Imperial College London

Responses of Ecological Communities to Deforestation: An Empirical and Theoretical Perspective

Benjamin Howes

Submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy in Life Sciences, Imperial College London, January 2024

Funding

Funded by the Natural Environment Research Council (NERC) as part of the Quantitative Methods in Ecology and Evolution (QMEE) doctoral training programme.

Supervision

Dr Cristina Banks-Leite, Imperial College London

Dr Manuela González-Suárez, University of Reading

Professor Henrik Jeldtoft Jensen, Imperial College London

Declaration of Originality

I declare that the work in this thesis was performed and prepared by myself under the supervision of Dr Cristina Banks-leite, Dr Manuela González-Suárez and Professor Henrik Jeldtoft Jensen. All sources of information are appropriately cited.

Whilst this work is my own, it has been a collaborative effort that could not have been done without the contributions of collaborators and colleagues. I am to publish all chapters in peer-reviewed journals, and therefore all chapters are written in the third person plural. Specific contributions by colleagues to each chapter are highlighted below.

Chapter 2

This chapter is under the second round of review in the journal Global Ecology and Biogeography, and the text shown here is a reproduction of this. This research was largely carried out under the supervision of Professor Matthew Betts (Oregon State University), who aided with the conception of the idea, and Dr Cristina Banks-Leite, Dr Manuela González-Suárez, and Dr Flavia Belloto-Trigo all contributed to text and structural edits, as well as giving general advice on analysis and methods. I carried out all data extraction and analysis, as well as writing the chapter, and the data used from IUCN and BirdLife were freely available to researchers.

Chapter 3

All data used in this chapter were provided by the authors of the BIOFRAG dataset, which is an open-access dataset for use by researchers. My supervisors Dr Cristina Banks-Leite and Dr Manuela González-Suárez supported text and structural edits, as well as advising on methods and analysis. I devised the idea, carried out all analyses, and wrote the chapter. All authors of BIOFRAG who provided data will be cited when the chapter is published.

Chapter 4

This chapter is published as an open-access research paper in the journal Natural Sciences (DOI: 10.1002/ntls.20220027) and is reproduced here under a CC BY 4.0 DEED Creative Commons licence. Data used in this chapter were provided by Dr Alexandre Uezu, Dr Jack H. Hatfield, Dr José C. Morante-Filho, Dr Luis dos Anjos and Dr Pedro F. Develey. The idea, analysis, and writing for the chapter were carried out entirely by myself, though feedback and edits were provided by my supervisors Dr Cristina Banks-Leite, Dr Manuela González-Suárez and Professor Henrik Jeldoft Jensen.

Chapter 5

The novel code used in this chapter was entirely written by myself, and I devised the idea, analysis, and wrote the paper. Edits and feedback were provided by my supervisors Dr Cristina Banks-Leite, Dr Manuela González-Suárez and Professor Henrik Jeldoft Jensen.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

Firstly, I would like to thank my supervisors Dr Cristina Banks-Leite, Dr Manuela González-Suárez and Professor Henrik Jeldoft Jensen who have provided constant support throughout my PhD, both academically and emotionally. It would also be remiss of me not to thank the countless other collaborators who helped me through this academic journey: Professor Matthew Betts for letting me work at Oregon State University with him without really having any plan of what I was going to do; Professor Robert Ewers for inviting me into his lab in Cristina's absence; and Professor Samraat Pawar for teaching me metabolic theory (and the high fives).

I would also like to thank all of the past and present CALE lab members for making my time spent in the office fun, though maybe not very productive: Jenna, Flavia, Elise, Rhys, Natasha, and Hans. I would also like to thank everyone who made my time at Silwood Park (and Slodge) so enjoyable, there are really too many people to list them all, but special mention goes to Joe, Deraj, Chris, Andres, and Josh.

Main Abstract

Habitat loss, specifically deforestation, is the biggest threat to global biodiversity. Despite this, we still have relatively little understanding of how ecological communities vary in their response to deforestation. where the most sensitive communities are, and what drives this sensitivity. In this thesis, I answer these questions, amongst others, at the global and local scale. In **Chapter 2** and **Chapter 3** I investigate the causes of global variation in community sensitivity, using range maps of terrestrial vertebrates to identify latitudinal patterns in the distribution of forest species, and community datasets to find thresholds in responses of communities to deforestation. In **Chapter 4** and **Chapter 5** I explore variation in community sensitivity to forest loss at the local scale, focusing on changes in species interactions caused by forest loss, and responses of communities to fragmentation. In **Chapter 2** I found that proportions of forest species decreased from the equator to the poles, even once variables such as forest cover were accounted for. This finding was corroborated in **Chapter 3** where I identified that communities at the equator were more sensitive to deforestation than those at the poles. I additionally found that communities in highly disturbed regions were less sensitive than those in undisturbed regions. At the local level I found that species interactions altered as deforestation occurred, and that they were most negative in intermediately forested areas. I further found that the spatial configuration of deforestation can impact community responses. A few large patches maintain community composition, while many small patches increase species richness. Overall the results from my thesis suggest that communities do vary in their response to deforestation and that communities in undisturbed areas close to the equator are most sensitive.

Table of Contents

Funding **2**
Supervision **2**
Declaration of Originality **3**
Chapter 2 3
Chapter 3 3
Chapter 4 4
Chapter 5 4
Copyright Declaration **5**
Acknowledgements **6**
Main Abstract **7**
Content **8**
Figures **10**
Tables **11**
Chapter 1: Main Introduction **13**
Forests and Deforestation 13
Species Responses to Deforestation 14
Ecological Communities 15
Community Responses to Deforestation 15
Species Interactions and Deforestation 17
Forest Loss versus Fragmentation 18
Complex System Modelling 19
Summary 21
Chapter 2: A Global Latitudinal Gradient in the Proportion of Terrestrial Vertebrate Forest Species **22**
Abstract 22
Introduction 23
Methods 28
Results 34
Discussion 39
Chapter 3: Global Variation in Forest Thresholds for Community Persistence **44**
Abstract 44
Main Text 45
Methods 55
Chapter 4: Deforestation Alters Species Interactions **64**
Abstract 64
Introduction 65
Methods 67
Results 78
Discussion 80
Chapter 5: Fragmentation increases species richness but disturbs community composition **87**

Abstract..... 87
Introduction..... 88
Methods..... 91
Results..... 101
Discussion..... 104
Chapter 6: Main Discussion..... 107
Overview..... 107
Community Responses to Deforestation..... 107
Implications for Landscape-Level Conservation and Restoration..... 109
Complex Systems for Ecological Research..... 111
Future Work..... 113
Summary..... 115
References..... 116
Supplementary Material..... 130
Chapter 2: A global latitudinal gradient in the proportion of terrestrial vertebrate forest species..... 130
Chapter 3: Global Variation in Forest Thresholds for Community Persistence..... 133
Chapter 4: Deforestation Alters Species Interactions..... 136
Chapter 5: Fragmentation increases species richness but disturbs community composition..... 149

Figures

Chapter 2: A Global Latitudinal Gradient in the Proportion of Terrestrial Vertebrate Forest Species

Figure 2.1. Global Distribution of the Proportion of Forest Species…… 35
Figure 2.2. Latitudinal Gradient in the Proportion of Forest Species and Richness...36
Figure 2.3. Effect of Latitude on the Proportion of Forest Species……37
Figure 2.4. Effect of All Variables on the Proportion of Forest Species……38

Chapter 3: Global Variation in Forest Thresholds for Community Persistence

Figure 3.1. Location of Studies……47
Figure 3.2. Position of Thresholds……..48
Figure 3.3. Theoretical Framework of Community Intergrity Change……...50
Figure 3.4. Position of Threshold Types……52

Chapter 4: Deforestation Alters Species Interactions

Figure 4.1. Methodological Framework……..77
Figure 4.2. Interspecific Interaction Changes……79
Figure 4.3. Pairwise Interaction Changes……..80
Figure 4.4. Theoretical Cup and Ball Diagram……..84

Chapter 5: Fragmentation increases species richness but disturbs community composition

Figure 5.1. Fragmented Landscape Configurations……98
Figure 5.2. Effect of Fragmentation on Species Richness and Stability……102
Figure 5.3. Effect of Fragmentation on Community Composition……103

Supplementary Material

Chapter 2: A Global Latitudinal Gradient in the Proportion of Terrestrial Vertebrate Forest Species

Figure S2.1. Hand Wing Index of Forest and Non-Forest Species……131
Figure S2.2. Model Diagnostic Checks……132

Chapter 3: Global Variation in Forest Thresholds for Community Persistence

Figure S3.1. Erroneous Piecewise Regression……..135
Figure S3.2. Visualisation of Piecewise Regressions……135

Chapter 4: Deforestation Alters Species Interactions

Figure S4.1. Co-Occurrence Method Comparison……138

Figure S4.2. Example Centerting Interactions and Forest Cover……141

Chapter 5: Fragmentation increases species richness but disturbs community composition

Figure S5.1. Effect of Fragmentation on Abundance Community Composition……150

Tables

Chapter 2: A Global Latitudinal Gradient in the Proportion of Terrestrial Vertebrate Forest Species

Table 2.1. Hypotheses and Rationale of Factors Creating a Latitudinal Gradient in the Proportion of Forest Species……...25

Table 2.2. Comparison of R^2 between Latitude-Only and Full Models…….39

Chapter 3: Global Variation in Forest Thresholds for Community Persistence

Chapter 4: Deforestation Alters Species Interactions

Table 4.1. Coefficient Estimates……..78

Chapter 5: Fragmentation increases species richness but disturbs community composition

Table 5.1. Model parameters……...96

Table 5.2. Parameter Values used in Simulations……97

Supplementary Material

Chapter 2: A Global Latitudinal Gradient in the Proportion of Terrestrial Vertebrate Forest Species

Table S2.1. Model Coefficient Tables……...132

Chapter 3: Global Variation in Forest Thresholds for Community Persistence

Table S3.1. Sensitivity Analysis of Threshold Model…….133

Table S3.2. Sensitivity Analysis of Threshold Type Model……...134

Table S3.3. Overview of Studies……...134

Table S3.4. Number of Control Sites Per Study……134

Table S3.5. Scale of Effect for each Study……..134

Chapter 4: Deforestation Alters Species Interactions

Table S4.1. Study Information……...136

Table S4.2. Occurrence Fit……137

Table S4.3. Abundance Fit…….138

Table S4.4. Co-Occurrence Method Comparison……..139
Table S4.5. Model Sensitivity Analysis……140
Table S4.6. Average Species Interaction……141
Chapter 5: Fragmentation increases species richness but disturbs community composition

Chapter 1: Main Introduction

Forests and Deforestation

Forests cover approximately thirty percent of the world's land area (FAO 2020), yet they provide habitat for approximately 75% of terrestrial vertebrates (FAO 2020). The tropical forests of the equator are particularly biodiverse, harbouring 62% of the world's terrestrial vertebrate species, 29% of which are endemic to such areas (Pillay *et al.* 2022). As well as providing habitat for highly biodiverse ecological communities, forests also provide key ecosystem services such as carbon sequestration, wood production, and water purification (Cardinale *et al.* 2012; Oliver *et al.* 2015), which are maintained by the species that inhabit them, by for example dispersing seeds and limiting herbivory (Bunker *et al.* 2005; Cardinale *et al.* 2012; Paquette & Messier 2011; Şekercioğlu *et al.* 2004).

Despite their global importance for biodiversity and humanity, contemporary global forest cover has continually declined, with half of all forest loss in the last century (Ritchie & Roser 2023b). In fact, much of the temperate forests of Europe, North America, and Asia have been almost completely eradicated (Ritchie & Roser 2023b). Deforestation is predominantly caused by anthropogenic activities, whereby forests are removed to be replaced by new land-use types, typically agriculture, pastures or urban infrastructure (Curtis *et al.* 2018), though natural processes such as fires and storms can similarly deforest areas (Betts *et al.* 2019; Curtis *et al.* 2018). Global deforestation continues to this day, and approximately 10 million hectares of forests are removed each year (Ritchie & Roser 2023a), most notably from South America and Southeast Asia (Ritchie & Roser 2023a).

Species Responses to Deforestation

Species which inhabit forests are understandably expected to be affected by deforestation, and many studies have looked to identify how the occurrence and abundance of species vary with deforestation (Betts *et al.* 2010, 2014; Furlani *et al.* 2009; Giam 2017), and which species are most vulnerable (Hatfield *et al.* 2018b; Henle *et al.* 2004; Keinath *et al.* 2017; Quesnelle *et al.* 2014; Weeks *et al.* 2023). Generally, studies have found high interspecific variation in responses to deforestation, and species at high trophic levels with small population sizes; large body masses; poor dispersal ability; and high habitat specialisation have been identified as the most sensitive to deforestation (Henle *et al.* 2004; Keinath *et al.* 2017), whilst well dispersing generalist species are least affected. There is also growing evidence of significant intraspecific variation in responses to deforestation. Studies have suggested that populations of species far from their range edge (Banks-Leite *et al.* 2022; Bellotto-Trigo *et al.* 2023; Orme *et al.* 2019) and close to the centre of their environmental niche (Williams & Newbold 2021) are less affected by forest loss than those at the edge of their geographic or environmental space.

In addition to exploring which species are most sensitive to deforestation, several studies have focused on the shape of the relationship between forest cover and incidence or abundance (Ávila-Gómez *et al.* 2015; Betts *et al.* 2007, 2010). These studies have often identified a nonlinearity to this relationship, where the negative relationship between forest cover and incidence significantly increases after a certain level of deforestation has occurred, this is often called a threshold in forest cover (Ávila-Gómez *et al.* 2015; Betts *et al.* 2007, 2010). The presence of such a threshold is of particular interest to conservationists and policymakers, as it could potentially be used as a trade-off between species loss and economic gain, where a certain level of forest cover could be maintained to conserve certain target species.

Ecological Communities

Ecological communities are generally defined as populations of different species that inhabit the same geographic space and are connected directly and indirectly through interactions (Looijen & van Andel 1999). Interactions define the composition of ecological communities, determine the richness of the community and, when in a stable state, ensure their stability (Case & Taper 2000; Lurgi *et al.* 2016; Ratzke *et al.* 2019). Interactions can take a multitude of forms; predation, competition and mutualism to name a few, all of which directly or indirectly impact the fitness and therefore the continued survival of other species in the community. This interconnectedness between species makes studying ecological communities challenging, as changes in the occurrence or abundance of one species can affect all other species in the community (Berlow 1999; Hagen *et al.* 2012; Valiente-Banuet *et al.* 2015).

Community Responses to Deforestation

Similarly to species, it is expected that ecological communities will vary in their response to deforestation (Carrara *et al.* 2015; Halstead *et al.* 2019; Melo *et al.* 2018; Prist *et al.* 2012; Rutt *et al.* 2019; Watson *et al.* 2005). Contrary to species level, identification of which communities are most sensitive to deforestation, and the mechanisms that drive this are understudied and largely unknown. Community level analysis is however arguably more important and relevant to understanding how biodiversity changes with deforestation since species do not exist in a vacuum, and are rather impacted directly and indirectly via interactions with other species which inhabit the same space as them (Berlow 1999; Case *et al.* 2005; Case & Taper 2000; Hagen *et al.* 2012). For this reason, predictions of community level change based on summations of independent species responses to environmental perturbations are not appropriate (Ewers *et al.* 2009; Ferrier & Guisan 2006), and rather analyses that focus on the entire community's response are required.

Though large-scale studies on community responses to deforestation are scarce, theoretical expectations have been made based on the extrapolation of species-level responses. Communities which are composed of species which are sensitive to deforestation are thought to be more vulnerable to deforestation than communities composed of disturbance-adapted species (Betts *et al.* 2017; Boesing *et al.* 2018; Swift & Hannon 2010). For instance, communities with high proportions of forest specialist species and poor dispersers will be more sensitive to deforestation than communities with well-dispersing generalist species (Hatfield *et al.* 2018a; Thomas 2000). Whilst this is likely to be the case, identifying which communities are composed of species that are either sensitive or tolerant to deforestation, and thus most need conservation efforts, is a more difficult task, yet to be solved.

