Therefore, FUSE may be a better metric for prioritizing species with important roles in maintaining ecosystem functioning. Priority is given to species that exhibit functional specialization—meaning they occupy extreme values in trait space and influence the overall volume of trait space—and those that are functionally unique—occupying isolated positions in trait space and exhibiting low functional redundancy (Pimiento et al., 2020; Griffin et al., 2020)). If FUSE species are lost, their unique functional roles within the ecosystem cannot be easily replaced, as there is no other species with the exact suite of traits to compensate for their role (low functional redundancy; Mouillot et al., 2013). Prioritizing FUSE species is similar to prioritizing rare species, which are often threatened, whose loss negatively affects assemblage structure and function because of their unique trait combinations (Mouillot et al. 2013; Leitão et al., 2016). By focusing on the identification and protection of FUSE species, which can be complementary to the EDGE species approach, we can highlight and safeguard species that perform essential and distinct ecological functions, which are vital for the regeneration and health of ecosystems.

Nonetheless, the critical need to assess the effectiveness of protected areas in preserving all aspects of biodiversity is tempered by the stark reality that not all of these areas offer the same levels of protection. The IUCN categorizes protected areas into six distinct classes, each characterized by its specific conservation objectives, management strategies, and permitted activities. These categories, labeled I-VI, present a wide range of approaches, from the strictest forms of protection (I-IV) to those that allow for sustainable human use and resource management (V-VI) (Dudley, 2008). In addition to these IUCN categories, it is also important to acknowledge the significance of community-based resource management areas, which often play a crucial role in biodiversity conservation through local stewardship and engagement (Porter-Bolland, et al., 2012; Watson et al., 2016; Gonzalez et al., 2023). Understanding the nuances and varying degrees of protection provided by these categories is paramount as it directly influences the degree to which these areas can fulfill their roles in protecting biodiversity in the present and sustaining it into the future. The inclusion of community-managed areas adds another layer of complexity to this assessment, as they represent a collaborative and locally driven approach to conservation that complements traditional protected areas (Watson et al, 2016). By comprehensively evaluating the protection status and management strategies across all these dimensions, we can gain a holistic view of how well these areas collectively contribute to the protection of biodiversity, including FD and conservation of FUSE species critical for ecosystem health and resilience.

This study aims to assess the patterns of functional and taxonomic diversity of frugivorous birds and mammals in the Tropical Andes and assess their level of protection. By taking an integrative approach that considers multiple dimensions of biodiversity, including the importance of frugivores, I aim to 1) quantify the spatial distribution of taxonomic and functional diversity for species exhibiting frugivory; 2) evaluate how well existing PAs align with areas of high frugivorous species diversity (e.g., taxonomic diversity, functional diversity, and FUSE species); and 3) evaluate how well community managed areas capture these different aspects of diversity. For this analysis, I rely on a comprehensive trait database, Frugivoria, containing ecological, life-history, morphological, and geographical traits for frugivorous mammals and birds in the Central and South American moist forests (Gerstner et al., 2023). This study contributes valuable insights for optimizing conservation strategies and fostering the long-term health and resilience of the Tropical Andes ecosystem. By considering both taxonomic and functional diversity and integrating cutting-edge prioritization approaches utilizing functional diversity, such as FUSE species identification, we can help preserve the irreplaceable diversity and ecological integrity of this region, aligning with the objectives of the 30X30 initiative.

3.3 Methods

3.3.1 Overview

In this study, I assess the congruence of taxonomic and functional diversity of birds and mammals exhibiting frugivory in the Tropical Andes. I calculated both taxonomic and functional diversity for both taxa and generated areal estimates of regions with the highest levels of biodiversity. Within these areas, I assessed the alignment and mismatch between dimensions of biodiversity and their overlap with PAs. Furthermore, I examined the spatial distribution and level of protection of Functionally Unique, Specialized, and Endangered (FUSE) species and the degree of forest integrity within each IUCN park category.

3.3.2 Species subset and spatial data

To assemble the species dataset for the Tropical Andes, I collected IUCN species lists for the classes "Aves" and "Mammalia" from the major countries in the region, namely Venezuela, Colombia, Ecuador, Peru, and Bolivia, using the R package 'rredlist' and the function 'rl_sp_country' (IUCN, 2022). By merging these IUCN species lists with the Frugivoria trait dataset (Gerstner et al., 2023) – a dataset for birds and mammals exhibiting frugivory and found in tropical moist forests of Central and South America – I retained only the species present in both datasets, resulting in our final dataset.

To characterize the spatial distribution of each species, I acquired range maps from the IUCN Spatial Dataset (IUCN, 2023) and BirdLife International (Bird Life International, 2023). Following the approach of Gerstner et al. (2023), I focused on species with presence codes of 'Extant', 'Probably Extant', and 'Possibly Extant', which encompass inferred areas where the species is likely to occur. Additionally, I considered all 'origin' designations (e.g., native, re-introduced, introduced, etc.) and 'seasonality' codes (e.g., resident, breeding season, non-breeding season, passage) as these represent important aspects of the species' range where seed dispersal services are likely to be performed.

I refined the species range maps to reduce uncertainty in species distributions and better align with the actual area of occupancy, where the species likely occurs, following the method suggested by Brooks et al. (2019) to determine the "area of habitat" (AOH). Quantifying AOH involved rasterizing all species range maps to a resolution of 1 km^2 to match the resolution of the environmental data used for species-specific subsetting based on elevational ranges and habitat types. Whenever available, I manually extracted elevational ranges from IUCN species accounts (IUCN, 2022), and the Frugivoria dataset (Gerstner et al., 2023) provided us with a comprehensive list of habitat types for each species. I used the Shuttle Radar Topography Mission (SRTM; Farr et al., 2007) digital elevation model (SRTM30) and a global map utilizing the IUCN habitat classification scheme (Jung et al., 2020), to refine species range maps based on their elevational tolerances and habitat types. This meticulous refinement process ensured a more precise representation of species distributions, enhancing the reliability of our diversity estimates.

Subsequently, using the R package 'letsR', I created separate presence/absence matrices (PAM) for birds and mammals where each assemblage was delineated by a 10 km grid cell. I chose this spatial grain to account for residual uncertainty in the distributional limits of species based on range maps which can under or overestimate the range limits of species (Hurlbert and Jetz, 2007; Jetz et al., 2008), however, it is finer than many studies conducting similar analyses with range map-based data (Borges et al., 2021; Brum et al., 2017; González-Maya et al., 2017; Herrera, 2017). To maintain focus on the Tropical Andes, I removed any assemblages from the PAM that occurred outside of the region, precisely delineated using the shapefile obtained through the Critical Ecosystem Partnership Fund. Our final subset, used for diversity calculations, included 1,052 species exhibiting frugivory (344 mammals; 708 birds).

3.3.3 Trait data

To quantify functional diversity, I selected six traits in the Frugivoria dataset for both birds and mammals (Gerstner et al., 2023). These traits were chosen based on their comprehensive coverage of each species' ecological niche and behavior. The selected traits include body mass, diet type, diet breadth, foraging strategy, habitat breadth, and generation time, which likely play crucial functional roles and reflect a species' biology and ecological significance. For example, body mass and foraging strata align with food web structure; body mass can indicate the amount of resources a species consumes and releases (Cooke et al., 2019) and is linked to dispersal capabilities (Jenkins et al., 2007); foraging strategy can influence the dynamics of species interactions and energy flow within ecological communities (Dell et al., 2014). Diet type is essential for services like pollination and seed dispersal (Jordano et al., 2007; Ripple et al., 2017, 2015), while diet breadth indicates the degree of dietary specialization a species possesses. Habitat breadth reveals a species' ability to utilize and compete in multiple environments, serving as a measure of habitat plasticity, which can influence how well species adapt to environmental changes (Luck et al., 2013). Lastly, generation time, representing the turnover rate of breeding individuals in a population, provides insights into a species' potential for recovery after disturbances (Cooke et al., 2020; Newbold et al., 2013) and is a proponent for determining extinction risk (Bird et al., 2020; Jonasson et al., 2022; Pacifici et al., 2013).

To address the issue of missing values for generation time (n = 29; mammals, n = 27; birds, n = 2), I employed multivariate imputation by chained equations using the R package 'mice' (version 3.16.0). I applied three imputation methods, namely Predictive Mean Matching

(PMM), Random Forest (RF), and Categorical and Regression Trees (CART). For each method, I conducted 1000 iterations, randomly inserting NAs into the generation time dataset (10%) and calculating the differences between the real and imputed values. After assessing the performance of each method, I found that CART yielded values closest to the real values, exhibiting an average standard deviation of differences of 1.20 ± 1.87. Consequently, I selected CART as the most suitable method for imputing missing generation time data in our analysis and performed a final iteration of imputation for mammals and birds with 20 iterations each. I randomly selected one of the 20 possible bird datasets and one of the 20 possible mammal datasets and used those for all subsequent analyses.

3.3.4 Diversity calculations

Taxonomic diversity (TD) was evaluated by summing the total number of species present within each 10 km^2 grid cell. For all functional diversity (FD) analyses, I utilized the 'mFD' package (Magneville et al., 2022) in R to calculate functional diversity as functional dispersion (FDis; Laliberté and Legendre, 2010), a multidimensional measure capturing species dispersion in multidimensional trait space. FDis computes the average distance of species from the centroid of the community, which in this case, is defined as a 10 km grid cell. It offers the flexibility to be weighted by abundance or, in the absence of abundance data, all species can be equally weighted. Importantly, FDis is not influenced by species richness, unlike functional richness (FRic), which is sensitive to outliers (Laliberté and Legendre, 2010). This ensures that the number of species does not introduce bias into the calculation of functional diversity values.

For birds and mammals, most traits shared common units. However, the treatment of foraging strata (Wilman et al., 2014; Gerstner et al., 2023) differed. For mammals, it was a categorical variable (e.g., ground, scansorial, canopy, etc.) while for birds, it was a fuzzy trait with values of percentage allocation of time spent in each strata type represented in multiple columns. To prevent the over-weighting of the "forest strata" trait for birds, I categorized the main trait category as "forest strata," and the different columns for strata types were weighted accordingly. This ensured that the "forest strata" category carried the same weight as the other trait variables during the analysis.

I used the Gower distance to compute functional distance matrices separately for birds and mammals because it is well-suited to handle both categorical and continuous traits (de Bello et al., 2021). Subsequently, a Principal Coordinates Analysis (PCoA) was conducted on each functional distance matrix – one for each taxa. The number of axes for each group was selected based on their ability to preserve the original trait dissimilarity matrix while facilitating graphical interpretation (Maire et al., 2015; Villéger et al., 2008). For mammals, this resulted in a 4-dimensional PCoA space, while for birds, the lowest mean absolute deviation (MAD) value was obtained in a 5-dimensional space. Due to the similarity in MAD values for both 4D and 5D spaces, I opted to proceed with the 4D space, following the recommendation by Magneville et al. (2022) due to the added computation time for an additional PCoA axis.

3.3.5 Spatial Assessments

With the TD and FD rasters, I quantified the spatial overlap between TD and FD of frugivorous birds and mammals, frugivores more generally within the Tropical Andes region, as well as their overlap with PAs. I completed this analysis for mammals, birds, and all frugivores combined. I first identified the grid cells representing the highest 10%, 20%, and 30% of both TD and FD per taxa in the study area, representing areas with the highest levels of diversity. Having determined these diversity peaks, I calculated the total area of each thresholded extent for TD and FD and assessed areas of overlap between the two biodiversity dimensions. This step allowed us to identify the spatial extent of these critical biodiversity peaks and evaluate areas for multidimensional biodiversity conservation. For an assessment across all frugivores, I assessed areas of overlap between diversity dimensions where the thresholds coincided for both mammals and birds.

Subsequently, I evaluated the amount of high diversity areas protected by PAs. I overlaid the thresholded peak diversity extents (high TD and FD) with the boundaries of PAs from the World Database on Protected Areas (WDPA, 2023) to determine the extent of overlap. By quantifying the protected area for different TD and FD thresholds, I gained insights into the status of taxonomic and functional diversity within PAs. To explore the alignment between TD and FD across the study area, I ran linear regression on FD vs. TD, individually for both birds and mammals across the entire study area, using each grid cell value. Further, I calculated the proportional frequency of diversity values in PAs versus surrounding areas to understand how well PAs capture different dimensions of diversity for different taxa.

To identify FUSE species (Pimiento et al., 2020), I utilized the 'fuse' function available in the 'mFD' package (Magneville et al., 2022). To understand the distribution of FUSE species, I summed their respective AOH (area of habitat) maps, where presence was denoted as 1 and absence as 0. To quantify the level of FUSE species in each PA, I took the mean of the counts of FUSE species within each IUCN PA type (I-VI) and also computed the proportion of FUSE species out of all co-occurring species. This analysis allowed us to pinpoint areas with elevated concentrations of FUSE frugivores, thus highlighting potential FUSE hotspots. I also assessed whether the highest 10% of FUSE mammals and birds overlapped with the highest levels of FD for those taxa to understand if the distribution of FUSE species represents the overall distribution of high FD well.

To evaluate the degree of anthropogenic influence within these PAs which can reflect habitat quality, I extracted Forest Integrity Index (FII) values (Grantham et al., 2020) for each PA and subsequently computed the average FII of each IUCN PA category.

Further, to understand the extent to which more stringent IUCN PAs serve as sanctuaries of FD, TD, and FUSE species, I quantified the differences in FD, TD, FUSE species count, proportional composition of FUSE species, and FII between strict (I-IV; do not permit sustainable use of resources) and less-strict (V-VI; allow recreation and sustainable use of resources) PA categories. This evaluation was conducted using a Kruskal-Wallis Test, followed by Dunn tests (Dunn, 1964) for multiple pairwise comparisons.

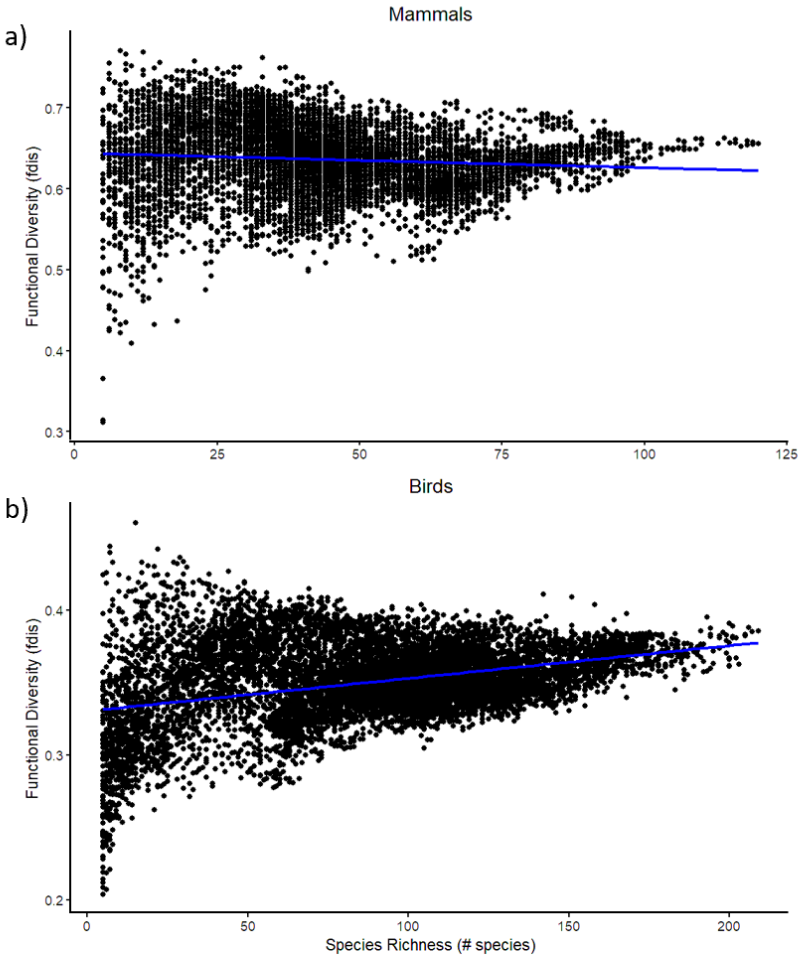
Finally, to assess the effectiveness of community-managed areas in capturing biodiversity across multiple dimensions, I utilized Mann-Whitney U tests. Specifically, I compared the average levels of FD, TD, and FUSE species in community-managed areas classified under IUCN PA category VI with other PAs within the same category.

3.4 Results

3.4.1 Spatial distribution of diversity dimensions

Taxonomic and functional diversity –Our analysis revealed distinctive patterns between taxonomic and functional diversity, for birds and mammals. Specifically, I found a positive correlation between species richness and functional diversity among birds ($\beta = 2.257e\text{-}04$, $p < 0.001$, $R2 = 0.129$; Figure 3.1b), and a negative correlation among mammals ($\beta = -1.815e\text{-}04$, $p < 0.001$, $R2 = 0.007$; Figure 3.1a).

Figure 3.1: Correlation between functional diversity, quantified as functional dispersion (fdis), and species richness (total species count) within each 10 km grid cell across the study area.



Despite these correlations, the strength of these relationships remained modest, as a large portion of the variability in functional diversity (87.1% for birds and 99.3% for mammals) remained unexplained by taxonomic diversity. Subsequent spatial mapping of these biodiversity dimensions unveiled a complex landscape of spatial alignments and mismatches. For birds and mammals, the regions displaying the highest levels of taxonomic diversity were predominantly located at the lower elevations (500 - 1200 m) of the Tropical Andes. This phenomenon occurs in a critical transitional zone where both lowland and montane species can coexist (Lomolino, 2001; Rahbek, 1997) and is indicated by the prevalence of green and yellow hues in Figure 3.2a.

This heightened taxonomic diversity highlights these lower-elevation areas as key reservoirs of species richness. Conversely, areas characterized by high functional diversity were predominantly associated with montane regions at higher elevations (greater than 1200 m) (Figure 3.3a).

Our analysis of distinct upper diversity thresholds (pixels in the highest 10%, 20%, and 30% of diversity) uncovered large spatial mismatches among dimensions of biodiversity. For birds, regions with the highest 30% of diversity (encompassing both functional and taxonomic dimensions) had only 10.41% overlap (roughly 95,628 km^2; Table 3.1; Figure 3A.3c). This overlap was mainly in the western portions of Venezuela and the eastern areas of the Andes in Ecuador, southern Peru, and Bolivia (Figure 3A.3c). Similarly, areas characterized by the upper 10% of diversity exhibited minimal overlap of just 0.68% (6,314 km^2; Table 3.1; Figure 3.4a) and were largely confined to the eastern Andes in Ecuador and northern Bolivia (Figure 3.4a). For mammals, only 0.038% (332 km^2) overlap occurred within the highest 10% of both diversity dimensions (Table 3.1; Figure 3.4b), confined to specific locations in southern Peru and northern Bolivia (Figure 3.4b). Similarly, only 3.02% (26,503 km^2) spatial overlap occurred between the upper 30% of functional and taxonomic diversity, mainly in northwestern Ecuador, southern Peru, and northern to mid-Bolivia (Figure 3A.2c). Between bird and mammal diversity, only 6.19% (53,837 km^2) of bird diversity overlapped with mammal diversity for the highest 30% threshold, primarily in southern Peru and northern Bolivia (Figure 3A.4c). However, for the highest 10% threshold, no overlap was found, indicating that the spatial distribution of the highest values of taxonomic diversity (TD) and functional diversity (FD) between mammals and birds does not align (Table 3.1; Figure 3.4c).

Figure 3.2: Overview of spatial and statistical patterns in mammalian taxonomic diversity across the entire study region and within designated protected areas (PA) categorized as I – VI. Panel a) Depicts the distribution of bird and mammal taxonomic diversity (TD) for frugivorous species, measured as species count per 10 km grid cell. Panel b) Illustrates the average TD within individual PAs (represented as the centroid of each PA). Panel c) Shows the frequency of TD values (normalized to 1 for comparison) throughout the study area (light blue), with dark blue regions indicating the relative distribution of TD within PAs. The dotted red line indicates the highest 10% of TD values.

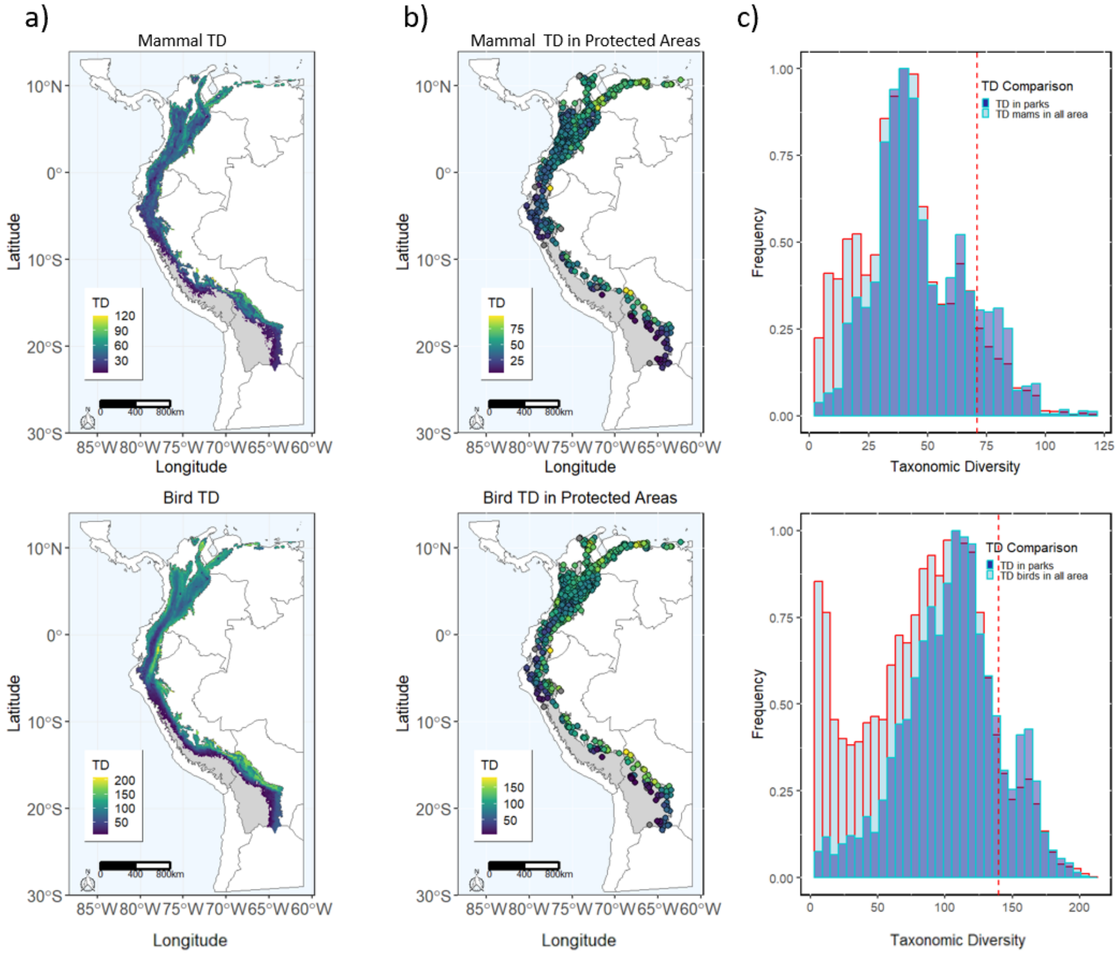
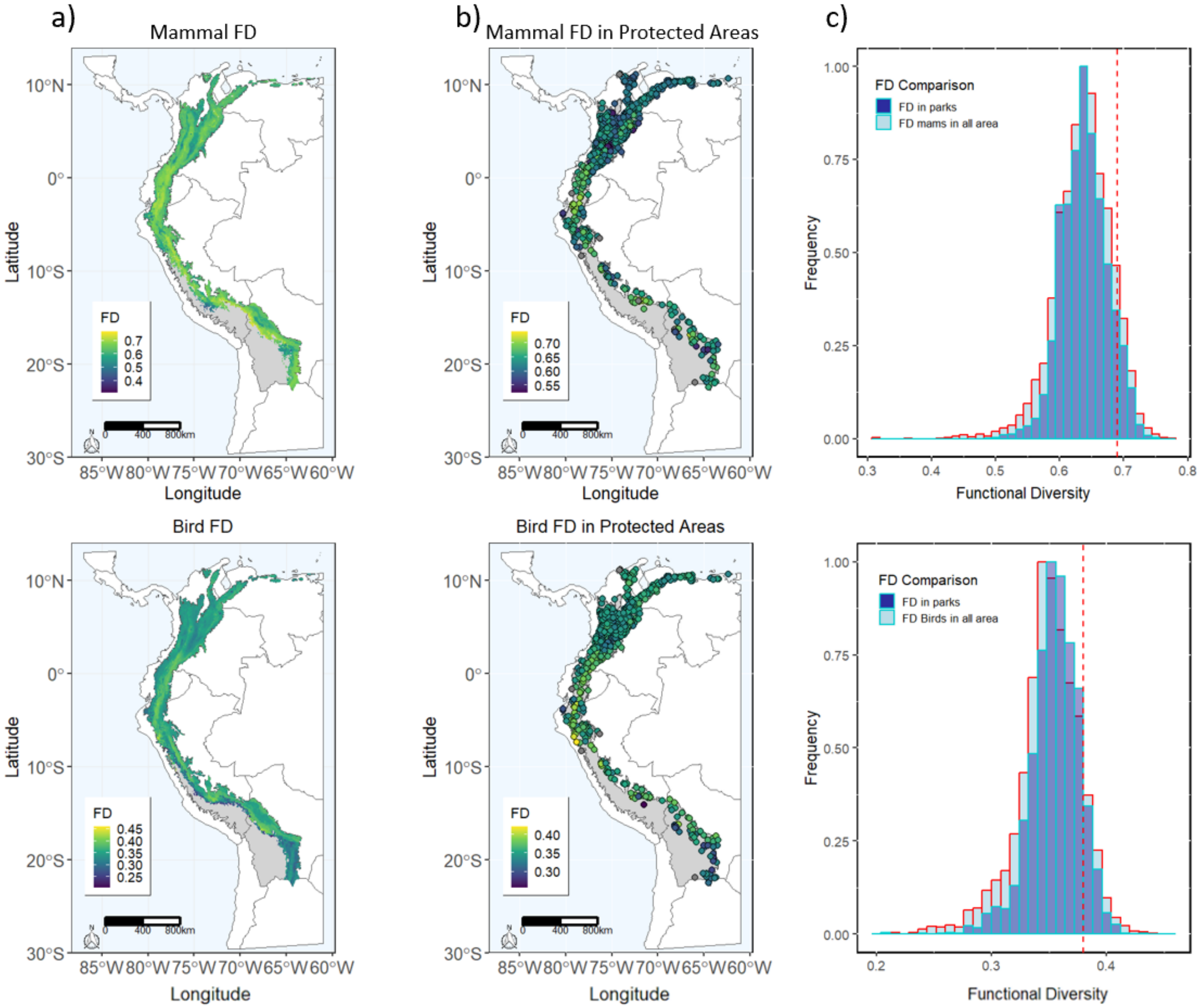


Figure 3.3: Overview of spatial and statistical patterns in mammalian taxonomic diversity across the entire study region and within designated parks and reserves categorized as I - VI. Panel a) Depicts the distribution of bird and mammal functional diversity (FD) for frugivorous species, measured as species count per 10 km grid cell. Panel b) Illustrates the average FD within individual protected areas (PA) (represented as the centroid of each PA). Panel c) Shows the frequency of FD values (normalized to 1 for comparison) throughout the study area (light blue), with dark blue regions indicating the relative distribution of FD within PAs. The dotted red line indicates the highest 10% of FD values.



FUSE species –Our FUSE analysis identified a total of 86 FUSE mammal species and 110 FUSE bird species (Figure 3.5), 43% of which were endemic to the Tropical Andes (Figure 3.5). For the top 10% of FUSE species richness, the areas exhibiting the greatest FUSE species richness did not strongly coincide with areas containing the highest levels of functional diversity

(FD). Specifically, among regions with the top 10% of FUSE species richness or FD, a mere 5.8% of the area overlapped for mammals, and 7.8% overlapped for birds. These results indicate that if conservation measures in this area aim to protect high FD, most areas will not simultaneously protect high FUSE species diversity.

3.4.2 Protected Area assessment

I found that only 15.5% of the Tropical Andes region is protected by PAs with IUCN designations I-VI. Overall, in the areas with the highest levels of diversity (the highest 10, 20, and 30% of diversity values), taxonomic diversity had a slightly higher level of fractional area and protection than functional diversity (TD: 39.8%; FD: 35.4%). For birds, the highest 10% of areas with the greatest functional diversity had slightly greater coverage (37.7%; 89,563 km^2) compared to the area with the highest taxonomic diversity (33.48%; 106,097 km^2) (Table 3.1). However, this coverage dropped to 31.6% in regions where these two diversity dimensions overlapped. For mammals, 34.1% (23,512 km$^{2)}$ of the areas with the highest 10% functional diversity were protected, whereas 42.21% (63,891 km^2) of the areas with the highest 10% taxonomic diversity had protection (Table 3.1). For frugivores, functional diversity had more protection compared to taxonomic diversity at the highest 10% threshold. However, at this threshold there was no protection in the areas where the two frugivore diversity dimensions overlapped, or where the two mammals' diversity dimensions overlapped (Table 3.1). On average, areas of multidimensional diversity were only 22% protected.

Table 3.1: Analysis of upper diversity thresholds at the 10%, 20%, and 30% peaks. This table showcases computed spatial metrics for functional and taxonomic diversity regions, including their intersections within individual taxonomic groups and combined overlapping areas for both frugivorous species. It also indicates the degree of protection offered by IUCN categories I-VI protected areas.