Large-scale community level analyses have historically been difficult to undertake, most notably due to data limitations. Such analysis requires empirical data on community composition from across the globe as well as accompanying land use data to understand how communities respond to deforestation. Over the last decade, however, these aforementioned obstacles have largely been overcome. Remote sensing using satellite imagery has allowed for the classification of land use types and estimation of key landscape variables such as forest cover at high resolutions (Friedl, Mark & Sulla-Menashe, Damien 2022; Gorelick *et al.* 2017; Hansen *et al.* 2013). Additionally, large-scale empirical datasets of species occurrences have become commonplace via collaborative efforts such as the PREDICTS (Hudson *et al.* 2017), BIOFRAG (Pfeifer *et al.* 2014), and BIOTIME (Dornelas *et al.* 2018) databases which collate many local studies to create a collection of high-quality datasets over large spatial and temporal extents. Lastly, databases of species ranges and traits such as those from IUCN (“The IUCN Red List of Threatened Species” 2023b), BirdLife (“BirdLife Data Zone” 2021), and Avonet (Tobias *et al.* 2022) have been created and expanded, opening up new avenues of analysis.

Understanding which communities are most sensitive to deforestation, and where these occur, is paramount to mitigating the negative effects of forest loss. I utilise these aforementioned advances in community and land use data to answer this question in my first two chapters, **Chapter 2** and **Chapter 3**. In **Chapter 2** I explore the global distribution of forest specialist species using range maps from IUCN ("The IUCN Red List of Threatened Species" 2023b) and BirdLife ("BirdLife Data Zone" 2021), combined with many remotely sensed land use data. I identify the evolutionary and ecological variables which drive the distribution and highlight areas which are likely to contain communities which are particularly sensitive to deforestation. In **Chapter 3** I use the BIOFRAG dataset (Pfeifer *et al.* 2014) to assess the relationship between forest cover and community composition, specifically exploring thresholds in forest cover after which community composition is strongly impacted. To do this I combine empirical data of the species present in approximately 100 global studies with a global forest cover dataset (Hansen *et al.* 2013), and then assess how the position of thresholds varies with regional factors; regional forest cover and latitude. Ultimately, identifying where communities are most sensitive to forest loss.

Species Interactions and Deforestation

Whilst deforestation is known to alter the identity of species within communities, it could also impact other aspects of community structure and function, such as interactions between species (Valiente-Banuet *et al.* 2015). Theoretical work has found that changes in species interactions can destabilise communities (Ratzke *et al.* 2019; Thebault & Fontaine 2010), causing changes to abundances and potential species extinctions. One way in which interactions can most commonly be impacted is via environmental perturbations (Callaway *et al.* 2002; Dangles *et al.* 2013). Changes in environmental conditions have been found to both add or remove interactions, via the addition or removal of species themselves (Kofron & Chapman 1995) or via changes in the effect size or direction (positive or negative) of existing interactions (Chamberlain *et al.* 2014). Whilst changes to species interactions under some perturbations like changes in temperature and salinity are relatively well understood

(Callaway *et al.* 2002; Dell *et al.* 2014; Grigaltchik *et al.* 2012), no study has looked at how the most common form of environmental change, deforestation, could alter interspecific interactions.

Empirically estimating interactions is labour intensive and time consuming, and therefore most estimations of interactions via observational studies are done on a small number of species under lab conditions (García *et al.* 2023; Lechón-Alonso *et al.* 2021). Ecological communities are however usually composed of many hundreds of species, making estimations of interactions for entire communities via observational techniques near impossible; and therefore assessing how they change with environmental perturbations even more improbable. As such modelling approaches which simulate ecological communities and interactions are required to investigate the impacts of deforestation on interactions within ecological communities.

Forest Loss versus Fragmentation

Thus far I have focused on deforestation as a single process; loss of forested area. The loss of forested area is however usually accompanied by two other processes: fragmentation and degradation (Banks-Leite *et al.* 2022). Fragmentation refers to the breaking apart of forested areas into disparate poorly connected patches (Fahrig 2017; Fahrig *et al.* 2019), whilst degradation is the reduction in forest quality via processes such as pollution and selective logging (Vásquez-Grandón *et al.* 2018). While the reduction in total forested area is almost unanimously accepted as being bad for biodiversity leading to the loss of rare species, reduction in species richness, and alterations in community composition (Arroyo-Rodríguez *et al.* 2020; Buchmann *et al.* 2013; Cushman 2006), the impact of fragmentation in isolation (fragmentation *per se*) is debated (Bellotto-Trigo *et al.* 2023; Fahrig 2017, 2021; Fahrig *et al.* 2019; Fletcher *et al.* 2018).

Whilst many studies have attempted to find a general response of ecological communities to forest fragmentation using empirical data, they are usually hampered by limitations. The

most notable among these are the varying properties of species within different communities, difficulty in separating the effects of forest loss and fragmentation, and the biodiversity metric used to assess the effect of fragmentation (Cushman *et al.* 2012; Ewers & Didham 2006; Fleishman *et al.* 2006; Hillebrand 2004). As such, it is unlikely that a conclusive result will be found on the effect of fragmentation alone on communities when using empirical data. Rather I propose that a modelling approach would be more suitable, where the properties of every species in the community are known, the effects of forest loss and fragmentation can be completely separated using synthetic landscapes, and a variety of biodiversity metrics can be used since data collection is not a limitation.

Complex System Modelling

Complex systems are commonly defined as systems composed of many linked individual components whose total state or behaviour is almost impossible to predict from the individual components alone, but together form some level of organisation (Anand et al. 2010; Riva et al. 2023). Complex systems modelling aims to simulate self-organisation via links between individual components by creating relatively simplified rules that govern processes at the scale of the individual component, which in turn create emergent properties of the system at the larger scale (Riva *et al.* 2023). Complex systems modelling has been used to study the responses of systems to external perturbations and to identify processes which maintain and organise the system itself (Bentham 2011; Cairoli *et al.* 2014; Christensen *et al.* 2002).

Ecological communities are quintessential complex systems (Riva *et al.* 2023), in that they are formed of individual components (individuals) which are linked together (via species interactions) and self-organise to stable states so that their total system behaviour cannot be predicted from the individual components (Pimm 1984). Complex systems modelling is a growing area of ecology (Riva *et al.* 2023) though its uses have yet to be fully embraced by the research community (Anand *et al.* 2010).

In this thesis I focus on the use of a complex systems modelling approach called the Tangled Nature Model (hereafter TaNa). The TaNa was originally created to model the coexistence of species in ecological communities over evolutionary time frames (Christensen *et al.* 2002), but has since been developed for use over shorter ecologically relevant time frames (Brinck 2018; Brinck & Jensen 2017), where it has been shown to recreate realistic species abundance distributions, species-area relationships, and trophic networks (Christensen *et al.* 2002; Laird *et al.* 2008; Laird & Jensen 2006). The unique aspect of the TaNa for ecological modelling is its emphasis on species interactions as the predominant driver of community dynamics (Anderson & Jensen 2005; Christensen *et al.* 2002). I utilise this novel aspect of the TaNa in **Chapter 4** where I investigate how deforestation impacts species interactions by combining simulated data from the TaNa with empirical data from five studies in the Atlantic Forest of Brazil. To do this I simulate comparable communities in the TaNa to the empirical data and expose them to deforestation, extracting and assessing how interactions between species vary before and after deforestation. I then compare these to interspecific interactions from the empirical dataset which I estimate using co-occurrence methods.

Whilst the TaNa focuses on interspecific interactions, species in the model also have key traits known to affect real-world organisms' responses to deforestation, such as dispersal (Ewers & Didham 2006; Niebuhr *et al.* 2015), and simulates key ecological processes which influence community stability and persistence, like immigration (Pardini *et al.* 2010; Püttker *et al.* 2011). As a model with a simple set of rules, the TaNa is computationally efficient and highly tractable; it is easy to identify how changes in parameters (such as species traits) influence the outcome of the model. I utilise these properties of the TaNa in **Chapter 5** to examine the use of complex systems modelling approaches to explore the impact of fragmentation *per se* on ecological communities. To do this I generate ecological communities with varying dispersal abilities and immigration rates and assess how they are impacted by the effects of fragmentation in isolation. I additionally explore how the

interpretation of whether fragmentation has overall positive or negative effects on biodiversity can vary based on the biodiversity metric used.

Summary

Overall this thesis contributes to our understanding of which communities are most affected by deforestation, both in terms of location, composition, and properties. It also highlights the use of complex systems to explore the causes and impacts of deforestation, including uncovering underlying impacts of deforestation which may not be immediately apparent, and disentangling factors to identify mechanistic causes.

Chapter 2: A Global Latitudinal Gradient in the Proportion of Terrestrial Vertebrate Forest Species

This chapter has been submitted and undergone a first round of peer-review in the journal Global Ecology and Biogeography. It is therefore presented here with the appropriate format and styling for this journal.

Abstract

Aim

Global patterns in species distributions such as the latitudinal biodiversity gradient are of great interest to ecologists and have been thoroughly studied. Whether such a gradient holds true for the proportion of species associated with key ecotypes such as forests is however unknown. Identifying a gradient and ascertaining the factors causing it could further our understanding of community sensitivity to deforestation and uncover drivers of habitat specialisation. The null hypothesis is that proportions of forest species remain globally consistent, though we hypothesise that proportions will change with differences in ecotype amount, spatial structure, and environmental stability. Here we study whether the proportion of forest species follows a latitudinal gradient, and test hypotheses for why this may occur.

Location

Worldwide

Time Period

Present

Major Taxa Studied

Terrestrial vertebrates

Methods

We combined range maps and habitat use data for all terrestrial vertebrates to calculate the proportion of forest species in an area. We then used data on the global distribution of current, recent historical, and long-term historical forest cover, as well as maps of global disturbances and plant diversity to test our hypotheses using generalised linear models.

Results

We identified a latitudinal gradient in the proportion of forest species whereby the highest proportions occurred at the equator and decreased polewards. We additionally found that the proportion of forest species increased with current forest cover, historical deforestation, plant structural complexity, and habitat stability. Despite the inclusion of these variables, the strong latitudinal gradient remained, suggesting additional causes of the gradient.

Main Conclusions

Our findings suggest that the global distribution of the proportion of forest species are a result of recent ecological, as well as long-term evolutionary factors. Interestingly, high proportions of forest species were found in areas that experienced historical deforestation, suggesting a lagged response to such perturbations and potential extinction debt.

Introduction

Global biodiversity is under threat from anthropogenic-driven habitat loss (Giam 2017; IPBES 2019), and conservation action is required to minimise the negative effects of this (Cullen 2012; Dinerstein *et al.* 2017). Identifying patterns in the distribution and diversity of species is a key step to prioritising areas to protect to achieve conservation goals (Gaston 2007). The global distribution and diversity of species have generally been well studied, and the latitudinal biodiversity gradient (LBG) is one of the most pervasive and longest-studied

phenomena in ecology dating back to Humboldt (Humboldt 1849), Darwin (Darwin 1859), and Wallace (Wallace 1878). Current research on this pattern continues to emphasise its generality across nearly all terrestrial and marine taxa (Hillebrand 2004; Menegotto & Rangel 2018; Smith *et al.* 2017; Willig *et al.* 2003) and different measures of diversity (Schumm *et al.* 2019; Smith *et al.* 2017). Such measures of diversity may however be inefficient when designating protected areas, as they fail to incorporate the expected response of communities to habitat loss – high-diversity communities will not necessarily be more negatively affected by habitat loss than low-diversity communities. Therefore, identification of whether a latitudinal gradient exists in the proportion of species associated with key ecotypes, and understanding the mechanisms which cause this, is not only critical to furthering our understanding of biodiversity distributions but could also inform conservation.

Forests are perhaps the world's most important terrestrial ecotype, covering 30% of the earth's land surface, yet providing habitat for approximately 80%, 75% and 70% of terrestrial amphibian, bird, and mammal species respectively (FAO 2020). Despite this, forests are the habitat type experiencing the fastest loss, with a third of global forest cover already lost (Ritchie & Roser 2023b). It is, therefore, no surprise that the protection of forests, and the ecological communities within them, is a vital part of global plans to mitigate biodiversity loss (Arroyo-Rodríguez *et al.* 2020; Dinerstein *et al.* 2017; Waldron *et al.* 2020). Species however vary in their sensitivity to forest loss (Betts *et al.* 2014; Henle *et al.* 2004; Keinath *et al.* 2017), with species that specifically rely on resources and conditions provided by forests (forest species) expected to be the most sensitive (Henle *et al.* 2004; Keinath *et al.* 2017). Communities with particularly high proportions of forest species would be expected to be most strongly affected by deforestation, and it could therefore be proposed that areas with high proportions of sensitive species should be protected above areas that simply have the most species. Due to their global biodiversity and continued anthropogenic threat, we focus on identifying global patterns in the proportion of forest species, as well as the potential

drivers of this pattern, particularly focusing on whether the proportion of forest species follows a latitudinal gradient.

For a latitudinal gradient in the proportion of forest species to occur, the richness of forest species must decrease disproportionately towards the poles compared to generalist species associated with no particular ecotypes, or species associated with other ecotypes (e.g. grasslands, deserts). A number of factors would be expected to create such a pattern, including the global distribution of forests, latitudinal variation in the structural complexity of forests (Ehbrecht *et al.* 2021), more disturbances towards the poles (Betts *et al.* 2019), higher stability at the equator (Brown 2014; Pontarp *et al.* 2019; Wiens & Graham 2005), more species creating a 'niche packing' effect (Macarthur, 1965; Sánchez González *et al.*, 2023), and presence of extreme environments (Rivas-Salvador *et al.* 2019). An overview of each hypothesis, its rationale, and our expectations are outlined in Table 2.1.

Table 2.1. Outline of hypothesis expected to affect the global distribution of the proportion of forest species, giving rationale as to how and why it is expected to create a latitudinal gradient, and linking to similar hypotheses within the LBG literature as defined in Fine (2015).

Hypothesis Name	**Rationale**	**Latitudinal Expectation**	**LBG Analogue**
Current Forest Cover	Increased proportion of total area available to forest species, and decreased area available for species associated with other ecotypes (Matthews *et al.* 2014; Pimm & Askins 1995).	**Increase the proportion of forest species as latitude decreases** since contemporary forest cover is skewed towards the equator (Hillebrand 2004; Saupe 2023; Willig *et al.* 2003).	Geographic area
Historial Deforestation	Areas with historically high proportions of forest cover are expected to have higher	**No effect on the latitudinal distribution of the**	Geographic area

	proportions of forest species due to the increased forested area previously available to them and known lagged responses of species to environmental changes (Chen & Peng 2017; Liao *et al.* 2022; Tilman *et al.* 1994).	**proportion of forest species** as historical forest cover was relatively uniformly spread globally (Betts *et al.* 2017).	
Forest Stability	Long-term persistence of environmental conditions is known to increase speciation rates and is a central hypothesis of the traditional LBG (Brown 2014; Pontarp *et al.* 2019; Wiens & Graham 2005). Increased speciation rates in forested environments would be expected to increase the number of forest species, and thus increase the proportion of forest species in the area.	**Increase the proportion of forest species as latitude decreases** since the tropical forests of the equator have generally been stable over the last 55 million years (Fine & Ree 2006), whilst forests at higher latitudes have changed between tropical, temperate, and boreal conditions (Fine & Ree 2006).	Climatic stability, Time
Forest Complexity	More species occur in areas with more structurally complex forests, due to the greater number of available niches. More species in forested environments means more forest species, which increases the proportion of forest species in the area.	**Increase the proportion of forest species as latitude decreases** since alpha diversity of plants in forests, and therefore structural complexity, increases towards the equator (Sabatini *et al.* 2022).	Spatial Heterogeneity
Natural Disturbances	Higher numbers of generalist species have been found in disturbed regions (Betts *et al.* 2019), likely due to the lack of stable habitat types, which increases extinction rates of specialist species. Hence the proportion of forest species is expected to decrease in areas which have experienced natural disturbances.	**Increase the proportion of forest species as latitude decreases** since natural disturbances are skewed towards higher latitudinal regions (Betts *et al.* 2019)	Climatic stability

Altitude	Extreme environments such as mountains increase the specialisation of species (Rivas-Salvador *et al.* 2019), as well as creating physiological barriers to dispersal which further increase speciation and niche partitioning (Janzen 1967). Both of these factors could increase the number of forest species in an area, and thus the proportion of forest species.	**Increase the proportion of forest species as latitude decreases** as extreme altitudes are skewed towards low latitude regions.	Narrow physiological tolerance
Niche Packing	As species richness increases, niches must be further partitioned and therefore the proportion of specialist species increases (Macarthur 1965; Sánchez González *et al.* 2023). If these species inhabit forested areas, then we would expect a higher proportion of forest specialist species in areas with higher species richness.	**Increase the proportion of forest species as latitude decreases** since species richness is known to follow a latitudinal gradient whereby the number of species increases from the poles to the equator (Allen & Gillooly 2006; Willig *et al.* 2003)	Niche Width

In summary, we predict that the proportion of forest species will increase with the amount of current and historical forest cover in an area, as well as the stability and complexity of the forests, where we expect higher proportions of forest species in areas which have had few disturbances and have remained the same forest type. We additionally expect factors which increase specialisation and speciation to increase the proportion of forest species towards the equator, specifically total species richness and altitude. A combination of these factors is predicted to create a latitudinal gradient in the proportion of forest species, likely to be most heavily driven by the disproportionately high amounts of forest at the equator. To test these expectations we use range and habitat use data from IUCN (“The IUCN Red List of Threatened Species” 2021a) and Birdlife International (“BirdLife Data Zone” 2021) to map

the global distribution of the proportion of forest species (described as those only occurring in forest habitats). We then use generalised additive mixed models to test the aforementioned hypotheses that could cause such a pattern using data on current forest cover, recent deforestation, geological forest cover, natural disturbances, plant diversity, altitude, and species richness. We find a clear latitudinal gradient in the proportion of forest species, which is associated with, but not solely explained by current global forest cover, recent deforestation, total species richness, and geological forest stability.