Taxon	**Threshold (Percent)**	**Functional Diversity Area**	**Taxonomic Diversity Area**	**Overlap Area**	**Percent full overlap**	**Functional Diversity Area Protected**	**Taxonomic Diversity Area Protected**	**Overlap Area Protected**	**Percent Functional Diversity Area Protected**	**Percent Taxonomic Diversity Area Protected**	**Percent Overlap Protected**
bird	10	73113.24	73694.82	6314.33	0.69	27583.63	24675.72	1994.00	37.73	33.48	31.58
bird	20	154368.60	157525.80	39464.53	4.30	53505.60	62395.50	10551.57	34.66	39.61	26.74
bird	30	239196.60	251326.80	95628.79	10.41	89563.72	106097.30	37553.62	37.44	42.21	39.27
mammal	10	69042.16	68460.58	332.33	0.04	23512.55	31405.46	0.00	34.06	45.87	0.00
mammal	20	149965.20	149217.50	4901.91	0.56	47025.11	63891.00	664.67	31.36	42.82	13.56
mammal	30	229891.30	231802.20	26503.55	3.02	80673.81	94548.71	6231.24	35.09	40.79	23.51
frugivores	10	70786.91	37221.28	0.00	0.00	23927.97	11216.24	0.00	33.80	30.13	0.00
frugivores	20	150463.70	147638.90	15785.81	1.82	56247.34	59986.09	4320.33	37.38	40.63	27.37
frugivores	30	229725.10	233463.90	53837.93	6.19	84827.97	100115.30	17032.06	36.93	42.88	31.64

Figure 3.4: Intersection of the highest 10% taxonomic and functional diversity distributions. Panel a) illustrates diversity overlap for birds, Panel b) displays diversity overlap for mammals, and Panel c) depicts the regions of shared above-threshold TD and FD between birds and mammals. Notably, no areas of TD/FD overlap are observed between birds and mammals at this threshold.

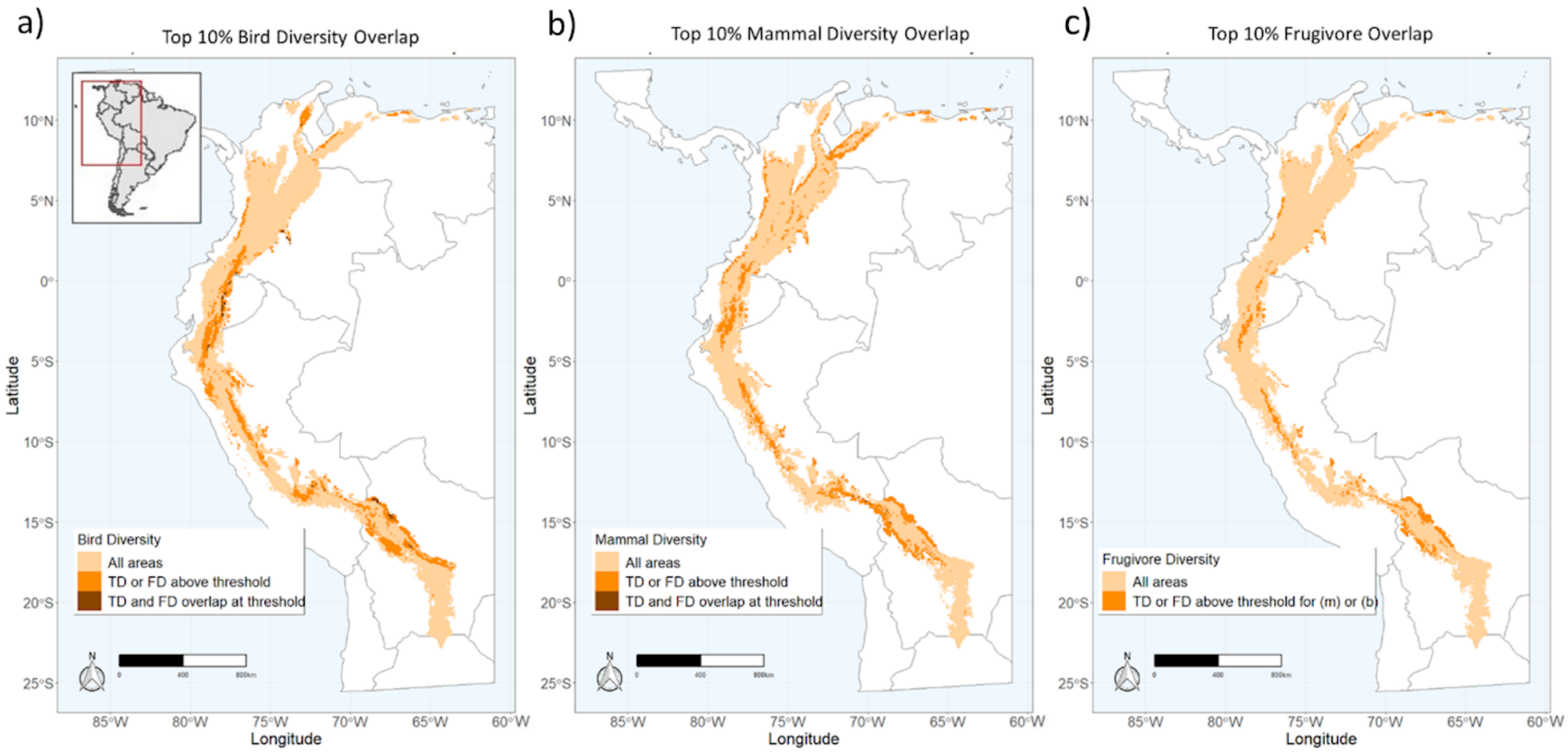


Figure 3.5: Top 10 endemic Andean species with the Highest Functionally Unique, Specialized, and Endangered (FUSE) Scores. Colored circles indicate the IUCN threat status for each species corresponding to their respective FUSE scores.

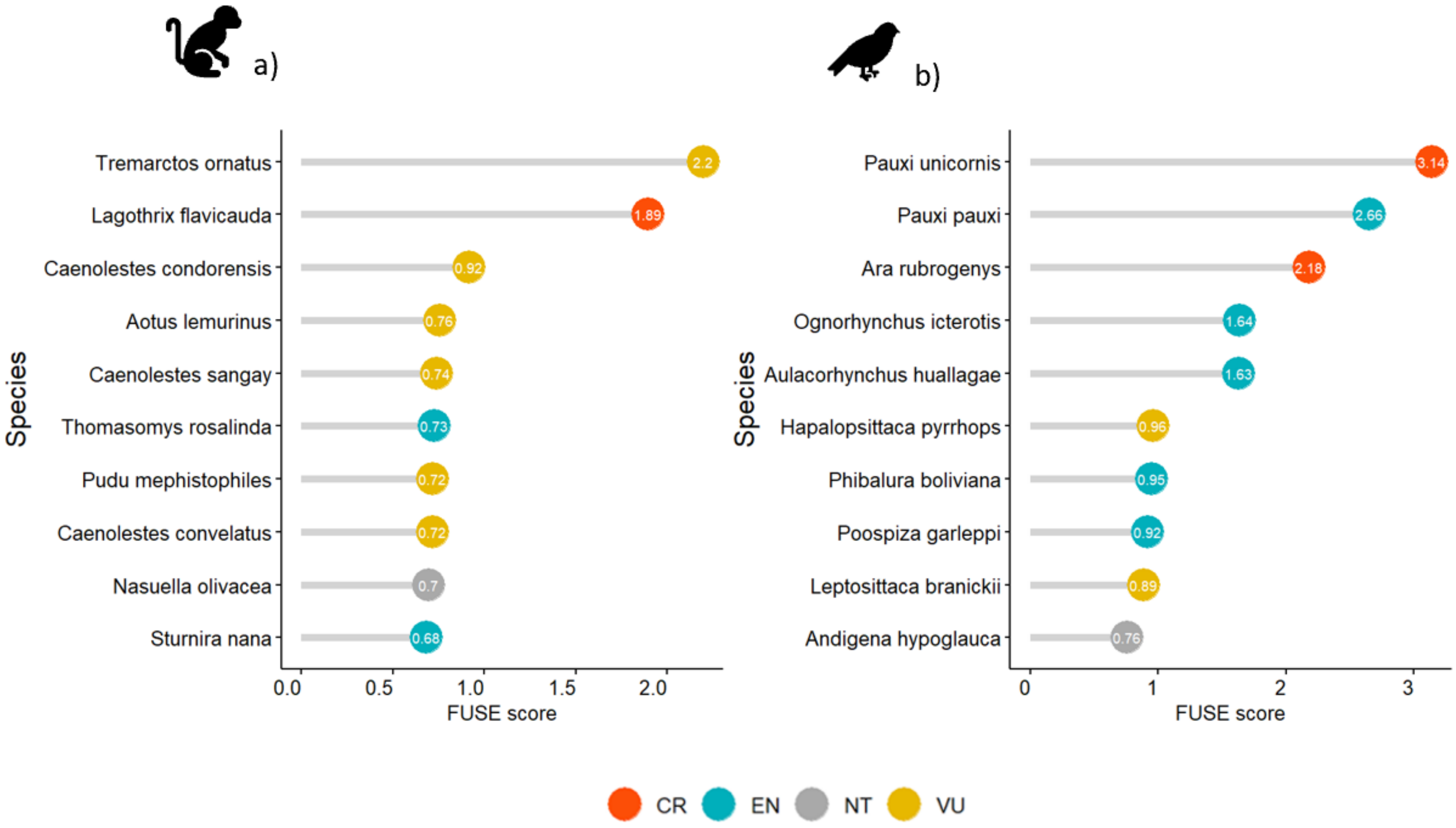


Figure 3.6: Functionally Unique, Specialized, and Endangered (FUSE) Species and Forest Integrity Across IUCN PA Categories. Panel a) illustrates the average count of FUSE species for mammals and birds within each protected area (PA) type. Panel b) depicts the average forest integrity index (Grantham et al., 2020) calculated for each PA category. Whiskers on both plots represent the standard deviation of values within each category.

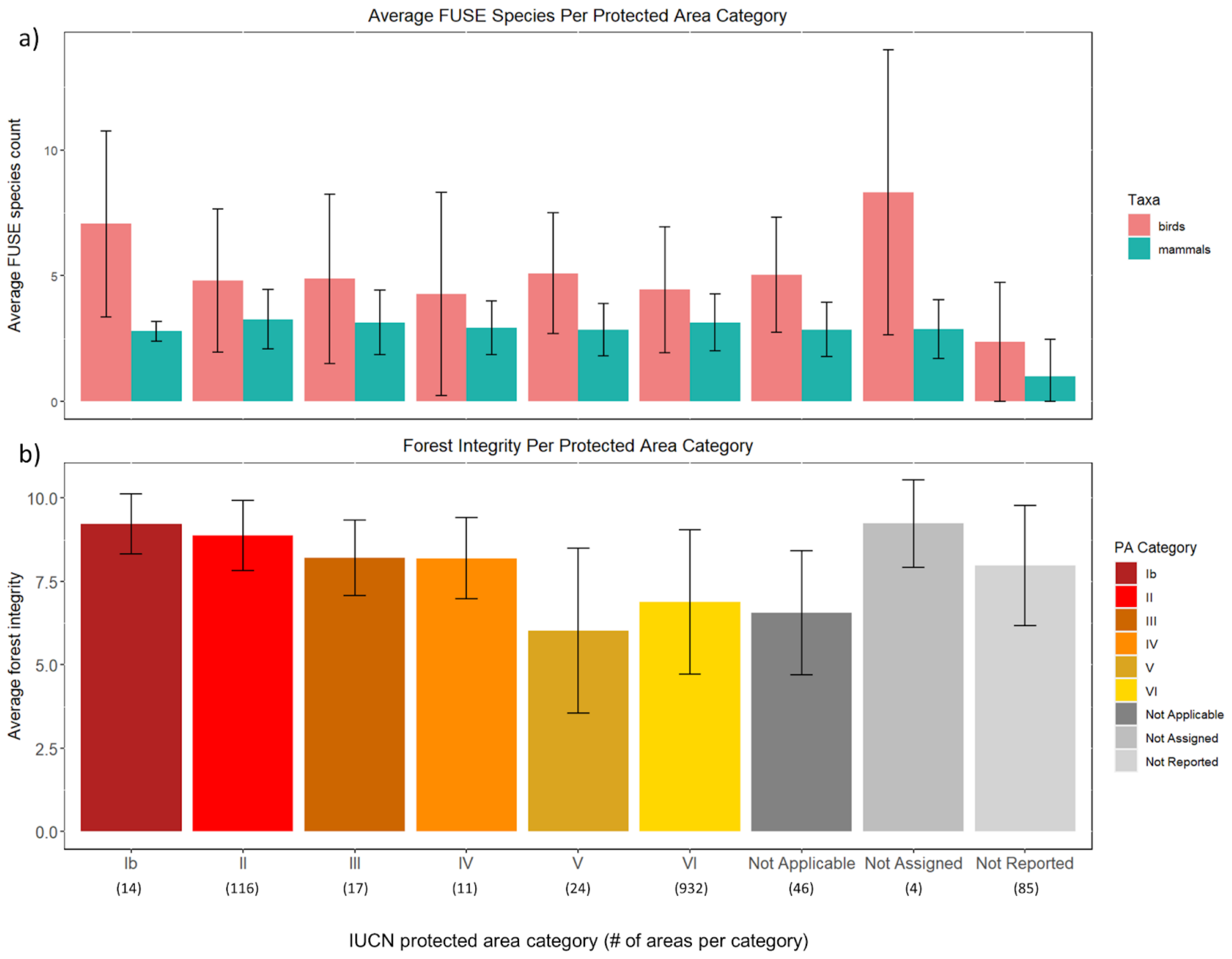
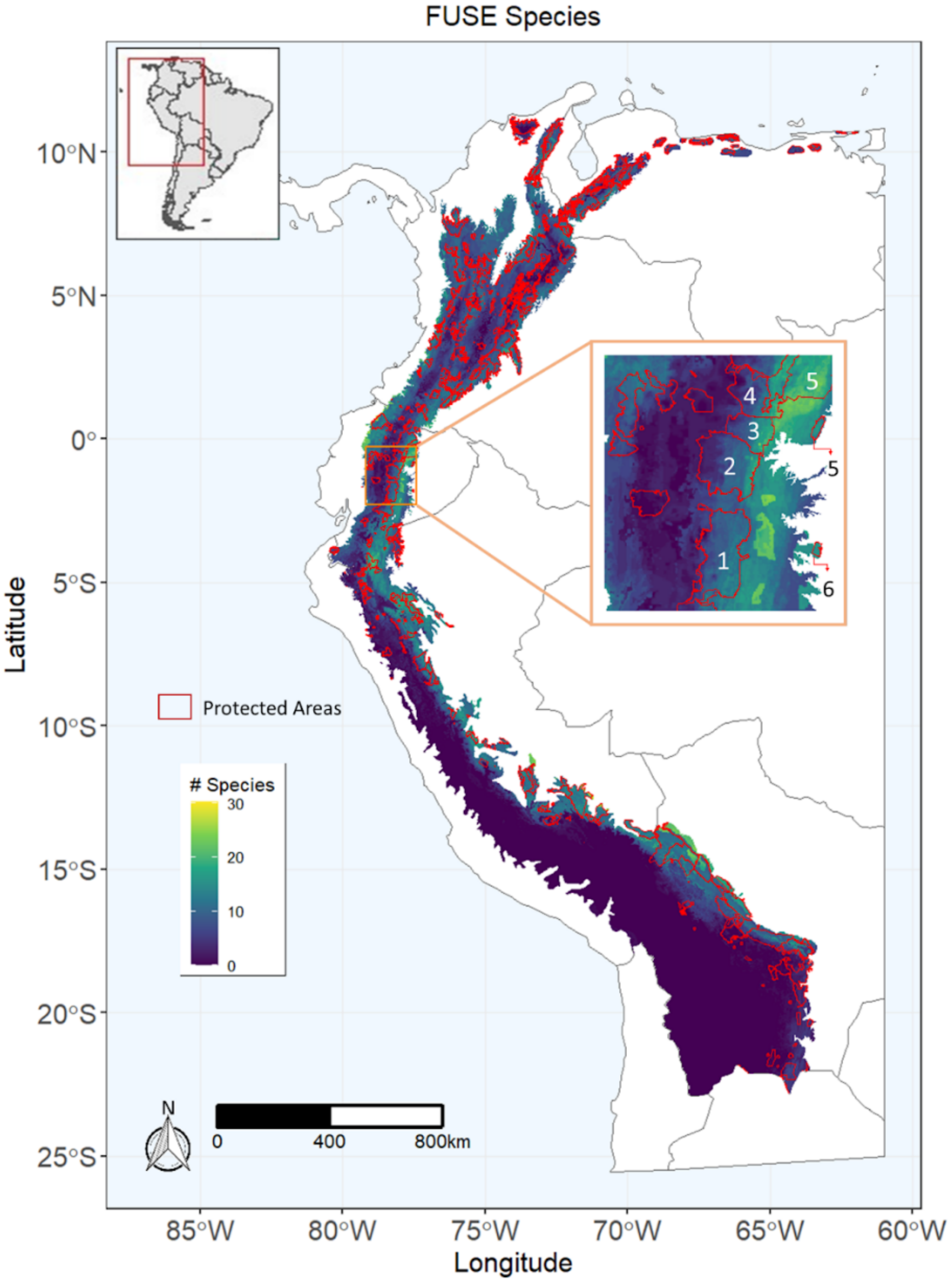


Table 3.2: Provides an overview of diversity metrics computed across various IUCN protected area categories. This includes the count and percentage composition of FUSE (Functionally Unique, Specialized, and Endangered) species, as well as the average functional and taxonomic diversity for each taxon. The table also incorporates the average forest integrity index (FII; Grantham et al. 2020). Notably, the table highlights the more stringent IUCN PA categories (I-IV) and metrics for which the values surpassed at least one of the less strict categories (V-VI), with significant statistical differences being indicated in bold. Values in bold accompanied by an (*) signify statistically significant values that are lower than those of less strict categories.