Methods

Map Projections

All map data in this study were projected to Behrmann Equal Area projection at the native resolution of 96x96km pixels, including conversion to raster format if the data were not already of this type. For brevity, any mention of projecting maps in the methods refers to Behrmann Equal Area projection at 96x96km pixel size. We chose to use an equal area projection to ensure that the proportion of species per area is comparable across the globe.

Proportion of Forest Species

We obtained distribution range maps and habitat data for four taxa: amphibians, birds, mammals, and reptiles from IUCN (“The IUCN Red List of Threatened Species” 2021a) and BirdLife International (“BirdLife Data Zone” 2021). For a small number of species habitat data were not available (0.07%), these were removed from further analysis. We also removed areas of the distribution range classified as presence uncertain; non-native origin and seasonality descriptors other than “Resident” or “Breeding”, which removed a further 1.3% of the initial species. Lastly, we removed all marine species which left a final total species count of 6,963 amphibian, 10,041 bird, 5,619 mammal, and 9,477 reptile species.

We then designated species as forest or non-forest based on the habitats that were defined as suitable for them by the IUCN. Forest species only had forest habitats (category 1 in the IUCN Habitats Classification Scheme) described as suitable, whilst non-forest species either

had a mixture of forest and non-forest habitat types, or only non-forest habitat types defined as suitable. We then projected the ranges of all species and used these to create maps of the total species richness and the total forest species richness of each taxon in each 96x96km pixel. From these rasters we then calculated the proportion of forest species in each pixel by dividing the forest species richness by the total species richness.

Current Forest Cover (Proportion of Current Forested Land Area)

We used the Hansen tree cover dataset (Hansen *et al.* 2013), which describes the approximate percentage of tree cover in a 30x30m pixel for the year 2000. While these data are now over 20 years old, they still represent the best estimate of global tree cover and are unlikely to affect our results as the IUCN and BirdLife range maps were created over a similarly coarse temporal and spatial scale. We reclassified pixels into forest or non-forest categories using a 70% cut-off value (i.e. pixels with at least 70% tree cover were classified as forest). A cut-off of 70% was chosen as a conservative categorisation of forest as the definition of a forest varies nationally between datasets (e.g 60% of tree cover is determined as forest in Brazil's mapBiomas, 20% in the United States USGS National Land Cover Database, and 70% in the Copernicus Global Land Cover Dataset) (Buchhorn *et al.* 2020; Souza *et al.* 2020; Yang *et al.* 2018). The Hansen dataset also contained a land mask that classifies pixels as land or not land (i.e. oceans). We projected both the forest and land data to our larger 96x96km pixel size, summing their total area to calculate a total forest area and total land area per 96x96km pixel. We then divided the total forest area by the total land area to calculate the proportion of current forest cover per pixel, and the total land area by the total area of our larger pixels to calculate the proportion of land area per pixel.

Historical Deforestation (Historical Proportion of Deforested Land Area)

The historical proportion of deforested land area was calculated using a map of predicted forest cover prior to human activities, which had had the areas of current forest cover removed, this map was also already categorised as forest or non-forest , available from Betts *et al.* (2017) at 0.009° pixel resolution. This map was generated using a random forest

regression model which predicted the distribution of historical forest as a function of 19 bioclimatic variables from the WorldClim database ("WorldClim" 2024), and the coordinates of forest loss from the Hansen tree cover dataset (Hansen *et al.* 2013). The maps were projected and the total deforested area was calculated by summing the area of the forested pixels. The historical proportion of deforested land area was calculated by dividing the total deforested area by the total land area. Note that due to the different initial pixel sizes of the land area raster and the deforested raster some pixels contained values greater than one. Most of these were removed during the 'Data Preparation' process described below, but 85 of 5,434 pixels for amphibians (1.6%, mean value > 1 (sd) = 1.13 ±0.317), 125 out of 9,571 pixels for birds (1.3%, mean value > 1 (sd) = 1.13 ±0.305), 125 out of 9,435 pixels for mammals (1.3%, mean value > 1 (sd) = 1.13 ±0.305), and 101 out of 6,125 pixels for reptiles (1.6%, mean value > 1 (sd) = 1.05 ±0.087) remained with values above one, and as such these were all set to a value of one since the proportion of land area deforested cannot by greater than the total land area present.

Natural Disturbances

Maps of historical naturally-occurring disturbance areas caused by fire (Lavorel *et al.* 2007), glaciation (University of Geneva, Switzerland *et al.* 2001), and storms ("Location of tropical cyclones" 2024) were extracted from relevant sources where they were generated by expert opinion (Lavorel *et al.* 2007; University of Geneva, Switzerland *et al.* 2001) and satellite observation ("Location of tropical cyclones" 2024). These layers were projected and then combined by summing the occurrence of disturbances to create a single raster including all three disturbance types, with pixel values between 0 - 3 depending on the number of disturbances in the pixel.

Maximum Geological Forest Time and Geological Forest Stability

Maps of predicted forest type (tropical, temperate, or boreal) from five geological epochs (Holocene, Pliocene, Miocene, Oligocene, and Eocene) were hand digitised from Fine & Ree (2006) into five vector layers. These layers were projected and pixels were then given values

based on two measures of temporal stability of the forest type present. Firstly, we calculated the ‘maximum geological forest time’, which we define as the approximate amount of time the pixel could have remained the same forest type since the Holocene (0 mya). The period assigned to each of the epochs is as follows, Holocene = 0mya, Pliocene = 5mya, Miocene = 23mya, Oligocene = 33mya, and Eocene = 55mya. For example, a pixel not classified as tropical forest in the Eocene but then classified as tropical forest throughout the Oligocene, Miocene, Pliocene and Holocene would have a value of 33 (the Oligocene started roughly 33 million years ago), which represents how long the area could have been continuously covered by tropical forest. Secondly, we calculated the ‘geological forest stability’, which we define as the number of layers the pixel was classified as the same forest type as it is currently (Holocene, 0mya). For example, if a pixel was categorised as tropical forest in the Eocene and Oligocene, temperate in the Miocene, and tropical forest in the Pliocene and Holocene the pixel would be given a value of 4.

Plant Alpha Diversity

Raster maps of the global alpha diversity of flora were obtained from Sabatini *et al.* (2022) at 0.04° pixel resolution and projected by averaging the plant alpha diversity across the larger 96x96km pixels.

Altitude

We obtained a raster map of global altitude from ETOPO 3 (NOAA National Centers for Environmental Information 2022) at 60-arc second resolution. This was projected by averaging the altitude of pixels within the larger 96x96km pixel.

Latitude (Distance from Equator 1000km)

To define latitude, we calculated the distance from each pixel to the equator (described in 1000kms for ease of interpretation).

Forest Ecoregions

Lastly, we cropped all of our maps to only include data for ecoregions which contain forest biomes based on the RESOLVE 2017 dataset (Dinerstein *et al.* 2017). This was done to ensure we were comparing areas in which forest species would be expected to occur and to aid with the interpretation of results. For example, differences in plant diversity between two forest ecoregions suggest changes in structural complexity and the number of niches available to forest species. On the other hand, comparing differences in plant, and therefore structural, diversity between a forest and grassland ecoregion has far less ecological meaning in the context of the proportion of forest species. We also assigned each pixel to the forest ecoregion it occurred in which was later used as a random effect in our models.

Quantification and Statistical Analysis

All data quantification, spatial analysis, statistical analysis, and visualisation were conducted in R version 4.2.2 (R Core Team, 2023) using the mgcv package version 1.9.0 (Wood, 2011) and the ‘bam’ function for generalised additive mixed effect models, the DHARMa package version 0.4.6 (Hartig, 2022) was used for model evaluation of overdispersion and model fit, the performance package version 0.10.8 (Lüdecke *et al.*, 2021) was used to assess multicollinearity, and the ncf package version 1.3.2 (Bjornstad, 2022) was used to assess spatial autocorrelation using correlograms.

Data Preparation

Before analysing the proportion of forest species in each pixel, we removed pixels which had a small amount of land within them (<10% land area).

Modelling the Proportion of Forest Species

We modelled the proportion of forest species for each taxon individually, as well as for all taxa combined, using two sets of explanatory variables. All models used binomial errors and the logit link function, since our response variable, the proportion of forest species, is a set of trials (total number of species in a pixel), and successes (the number of forest specialist

species in a pixel). We first used a simple model which predicted the proportion of forest species in a pixel using only latitude (distance of the pixel from the equator) as an explanatory variable. Secondly, we used a more complex model which used all of the previously described variables (Current Forest Cover, Historical Deforestation, Naturally Disturbed Areas, Maximum Geological Forest Time, Geological Forest Stability, Plant Alpha Diversity, Altitude, Total Species Richness, and Latitude) to predict the proportion of forest species in a pixel. In both cases, we ran generalised additive mixed models, with explanatory variables scaled and centred to allow for comparison of effect sizes. Though we used generalised additive mixed models, the fixed effects were included as linear predictors and outputs from them can be interpreted as if from a generalised linear mixed effect model. Models including all taxa combined included random intercept terms for taxa. The structure of all models can be seen in the ‘model structure‘ section of the supplementary, and model variables were chosen based on our prior hypotheses and were not included or excluded based on model comparisons such as AIC.

For both sets of models we accounted for spatial autocorrelation since preliminary models which did not account for it showed spatial clustering of residuals and therefore lacked independence between sampling points (Legendre 1993). We used two methods to account for spatial autocorrelation. Firstly, we included a Gaussian process which fits a smooth spatial trend by modelling the way the correlation between pairs of observations (proportion of forest species in a pixel) varies as a function of the distance in space between them (Simpson 2018). Several functions can be used to model the correlation between observations and we chose to use the Matern family since it is the most commonly used, though all family choices appeared to perform similarly well in observational comparisons of correlograms. Secondly, we included a random intercept term for ecoregion which improved model fit and slightly further reduced spatial autocorrelation. Diagnostics for all models including comparisons of spatial autocorrelations between models with and without the Gaussian process can be seen in S1.

Proportion of Forest Species vs Total Species Richness

To highlight the differences between total species richness and our measure of the proportion of forest species we plotted the proportion of forest species and total species richness against absolute latitude, as well as plotting the proportion of forest species against total species richness to visualise the relationship.

Results

Mammals had the highest proportion of forest species (N = 2,166, 39%), followed by reptiles (N = 3,086, 33%), birds (N = 3,011, 30%), and amphibians (N = 1,735, 25%). The highest proportion of forest species for all taxa occurs in the equatorial tropical forests, with particularly high proportions in South America and Southeast Asia (Figure 2.1, 2.2A). All taxa followed a similar latitudinal pattern, whereby the proportion of forest species decreased with latitude (effect size of latitude, all taxa = -0.96, amphibians = -0.72, birds = –0.97, mammals = –0.94, reptiles = –1.01) (Figures 2.1, 2.3). For example, our models predict that 33.5% of species at the equator are forest species, which decreases to 9.3% at 30° latitude (tropics) and to 4.1% at 50° latitude (temperate). The proportion of forest species generally follows the same trend as total species richness (Figures 2.2A, 2.2B) which similarly decreases from the equator to the poles.

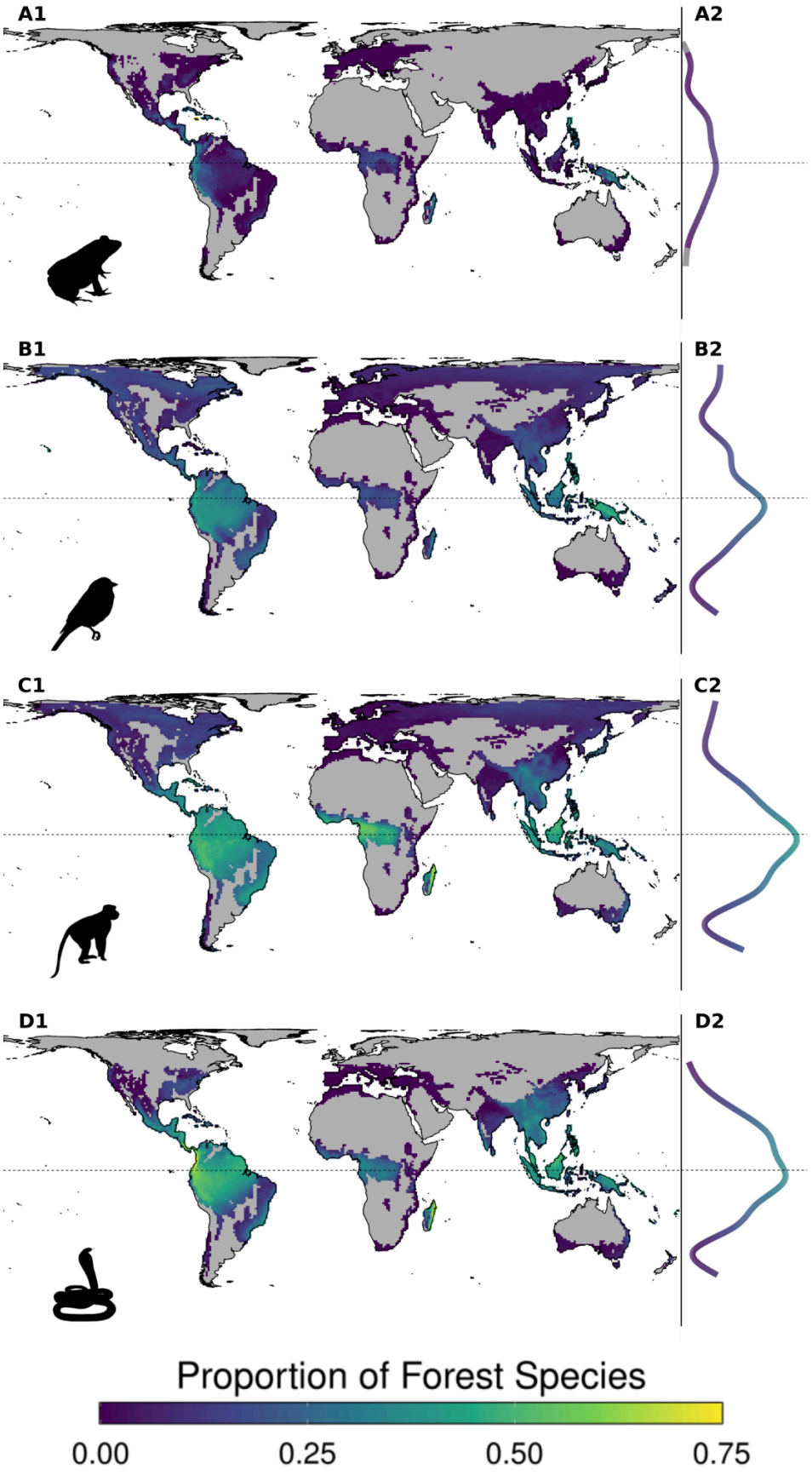

A1
A2
B1
B2
C1
C2
D1
D2
Proportion of Forest Species
0.00
0.25
0.50
0.75

Figure 2.1. Global distribution of the proportion of forest species of amphibians (A1), birds (B1), mammals (C1), and reptiles (D1) in forest ecoregions. All taxa show a peak in the proportion of forest species at the equator, and a gradual decrease towards the poles. Figures A2:D2 represent the relationship between the proportion of forest species and latitude generated by a GAM fit, with the proportion of forest species shown on the x-axis and by the colour scheme; grey represents erroneously predicted values below 0%. We calculated the proportion of forest species for pixels with at least 10 species of a taxa and that contained at least 10% land cover.

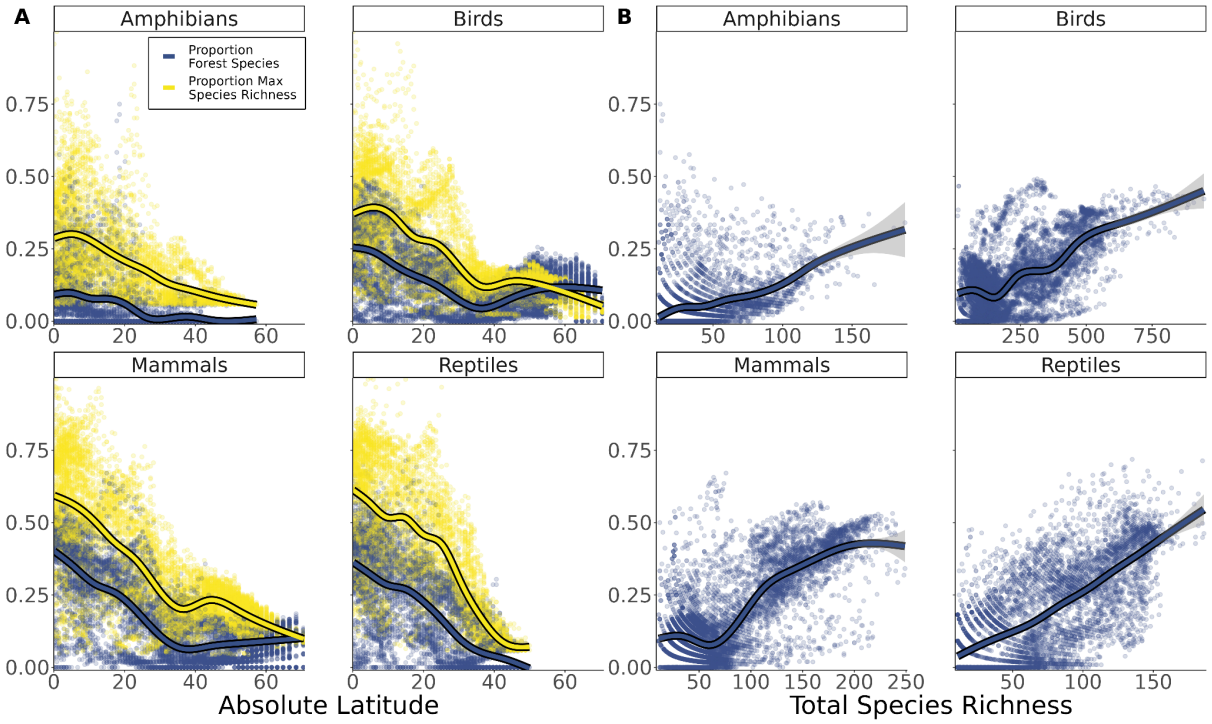


Figure 2.2. Both the proportion of forest species and total species richness decreased with absolute latitude (**A**), and the proportion of forest species generally increased with total species richness (**B**). (**A**) and (**B**) are both generalised additive model fits. Proportion of maximum species richness is used in (**A**) to bound values between 0 - 1 and allow for plotting on the same scale as the proportion of forest species. This was calculated per taxa as the total species richness of a pixel divided by the maximum total species richness of any pixel.

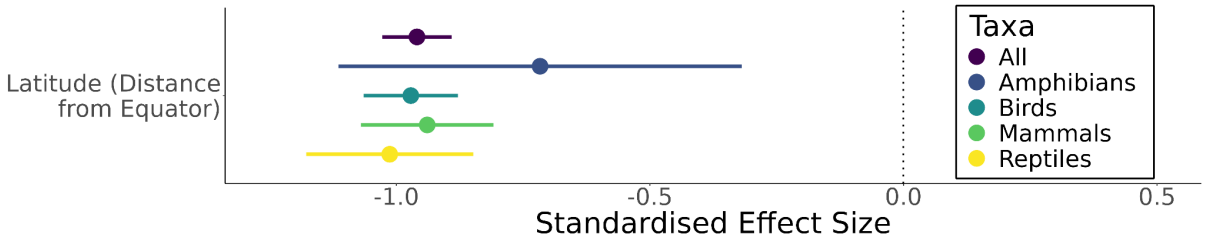


Figure 2.3. Proportion of forest species significantly decreases with latitude for amphibians, birds, mammals and reptiles. Results of generalised additive mixed models using latitude as the only explanatory variable, and accounting for spatial autocorrelation. Latitude has been scaled and centered to standardise the effect size and allow for comparison between models, Predictors can be interpreted as significant if the 95% confidence intervals do not cross 0.

As expected, for all taxa the proportion of forest species increased with current forest cover, historical deforestation, geological forest stability, and total species richness (Figure 2.4). Additionally, the proportion of forest species generally increased with altitude, though this was not significant for mammals. We found no consistent relationships between natural disturbances or plant alpha diversity and the proportion of forest species, and a generally negative, albeit weak, relationship between the maximum amount of time an area had remained the same type of forest (maximum geological forest time) and the proportion of forest species. Even with these explanatory variables, the proportion of forest species of all taxa, apart from amphibians, still showed a strong negative relationship with increasing latitude, though the effect was slightly dampened (change in effect size of latitude compared to latitude-only models, all taxa = 0.37, amphibians = 0.54, birds = 0.28, mammals = 0.45, reptiles = 0.50) (Figure 2.4), and the variation explained by the models only increased marginally (Table 2.2).

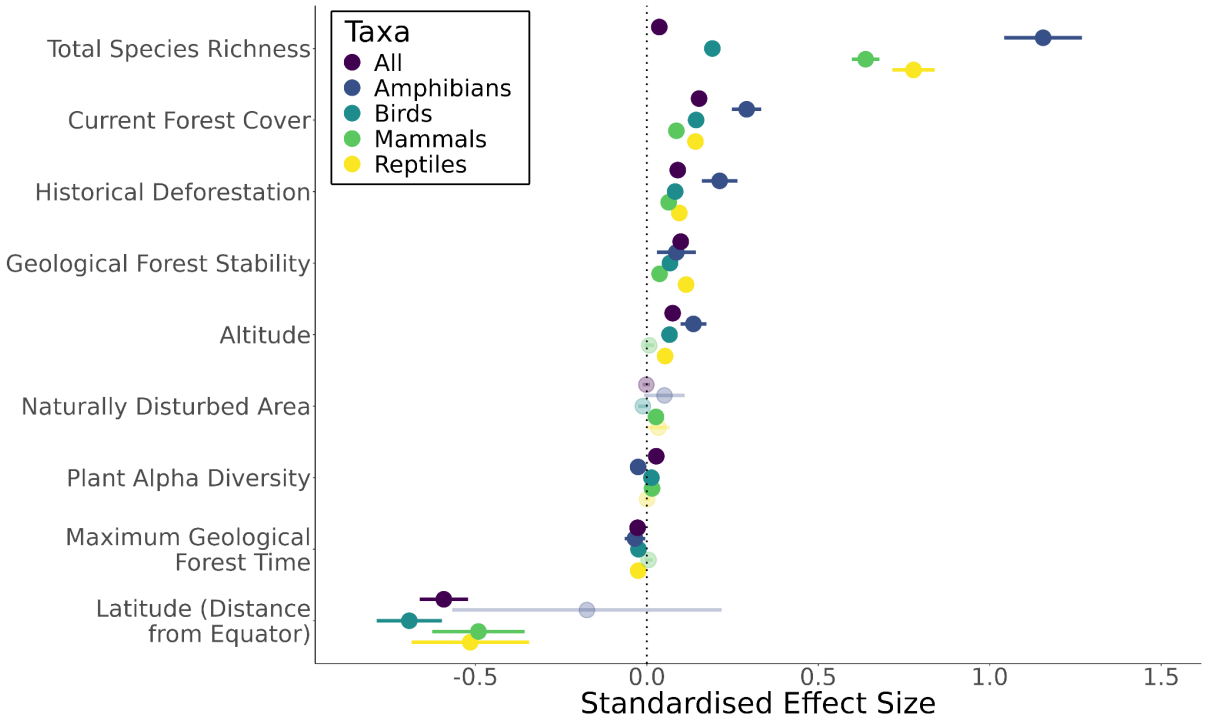


Figure 2.4. Proportions of forest species ubiquitously increases with current forest cover, historical deforestation, geological forest stability, and total species richness, whilst generally increasing with altitude. On the other hand proportions of forest species generally decreased with latitude and maximum geological forest time, whilst results were mixed for naturally disturbed areas and plant alpha diversity Results of generalised additive mixed models using all explanatory variables, while accounting for spatial autocorrelation. All variables have been scaled and centered to standardise effect sizes and allow for comparison between variables and models. Predictors can be interpreted as significant if the 95% confidence intervals do not cross 0. Non-significant predictors are denoted by transparent points and intervals.

Table 2.2. R^2 of generalised additive mixed models using latitude as the only explanatory variable, and models including all explanatory variables. Note that the R^2 value is artificially high due to the explanatory power of terms accounting for spatial autocorrelation and random effects.

Taxa	Latitude-Only Model R^2	Full Model R^2
All	0.86	0.87
Amphibians	0.84	0.86
Birds	0.93	0.96
Mammals	0.95	0.96
Reptiles	0.93	0.95

Discussion

We found a strong negative relationship between latitude and the proportion of forest species of all terrestrial vertebrates. Surprisingly, the strong effect of latitude on the proportion of forest species remained for all taxa except amphibians once we accounted for key variables expected to influence the global distribution of forest species. Other than latitude, the total species richness of the community, as well as the amount of forest cover, and historical deforestation in the area most strongly affected the proportion of forest species. Interestingly, the effect of historical deforestation was almost as strong as that of current forest cover; regions with higher amounts of forest historically tend to still have greater proportions of forest species – even long after that forest has been removed. This suggests a lagged response of forest species to the removal of habitat. In contrast to our expectations, we found a generally negative effect of maximum geological forest time, as well as non-significant and mixed effects of plant diversity and the presence of natural disturbances on the proportion of forest species, though the positive influence of geological

forest stability suggests that long-term stability of a habitat type increases specialisation towards that habitat to some extent.

As expected, we found that the proportion of forest species in an area increased with the amount of forest cover. The species-area relationship is a well-studied phenomenon in ecology (Matthews *et al.* 2014; Pimm & Askins 1995), and the higher proportion of forest species in areas with more forest cover is likely a direct result; more area of a habitat type at the regional scale increases the pool of species associated with that habitat (Halstead *et al.* 2019; Leibold *et al.* 2004), and decreases the number associated with other habitat types. We similarly found that the proportion of forest species increased with the amount of deforestation that had historically occurred in the area. This suggests that forest species which occurred in the area before deforestation still at least partially remain; likely in smaller forest remnants. Whether the high proportions of forest species in such areas will remain long-term is unknown, however, previous studies have found lagged responses of species to habitat loss (Chen & Peng 2017; Liao *et al.* 2022; Tilman *et al.* 1994), which only increase as the size of the area increases (Haddad *et al.* 2015), meaning that this may be a case of extinction debts that are yet to be paid, and could take centuries to do so (Chen & Peng 2017).

Factors that were expected to increase the specialisation of species were similarly found to generally increase the proportion of forest species in an area. Total species richness had consistently strong effects on the proportion of forest species, supporting the ‘niche packing’ hypothesis that species are required to specialise in more specific niches to maintain a stable community with high species richness (Macarthur 1965; Sánchez González *et al.* 2023). This is further corroborated by the weak effect of total species richness in our model including all taxa, indicating that niche-packing is working at an intra-taxa level, in which species are more likely to be competing for shared resources. For example, the total number of amphibian species does not affect the proportion of forest bird species in an area. Contrastingly, altitude had a relatively weak but generally positive effect on the proportion of

forest species, potentially stemming from increased speciation due to dispersal limitations (Janzen 1967) and increased specialisation to survive in relatively extreme environments (Rivas-Salvador *et al.* 2019).

On the other hand, we found that hypotheses that most closely match those in the traditional latitudinal biodiversity gradient literature, such as more stable environments increasing speciation (Allen & Gillooly 2006; Brown 2014; Mittelbach *et al.* 2007), tended to be less supported. Neither the presence of natural disturbances nor the amount of time an area had continuously been the same forest type consistently affected the proportion of forest species. The lack of relationship between the proportion of forest species and these variables may be due to the ability of species to move, find refugia, and then repopulate the disturbed or altered area when it returns to a habitable state; similar to species repopulating areas after modern disturbances (Haney *et al.* 2008; Kotliar *et al.* 2007). This is further supported by the positive relationship between geological forest stability and the proportion of forest species, whereby a higher proportion of forest species were found in areas which had generally contained the same habitat type, even if it was not temporally continuous. This suggests that forest species can persist in areas when the type of forest present changes, either by moving to nearby refugia and repopulating when the forest type changes again to become favourable (Sommer & Nadachowski 2006; Waltari *et al.* 2007), or simply adapting to the new conditions.

Perhaps the most obvious differences between forests across latitudes are their plant diversity and structural complexity; with diversity and structural complexity increasing towards the equator. Higher plant diversity, and therefore structural complexity (Sabatini *et al.* 2022), should create more niches for forest species (Bracewell *et al.* 2018; Brown 2014; Cao *et al.* 2021) and therefore increase the proportion of forest species found in an area. However, while we found that the proportion of forest species did generally increase with plant diversity, the effect was extremely weak and variable between taxa, and did not account for the observed latitudinal gradient. In fact, latitude had one of the strongest effects

on the proportion of forest species (only total species richness had a larger effect for some taxa) after all other variables had been included, which indicates that the gradient is additionally being driven by unmeasured factors.

We focused our analysis on the effects of habitat type and environmental perturbations on the proportion of forest species in an area, however, traits of the species themselves may explain the observed latitudinal gradient. A plausible hypothesis borrowed from the latitudinal biodiversity gradient literature is the “Out of the Tropics” hypothesis. This hypothesis postulates that species richness decreases with latitude because species require greater dispersal ability to move from the equatorial refugia of the last glacial maxima to the relatively new habitat of the high latitudes (Brown 2014; Jablonski *et al.* 2006). This hypothesis is supported for the LBG in birds by the latitudinal gradient in hand wing index; a widely used proxy for dispersal ability (Sheard *et al.* 2019, 2020; Weeks *et al.* 2023), and could explain our observed latitudinal pattern in the proportion of forest species, if forest species are generally poorer dispersers than generalist or open area species. Whilst a recent analysis found no significant differences in the hand wing index of closed versus open-area species (Sheard *et al.* 2019), analysis of our categorisation of species into forest or non-forest suggests that forest species are indeed poorer dispersers (Figure S2.1), and that this therefore may be a viable alternative hypothesis as to why the proportion of forest species varies with latitude.

Our findings were generally consistent across birds, mammals, reptiles, and in models which included all taxa, however, results for amphibians varied considerably. The most interesting of these was the lack of support for a relationship between latitude and the proportion of forest amphibian species. This in large part appeared to be caused by an increased effect of total species richness, current forest cover, historical deforestation, and altitude on the proportion of amphibian species compared to that of the other studied taxa, explaining much of the observed latitudinal variation. This discrepancy was likely due to the smaller latitudinal range over which amphibians occur compared to the other studied taxa (Figure 2.1), with

very few data points at high latitudes, as well as the smaller range sizes of amphibian species which could be entirely encompassed in a single pixel. Despite this, the direction and relative strength of effect of variables were quite consistent between taxa. Similarly, we believe such a pattern is likely to hold across proportions of species associated with other ecotypes, since total species richness and latitude had by far the strongest effect, though differences in the distributions of other ecotypes could dampen or exacerbate the latitudinal gradient (Dinerstein *et al.* 2017), as could their respective stability over evolutionary timeframes (Jetz & Fine 2012).