PA Category	**# Protected Areas in category**	**Average % composition FUSE frugivores**	**Average Mammal FUSE species**	**Average Bird FUSE species**	**Average Bird FD**	**Average Bird TD**	**Average Mammal FD**	**Average Mammal TD**	**Average Forest Integrity Index**
Ib (strict nature reserve)	7	6.93	2.79	7.06	0.36	100.59	0.65	35.90	**9.22**
II (wilderness area)	116	5.61	**3.27**	4.81	0.35	98.96	0.64	45.52	**8.88**
III (natural monument or feature)	17	6.01	3.14	4.88	0.36	**119.31**	0.64	**58.28**	**8.20**
IV (habitat/species management zone)	11	6.13	2.93	4.27	**0.38**	77.17	0.66	**34.23***	**8.19**
V (protected landscape)	24	5.36	2.86	5.10	0.36	102.00	0.64	49.09	6.02
VI (PA with sustainable use of resources)	932	5.49	3.14	4.45	0.35	93.47	0.64	44.17	6.88
Not Applicable	46	5.37	2.86	5.04	0.35	102.09	0.63	47.82	6.56
Not Assigned	4	6.58	2.88	8.31	0.36	116.79	0.64	48.32	9.23
Not Reported	85	2.47	1.00	2.37	0.33	52.20	0.64	27.75	7.98

Figure 3.7: Distribution of Functionally Unique, Specialized, and Endangered (FUSE) frugivore species. This map presents the aggregated binary Area of Habitat (AOH) maps for FUSE species found within the Tropical Andes. The outlined red areas correspond to all protected areas sourced from the World Database on Protected Areas (WDPA; protectedplanet.com). The zoom box highlights an Ecuadorian region with significant FUSE species numbers adjacent to parks lacking protection. The labeled parks, accompanied by their IUCN park category in parentheses, are as follows: 1: Sangay (II); 2: Llanganates (II); 3: Colonso Chalupas (Ib); 4: Antisana (II); 5: Sumaco Napo-Galeras (Level II); 6: Pueblo Originario Kichwa de Sarayaku (Not Assigned; Indigenous governance).



For FUSE species, the percent composition of FUSE frugivores in category “Not Assigned” had the second highest average values (Table 3.2; Figure 3.6a), while the highest was in strict nature reserves Ib. Among birds, the Ib,V, and Not Assigned categories had the highest average numbers of FUSE species (Figure 3.6a). In contrast, for mammals, categories II, III, and VI had the highest numbers of FUSE species (Figure 3.6a). Despite this, overlaying FUSE species distributions showed that some high diversity areas lack protection from established PAs (Figure 3.7).

When evaluating whether diversity varied between strict and less strict IUCN PA categories, I only detected differences for some categories (Table 3.2). Category III had higher values of average TD than VI for birds and mammals (Kruskal-Wallace test, $p < 0.05$ for birds and mammals, respectively), while category IV had higher average FD for birds when compared with category VI (Kruskal-Wallace test, $p < 0.05$, TD mammals and FD birds, respectively). Interestingly, for mammals, category VI had higher average than category IV (Kruskal-Wallace test, $p < 0.05$, TD mammals). There were no differences among mammal FD, average count of FUSE birds, or proportions of FUSE species within strict and less-strict categories, however category II was higher than V and VI for FUSE mammals (Kruskal-Wallace test, $p < 0.05$, respectively). For FII, all strict PA categories were higher than less-strict categories (Kruskal-Wallace test, $p < 0.05$ for all comparisons).

When comparing the levels of diversity in community managed areas and other areas within the same IUCN PA category (VI), significant differences were detected for FD and TD (Mann-Whitney U test $p < 0.05$; both mammals and birds respectively). FD was significantly higher for mammals and birds within 32 community managed areas within PA category VI than other PAs within the same category, and this was more apparent for mammals (~5% increase of

FD within community managed areas). Interestingly, TD was significantly lower within community managed areas for both birds and mammals. There was no significant difference between levels of FUSE species in and outside community managed areas, except for FUSE mammals, where the number of FUSE species was lower inside community managed areas.

3.5 Discussion

This analysis unveiled a weak relationship between TD and FD (Figure 3.1) and large spatial mismatches between their geographic distributions and levels of protection, shedding light on the challenges associated with adopting an integrated approach to biodiversity conservation. Further, there were heightened levels of TD in lower-elevation areas within the Tropical Andes, serving as key reservoirs of species richness. Conversely, areas characterized by the highest levels of functional diversity were predominantly associated with montane regions situated at higher elevations (greater than 1200 m) (Figure 3.3a; Figure 3A.1). When assessing levels of protection, the results showed high biodiversity areas, both taxonomically and functionally, were only about 37% protected when considered independently with protection dropping to 0% for mammals when focusing on areas where diversity dimensions overlapped. Further, I found that high numbers of FUSE mammals, levels of bird FD, bird TD, and mammal TD were elevated in some strict protected areas when compared with less strict areas, highlighting the importance of certain strict protective measures in the maintenance of biodiversity. I also found that community managed areas within IUCN category VI harbored higher FD than other PAs within the same category, highlighting these areas as potential strongholds for this dimension of diversity.

The weak diversity relationships and spatial mismatches align with findings from other studies conducted across diverse regions and taxa. For instance, similar patterns of mismatches occur for birds in France (Devictor et al., 2010) and the Brazilian Cerrado (Borges et al., 2021), as well as for plants in China (Xu et al., 2019), global mammals (Brum et al., 2015), and rodents in Mexico (Martín-Regalado et al., 2019). These incongruencies between dimensions of diversity can occur for multiple reasons that differ based on the historic processes and ecological constraints of the taxa and region of study. For example, for birds, there was a weak positive relationship between TD and FD. Similar to Dehling et al. (2014), the highest species richness for birds occurred in lower elevations of the Tropical Andes, suggesting that for frugivorous birds, at these lower elevations there might be high interspecific competition leading to niche differentiation and therefore slightly higher FD (Figure 3A.1a). Conversely, for mammals, TD had a very weak negative association with FD, indicating that while certain lower elevation areas boast a greater abundance of species, frugivorous mammals in the Tropical Andes may exhibit slight "niche packing" at lower elevations leading to functional redundancy (Mouchet et al. 2010), with higher FD at intermediate elevations possibly due to increased levels of spatial heterogeneity (Safi et al. 2011) (Figure 3A.1b). It is also important to note that these relationships were tenuous and therefore caution should be taken when generalizing patterns as they have been shown to differ among mountain ranges and even within individual mountains (Montaño-Centellas et al., 2020). However, despite these weak patterns, the overall highest values of FD occurred at intermediate elevations for both taxa (Figure 3A.1a & b), differing from the highest TD.

Understanding the potential spatial trade-offs that arise when prioritizing one dimension of diversity over another is a crucial consideration when expanding existing protected areas (PAs) or designing and planning new ones (Cadotte and Tucker, 2018; Faith and Walker, 2002). This becomes particularly significant given that the criteria set by the IUCN for identifying Key

Biodiversity Areas (KBAs), which are locations contributing significantly to the global persistence of biodiversity, recognize the importance of taxonomic diversity (TD) and phylogenetic diversity. However, functional diversity (FD) is not explicitly considered in these criteria (IUCN, 2016). Consequently, international standards for identifying regions with the highest diversity for spatial conservation may not fully account for critical areas characterized by high FD, which has substantial implications for the maintenance of ecosystem functioning and the delivery of ecosystem services. The mismatches found here and in other studies further support this assertion and indicate that focusing solely on one diversity dimension may lead to conservation strategies that fail to adequately capture biodiversity (Brum et al., 2017; Devictor et. al., 2010; Strecker et al., 2011; Mazel et al. 2018; Borges et al. 2021). In the context of these disparities, our study further revealed that overall protection of high diversity areas is low (~30%), with the highest levels of functional diversity (FD) experiencing lower protection than those with high taxonomic diversity (TD), albeit by a modest margin (a difference of 4.4%) (Table 3.1). Despite the limited geographical extent of PAs in the Tropical Andes (15.4% of area with protection levels I-VI), they cover a representative proportion of the overall diversity for frugivores (Figures 3.2c & 3.3c). However, there was a noticeable lack of protection afforded to regions characterized by the highest levels of FD (Figure 3.3c). This underscores the potential for targeted conservation efforts aimed at protecting areas with the highest levels of FD.

Despite the pronounced disparities in taxonomic and functional diversity distributions, our study uncovered instances of minimal convergence between these dimensions for mammals and birds. For instance, when examining the upper 10% of overlapping diversity areas for mammals, there was a mere 0.038% of overlap (Figure 3.4b), primarily concentrated in southern Peru and northern Bolivia, with no corresponding protection under IUCN designations I-VI

(Table 3.1). In contrast, the outlook for birds was more optimistic, with 31.6% of this overlapping diversity area receiving some level of protection (Table 3.1). These shared areas are pivotal for conservation strategies, particularly for approaches that aim to optimize biodiversity outcomes, especially for mammals. The co-occurrence of elevated species richness and diverse ecological roles in these areas underscores their heightened significance for maintaining ecosystem functioning. It is important to note, however, that the diversity thresholds used here reflect the most stringent criteria for evaluating overlapping diversity and managers may choose to look at a broader range of diversity thresholds to achieve their goals. Further, there was no overlap between the highest levels of diversity when considering frugivores (combined birds and mammals; Figure 3.4c), indicating that measures taken to conserve the highest levels of frugivore diversity should consider birds and mammals separately. For areas containing less strict peaks of diversity (Figure 3A.4c) it may be prudent to consider spatial prioritization of frugivores together. However, when prioritizing area for conservation action based on the overlap of functional and taxonomic diversity within taxa, or considering each dimension of diversity separately, critical roles of unique and specialized species that are crucial for maintaining ecosystem functioning could be missed (Pimiento et al., 2020; Kosman et al., 2019; Griffin et al. 2020). While functional dispersion gauges within-community diversity, it overlooks the distinctive and specialized contributions of individual species within the broader community (Laliberte & Legendre, 2010; Kosman et al., 2019). For example, a recent study found that while scarlet macaws are widespread and not of conservation concern, they are functionally singular and thus should be prioritized for conservation (Kosman et al., 2019). In contrast, South American tapirs are large-bodied and highly threatened frugivores with outsized roles as long-distance seed dispersers. Subsequently, their loss would have cascading effects for ecosystem function (O’Farrill et al., 2013). Future management might be interested in prioritizing these functionally singular and threatened species whose role cannot be replaced if the species is lost from the community (Griffin et al., 2020).

Evaluation of the relationship between high FUSE areas and areas with high FD revealed that for both birds and mammals these areas were not closely aligned. There was limited spatial overlap between the highest 10% of FUSE richness and FD for both groups, indicating that conservation efforts focused on maximizing FD might not inherently encompass regions with a high abundance of FUSE species. This discrepancy may be attributed to the metric used to calculate functional dispersion in this study. Specific FUSE species, which possess extreme traits, can be outliers within their respective communities. Consequently, even if a community harbors one or two FUSE species with highly distinct traits, their greater distance from the centroid of trait space is averaged out when calculating FD. These findings underscore the need for explicit consideration of FUSE species when evaluating complementarity for establishing future spatial conservation priorities. In the past decade, it has been recommended that EDGE species be included in the standards for assigning KBAs under Criteria A (threatened biodiversity; Brooks et al. 2015; Faith et al. 2015), similarly, FUSE species should be explicitly considered in this framework, as these species directly tie to ecosystem function and can help maintain healthy ecosystems.

Our assessment of both FUSE species and overall diversity yielded valuable insights into the efficacy of PAs in capturing diversity. The highest average FD and TD were always found in more strict categories (I-IV) (Table 3.2). Similarly, these categories had the highest average proportions and counts of FUSE species (Table 3.2; Figure 3.6a). Specifically, FUSE frugivores had the highest percent composition in Ib (strict nature reserves), potentially due to these areas being relatively intact and undisturbed by humans. However, when evaluating stricter categories with statistically significant differences from less strict categories, category II (National Parks) was more important for capturing FUSE mammal diversity, while category III (Natural Monuments or Features) better captured TD across taxa (Table 3.2). Category IV (habitat/species management zones), however, only captured the FD of birds, possibly benefiting from species specific management strategies. Conversely for this same category (IV), mammals showed significantly less TD than in less strict PAs (specifically VI), possibly indicating that VI areas, which allow sustainable use of resources, can be a successful strategy when the goal is to conserve higher numbers of mammalian frugivores. Overall, the strict categories of PAs inadequately represented the functional diversity (FD) of mammals and FUSE birds when compared to PAs with less stringent measures. This underscores the necessity for the development of novel and stringent protective protocols that specifically consider areas of complementarity in terms of high mammal FD and FUSE birds.

Importantly, all the stricter PAs had higher levels of forest integrity (Figure 3.6b; Table 3.2), implying a promising connection between robust protective measures and the quality of habitat, which is especially important for threatened FUSE species. Some FUSE species have already had measures put into place for their species-specific conservation. For example, the helmeted curassow (*Pauxi pauxi*), an endemic and endangered bird with one of the highest FUSE scores (2.66; Figure 3.5b), has a network of reserves relegated to its protection in the Eastern Cordillera of Colombia (Proaves, 2023). Being a large bodied frugivore, this species likely has an important role in dispersing large seeds, which contributes to maintenance of forest structure (Sales et al., 2021). Further, the Wildlife Conservation Society is actively partnering with institutions in Peru to aid in the protection of two FUSE mammals, the Andean Bear (*Tremarctos* *ornatus*) (Figure 3.5a). and white-lipped pecary (*Tayassu pecari*) (WCS, 2023) Maintaining both the quality and quantity of habitat for FUSE species is especially important because of the known correlations between habitat loss and the decline of species with unique traits. For instance, previous research has highlighted that the species richness of frugivorous species may be related to body size, a key functional trait, where loss of larger bodied species disrupts seed dispersal networks and changes forest composition (Bomfim et al., 2018; Fricke et al., 2022; Lim et al., 2020; Sales et al., 2020, 2021). Safeguarding these unique and ecologically vital species is paramount to the functioning of the Tropical Andes, underscoring the need to protect and restore habitats that support them.