The observed pattern of increasing proportions of forest species towards the equator may explain the common finding that ecological communities in forests near the equator are less resilient to habitat loss than those nearer the poles (Betts *et al.* 2019; Melo *et al.* 2018). Since forest species are less likely to be able to persist after deforestation and land-use change than generalist species (Hansen & Urban 1992; Swift & Hannon 2010). This finding would suggest that protected areas should be prioritised close to the equator, as should efforts to minimise deforestation which will have a strong negative effect on forest species (Anjos *et al.* 2019; Sverdrup-Thygeson *et al.* 2017; Vargas-Cárdenas *et al.* 2022). Additionally, our results suggest that reforesting historically deforested areas is a conservation priority since they contain disproportionately high proportions of forest species, which may be declining towards local extinction (Liao *et al.* 2022; Tilman *et al.* 1994). Ultimately, we report a striking relationship between the proportion of forest species and latitude, which can only partially be explained by key factors such as current and historical forest cover. It is likely that this pattern, similarly to the latitudinal biodiversity gradient, is driven by a plethora of factors, and uncovering these may shed light on drivers of species distributions.

Chapter 3: Global Variation in Forest Thresholds for Community Persistence

This chapter has been prepared for submission in a Nature or Science journal, and has therefore been written in an appropriate format for this.

Abstract

Area-based conservation initiatives such as 30 by 30 are perhaps the most promising methods for halting the global biodiversity crisis. However, the amount of habitat needed to maintain biodiversity is not well defined and is likely to vary globally. Thresholds in the relationship between biodiversity and habitat loss have previously been identified as useful methods to identify the optimal trade-off between habitat loss and conservation, but there is little understanding of how such thresholds vary globally. Using 97 faunal studies on a global scale, we identified the amount of forest after which the relationship between forest cover and community integrity significantly changes; the threshold. We found that thresholds were not globally consistent and that there are two distinct types of thresholds, a Decelerating and an Accelerating Threshold. We found communities in highly forested regions near the equator were almost completely modified after only losing 25% of forest cover, whilst in highly deforested regions some similarity to the natural community could be maintained after 75% deforestation or more. Our results suggest undisturbed highly forested regions should be prioritised for protection as they contain the most sensitive species, whilst communities in deforested regions are more resilient to deforestation, likely due to shifting baselines.

Main Text

Habitat loss is the greatest threat to biodiversity globally, largely causing the current biodiversity crisis (IPBES 2019; Mace *et al.* 2005; Zakaria *et al.* 2016). Area-based conservation efforts provide perhaps the best opportunity to conserve biodiversity, yet the amount of area needed to maintain biodiversity has been a constant source of debate (Allan *et al.* 2022; Fahrig 2001; Soulé & Sanjayan 1998; Williams *et al.* 2022). Recently leaders from 196 nations committed at COP15 (Maxwell *et al.* 2020; Waldron *et al.* 2020) to protect 30% of the surface of the earth for nature conservation by 2030 (Waldron *et al.* 2020), an initiative known widely as 30 by 30 (Maxwell *et al.* 2020; Waldron *et al.* 2020). Whilst some studies have found that maintaining 30% of areas as habitat reduces local extinctions and maintains biodiversity (Arroyo-Rodríguez *et al.* 2020; Hannah *et al.* 2020; Melo *et al.* 2018; Mogg *et al.* 2019), many others have suggested that more or less habitat is required to avert biodiversity loss (Betts *et al.* 2007, 2010; Boesing *et al.* 2018; Carlson 2000; Morante-Filho *et al.* 2015).

Studies that have estimated the amount of habitat required to maintain biodiversity have often used ecological thresholds (Banks-Leite *et al.* 2014; Estavillo *et al.* 2013; Swift & Hannon 2010), where the relationship between biodiversity and habitat amount significantly changes after an amount of habitat has been lost. This threshold potentially indicates an optimal trade-off amount of habitat loss which maintains biodiversity whilst allowing for the development of more economically valuable land uses (Banks-Leite *et al.* 2014). Some studies have suggested that the position of this ecological threshold is globally consistent, and occurs at approximately 30% habitat amount (Arroyo-Rodríguez *et al.* 2020; Banks-Leite *et al.* 2021), indicating that all areas worldwide could be relatively heavily modified whilst still maintaining biodiversity, and implying that globally ecological communities are identical in their response to habitat loss.

Species have been found to differ in their response to habitat loss, due to their traits and habitat requirements, with habitat specialist species and poor dispersers being the most

sensitive to habitat change (Betts *et al.* 2017, 2019; Bregman *et al.* 2014; Keinath *et al.* 2017; Weeks *et al.* 2023). As such ecological communities formed of many habitat specialist species and those with large numbers of poorly dispersing species are expected to require more pristine conditions than generalist mobile species. Studies have identified a latitudinal gradient in the dispersal capability of species, with generally more poorly dispersing species at the equator, and well dispersing species towards the poles (Salisbury *et al.* 2012; Sheard *et al.* 2019). Similarly, studies have found that more habitat specialist species occur at the equator (Howes *et al*. in prep; Chapter 2), and in regions which have experienced little to no anthropogenic disturbance; the extinction filter hypothesis (Betts *et al.* 2019). As such it would be expected that the most sensitive communities exist in undisturbed regions at the equator, and therefore ecological thresholds should be highest in these areas.

We used a global dataset to identify 1) Are ecological thresholds globally consistent?, and 2) Does the position of ecological thresholds vary with latitude and regional disturbance? To answer these questions we used the BIOFRAG dataset (Pfeifer *et al.* 2014) containing 97 studies (arthropods: 26 studies; birds: 41 studies; herpetofauna: 15 studies; mammals: 15 studies) in combination with a global forest cover dataset (Hansen *et al.* 2013). For each study, we identified communities in forested landscapes, which were those at, or close to, 100% forest cover. Then we calculated the Sørensen similarity between the community in these forested landscapes and the communities in deforested landscapes, giving a value of similarity between zero and one, which we term community integrity. To identify thresholds in each study we used mixed-effect piecewise regression models, where we modelled community integrity as a function of forest cover, where the relationship between forest cover and community integrity could change at a given forest cover percentage; the ecological threshold. Lastly, we ran mixed-effect beta regression models where we modelled the previously identified position of the ecological threshold as a function of the latitude of the study, and the proportion of land area that had been deforested (hereafter ‘regional deforestation’).

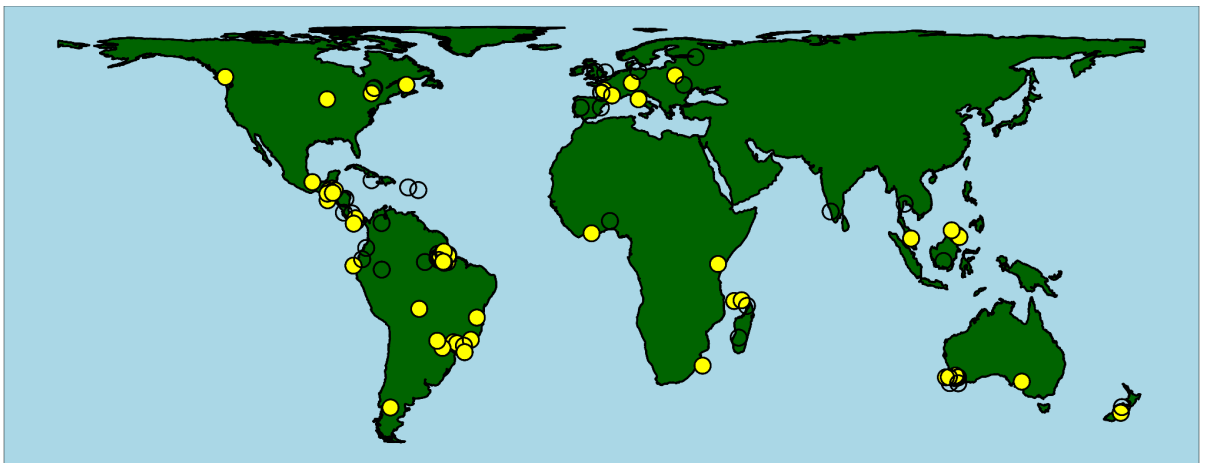


Figure 3.1. Global placement of the 97 studies. Points filled with yellow are studies for which thresholds were identified and used in the analysis; 53 studies. Points with no fill are studies for which thresholds were not identified and therefore were not used in the analysis; 44 studies. Points are slightly jittered to show studies whose points overlap.

We found that thresholds were not globally consistent and that they occur at forest cover percentages ranging from 6% to 97%, with a mean of 64%. Overall 53 (55%) of studies exhibited a threshold in the relationship between community integrity and forest cover, and the highest thresholds occurred in areas of low regional deforestation. While thresholds decreased as regional deforestation increased (-1.81, 95% CI = (-2.50, -1.13)), the strength of the effect was weakened as latitude decreased (-0.07, 95% CI = (-0.11, -0.04)), and thresholds towards the equator were almost universally high (Figure 3.2). Latitude itself had a weak effect on threshold position (0.05, 95% CI = (0.03, 0.06)), which varied with regional deforestation. Not all of the identified thresholds however exhibited the same shape of relationship between community integrity and forest cover (Figure 3.3A) meaning that the implications of the threshold position are not consistent.

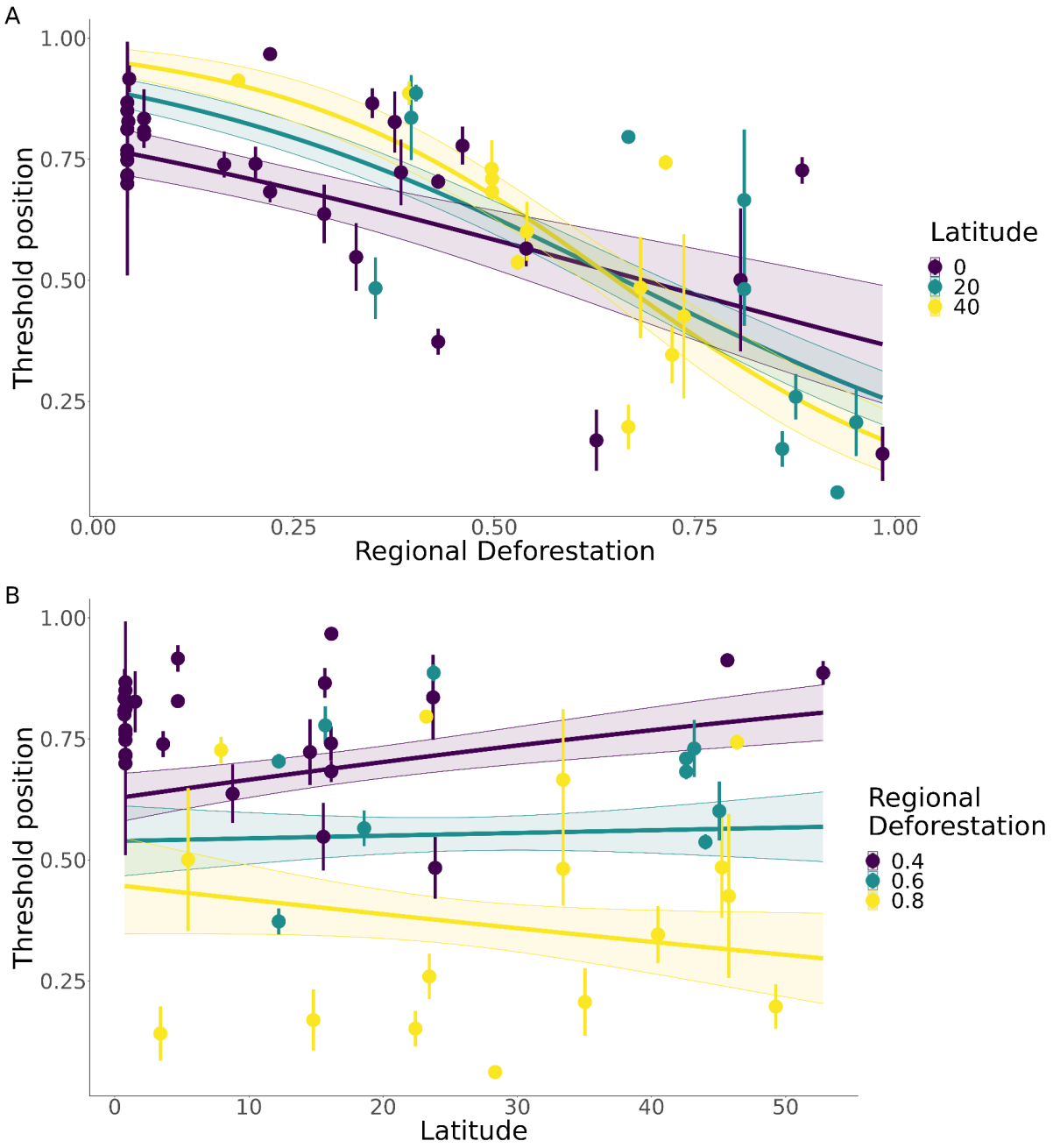


Figure 3.2. The position of thresholds decreases with regional deforestation, though thresholds at low latitudes stay relatively high even when regional deforestation is high (**A**). Threshold position decreases with latitude when regional deforestation is high, but increases with latitude when regional deforestation is low (**B**). Predictive lines based on beta generalised linear mixed effect model, and shaded areas indicate 95% confidence intervals, whilst the colour of the line and shaded areas indicate latitude (**A**), or regional deforestation (**B**). Points are the position of the thresholds in the 53 studies, and lines around the points

represent the 95% confidence interval of the threshold, which was propagated into the model using weights. Points are coloured by the latitude (**A**), or regional deforestation (**B**) value they are closest to.

We found two types of thresholds: the “Accelerating” and the “Decelerating” thresholds, found in 49% and 51% of the studies respectively. The “Accelerating Threshold”, signified a significant increase in the strength of the relationship between community integrity and forest cover after the threshold (as forest cover is lost) (Figure 3.3A; yellow). The second was a “Decelerating Threshold” which signified a significant decrease in the strength of the relationship between community integrity and forest cover after the threshold (as forest cover is lost) (Figure 3.3A; purple). Decelerating Thresholds occurred at significantly higher percentages of forest cover than Accelerating Thresholds (Decelerating = 0.73, 95% CI = (0.68, 0.76), Accelerating = 0.50, 95% CI = (0.44, 0.55)), and they were only weakly affected by regional deforestation, especially at low latitudes where Decelerating Thresholds occurred at consistently high levels of forest cover (Figure 3.4B). Accelerating Thresholds on the other hand were strongly affected by regional deforestation (-3.05, 95% CI = (-2.16, -3.95)), and the highest thresholds occurred at low levels of regional deforestation across all latitudes (Figure 3.4A).

Based on our results we theorise that all communities follow a similar pattern of response to deforestation, where the community first experiences a Decelerating Threshold, and then an Accelerating Threshold as habitat is lost. We suggest this as Decelerating Thresholds occur at higher percentages of forest cover than Accelerating Thresholds, and due to the shape of the relationship between community integrity and forest cover seen in studies with high-quality biodiversity data across the gradient of deforestation (Figures 3.3B and 3.3C). The Decelerating Threshold likely represents the end of the loss of forest specialist species, whilst the Accelerating Threshold represents the beginning of the loss of generalist species.