However, certain pockets of substantial FUSE species richness remain beyond the protection of established PAs and are important regions to prioritize for conservation. For instance, three disjunct areas are situated 10 km east of a chain of national parks within Ecuador, including Sangay National Park and Llanganates National Park (both IUCN category II; Figure 3.7: labels 1 & 2) in the eastern Cordillera (Figure 3.7). These areas host high levels of FUSE species, surpassing 8% of species in some regions. These neighboring parks may serve as important source populations for many FUSE species found in these areas and bolster local diversity (Hansen, 2011). For example, Sangay National Park is one of three areas with the highest concentration of Andean bears (*Tremarctos ornatus)* (UNESCO, 2023), an important seed disperser with the IUCN status near threatened (IUCN, 2023), and a species with the highest FUSE score for mammals endemic to the Tropical Andes (Figure 3.5). Similarly, the same area is home to the yellow throated toucan (*Ramphastos ambiguus*), another near threatened FUSE species. The importance of conserving areas adjacent to these parks extends beyond their inherent high FUSE diversity; they may serve as vital corridors (Cushman et al., 2009), buffer zones (Weisse and Naughton-Treves, 2016), and essential components in ensuring the resilience and connectivity of the broader region.

While nearby PAs may serve as source populations, extensions of parks or the inclusion of complementary areas could play a pivotal role in ensuring the sustained viability and resilience of FUSE species. Creating entirely new parks demands substantial resources and presents logistical challenges. In contrast, a more pragmatic strategy may involve assessing complementarity and augmenting or extending existing reserves to comprehensively encompass various dimensions of biodiversity (i.e., high FD, high TD, or high proportions of FUSE species distributions), however this may benefit dimensions of biodiversity differently. The extension of PA boundaries can have a steep increase in protection for functional diversity, even if species richness does not increase markedly (Pollock et al. 2017).

Noteworthy hotspots of FUSE frugivores emerged within the IUCN PA category labeled as "Not Assigned". This category exists in large part because data suppliers for the WDPA (WDPA, 2023) chose not to assign an official IUCN PA category (I-VI). This is likely because they do not officially adhere to IUCN standards and guidelines and is similar for the other PA category “Not Reported” (Dudley, 2008) Of the IUCN PA categories, “Not Assigned” showed the highest count and proportion of FUSE species (6.58%; Table 3.2). This observation can be attributed to three conservation zones designated in Ecuador and Peru. Two of these zones have been strategically established for and by indigenous peoples to maintain biodiversity and provide essential ecosystem services to their local communities (Figure 3.6a). One of these areas was the Territorio del Pueblo Shuar Arutam (6.49% FUSE; avg # FUSE mammals: 1.8; avg # FUSE birds: 8.55; FII: 9.5; Table 3.2), nestled in the Ecuadorian Andes (ICCA, 2023). Another notable example was the Pueblo Originario Kichwa de Sarayaku (5.41% FUSE; avg # FUSE mammals:

4.24; avg # FUSE birds: 11.34; FII - 9.9; Table 3.2), situated in the Andean foothills of Ecuador (see Figure 3.7, label 6). Remarkably, this area boasts an impressive 95% of untouched primary forest, as detailed on their official website (Sarayaku, 2022). Moreover, the Nieva River region (9.48% FUSE; avg # FUSE mammals: 3.49; avg # FUSE birds: 13.07; FII: 9.7) within Peru, falls under the stewardship of the National System of Natural Areas Protected by the State (SINANPE). Its mission is to ensure ecological processes continue, preserving high biodiversity in the Nieva River basin for the benefit of nearby communities, and to create a conservation corridor in the Peruvian Yungas (GOB, 2023).

Community-managed reserves are essential for conservation of biodiversity and ecosystem functions (Alves-Pinto et al., 2021; Esmail et al., 2023). This study confirms the importance of community-managed reserves—functional diversity was higher in the 32 community-managed reserves classified under the IUCN park designation VI compared to other PAs within the same category. Specifically, levels of FD for birds and mammals were higher in community managed areas (bird - community 0.36, other VI PAs 0.35; mammals - community 0.67, other VI PAs 0.64), with the most pronounced difference being for mammals (average of increase of ~5%). Conversely, TD was significantly lower in these areas, indicating that community managed reserves are important for maintaining FD specifically. This is especially important given that strict PAs did not adequately support this dimension of diversity. The connection between FD and community-managed areas can be attributed to the profound impact of their stewardship practices that are intricately woven into the fabric of their traditions, perspectives, knowledge systems, and their connection to their landscapes (Kennedy et al., 2023). These communities' stewardship often extends beyond mere protection and embraces traditional practices that may actively promote and enhance biodiversity (Velho et al., 2016), and specifically the functional diversity of their ecosystems, however this relationship remains relatively unexplored. For instance, indigenous and local communities, through their intimate understanding of their environments, may favor traditional land management techniques that preserve the diverse array of species and ecological roles within their ecosystems (IPBES, 2022). This includes maintaining habitats for a range of plant and animal species, ensuring that key ecological functions, such as pollination and seed dispersal (IPBES, 2022), which have the potential to shape landscapes both historically and sustain them into the future (Ribeiro et al., 2014). These practices, deeply embedded in their cultural traditions and knowledge systems, may contribute significantly to the richness of functional traits and roles present within the ecosystem, thus elevating FD.

These examples underscore the crucial role that community-managed lands can play in the conservation of birds and mammals exhibiting frugivory. To protect 30% of land by 2030, it is unlikely this can occur through the expansion of PAs alone (Cook, et al. 2022; Dudley et al., 2018). The designation of other effective conservation measures (OECMs; CBD, 2018) are a promising way to achieve this goal as much biodiversity is occurring on private, community, or Indigenous Lands (Alves-Pinto et al., 2021; Garnett et al., 2018; O’Bryan et al., 2021). Our results show the importance of including OECMs in the 30X30 target and support the need to foster partnerships that empower local communities to sustainably manage and protect their lands as a way of enhancing biodiversity, particularly for frugivorous species. Collaborative engagement with indigenous peoples and local communities aligns conservation efforts with local values, enhancing the chances of success in protecting and maintaining biodiversity. This alignment between indigenous and local knowledge and modern conservation objectives could catalyze more inclusive and robust conservation strategies within the 30X30 initiative (Cook et al. 2022).

As this study and previous findings reveal, the pursuit of comprehensive biodiversity conservation of frugivorous species may entail biodiversity trade-offs (Faith & Walker, 2002; Cadotte & Tucker et al., 2018). Striking the right balance requires open engagement with partners, and integrating ecological insights, societal values, and effective governance mechanisms. This type of prioritization has been proven to conserve phylogenetic diversity (Bennett et al., 2014; Carvalho et al., 2017; Forest et al., 2007; Robuchon et al., 2021; Velazco et al., 2021), however there are few examples for functional diversity –likely because conservation outcomes related to maintenance of ecosystem functioning are difficult to quantify (Cadotte et al., 2011). Thus, devising inclusive strategies to address these biodiversity trade-offs and measuring success of initiatives for species providing essential ecosystem services such as frugivores is an essential endeavor for policymakers and conservation practitioners alike. Ultimately, this holistic approach will facilitate a more balanced and inclusive approach to protect the future of biodiversity and ecological functions.

3.6 Data Availability Statement

All the code used in this study, including data processing, analysis, and visualizations, is publicly available on GitHub at https://github.com/bioXgeo/neotropical_frug_diversity.

REFERENCES

Alves-Pinto, H., Geldmann, J., Jonas, H., Maioli, V., Balmford, A., Ewa Latawiec, A., Crouzeilles, R., Strassburg, B., (2021). Opportunities and challenges of other effective area-based conservation measures (OECMs) for biodiversity conservation. Perspectives in Ecology and Conservation, 19, 115–120. https://doi.org/10.1016/j.pecon.2021.01.004

Bennett, J.R., Elliott, G., Mellish, B., Joseph, L.N., Tulloch, A.I.T., Probert, W.J.M., Di Fonzo, M.M.I., Monks, J.M., Possingham, H.P., Maloney, R., (2014). Balancing phylogenetic diversity and species numbers in conservation prioritization, using a case study of threatened species in New Zealand. Biological Conservation, 174, 47–54. https://doi.org/10.1016/j.biocon.2014.03.013

Bird, J.P., Martin, R., Akçakaya, H.R., Gilroy, J., Burfield, I.J., Garnett, S.T., Symes, A., Taylor, J., Şekercioğlu, Ç.H., Butchart, S.H.M., (2020). Generation lengths of the world's birds and their implications for extinction risk. Conservation Biology, 34, 1252–1261. https://doi.org/10.1111/cobi.13486

Bird Life International (2023). Partnership for nature and people. URL http://www.birdlife.org/ (accessed 2.03.23).

Bomfim, J., Guimarães Jr, P.R., Peres, C.A., Carvalho, G., Cazetta, E., (2018). Local extinctions of obligate frugivores and patch size reduction disrupt the structure of seed dispersal networks. Ecography, 41, 1899–1909. https://doi.org/10.1111/ecog.03592

Borges, F.J.A., Fortunato, D. de S., Loyola, R., (2021). Critical areas for retaining multiple dimensions of bird diversity in the Cerrado. Journal for Nature Conservation, 64, 126079.

Brum, F.T., Graham, C.H., Costa, G.C., Hedges, S.B., Penone, C., Radeloff, V.C., Rondinini, C., Loyola, R., Davidson, A.D., (2017). Global priorities for conservation across multiple dimensions of mammalian diversity. Proceedings of the National Academy of Sciences. 114, 7641–7646. https://doi.org/10.1073/pnas.1706461114

Cadotte, M.W., Carscadden, K., Mirotchnick, N., (2011). Beyond species: functional diversity and the maintenance of ecological processes and services. Journal of Applied Ecoly, 48, 1079–1087. https://doi.org/10.1111/j.1365-2664.2011.02048.x

Cadotte, M.W., Tucker, C.M., (2018). Difficult decisions: Strategies for conservation prioritization when taxonomic, phylogenetic and functional diversity are not spatially congruent. Biological Conservation, 225, 128–133. https://doi.org/10.1016/j.biocon.2018.06.014

Carvalho, S.B., Velo-Antón, G., Tarroso, P., Portela, A.P., Barata, M., Carranza, S., Moritz, C., Possingham, H.P., (2017). Spatial conservation prioritization of biodiversity spanning the evolutionary continuum. Nature Ecology & Evolution, 1, 151. https://doi.org/10.1038/s41559-017-0151

CBD. (2018). Decision Adopted by the Conference of the Parties to the Convention on Biological Diversity: Protected Areas and Other Effective Area-based Measures. Sharm El-Sheikh, Egypt, pp. 1–19.

Cook, C.N., (2022). Progress developing the concept of other effective area-based conservation measures. Conservation Biology, e14106. https://doi.org/10.1111/cobi.14106

Cooke, R.S.C., Bates, A.E., Eigenbrod, F., (2019). Global trade-offs of functional redundancy and functional dispersion for birds and mammals. Global Ecology and Biogeography, 28, 484–495. https://doi.org/10.1111/geb.12869

Cooke, R.S.C., Eigenbrod, F., Bates, A.E., (2020). Ecological distinctiveness of birds and mammals at the global scale. Global Ecology and Conservation, 22, e00970. https://doi.org/10.1016/j.gecco.2020.e00970

Cushman, S.A., McKELVEY, K.S., Schwartz, M.K., (2009). Use of Empirically Derived Source-Destination Models to Map Regional Conservation Corridors. Conservation Biology, 23, 368–376. https://doi.org/10.1111/j.1523-1739.2008.01111.x

de Bello, F., Botta-Dukát, Z., Lepš, J., Fibich, P., (2021). Towards a more balanced combination of multiple traits when computing functional differences between species. Methods in Ecology and Evolution. 12, 443–448. https://doi.org/10.1111/2041-210X.13537

Dehling, D.M., Töpfer, T., Schaefer, H.M., Jordano, P., Böhning-Gaese, K., Schleuning, M., (2014). Functional relationships beyond species richness patterns: trait matching in plant–bird mutualisms across scales. Global Ecology and Biogeography, 23, 1085–1093. https://doi.org/10.1111/geb.12193

Dell, A.I., Pawar, S., Savage, V.M., (2014). Temperature dependence of trophic interactions are driven by asymmetry of species responses and foraging strategy. Journal of Animal Ecology, 83, 70–84. https://doi.org/10.1111/1365-2656.12081

Devictor, V., Mouillot, D., Meynard, C., Jiguet, F., Thuiller, W., Mouquet, N., (2010). Spatial mismatch and congruence between taxonomic, phylogenetic and functional diversity: the need for integrative conservation strategies in a changing world. Ecology Letters, 13, 1030–1040. https://doi.org/10.1111/j.1461-0248.2010.01493.x

Dudley, N. (Editor), (2008). Guidelines for Applying Protected Area Management Categories. Gland, Switzerland: IUCN. x + 86pp. WITH Stolton, S., P. Shadie and N. Dudley (2013). IUCN WCPA Best Practice Guidance on Recognising Protected Areas and Assigning Management Categories and Governance Types, Best Practice Protected Area Guidelines Series No. 21, Gland, Switzerland: IUCN.

Dudley, N., Jonas, H., Nelson, F., Parrish, J., Pyhälä, A., Stolton, S., Watson, J.E.M., (2018). The essential role of other effective area-based conservation measures in achieving big bold conservation targets. Global Ecology and Conservation, 15, e00424. https://doi.org/10.1016/j.gecco.2018.e00424

Dunn, O.J., (1964). Multiple Comparisons Using Rank Sums. Technometrics 6, 241–252. https://doi.org/10.1080/00401706.1964.10490181

Esmail, N., McPherson, J.M., Abulu, L., Amend, T., Amit, R., Bhatia, S., Bikaba, D., Brichieri-Colombi, T.A., Brown, J., Buschman, V., Fabinyi, M., Farhadinia, M., Ghayoumi, R., Hay-Edie, T., Horigue, V., Jungblut, V., Jupiter, S., Keane, A., Macdonald, D.W., Mahajan, S.L., McVey, A., Moehrenschlager, A., Nelson, F., Noshirwani, M.M., Ntiamoa-Baidu, Y., Postigo, J.L., Rakotondrazafy, V., Rao, M., Roe, D., Huelsz, J.A.S., Stolton, S., Tawake, A. & Wintle, B. (2023). What's on the horizon for community-based conservation? Emerging threats and opportunities. Trends in Ecology & Evolution, 38, 666–680.