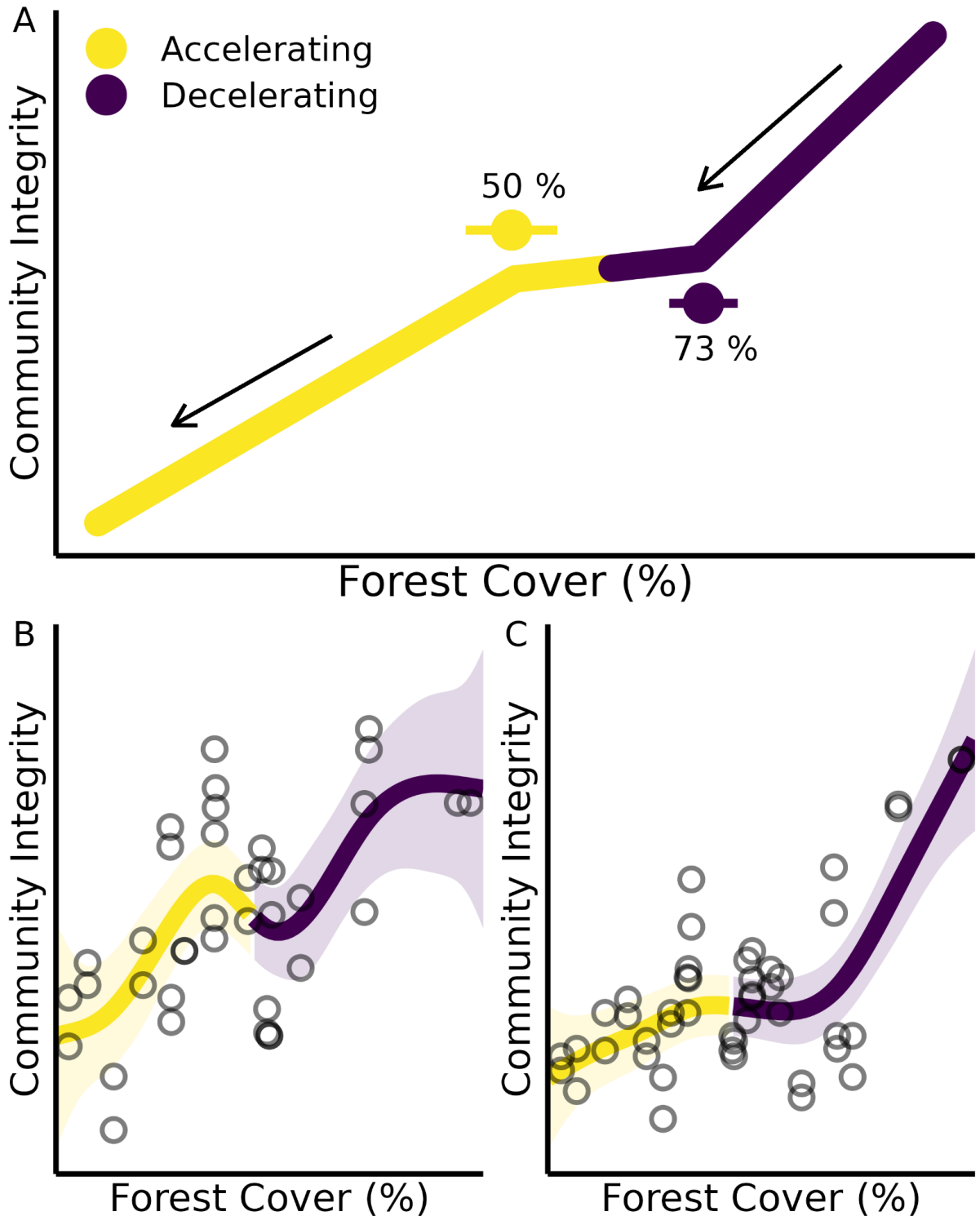
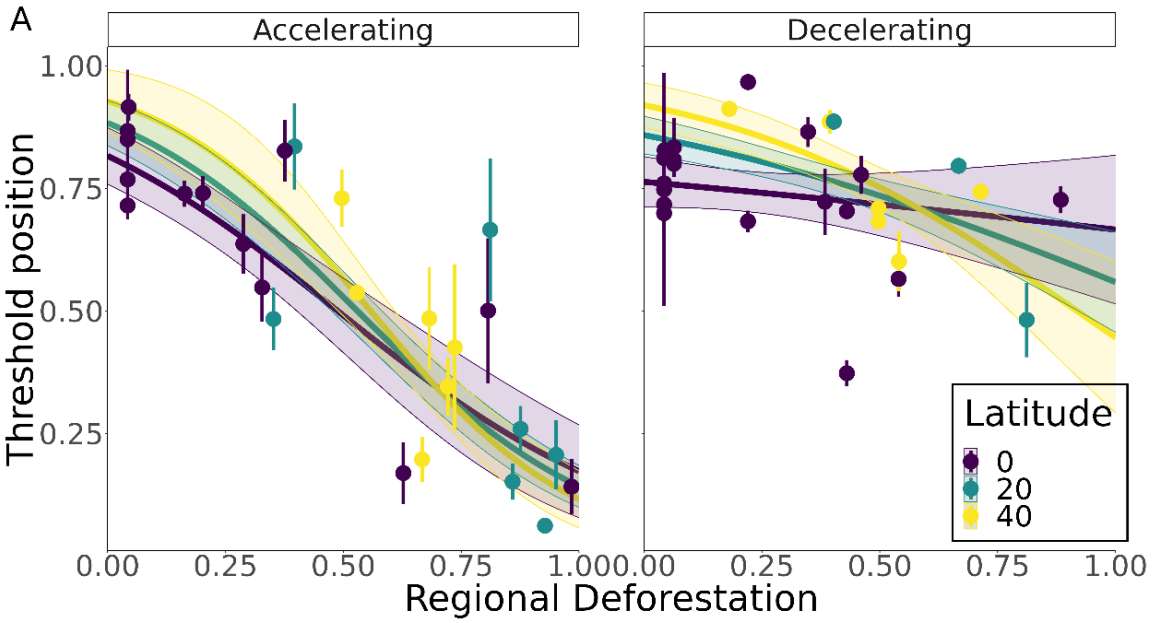
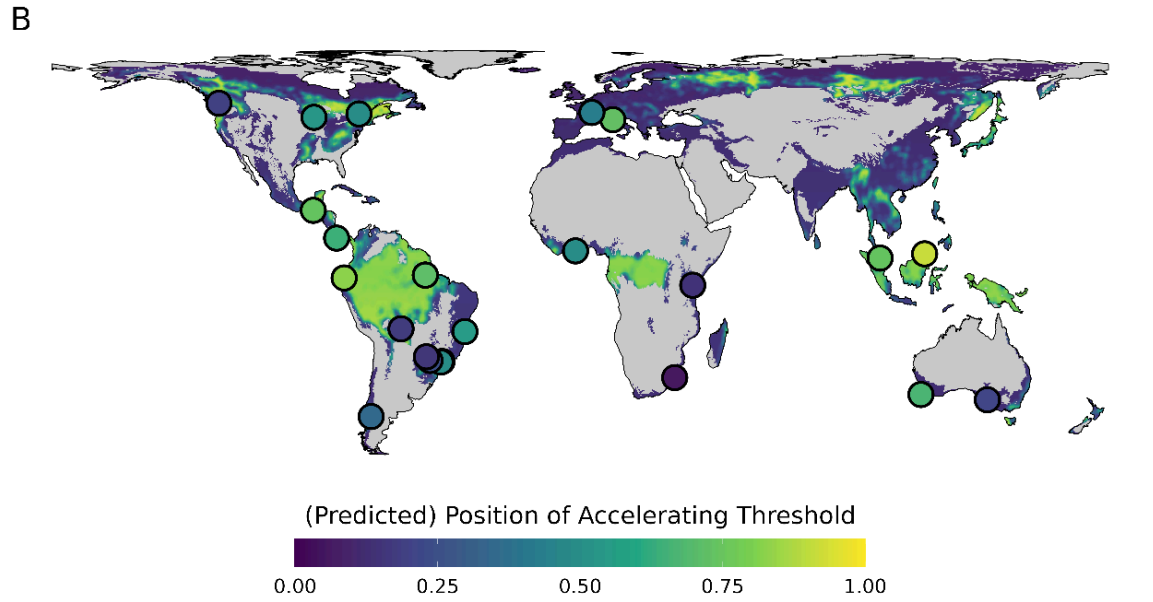


Figure 3.3. Decelerating Thresholds occur at higher percentages of forest cover than Accelerating Thresholds; points and accompanying lines are model predictions and 95% CIs for the landscape level forest cover (%) at which Accelerating and Decelerating thresholds occur. We suggest a theoretical framework from this (**A**), in which all communities

experience a Decelerating Threshold (Purple) and then an Accelerating Threshold (Yellow). We suggest that the community integrity of sites initially decreases rapidly with a reduction in forest cover compared to fully forested sites, before levelling out and remaining relatively stable, before once again decreasing rapidly as very low levels of forest cover are passed. Panels (**B**) and (**C**) are generalised additive mixed effect models fit to community integrity data from two studies in the dataset which show the proposed pattern. See methods for details of generalised additive mixed effect modelling.


A
Accelerating
Decelerating
Threshold position
1.00
0.75
0.50
0.25
0.00
0.25
0.50
0.75
1.00
Regional Deforestation
Latitude
0
20
40


B
(Predicted) Position of Accelerating Threshold
0.00
0.25
0.50
0.75
1.00

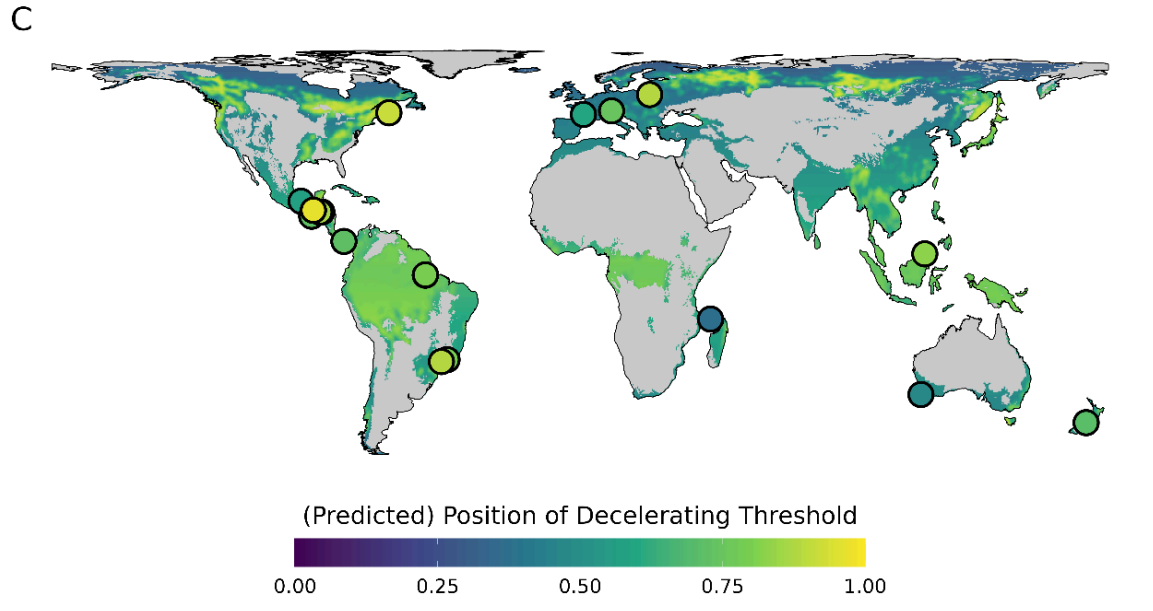

C
(Predicted) Position of Decelerating Threshold
0.00
0.25
0.50
0.75
1.00

Figure 3.4. The position of Accelerating Thresholds significantly increases as regional deforestation decreases. The position of Decelerating Thresholds is significantly less affected by regional deforestation than Accelerating Thresholds, and those close to the equator are universally high regardless of the level of regional deforestation. (**A**) Predictions from a generalised linear mixed effect beta regression using regional deforestation, latitude, threshold type, and their interactions to predict the placement of Accelerating and Decelerating Thresholds. Shaded areas indicate 95% confidence intervals, points are positions of threshold in the 53 studies with lines indicating the 95% confidence intervals of the thresholds. The colour of shaded areas and lines represents the latitude at which the predictions were made, whilst the points are coloured by the latitude the study was closest to. (**B**) Predictions from generalised linear mixed effect beta regression with threshold type set to Accelerating across the forest ecoregions of the world (see Map Predictions in methods), with pixel colour indicating the predicted position of Accelerating Thresholds. Points are positions of studies with Accelerating Thresholds and fill colour is the position of the Accelerating Threshold for that study. (**C**) Predictions from generalised linear mixed effect beta regression with threshold type set to Decelerating across the forest ecoregions of the world (see Map Predictions in methods), with pixel colour indicating the predicted position of Decelerating Threshold. Points are positions of studies with Decelerating Thresholds and fill colour is the position of the Decelerating Threshold for that study.

Our results suggest that any level of deforestation in any region will immediately cause changes in community integrity, despite differences in the latitude and regional disturbance of the area, since Decelerating Thresholds occur at almost universally high levels of forest cover. As such protection of fully forested areas is vital in all parts of the world to maintain the full diversity of biodiversity, and regions which are depauperate of such environments likely contain modified communities. While we found that deforestation of any kind causes changes to community integrity, our results concerning Accelerating Thresholds suggest that communities in undisturbed regions are particularly sensitive and that communities may

have passed the Decelerating and Accelerating Thresholds after only a 25% loss of forest cover (Figures 3.3A and 3.4B), suggesting limited scope for an “acceptable” amount of deforestation in such regions. Communities in disturbed regions which have likely lost most of their sensitive species (Betts *et al.* 2017; Pardini *et al.* 2010) were however surprisingly resistant to forest loss, maintaining a somewhat similar community of species to that of a fully forested area. In such regions, there may be a feasible trade-off between deforestation and biodiversity maintenance, as Accelerating Thresholds often only occurred after 75% deforestation or more (Figure 3.4A), at least if the aim is to maintain biodiversity close to that of fully forested areas in the region. Our findings are however likely a result of shifting baselines (Pardini *et al.* 2010; Tabarelli *et al.* 2010), where communities even in fully forested areas of highly disturbed regions are heavily modified by anthropogenic activities and are not representative of natural states.

Our results show that policies such as 30 by 30 are not enough to maintain the full scope of biodiversity (0 Decelerating Thresholds and 7 Accelerating Thresholds occurred below 30% forest cover), but may largely conserve biodiversity in highly disturbed regions. As such we suggest that the 30% of protected area proposed by 30 by 30 should not be evenly distributed, with disproportionately large amounts of protection focused in undisturbed regions which likely require close to 100% forest cover to maintain biodiversity. On the other hand, disturbed regions near the poles may well maintain large amounts of their current biodiversity with less than 30% total protected area. The enactment of such global biodiversity conservation policies is a promising step towards averting the upcoming biodiversity crisis, but we must ensure that such policies are not blindly implemented, and taking simple steps to tailor application to local conditions could optimise their effectiveness and ultimately avert unnecessary biodiversity loss.

Methods

Overview

We obtained community composition data from 97 studies focusing on the effect of habitat loss and fragmentation on Arthropod (N = 26), Bird (N = 41), Herpetofauna (N = 15), and Mammal (N = 15) communities, and paired this with forest cover data. We then fitted mixed effect piecewise regression models to the data to identify thresholds in landscape level forest cover, after which the relationship between forest cover and community integrity significantly changes. We then used regional level variables about the amount of deforestation that had occurred in the region, as well as the latitudinal position of the communities, to explain the position of the previously identified thresholds. Lastly, we explore the different types of thresholds identified, including how their position is affected by our explanatory variables.

Data

Data on site location and species occurrence were obtained from 93 independent studies in the BIOFRAG dataset (Pfeifer *et al.* 2014), as well as four independently sourced data sets (Orme *et al.* 2019), totalling 97 studies, including data on Arthropods (N = 26), Birds (N = 41), Herpetofauna (N = 15), and Mammals (N = 15). All studies recorded the occurrence of species at between 6 - 754 georeferenced locations (sites) (Table S3.3). All analyses were performed in R 4.0.3 (R Core Team 2023) and scripts are available on GitHub.

Forest Data

Global forest cover data were obtained by binarising tree cover data from Hansen et al. (2013); 30m wide raster pixels containing 70% tree cover or more were classed as forest, and all other pixels were classed as non-forest. 70% tree cover was chosen as a conservative categorisation of a forest, as classification of what constitutes a forest varies nationally (e.g 60% of tree cover is determined as forest in Brazil's mapBiomas, 20% in the

United States USGS National Land Cover Database, and 70% in the Copernicus Global Land Cover Dataset) (Buchhorn *et al.* 2020; Souza *et al.* 2020; Yang *et al.* 2018).

Site Level Metrics

Using the binarised forest cover data, the percentage of forest cover around every site in each study was calculated within circular buffers from 100m to 5000m radii, in 100m increments, which we consider to be landscape scales, using the landscapemetrics (Hesselbarth *et al.* 2019) package. Forest cover data was obtained from a variety of buffer sizes as it is expected that taxa, and species within taxa, will respond at different scales to environmental changes (Holland *et al.* 2004, 2005). Forest cover percentage was calculated around each site using binarised pixels as opposed to the raw tree cover percentage to differentiate between fragmented dense forest, and sparse tree cover (e.g. 20% tree cover could represent a savannah, while our measure of 20% forest cover represents a small area of dense forest). Pixels were projected to Behrmann Equal Area Projection at 30m pixel size so every landscape scale was equivalent in area across studies. Forest cover was also calculated for each site within a 50m radius which was considered the local scale.

We then categorised sites into forested and deforested individually for each landscape scale (buffer sizes), where forested sites at each landscape scale were those within 5% forest cover of the site with the highest forest cover at each landscape scale for each study, and had at least 75% forest cover at the local scale (50m). At least one scale from each study had at least one forested site, with the majority having forested sites at all scales (Table S3.4). Lastly, we calculated the distance between each forested site and all other sites by projecting their coordinates to Azimuthal Equal Distance Projection, and using st_distance from the sf (Pebesma 2018) package, the outputs of which were converted from metres to distance in kilometres by dividing by 1000.

Regional Variables

Using the same binarised forest cover data as above, we calculated forest cover at the study level as the percentage of forest pixels in a minimum convex hull around all sites in a study, including a 10km buffer (mean area = 4,120km^2, range = (389km^2, 17,312km^2)). We then subtracted this value from one to calculate the "regional deforestation" since all studies are present in forested ecoregions and are expected to have been almost completely covered by forests before anthropogenic impact. Using regional deforestation as opposed to forest cover also separates this regional variable from the landscape level forest cover that is calculated around each site. Lastly, we calculated the average absolute latitude (hereafter known as "latitude") of each study by taking the average absolute latitude of all sites within a study.

Community Integrity

Community integrity was calculated between the community composition at each forested site, and the community composition of every other site (including other forested sites), in each study at each landscape scale using the Sørensen dissimilarity metric, which was converted to similarity by subtracting the dissimilarity from one. Community integrity metrics were used as opposed to species richness as they capture changes in community composition due to species turnover, as well as species loss and gain. Sørensen dissimilarity was calculated using the vegan (Oksanen *et al.* 2020) packages 'vegdist' function.

Scale of Effect

We ran either a piecewise regression, or a mixed effect piecewise regression for each study at each landscape scale (100 - 5000m) using either the segmented or segmented.lme functions from the segmented package (Muggeo 2008; Muggeo *et al.* 2014) to identify the scale of effect at which landscape forest cover most strongly affected community integrity; totalling 50 piecewise models per study. We did this as different taxa, and even species within taxa are known to respond to changes in forest cover at different scales (Holland *et al.*

2004, 2005; Thornton & Fletcher Jr 2014). The response variable for each piecewise model was the arcsine square root transformed community integrity between deforested and forested sites, and the explanatory variables were the landscape forest cover of the deforested site at the previously identified scale of effect and the distance from the deforested site to the forested site. When a study included more than one forested site, we used mixed effect piecewise regression including random intercepts and thresholds per forested site to account for the nested structure of the data. In some circumstances, models including random terms for intercept and threshold failed to converge. In these cases, the models were simplified to include only random intercepts. All models included a single threshold in forest cover, after which the relationship between forest cover and community integrity changed. Lastly, we extracted the R^2 of all models for each study, and the model run at the landscape scale with the highest R^2 was designated as the scale of effect for landscape forest cover for that study.

Identifying Thresholds

We then ran three models (piecewise or mixed effect piecewise regression) for each study. The response, explanatory, and random variables used were the same as in the calculation of the scale of effect, with forest cover at the previously calculated scale of effect used. Once again each model allowed for a single threshold in the relationship between forest cover and community integrity, after which the relationship changed. The first model estimated the relationship between community integrity and forest cover (B_1) before the threshold (Ψ), and estimated the difference-in-slopes (B_2) after the threshold (Ψ) (equation 3.1). The second model assumed a null relationship between landscape forest cover ($B_1 = 0$) and community integrity before the threshold (Ψ), but estimated a difference-in-slopes (B_2) after the threshold (Ψ), and the third model estimated the relationship between landscape forest cover and community integrity before the threshold (Ψ), but assumed a null relationship between forest cover and community integrity after the threshold; this was done by inverting forest

cover in equation 3.2 as described in Muggeo (Muggeo 2008). We retained the piecewise regression model for each study with the lowest AIC for further analysis.

$$y = B_0 + B_1 x_1 \text{ if } x_1 < \Psi$$

$$y = B_0 + B_1 x_1 + B_2 (x_1 - \Psi) \text{ if } x_1 > \Psi$$

(eqn 3.1)

$$y = B_0 \text{ if } x_1 < \Psi$$

$$y = B_0 + B_2 (x_1 - \Psi) \text{ if } x_1 > \Psi$$

(eqn 3.2)

Where y is predicted community integrity between a forested and deforested site, B_0 is the intercept, B_1 is the slope of the relationship between forest cover (x_1) and similarity, B_2 is the difference-in-slopes before and after the threshold (Ψ). Note that terms for the distance between the forested and deforested site have not been included for simplicity.

We compared these piecewise models to linear models run using the same response, explanatory, and random effects using a method suggested in Arnold (2010). Namely, we only retained piecewise models if they had a lower AIC than their linear counterparts, and if the additional parameters were informative using 85% confidence intervals. We defined the additional parameter as the difference-in-slopes before and after the threshold, as it is not possible to test the threshold itself using confidence intervals.

Of the 97 studies, piecewise models converged for, and thus thresholds were obtained for 88 (91%). The scale of effect varied between 100m and 5000m, with a mean of 2600m (See Table S3.5 for R^2 of each model at each scale). Of the 88 studies for which piecewise

models ran, 80 had a lower AIC than their linear counterparts, and of those 80, 70 had a significant difference-in-slope parameter. These 70 piecewise regressions (Arthropods (N = 20), Birds (N = 30), Herpetofauna (N = 9), Mammals (N = 11)) were retained for use in further analysis (though see ‘Analysis Dataset’ for further filtering before final analysis) , specifically extracting the position of the threshold, the standard error around the threshold estimate, the slope of the relationship between forest cover and community integrity before the threshold, and the difference-in-slopes after the threshold.

Threshold Type

When visualising the 70 piecewise models, it was clear that studies could be broadly grouped into two categories: Accelerating Thresholds, in which the left slope is significantly positive (using the same 85% CI used in model selection), and the right slope is either zero, negative, or less positive than the left slope (37 studies); or a Decelerating Threshold, in which the right slope is significantly positive, and the left slope is either zero, negative, or less positive than the right slope (33 studies). Generally speaking, Accelerating Thresholds identified the forest cover after which the community composition of deforested sites started to quickly diverge from that of forested sites, whilst Decelerating Thresholds identified the forest cover after which community composition between deforested and forested sites slowed down or stopped changing. Plots of all studies and the threshold type associated with them can be seen in the supplementary material.

Two Thresholds

In addition to testing for a single threshold in forest cover using piecewise models, we also tested whether two thresholds in forest cover would improve the fit of the piecewise models. We ran a piecewise model per study, allowing for two thresholds in the relationship between forest cover and community integrity, with the forest cover used at the previously identified optimal scale of effect. These models were similarly compared using AIC to linear models, as well as against piecewise models with only a single threshold for forest cover. Only

studies with a single forested site could be tested for two thresholds, since it is not possible to test for two thresholds when using random effects in segmented.lme (Muggeo 2016, 2008). Due to this limitation, only 12 piecewise regression models with two thresholds could be run, and of these, three had lower AIC than their single threshold counterparts. This is most likely due to two main reasons. Firstly, the optimal scale of effect for forest cover was calculated based on piecewise regressions with a single threshold, thus biassing results towards finding a single threshold as opposed to two thresholds. Secondly, many studies did not have sites that covered the full gradient of deforestation, therefore likely "missing" either the Accelerating or Decelerating Threshold. As such these were not analysed further.

We instead visually explored the presence of two thresholds in the relationship between community composition and forest cover using generalised additive mixed effects model in the mgcv package (Wood 2011), since our ability to test for two thresholds using segmented.lme was limited by the lack of support for random effects. We used the arcsine square root transformed community integrity as our response variable, smoothing terms for forest cover and distance to the forested site as fixed effects, and random smooth terms for forested site, and the relationship between forested site and forest cover, and distance from the forested site. We arbitrarily chose to use forest cover at the 1000m scale for all studies, and present two examples of studies which followed our conceptual framework in Figure 3.3.

Analysis of Threshold Position

All analyses were carried out using the glmmTMB package (Brooks *et al.* 2017) and models were evaluated using the DHARMa package (Hartig 2022). Significance in all models was determined by 95% confidence intervals; where non-overlapping confidence intervals indicate statistical significance.

Analysis Dataset

We reduced the dataset used for all analyses to only include studies which had a forest cover range of at least 30% (site with highest forest cover % - site with lowest forest cover %

> 30%). This was done to ensure that thresholds were identified from an appropriately large range, and to limit the potential correlation between the forest cover range of a study and the position of the identified threshold e.g. if sites were only sampled between 10 - 20% forest cover a threshold could only be found in that range. We also removed a single study for which the piecewise regression model converged, but was clearly erroneous when plotted (Figure S3.1), and when the model was further inspected it was found to have an R^2 of 1 indicating issues with fitting. This left 53 thresholds for the analysis. Plots for these 53 piecewise regression models can be seen in supplementary material S3.2, and all analysis was also carried out on the full dataset without any filtering for forest range, and another data which was more heavily filtered for forest range (see Tables S3.1 and S3.2).

Variation in Threshold Placement

We then assessed the relationship between our explanatory variables and the position of the thresholds using mixed effect beta regression models. Beta regression was used as the position of thresholds was bounded between 0 and 1. The position of the threshold was the response variable, and regional deforestation and latitude were used as explanatory variables, including an interaction term between them. In addition, we included a random effect for taxa, since we did not have enough data points to examine differences between taxa directly, but wanted to account for variability. Lastly, we included the log of the inverse of the standard error of the threshold from the piecewise regressions as a weighting term to propagate through the error from the piecewise regressions, and more heavily weight the influence of threshold positions for which we had more certainty. The inclusion of fixed effects was determined by AIC, whilst random effects and weightings were chosen based on data structure and study design.

Threshold Type

We ran an additional mixed-effect beta regression model to assess how the different threshold types (Accelerating or Decelerating) were affected by our explanatory variables.

This model included regional deforestation, latitude, and threshold type as fixed effects, including interaction terms between all variables, the position of the threshold as the response variable, taxa as random intercept, and additionally included the same weighting term for threshold position as used in the Variation in Threshold Placement section. The inclusion of fixed effects was determined by AIC, whilst random effects and weightings were chosen based on data structure and study design.

Lastly we ran a generalised linear mixed effect model with binomial errors to assess the position of Accelerating and Decelerating Thresholds. This model used threshold position as the response variable, and threshold type (Accelerating or Decelerating) as the explanatory variable, including a random intercept for taxa.

Map Predictions

We predicted the position of Accelerating and Decelerating Thresholds across the world using the mixed effect beta regression model developed above. We also predicted the probability of finding an Accelerating Threshold (as opposed to a Decelerating Threshold) using the generalised linear mixed effect model with binomial errors explained above. We did this for both models on a global map of 96x96 km raster pixels, converted to Behrmann Equal Area projection so the area of the raster pixels was the same across latitudes. We calculated the regional deforestation of each pixel as a proportion of 30m forest pixels within each 96x96km pixel, using the Hansen (Hansen *et al.* 2013) dataset, converted to Behrmann Equal Area projection, and the same 70% forest versus non-forest classification detailed in the Forest Data section. We also calculated the absolute latitude of the centre of each pixel, and then used this calculated regional deforestation and latitude to predict the position of Accelerating and Decelerating Thresholds in each pixel, as well as the probability of finding an Accelerating Threshold. Lastly, the outputs were cropped to only include forest ecoregions as defined in the RESOLVE 2017 dataset (Dinerstein *et al.* 2017), since our results are specific to forested areas, and are not necessarily applicable to other ecoregions.

Chapter 4: Deforestation Alters Species Interactions

This chapter has been published in the journal Natural Sciences and is reproduced here with the formatting and journal requirements under a Creative Commons licence.

Abstract

Interspecific interactions are a major determinant of stability in ecological communities, and are known to vary with biotic and abiotic conditions. Deforestation is the primary driver of the ongoing sixth mass extinction, yet its effect on species interactions remains largely unexplored. We investigate how deforestation affects species interactions using a complex systems model and a co-occurrence dataset of 363 bird species from five regions across the Brazilian Atlantic Forest totalling 27,226 interactions. Both theoretical and empirical results show that interspecific interactions vary non-monotonically with forest cover, and are more positive than average in areas with higher forest cover, and to a lesser extent in highly deforested areas. Observed differences in interactions reflect both species turnover and changes in pairwise interactions. Our results point to changes in stability across the gradient of deforestation that may lead to varying community resilience to environmental perturbations.

Key Interdisciplinary Aspects

- Species interactions are expected to vary due to the biological, chemical, and physical changes caused by deforestation on their local environment
- We utilise empirical data as well as a mathematical complex systems approach to show that species interactions are more positive in highly forested areas

- We propose that alteration of species interactions caused by deforestation will affect the stability of communities and their resilience to future perturbations (e.g climate change)

Introduction

Species and ecological communities are experiencing high levels of environmental change and subsequent population decline as a result of anthropogenic activities (IPBES 2019; Malhi *et al.* 2008; Paiva *et al.* 2019; Tabor *et al.* 2018). Although many studies have observed a decline in biodiversity associated with anthropogenic disturbances, such as habitat loss and fragmentation, its effects on species interactions remain poorly understood (although see Magrach *et al.* 2014). Interspecific interactions are a key component determining the assembly and richness of ecological communities, ultimately ensuring their stability (Case *et al.* 2005; Qian & Akçay 2020). Changes to the magnitude and sign (i.e., positive or negative) of interactions among organisms have been shown to destabilise communities, potentially leading to shifts in species abundance and composition (Qian & Akçay 2020; Ratzke *et al.* 2019; Thebault & Fontaine 2010). As such, identifying if and how interspecific interactions respond to environmental change is vital when considering the long-term stability of ecological communities and their associated populations in this rapidly changing world.

Ecologists have long studied how interactions vary across gradients, specifically abiotic gradients such as altitude, temperature, and soil composition (Callaway *et al.* 2002; Chamberlain *et al.* 2014; Daskin & Alford 2012). Most of this literature has focused on sessile organisms (He *et al.* 2013; He & Bertness 2014), finding that interactions become more positive (facilitation and mutualism) as environments become more stressful (Bertness & Callaway 1994; Callaway *et al.* 2002; He *et al.* 2013). This phenomenon known as the stress gradient hypothesis (Callaway 2007), is well studied and generally accepted in sessile organisms, however whether it also applies to mobile organisms is less clear (Callaway

2007; He & Bertness 2014). Interactions between species have also been found to vary temporally, spatially, and with the presence, absence and density of other species, though the sign and magnitude of this variation is unclear (Billick & Tonkel 2003; Chamberlain *et al.* 2014; Pruitt & Ferrari 2011; Schädler *et al.* 2007; Sunahara & Mogi 2002; Thompson & Cunningham 2002).

The interactions exerted upon an individual, population, or species can change in two ways: through the addition or removal of interactions, or via shifts in existing interactions (Chamberlain *et al.* 2014; Kofron & Chapman 1995). The addition or removal of interactions is a relatively simple concept and most commonly occurs due to species turnover when species join or disappear from a community (Kofron & Chapman 1995). Shifts in existing interactions between pairs of species can be more complex and difficult to study, but there are examples describing these changes in the literature (Callaway *et al.* 2002; Chamberlain *et al.* 2014; Pruitt & Ferrari 2011). For instance, some shifts in interspecific and intraspecific interactions result from changes in population density, and generally interactions become more negative (competition) as population density increases (Fox & Morin 2001; Mönkkönen *et al.* 2004).