Faith, D.P., Reid, C. a. M., Hunter, J., (2004). Integrating Phylogenetic Diversity, Complementarity, and Endemism for Conservation Assessment. Conservation Biology, 18, 255–261. https://doi.org/10.1111/j.1523-1739.2004.00330.x

Faith, D.P., Walker, P.A., (2002). The role of trade-offs in biodiversity conservation planning: Linking local management, regional planning and global conservation efforts. Journal of Biosciences, 27, 393–407. https://doi.org/10.1007/BF02704968

Faith, D.P. (2015) Phylogenetic diversity, functional trait diversity and extinction: avoiding tipping points and worst-case losses. Philosophical Transactions of the Royal Society B: Biological Sciences, 370, 20140011.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D., (2007). The Shuttle Radar Topography Mission. Reviews of Geophysics, 45. https://doi.org/10.1029/2005RG000183

Forest, F., Grenyer, R., Rouget, M., Davies, T.J., Cowling, R.M., Faith, D.P., Balmford, A., Manning, J.C., Procheş, Ş., van der Bank, M., Reeves, G., Hedderson, T.A.J., Savolainen, V., (2007). Preserving the evolutionary potential of floras in biodiversity hotspots. Nature, 445, 757–760. https://doi.org/10.1038/nature05587

Fricke, E.C., Ordonez, A., Rogers, H.S., Svenning, J.-C., (2022). The effects of defaunation on plants' capacity to track climate change. Science, 375, 210–214. https://doi.org/10.1126/science.abk3510

Garnett, S.T., Burgess, N.D., Fa, J.E., Fernández-Llamazares, Á., Molnár, Z., Robinson, C.J., Watson, J.E.M., Zander, K.K., Austin, B., Brondizio, E.S., Collier, N.F., Duncan, T., Ellis, E., Geyle, H., Jackson, M.V., Jonas, H., Malmer, P., McGowan, B., Sivongxay, A., Leiper, I., (2018). A spatial overview of the global importance of Indigenous lands for conservation. Nature Sustainability, 1, 369–374. https://doi.org/10.1038/s41893-018-0100-6

Gerstner, B.E., Bills, P., Zarnetske, P.L., (2023). Frugivoria: A trait database for birds and mammals exhibiting frugivory across contiguous Neotropical moist forests. Global Ecology and Biogeography, 32, 1466–1484. https://doi.org/10.1111/geb.13716

GOB (2023). Zona Reservada Río Nieva - Informes y publicaciones - Servicio Nacional de Áreas Naturales Protegidas por el Estado - Plataforma del Estado Peruano. https://www.gob.pe/institucion/sernanp/informes-publicaciones/1945868-zona-reservada-rio-nieva

Gonzalez, A.M., Espejo, N., Armenteras, D., Hobson, K.A., Kardynal, K.J., Mitchell, G.W., Mahony, N., Bishop, C.A., Negret, P.J. & Wilson, S., (2023). Habitat protection and restoration: Win–win opportunities for migratory birds in the Northern Andes. Perspectives in Ecology and Conservation, 21, 33–40.

González-Maya, J.F., Martínez-Meyer, E., Medellín, R., Ceballos, G., (2017). Distribution of mammal functional diversity in the Neotropical realm: Influence of land-use and extinction risk. PLoS ONE, 12. https://doi.org/10.1371/journal.pone.0175931

Grantham, H.S., Duncan, A., Evans, T.D., Jones, K.R., Beyer, H.L., Schuster, R., Walston, J., Ray, J.C., Robinson, J.G., Callow, M., Clements, T., Costa, H.M., DeGemmis, A., Elsen, P.R., Ervin, J., Franco, P., Goldman, E., Goetz, S., Hansen, A., Hofsvang, E., Jantz, P., Jupiter, S., Kang, A., Langhammer, P., Laurance, W.F., Lieberman, S., Linkie, M., Malhi, Y., Maxwell, S., Mendez, M., Mittermeier, R., Murray, N.J., Possingham, H., Radachowsky, J., Saatchi, S., Samper, C., Silverman, J., Shapiro, A., Strassburg, B., Stevens, T., Stokes, E., Taylor, R., Tear, T., Tizard, R., Venter, O., Visconti, P., Wang, S., Watson, J.E.M., (2020). Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity. Nature Communications, 11, 5978. https://doi.org/10.1038/s41467-020-19493-3

Griffin, J.N., Leprieur, F., Silvestro, D., Lefcheck, J.S., Albouy, C., Rasher, D.B., Davis, M., Svenning, J.-C., Pimiento, C., (2020). Functionally unique, specialised, and endangered (FUSE) species: towards integrated metrics for the conservation prioritisation toolbox. https://doi.org/10.1101/2020.05.09.084871

Herrera, J.P., (2017). Prioritizing protected areas in Madagascar for lemur diversity using a multidimensional perspective. Biological Conservation, 207, 1–8. https://doi.org/10.1016/j.biocon.2016.12.028

Howe, H.F., Smallwood, J., (1982). Ecology of Seed Dispersal. Annual Review of Ecology and Systematics, 13, 201–228. https://doi.org/10.1146/annurev.es.13.110182.001221

Hurlbert, A.H., Jetz, W., (2007). Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. Proceedings of the National Academy of Sciences, 104, 13384–13389. https://doi.org/10.1073/pnas.0704469104

Isaac, N.J.B., Turvey, S.T., Collen, B., Waterman, C. & Baillie, J.E.M., (2007). Mammals on the EDGE: Conservation Priorities Based on Threat and Phylogeny. PLOS ONE, 2, e296.

ICCA (2023). *El Pueblo Shuar Arutam de la cordillera del condor*. ICCA Consortium. https://www.iccaconsortium.org/es/2017/03/31/el-pueblo-shuar-arutam-de-la-cordillera-del-condor-2/

IPCC (2022). *Climate Change 2022: Impacts, Adaptation, and Vulnerability.* Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press. Cambridge University Press, Cambridge, UK and New York, NY, USA, 3056 pp., doi:10.1017/9781009325844.

IUCN (2022). The IUCN Red List of Threatened Species. Version 2022-2. https://www.iucnredlist.org.

IUCN (2023). Spatial Data. The IUCN Red List of Threatened Species. Version 2022-2. https://www.iucnredlist.org. (Downloaded on 1.20.23)

IUCN (2016). A Global Standard for the Identification of Key Biodiversity Areas, Version 1.0 . First edition. Gland, Switzerland: IUCN.

Jenkins, D.G., Brescacin, C.R., Duxbury, C.V., Elliott, J.A., Evans, J.A., Grablow, K.R., Hillegass, M., Lyon, B.N., Metzger, G.A., Olandese, M.L., Pepe, D., Silvers, G.A., Suresch, H.N., Thompson, T.N., Trexler, C.M., Williams, G.E., Williams, N.C., Williams, S.E., (2007). Does size matter for dispersal distance? Global Ecology and Biogeography, 16, 415–425. https://doi.org/10.1111/j.1466-8238.2007.00312.x

Jetz, W., Sekercioglu, C.H., Watson, J.E.M., (2008). Ecological correlates and conservation implications of overestimating species geographic ranges. Conservation Biology,J. Soc. Conservation Biology, 22, 110–119. https://doi.org/10.1111/j.1523-1739.2007.00847.x

Jonasson, J., Harkonen, T., Sundqvist, L., Edwards, S.V., Harding, K.C., (2022). A Unifying Framework for Estimating Generation Time in Age-Structured Populations: Implications for Phylogenetics and Conservation Biology. The American Naturalist, 200, 48–62. https://doi.org/10.1086/719667

Jordano, P., García, C., Godoy, J.A., García-Castaño, J.L., (2007). Differential contribution of frugivores to complex seed dispersal patterns. Proceedings of the National Academy of Sciences, 104, 3278–3282. https://doi.org/10.1073/pnas.0606793104

Jung, M., Dahal, P.R., Butchart, S.H.M., Donald, P.F., De Lamo, X., Lesiv, M., Kapos, V., Rondinini, C., Visconti, P., (2020). A global map of terrestrial habitat types. Scientific Data, 7, 256. https://doi.org/10.1038/s41597-020-00599-8

Kennedy, C.M., Fariss, B., Oakleaf, J.R., Garnett, S.T., Fernández-Llamazares, Á., Fa, J.E., Baruch-Mordo, S. & Kiesecker, J., (2023). Indigenous Peoples' lands are threatened by industrial development; conversion risk assessment reveals need to support Indigenous stewardship. One Earth, 6, 1032–1049.

Kosman, E., Burgio, K.R., Presley, S.J., Willig, M.R., Scheiner, S.M., (2019). Conservation prioritization based on trait-based metrics illustrated with global parrot distributions. Diversity and Distributions, 25, 1156–1165. https://doi.org/10.1111/ddi.12923

Laliberté, E., Legendre, P., (2010). A distance-based framework for measuring functional diversity from multiple traits. Ecology, 91, 299–305. https://doi.org/10.1890/08-2244.1

Lamperty, T., Brosi, B.J., (2022). Loss of endangered frugivores from seed dispersal networks generates severe mutualism disruption. Proceedings of the Royal Society B: Biological Sciences, 289, 20220887. https://doi.org/10.1098/rspb.2022.0887

Leitão, R.P., Zuanon, J., Villéger, S., Williams, S.E., Baraloto, C., Fortunel, C., Mendonça, F.P., Mouillot, D., (2016). Rare species contribute disproportionately to the functional structure of species assemblages. Proceedings of the Royal Society B: Biological Sciences, 283, 20160084. https://doi.org/10.1098/rspb.2016.0084

Lim, J.Y., Svenning, J.-C., Göldel, B., Faurby, S., Kissling, W.D., (2020). Frugivore-fruit size relationships between palms and mammals reveal past and future defaunation impacts. Nature Communications, 11, 4904. https://doi.org/10.1038/s41467-020-18530-5

Lomolino, Mark.V., (2001). Elevation gradients of species-density: historical and prospective views. Global Ecology and Biogeography, 10, 3–13. https://doi.org/10.1046/j.1466-822x.2001.00229.x

Luck, G.W., Carter, A., Smallbone, L., (2013). Changes in Bird Functional Diversity across Multiple Land Uses: Interpretations of Functional Redundancy Depend on Functional Group Identity. PLOS ONE 8, e63671. https://doi.org/10.1371/journal.pone.0063671

Magneville, C., Loiseau, N., Albouy, C., Casajus, N., Claverie, T., Escalas, A., Leprieur, F., Maire, E., Mouillot, D. & Villéger, S. (2022) mFD: an R package to compute and illustrate the multiple facets of functional diversity. Ecography, 2022: e05904 .https://doi.org/10.1111/ecog.05904

Maire, E., Grenouillet, G., Brosse, S., Villéger, S., (2015). How many dimensions are needed to accurately assess functional diversity? A pragmatic approach for assessing the quality of functional spaces. Global Ecology and Biogeography, 24, 728–740. https://doi.org/10.1111/geb.12299

Martín-Regalado, C.N., Briones-Salas, M., Lavariega, M.C., Moreno, C.E., (2019). Spatial incongruence in the species richness and functional diversity of cricetid rodents. PLOS ONE 14, e0217154. https://doi.org/10.1371/journal.pone.0217154

Mazel, F., Pennell, M.W., Cadotte, M.W., Diaz, S., Dalla Riva, G.V., Grenyer, R., Leprieur, F., Mooers, A.O., Mouillot, D., Tucker, C.M. & Pearse, W.D. (2018). Prioritizing phylogenetic diversity captures functional diversity unreliably. Nature Communications, 9, 2888.

Montaño-Centellas, F.A., McCain, C., Loiselle, B.A., (2020). Using functional and phylogenetic diversity to infer avian community assembly along elevational gradients. Global Ecology and Biogeography, 29, 232–245. https://doi.org/10.1111/geb.13021

Mouchet, M.A., Villéger, S., Mason, N.W.H., Mouillot, D., (2010). Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. Functional Ecology, 24, 867–876. https://doi.org/10.1111/j.1365-2435.2010.01695.x

Mouillot, D., Graham, N.A.J., Villéger, S., Mason, N.W.H., Bellwood, D.R., (2013). A functional approach reveals community responses to disturbances. Trends in Ecology and Evolution. 28, 167–177. https://doi.org/10.1016/j.tree.2012.10.004

Newbold, T., Scharlemann, J.P.W., Butchart, S.H.M., Şekercioğlu, Ç.H., Alkemade, R., Booth, H., Purves, D.W., (2013). Ecological traits affect the response of tropical forest bird species to land-use intensity. Proceedings of the Royal Society B: Biological Sciences, 280, 20122131. https://doi.org/10.1098/rspb.2012.2131

O'Bryan, C.J., Garnett, S.T., Fa, J.E., Leiper, I., Rehbein, J.A., Fernández-Llamazares, Á., Jackson, M.V., Jonas, H.D., Brondizio, E.S., Burgess, N.D., Robinson, C.J., Zander, K.K., Molnár, Z., Venter, O., Watson, J.E.M., (2021). The importance of Indigenous Peoples' lands for the conservation of terrestrial mammals. Conservation Biology, 35, 1002–1008. https://doi.org/10.1111/cobi.13620

O'Farrill, G., Galetti, M., Campos-Arceiz, A., (2013). Frugivory and seed dispersal by tapirs: an insight on their ecological role. Integrative Zoology, 8, 4–17. https://doi.org/10.1111/j.1749-4877.2012.00316.x

Pacifici, M., Santini, L., Marco, M.D., Baisero, D., Francucci, L., Marasini, G.G., Visconti, P., Rondinini, C., (2013). Generation length for mammals. Nature Conservation, 5, 89–94. https://doi.org/10.3897/natureconservation.5.5734

Petchey, O.L., Gaston, K.J., (2006). Functional diversity: back to basics and looking forward. Ecology Letters, 9, 741–758. https://doi.org/10.1111/j.1461-0248.2006.00924.x

Pimiento, C., Leprieur, F., Silvestro, D., Lefcheck, J.S., Albouy, C., Rasher, D.B., Davis, M., Svenning, J.-C., Griffin, J.N., (2020). Functional diversity of marine megafauna in the Anthropocene. Scientific Advances, 6, eaay7650. https://doi.org/10.1126/sciadv.aay7650

Proaves (2023). *Pauxi Pauxi Proaves Reserve*. Fundación ProAves - por la conservación en el país de las aves. (2023, July 27). https://proaves.org/en/pauxi-pauxi-proaves-reserve/

Pollock, L.J., Thuiller, W., Jetz, W., (2017). Large conservation gains possible for global biodiversity facets. Nature, 546, 141–144. https://doi.org/10.1038/nature22368

Porter-Bolland, L., Ellis, E.A., Guariguata, M.R., Ruiz-Mallén, I., Negrete-Yankelevich, S. & Reyes-García, V., (2012). Community managed forests and forest protected areas: An assessment of their conservation effectiveness across the tropics. Forest Ecology and Management, 268, 6–17.

Rahbek, C., (1997). The relationship among area, elevation, and regional species richness in neotropical birds. The American Naturalist, 149, 875–902. https://doi.org/10.1086/286028

Ripple, W.J., Newsome, T.M., Wolf, C., Dirzo, R., Everatt, K.T., Galetti, M., Hayward, M.W., Kerley, G.I.H., Levi, T., Lindsey, P.A., Macdonald, D.W., Malhi, Y., Painter, L.E., Sandom, C.J., Terborgh, J., Van Valkenburgh, B., (2015). Collapse of the world's largest herbivores. Scientific Advances, 1, e1400103. https://doi.org/10.1126/sciadv.1400103

Ripple, W.J., Wolf, C., Newsome, T.M., Hoffmann, M., Wirsing, A.J., McCauley, D.J., 2017. Extinction risk is most acute for the world's largest and smallest vertebrates. Proceedings of the National Academy of Sciences, 114, 10678–10683. https://doi.org/10.1073/pnas.1702078114

Robuchon, M., Pavoine, S., Véron, S., Delli, G., Faith, D.P., Mandrici, A., Pellens, R., Dubois, G., Leroy, B., (2021). Revisiting species and areas of interest for conserving global mammalian phylogenetic diversity. Nature Communications, 12, 3694. https://doi.org/10.1038/s41467-021-23861-y

Ribeiro, M.B.N., Jerozolimski, A., Robert, P. de, Salles, N.V., Kayapó, B., Pimentel, T.P. & Magnusson, W.E. (2014) Anthropogenic Landscape in Southeastern Amazonia: Contemporary Impacts of Low-Intensity Harvesting and Dispersal of Brazil Nuts by the Kayapó Indigenous People. PLOS ONE, 9, e102187.

Sales, L., Culot, L., Pires, M.M., (2020). Climate niche mismatch and the collapse of primate seed dispersal services in the Amazon. Biological Conservation, 247, 108628. https://doi.org/10.1016/j.biocon.2020.108628

Sales, L.P., Kissling, W.D., Galetti, M., Naimi, B., Pires, M.M., (2021). Climate change reshapes the eco-evolutionary dynamics of a Neotropical seed dispersal system. Global Ecology and Biogeography, 30, 1129–1138. https://doi.org/10.1111/geb.13271

Sarayaku. (2022). *Sarayaku, El Pueblo del Medio Día*. Pueblo Originario Kichwa de Sarayaku. https://sarayaku.org/en/

Strecker, A.L., Olden, J.D., Whittier, J.B., Paukert, C.P., (2011). Defining conservation priorities for freshwater fishes according to taxonomic, functional, and phylogenetic diversity. Ecological Applications, 21, 3002–3013. https://doi.org/10.1890/11-0599.1

Velazco, S.J.E., Svenning, J.-C., Ribeiro, B.R., Laureto, L.M.O., (2021). On opportunities and threats to conserve the phylogenetic diversity of Neotropical palms. Diversity and Distributions, 27, 512–523. https://doi.org/10.1111/ddi.13215

Watson, J.E.M., Darling, E.S., Venter, O., Maron, M., Walston, J., Possingham, H.P., Dudley, N., Hockings, M., Barnes, M. & Brooks, T.M., (2016). Bolder science needed now for protected areas. Conservation Biology, 30, 243–248.

WDPA (2023). UNEP-WCMC. Protected areas map of the world, May 2023. Available at: www.protectedplanet.net

UNESCO (2023). *Sangay National Park*. UNESCO World Heritage Centre. https://whc.unesco.org/en/list/260/

Velho, N., Srinivasan, U., Singh, P. & Laurance, W.F. (2016) Large mammal use of protected and community-managed lands in a biodiversity hotspot. Animal Conservation, 19, 199–208.

Villéger, S., Mason, N.W.H., Mouillot, D., (2008). New Multidimensional Functional Diversity Indices for a Multifaceted Framework in Functional Ecology. Ecology 89, 2290–2301. https://doi.org/10.1890/07-1206.1

Wildlife Conservation Society. (2023). *WCS Peru*. WCS Peru. https://peru.wcs.org/en-us/WCS-Peru.aspx

Weisse, M.J., Naughton-Treves, L.C., (2016). Conservation Beyond Park Boundaries: The Impact of Buffer Zones on Deforestation and Mining Concessions in the Peruvian Amazon. Environmental Management, 58, 297–311. https://doi.org/10.1007/s00267-016-0709-z

Wilman, H., Belmaker, J., Simpson, J., de la Rosa, C., Rivadeneira, M.M., Jetz, W., (2014). EltonTraits 1.0: Species-level foraging attributes of the world’s birds and mammals. Ecology, 95, 2027–2027. https://doi.org/10.1890/13-1917.1

Xu, Y., Huang, J., Lu, X., Ding, Y., Zang, R., (2019). Priorities and conservation gaps across three biodiversity dimensions of rare and endangered plant species in China. Biological Conservation, 229, 30–37. https://doi.org/10.1016/j.biocon.2018.11.010

APPENDIX 3A: SPATIAL MISMATCHES AND ALIGNMENTS

Figure 3A.1: Correlation between functional diversity, quantified as functional dispersion (fdis), and SRTM elevation within each 10 km^2 grid cell across the study area.

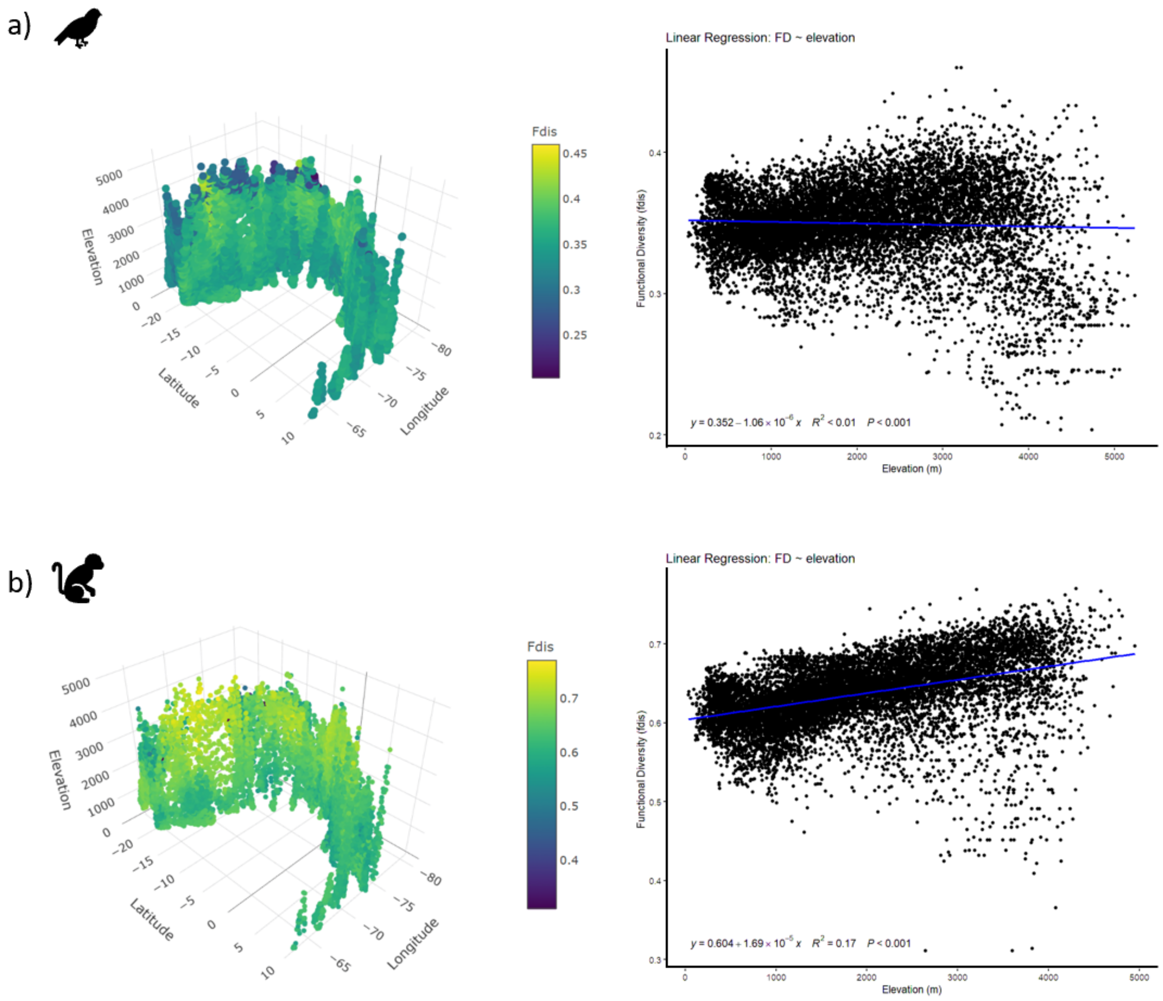


Figure 3A.2: Maps of Taxonomic Diversity (TD) and Functional Diversity (FD) of mammals across peak thresholds of diversity (10% and 30%). Panels a), b), d), and e) represent the highest 10% and 30% of diversity values for each dimension of biodiversity. Panels c) and f) show areas of overlap between both TD and FD.

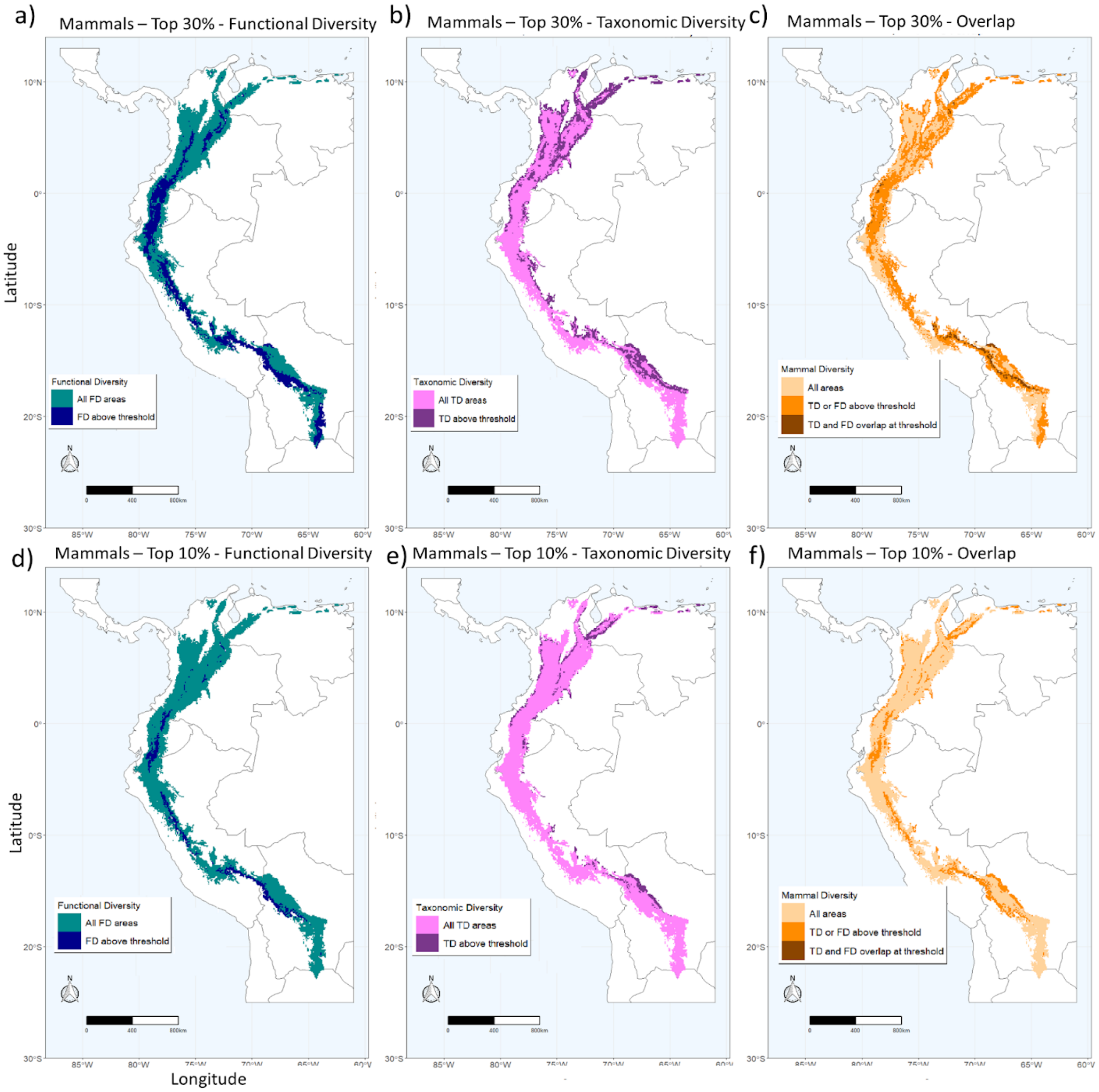
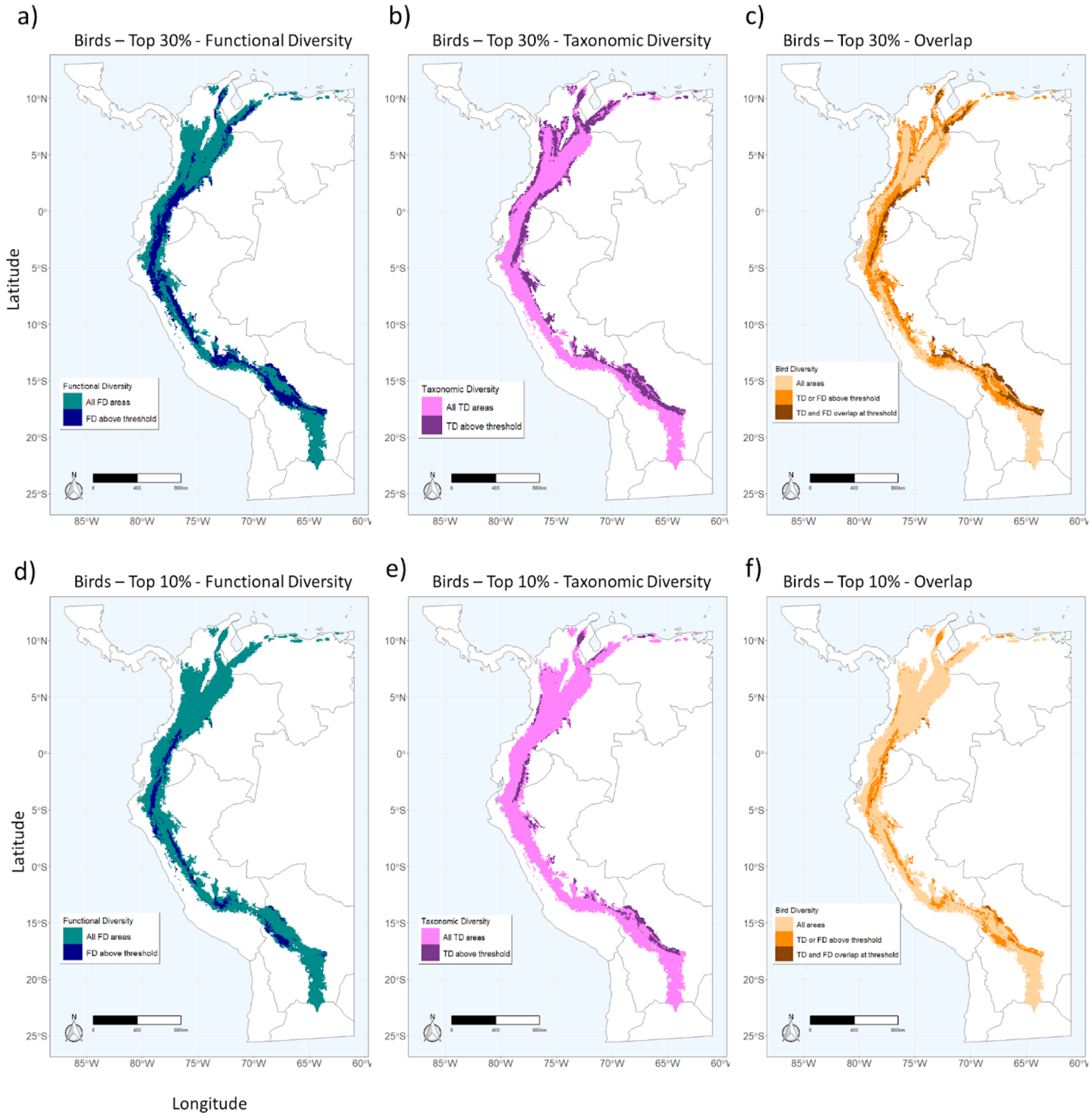


Figure 3A.3: Maps of Taxonomic Diversity (TD) and Functional Diversity (FD) of birds across peak thresholds of diversity (10% and 30%). Panels a), b), d), and e) represent the highest 10% and 30% of diversity values for each dimension of biodiversity. Panels c) and f) show areas of overlap between both TD and FD.



a) Birds – Top 30% - Functional Diversity

b) Birds – Top 30% - Taxonomic Diversity

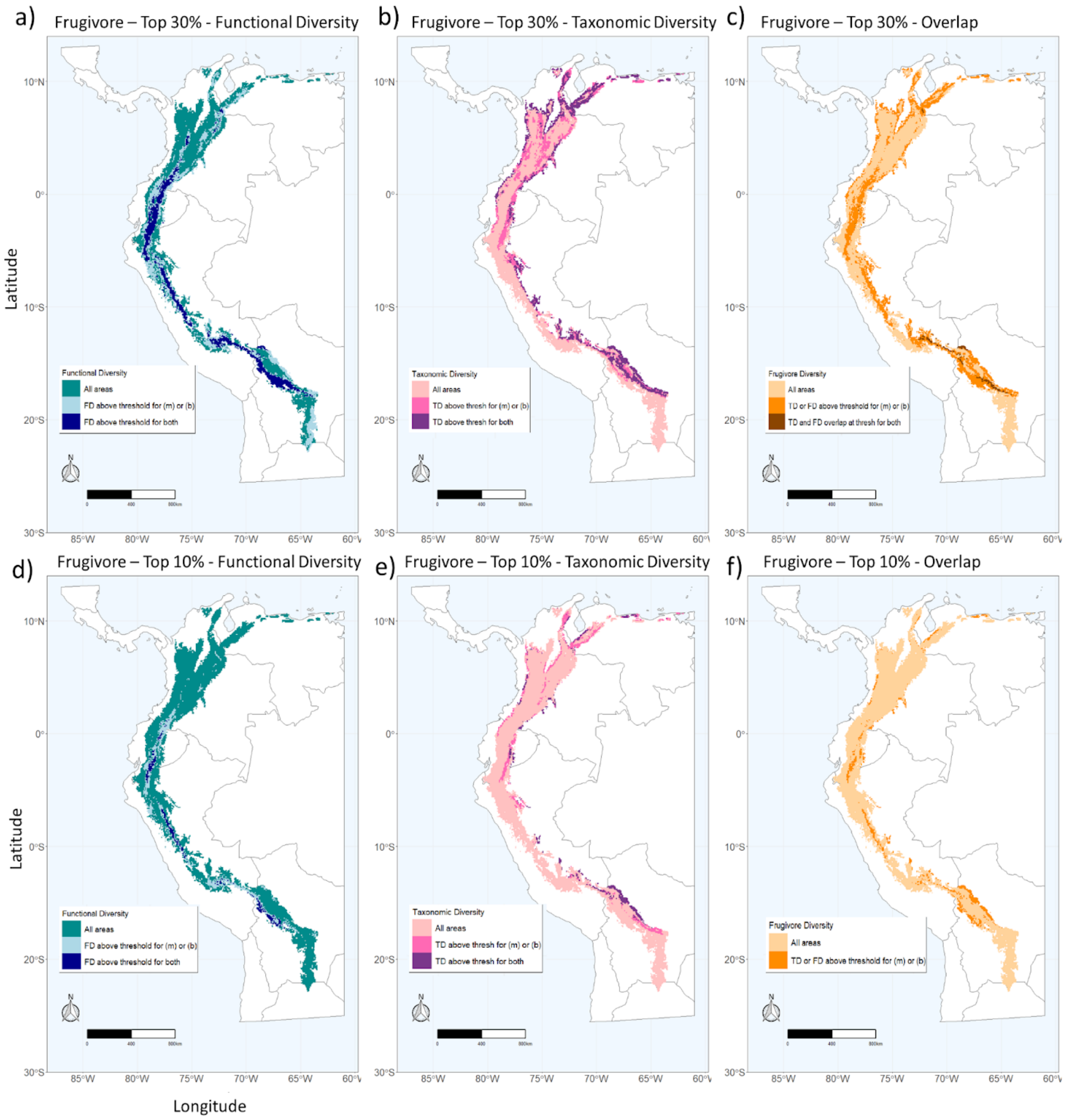
c) Birds – Top 30% - Overlap

d) Birds – Top 10% - Functional Diversity

e) Birds – Top 10% - Taxonomic Diversity

f) Birds – Top 10% - Overlap

Figure 3A.4: Maps of overlapping Taxonomic Diversity (TD) and Functional Diversity (FD) between mammals and birds (frugivores) across peak thresholds of diversity (10% and 30%). Panels a), b), d), and e) represent the highest 10% and 30% of diversity values overlapping for each taxa per dimension of biodiversity. Panel c) & f) shows overlapping areas between TD and FD across taxa.



CONCLUSIONS AND RECOMMENDATIONS

This research deepens our understanding of the geographic distributions of bird and mammal biodiversity across regions of the Neotropics to better inform conservation strategies. It encompasses the exploration of methods to enhance reliability of species distribution models for Colombian mammals and the assessment of spatial tradeoffs between regions of high taxonomic (TD) and functional diversity (FD) among frugivorous species in the Tropical Andes. To facilitate my analysis of FD, I also created an open-access functional trait dataset for frugivorous species named *Frugivoria*, which spans a substantial portion of moist forests in Central and South America. By illuminating these spatial dynamics, this research aims to enhance the field of tropical conservation by facilitating better-informed, holistic and inclusive assessments of biodiversity, thereby informing future spatial conservation priorities.

Chapter 1 collates a trait dataset with harmonized taxonomy called *Frugivoria*, containing a total of 45,216 traits and adding 23,661 new trait values for frugivorous mammals and birds in contiguous moist montane forests and adjacent moist lowland forests of Central and South America. This is the first functional trait dataset focused specifically on frugivores within these forests and expands traits once only available for mammals to birds. These tropical seed dispersers are essential for the maintenance of ecosystem functioning and forest structure. *Frugivoria* will help facilitate study of this important functional group in a time of global change. I recommend that future studies utilize this dataset to study patterns of functional diversity for frugivores and other questions central to macroecology. For example, *Frugivoria* can be used to understand community assembly, spatial patterns of biodiversity, trait distributions, and can help assess the vulnerability of species to environmental changes in moist montane and lowland regions of Central and South America. Specifically, for conservation, Frugivoria can be used to understand how species traits might relate to extinction risk. Of the species included in this dataset, 14.6% are classified as data deficient (DD) by the IUCN. Species that are DD are often not considered in biodiversity assessments (e.g., Key Biodiversity Areas (KBAs) or biodiversity hotspots), because they introduce uncertainty into estimates leading to underestimation of biodiversity for many regions. I recommend using *Frugivoria* to explore trait-based approaches for assigning IUCN threat statuses for DD frugivores, an emerging area of research and an important step towards their consideration in biodiversity assessments. Protecting species that have the important functional role of seed dispersal can help maintain ecosystem functioning and forest structure. Furthermore, *Frugivoria*'s open workflow aligns with FAIR data principles and the code to build its framework could easily be adapted to different taxa and study regions. I recommend the use of this workflow to facilitate the development of other trait datasets and contribute to ongoing efforts to increase the availability of functional trait datasets to further address the Raunkiæran shortfall, which will allow for the easier use of trait-based tools for conservation prioritization.

Chapter 2 demonstrates that the inclusion of geodiversity variables, reflecting spatial heterogeneity, into SDMs generally increased the predictive performance of the models, both statistically and often spatially, for mammals in Colombia. This finding highlights these geodiversity variables as a potential way to improve the geographic distribution estimates for species in areas with higher environmental heterogeneity, which can be used to prioritize areas for future sampling, and as a component for assessing species extinction risk and spatial conservation planning. Highest model performance typically occurred at spatial scales of 3-15 km, indicating that these spatial grains may be sufficient when incorporating geodiversity variables into SDMs for most species, however patterns were idiosyncratic for some species with highest model performance at coarser spatial grains. Due to scale-dependent and species-specific relationships with their environment, I recommend that multiple spatial grains be tested when incorporating geodiversity variables into SDMs, giving careful consideration to how a species may be interacting with its environment and the level of heterogeneity within the study area. This chapter also demonstrated that this method worked better for some biogeographic regions than others, with overfitting occurring in areas of low spatial heterogeneity as in the Amazon. I therefore recommend this approach to only be taken in the presence of moderate or high environmental heterogeneity to fully take advantage of information gained from these geodiversity variables. To ensure comparability of models within and across species, I intentionally refrained from implementing species-specific tuning of model parameters. However, it is important to note that within species distribution modeling (SDM) practices, species-specific tuning is considered a best practice for attaining models with an optimal level of complexity tailored to each species. Therefore, when applying this approach for estimating species distributions as part of a formal conservation assessment, I recommend applying species-specific tuning methods to choose appropriate model parameters which balances both complexity and model fit. Moreover, it should be acknowledged that this research focused on a subset of mammal species specific to Colombia because of the high heterogeneity of the region and the availability of expert vetted occurrences and range maps through BioModelos. However, expert validated occurrences and range maps are not always available for certain regions and taxa, making model validation difficult in many cases. Therefore, I want to stress the importance of initiatives like BioModelos that aim to increase the availability of these data for both the academic and conservation community. This will make similar studies more tenable and allow for the inclusion of additional taxa in regions characterized by high environmental heterogeneity, which may show different species-environment relationships with geodiversity. Improved understanding of this relationship and incorporating geodiversity predictors in SDMs, can lead to more robust and reliable predictions. More robust predictions are essential given the reliance on SDMs for conservation such as targeting species prioritization, guiding future sampling efforts, and informing biodiversity assessments.

Chapter 3 conducts a comparative analysis of the spatial distributions of TD and FD among frugivores and further evaluates their levels of protection and understand trade-offs when emphasizing specific dimensions of diversity. With the *Frugivoria* dataset from Chapter 1, the analysis in Chapter 3 reveals strong spatial mismatches between taxonomic and functional diversity in bird and mammal frugivores, with limited overlap between dimensions and taxa. Given that this mismatch is a common pattern seen in other studies also addressing spatial mismatches between facets of diversity, this seems to be the rule more than the exception. The existence of these mismatches necessitates the consideration of multiple dimensions of diversity for spatial conservation planning. Further, this chapter revealed that high biodiversity areas, both taxonomically and functionally, were only about 37% protected when considered independently. Protection plummeted to 0% for mammals when focusing on areas where diversity dimensions overlapped. Based on these findings, I recommend special attention to be given to mammals when protecting high levels of multidimensional diversity. However, due to generally low overlap among diversity dimensions within each taxon for the highest level of diversity, considering each facet separately when prioritizing conservation areas for maximum diversity might be prudent. This chapter also demonstrated that high numbers of FUSE mammals, levels of bird FD, bird TD, and mammal TD were elevated in some strict protected areas when compared with less strict areas. This highlights the importance of certain strict protective measures in the maintenance of biodiversity, however there was no statistical difference between protective measures for the FD of mammals as well as for FUSE birds. This finding further bolsters the earlier recommendation, advocating for focused conservation efforts for mammals, particularly to enhance the protection of heightened FD within the stricter categories of PAs, and stresses that attention also be given to FUSE birds. Future conservation prioritization for these groups as well as high biodiversity areas in general should assess complementarity for both established reserves and areas earmarked for potential expansion or the introduction of new protective measures. This is especially important given that current IUCN guidelines for establishing Key Biodiversity Areas (KBAs) do not explicitly consider FD and instead focuses on TD. Emphasizing the evaluation of FD is particularly crucial, as prioritizing this metric, specifically for frugivores, holds the promise of sustaining ecosystem functionality and overall health in the face of a dynamically changing climate. Further, we found that PAs with community-based management had higher levels of FD and other PAs within the same PA category. This finding stresses the importance of including areas with local governance into biodiversity targets, as diversity in these areas may specifically benefit from traditional land-use practices, knowledge systems, and a deep connection to their landscapes.

This work encountered persistent challenges, predominantly stemming from the COVID-19 pandemic and use of big data techniques. The progress of my dissertation was significantly hampered by the COVID-19 pandemic. Working in isolation made it challenging to collaborate effectively and discuss aspects of my research like R code with colleagues. Additionally, undergraduate students working on the *Frugivoria* project had limited availability to contribute to the project due to the disruptions caused by the pandemic. These difficulties were compounded by the resource-intensive nature of the work, particularly the need for high-performance computing (HPC) for data storage and processing in all three chapters of my dissertation and the need to collaborate with a data scientist to corroborate workflows. In both Chapters 1 and 3, the BirdLife International spatial dataset, a geodatabase of bird range maps, proved indispensable. However, given its substantial size, exceeding 3 gigabytes, this dataset posed difficulties in terms of storage, manipulation, and interaction. Furthermore, Chapters 1 and 2 were characterized by a substantial demand for data-intensive operations. Chapter 2 necessitated the generation of well over a thousand models, alongside their corresponding outputs, during the process of streamlining the automated workflow. Chapter 3 posed its own challenges, involving the processing of hundreds of range maps to generate AOH maps. Notably, for this same chapter, the computation of FD for a community defined by 10 $\mathrm{km}^{=}$ grid cells across the entire Tropical Andes required an extensive run time of approximately 48 hours, even when executed on MSU's HPC. Given these challenges and to promote transparency and reproducibility for all chapters, I shared the computational workflows and analyses on GitHub, ensuring that fellow researchers and conservation practitioners could replicate our methods and results. However, it is imperative to recognize that for studies wishing to compute similar metrics at finer scales or with an increased number of PCoA axes, the required time investment and demand for resources would increase exponentially. The emergence of cloud computing has garnered attention within the field of ecology, offering a potential solution to this issue. However, impediments such as cost, and the knowledge barriers associated with adopting cloud computing technologies can hinder their practical application for conservation practitioners in the field. To mitigate these barriers, the establishment and fostering of existing partnerships between academic institutions and conservation organizations on the ground could play a pivotal role. Such collaborations could work to close the divide between advanced research methodologies and applied conservation, facilitating the efficient and effective application of research outcomes to on the ground management, thereby reducing this research implementation gap. These partnerships, coupled with open-access tools like R and GitHub, contribute to the accessibility and reproducibility of ecological research, ultimately benefiting both the scientific community and conservation practitioners.

My intent for this dissertation was to enhance the understanding about the distribution of bird and mammal biodiversity within the Neotropics. I pursued these goals through comprehensive investigations spanning species distribution models, functional and taxonomic diversity assessments, and the creation of the *Frugivoria* functional diversity dataset. By shedding light on critical aspects of the distribution of biodiversity and offering a newly compiled dataset that not only enables fellow researchers to perform similar analyses, but also encourages the extension of these analyses, this study improves our capacity to make well-informed decisions pertaining to the prioritization of species and the identification of high-value conservation areas. As we navigate a rapidly changing world, it is my hope that the insights garnered from this work will contribute to more effective and holistic approaches to safeguarding biodiversity.