Habitat loss is one of the most significant drivers of biodiversity loss (Giam 2017; IPBES 2019). Despite this, research into its effects on interactions between species has been surprisingly scarce (although see Magrach *et al.* 2014). A number of factors already known to shift species interactions are associated with deforestation, including abiotic changes such as reduced humidity and increased soil erosion (Gandu *et al.* 2004; Zheng 2006); changes in quantity and type of resources (Benstead & Pringle 2004); and species turnover (Kofron & Chapman 1995). However, these factors can influence species interactions in different directions, and thus, the overall effect of deforestation on species interactions remains unclear. For forest-dwelling species deforestation reduces the resources available (Thuiller *et al.* 2006) and causes a ‘crowding effect’ (Hagan *et al.* 1996) both of which likely increase competition resulting in more negative interactions between species (Benstead & Pringle

2004; Mönkkönen *et al.* 2004; Thuiller *et al.* 2006). Contrastingly, abiotic changes associated with deforestation such as increased humidity and increased frequency of extreme weather events (Gandu *et al.* 2004; Wilson *et al.* 2005) can create a stressful environment for forest-dwelling species, and as described above, stressful conditions have been linked to more positive interactions. Deforestation can also lead to changes in species composition due to removal of shelter, nesting sites, and foraging substrate (Ferger *et al.* 2014; Jankowski *et al.* 2013; Tews *et al.* 2004) that will affect interactions via the addition and removal of species as well as changing the interaction types and strengths among species (Schädler *et al.* 2007).

Here we look to assess the effect of deforestation on interspecific interactions between species using an empirical dataset and a Complex systems modelling approach. We expect interactions to vary both as a result of the loss and gain of interactions via community turnover, and interaction changes between existing pairs of species likely driven by behavioural shifts caused by alterations in resource availability which we expect to cause more negative interactions in heavily deforested areas. We chose to use a joint empirical and modelling approach to increase reliability of results as the accuracy of methods used to infer interactions from empirical datasets are relatively unknown. Additionally, models corroborating results of empirical analysis suggest that findings are generalisable and not specific to the dataset used.

Methods

Tangled Nature Model

The modified Tangled Nature Model (hereafter known as “TaNa model”) used in this study is similar to that described by Brinck (2018), in that it is a short-term spatially explicit version of the original TaNa developed by Christensen *et al.* (2002). The TaNa model integrates a complex systems approach on ecological communities, whereby the properties of the community at the macroscopic scale emerge from the interactions and movement of

individuals at the microscopic scale. The dynamics of the TaNa model used in this study are in many ways similar to that of a neutral model, in that they are individual based and stochastic. Though unlike a neutral model the survival of individuals and therefore populations is highly influenced by interactions with cohabitants and the environment. TaNa models have been shown to accurately recreate realistic ecological results including species area relationships, trophic networks, and species abundance distributions (Brinck & Jensen 2017; Christensen *et al.* 2002; Laird *et al.* 2008). An overview of the TaNa model and detailed descriptions of the modification implemented in our version are given below.

Overview of TaNa model

Landscape

A landscape of 75 equally sized grid cells (15 x 5) was used, in which individuals were able to move up to a maximum dispersal distance of two cells at each time step. This size of landscape was used to ensure there were enough cells to create heterogeneous landscapes, whilst being computationally viable. To start, all cells were created as forest (100% forest cover), but at time $t0$we simulated deforestation by replacing a predetermined number of forest cells with deforested cells (0 - 90% forest cover).

Habitat type (forest or deforested) influenced the probability of reproduction differently depending on an individual's specialism value (see equation 4.4).

Species Traits

TaNa model set up includes a community of species, with each species differing in two distinct traits: specialism, and interactions with other species in the community. In our study we generated communities of 50 species for each run, as this was comparable to the number of species found per site in the empirical studies. Specialism is defined a priori by a parameter with a value between 0 and 1, where 0 represents a habitat generalist species and 1 represents a complete habitat specialist species (in our case, a forest specialist). The

specifics of how a species' specialism value affects its chance of reproduction can be seen in equation 4.4, but essentially a species with a high specialism value will have a higher fitness in a forested area than one with a low specialism value (all else being equal). Opposedly, the fitness of the same high specialism species will be more heavily reduced in a matrix area than a comparable low specialism species.

Interactions between species are defined a priori and stored in a matrix J of dimensions $M\ x\ M$where $M$is the total number of species Sin the species pool. Every species interacts with every other species, via the interaction link $J(S^{\alpha}, S^{\beta})$ with reciprocal interactions between species equal $J(S^{\alpha},\ S^{\beta})\ =\ J(S^{\beta},\ S^{\alpha})$ to most closely resemble interactions on a single trophic level. Interactions between species are randomly drawn from a normal distribution constrained between -1 and +1 with a mean of 0 to match the findings of Wootton & Stouffer (2016), that weak interactions are common, and strong interactions are rare. Full details of how interactions affect the probability of an individual reproducing can be found in equations 4.1, 4.2, and 4.3, but for simplicity, positive interactions increase the fitness of an individual, whilst negative interactions decrease the fitness.

The strength of interactions acting upon an individual at time t is given by:

$$J_f\ (S^{\alpha},\ t)\ =\ \sum_{S \in S} \frac{J(S^{\alpha}, S^{x}) n_f(S^{x}, t)}{N_f(t)} \qquad \text{(eqn 4.1)}$$

Where $J(S^{\alpha},\ S)$ is the interaction link between S^{α} and species S^{x} which represents any species, $n_f(S^{x},\ t)$ is the number of individuals of species S^{x}in grid cell f at time t and $N_f(t)$ is the total number of individuals in cell f at time t.

Model Dynamics

TaNa model simulated four dynamic processes: death, migration, immigration, and reproduction. At each time step, these dynamics are run sequentially cell by cell, starting from cell one, completing 50 iterations of each dynamic per cell.

Description of each dynamic in each cell and iteration:

I. Death: an individual is randomly chosen and removed from the cell with a constant probability of 0.15; a value used in previous TaNa studies and shown to have minimal impact on the model outcome (Arthur *et al.* 2017; Hall *et al.* 2002)

II. Migration: an individual is randomly chosen from the cell and moved to another cell up to its maximum dispersal distance (maximum dispersal distance of two). The probability that an individual will migrate is defined by a density dependent function, whereby the probability of migration increases as the population density of a cell increases, for full details see Brinck (2018).

III. Immigration: an individual selected randomly from the initial pool of 50 species is introduced to the cell with a probability of 0.05 before deforestation, and 0.001 after deforestation. The high level of immigration before deforestation ensured a good initial distribution of species through the landscape, whilst the lower level of immigration after deforestation was chosen to represent less frequent immigration in poorer habitat areas (Püttker *et al.* 2011).

IV. Reproduction: an individual is randomly chosen and reproduces with probability P_{off}, which is defined as follows:

$$P_{off}(S^{\alpha},\ t,\ f)\ =\ \frac{exp(\Xi(S^{\alpha},t,f))}{1+exp(\Xi(S^{\alpha},t,f))}\ \in\ (0,1) \qquad \text{(eqn 4.2)}$$

Where the reproductive capability Ξ is given by:

$$\Xi(S^{\alpha},\ t,\ f)\ =\omega_{J}J_{f}(S^{\alpha},\ t)\ +\ \mu\ -\ \frac{N_{f}(t)}{R_{f}} \qquad \text{(eqn 4.3)}$$

Where R_f is the carrying capacity of cell f (set to 10), ω_J is the weighting of interactions (set to 10), and the habitat effectμis given by:

$$\mu(S^{\alpha},\ t,\ f)\ =\pm\ \omega_{s}C(S^{\alpha}) \qquad \text{(eqn 4.4)}$$

Where μis positive when species S^{α} is in a forest cell f, and negative when S^{α} is in a deforested cell f, C is the specialism value of S^{α}, and ω_s is the weighting of specialism (set to 4). Cells deforest at the predetermined time $t0$so the sign of μis dependent on t.

Model Setup and Output

We defined 80 different communities, each formed by 50 species varying in their specialism and interactions. At model setup, every cell was populated by 100 individuals of randomly assorted species, the initial population size was chosen for computational efficiency and has no bearing on the final community. For each community 11 model scenarios were run to represent varying degrees of deforestation (occurring at time $t0$) and ranging from 0 to 100% at 10% increments. In total we ran 880 models. For comparison with an empirical dataset the 80 communities could be considered as broad study regions with overlapping species pools with the 11 landscape scenarios representing discrete sampling units (Figure 4.1).

All models were run for 2500 time steps, with deforestation after 500 steps ($t0 = 500$). We evaluated the following model output: the occurrence of each species in each cell at t = 2500 and for all species with at least 10 individuals present in the landscape we extracted interactions from the matrix J.

Empirical Data

Data Collection

We used bird occurrence data from five independent studies conducted in different regions of the Brazilian Atlantic Forest (Table S4.1, Figure 4.1). The Brazilian Atlantic Forest is characterised by heterogenous abiotic conditions, and elevation, encompassing tropical and subtropical climatic conditions, annual rainfall ranging from 1000mm - 4000m and elevation changes from 0 - 2900m. The vegetation structure is also varied, including areas of tropical rainforest, as well as deciduous, *Araucaria* pine, and *Lauraceae* dominated forests (Tabarelli *et al.* 2005). Each study provided point counts of bird species in multiple forest patches

which differed in size, each of which we refer to as a site (range 14 to 36 sites per study). Sites varied in the amount of surrounding forest cover, as well as the land use of the matrix which they were embedded in; pasture, agriculture and silviculture. Each study differed in sampling effort and observers but these variables were controlled for within the studies. In total, occurrence data on 363 bird species was gathered across 134 sites (Table S4.1).

The percentage forest cover of each site calculated at a 600m radius was taken from Orme *et al.* (2019), each of these 600m radius areas around sites can be thought of as equivalent to the full 75 cell landscape of the TaNa. Forest cover was calculated from the most recent high definition map available for each site, and a mixture of Quickbird, WorldView, SOS Mata Atlântica, and Instituto Florestal were used. In each of these maps forest cover included primary, secondary and semi-natural forest types. A buffer size of 600m was used as it captures the relatively restricted movement of most species in the studies (Hatfield *et al.* 2018a); better reflects the sampling design of the original studies; and is similar to that used in other studies (Banks-Leite *et al.* 2014; Boscolo & Metzger 2009).

Defining empirical interactions

Six methods for inferring interactions from co-occurrence data were used to estimate pairwise interactions from the point count data from the five empirical datasets from the Brazilian Atlantic Forest: correlation, partial correlation, GLMs, graphical lasso on log transformed counts, Gaussian copula graphical models, and Gaussian copula graphical models with forest cover as an environmental variable. All methods were run in R 4.0.3 (R Core Team 2020) following the methodology used in Harris (2016) for correlation, partial correlation, and GLMs; and the methodology used in Popovic *et al.* (2019) for graphical lasso on log transformed counts and Gaussian copula graphical models.

Specifically correlations were estimated using the cor function in base R which finds the sample correlation between species presence-absence matrices, which summarises their marginal association. Partial correlation using the corpcor package (Schafer *et al.* 2017); this

method is closely related to linear regression and summarises species conditional relationships. The GLM method includes the use of regularised logistic regression models which were run in base R, where the presence of a species is predicted using the presence or absence of the other species. This method produces two near identical interaction estimates which were averaged to produce a single interaction term between species pairs. Graphical lasso on log transformed counts used the glasso package (Friedman *et al.* 2019) which estimates a sparse inverse covariance matrix using a lasso penalty; fo full details see (Morueta-Holme *et al.* 2016). Lastly, the ecoCopula approach was carried out via the ecoCopula package (Popovic *et al.* 2019) and uses Gaussian copula graphic models, which look at direct and indirect associations in data (in our case species); see Popovic *et al.* (2019) for full details.

These methods were chosen because they either performed well in previous studies (Harris 2016; Popovic *et al.* 2019), or in our own performance tests (Table S4.2). Other methods such as Markov networks (Harris 2016) and Markov random fields (Clark *et al.* 2018) could not be tested due to data limitations; our data contained too many species for Markov networks to be computationally viable, and many rare or extremely common (but not ubiquitous) species which caused errors when running Markov random fields.

All methods estimate reciprocal pairwise interactions meaning they cannot accurately represent interactions between trophic levels (e.g. predation +/-), but outputs were on different scales, and most were found to differ significantly when analysed using a GLM and post-hoc comparison in emmeans (Lenth 2020) (Figure S4.1). However, all showed significant positive correlations (mean R^2 = 0.69, SD R^2 = 0.18) indicating that methods generally identified the same interactions as more or less positive. Thus, for comparison all estimates were re-scaled between -1/+1 using the largest absolute value for each method across all studies as the upper and lower limit. Then, we calculated the mean interaction across methods for each pair of species in each study.

Data preparation

Empirical and TaNa Model Data

Empirical data and TaNa model outputs both describe pairwise interactions in different sampling units (studies/landscapes), for which the forest cover of sites and landscapes is known or defined and both were prepared for analyses following the steps below.

Centering Forest Cover and Interactions

We are interested in assessing relative interaction changes as a result of the forest cover a species is found in. However, species naturally occur in different levels of forest cover, meaning that a study with 70% forest cover could be a “deforested” area for a species found primarily in fully covered areas, but a forested area for a species found primarily in highly deforested areas. As such, in order to compare shifts in interactions with forest cover across species with varying forest cover requirement, we centered forest cover for each species, which gave us a relative measure of whether the species was in a higher or lower forest cover area compared to its average (hereafter references to “higher” or “lower” forest cover or interactions are relative to the species average).

In the empirical dataset centered forest cover was calculated by firstly finding the arithmetic mean forest cover a species was found in across all sites of all studies. Then for each study the mean forest cover a species was found in was calculated, and the previously calculated overall forest cover mean for that species was subtracted from the study mean. In the TaNa model data, species are not comparable across communities, so all centering was done within communities. As such, centered forest cover was calculated by taking the mean forest cover across all landscapes a species occurred in, and then subtracting that from the forest cover of each individual landscape a species occurred in. Similarly, interactions were also centered for both the empirical and TaNa model data using the same methods but calculating the mean interaction a species was involved in across all studies or landscapes, and then subtracting that from each individual interaction a species was involved in. In the

empirical dataset Interactions between the same pairs of species were also centered by calculating the mean interaction that the specific pair of species were involved in and subtracting that value from each individual interaction between that pair of species. After doing this, for each pair of species interacting, we had a Centered Forest Cover for species 1 and for species 2 (where species 1 was the focal species in the analyses), Centered Interaction, and Centered Pair Interaction.

Phylogenetic filtering for the empirical data

The empirical dataset produced 121,996 pairs of interactions based on co-occurrence, but we focused our analyses on intra-trophic interactions due to the limitations of reciprocal inferred interactions by the co-occurrence methods, and to those species more likely to represent significant ecological interactions. We defined these as interactions between species with more shared evolutionary history, which has been found to closely correlate with shared traits and ecological requirements (Tucker *et al.* 2018). To define shared history we created a phylogenetic tree representing the 363 unique bird species in the dataset by trimming the open tree of life (Michonneau *et al.* 2016; Rees & Cranston 2017) with the R package ape (Paradis & Schliep 2019) and adding distance using the “compute.brlen” function and Grafen method. The distance between nodes of species were examined, and a consistent cut-off distance was chosen based on visual inspection of distances between known similar species. For example, only interactions between the predatory species *Accipiter striatus* and relatively similar predatory species *Harpagus diodon* and *Ictinia plumbea* were retained. Whilst interactions between the dove species *Leptotila rufaxilla* and all other dove and pigeon species (e.g. *Columbina talpacoti*, *Claravis pretiosa*, and *Leptotila verreauxi*) remained. This identified 8,842 closely related pairs of species, composed of 313 unique species from which our dataset described 27,226 pairs of interactions.

We also carried out our analysis without phylogenetic filtering and found qualitatively identical results (Table S4).

Statistical Analysis

We used linear mixed effect models for analysis of the empirical and TaNa model data. A first model analysed all data with Centered Interaction as the response variable and Centered Forest Cover Species 1, Centered Forest Cover Species 2 and their interaction as explanatory variables. A second model analysed only data from pairs of species recorded in multiple studies predicting the response variable Centered Pair Interaction by the predictors Centered Forest Cover Species 1, Centered Forest Cover Species 2 and their interaction. The random effect structure was chosen for each linear mixed model using the methods outlined in Zuur (2009) with all models including random slopes for species 1 and species 2 (equations for full models can be seen in the supplementary). All statistical analysis was carried out in R 4.0.3 (R Core Team 2020) using the lme4 package (Bates *et al.* 2015).

Method Framework

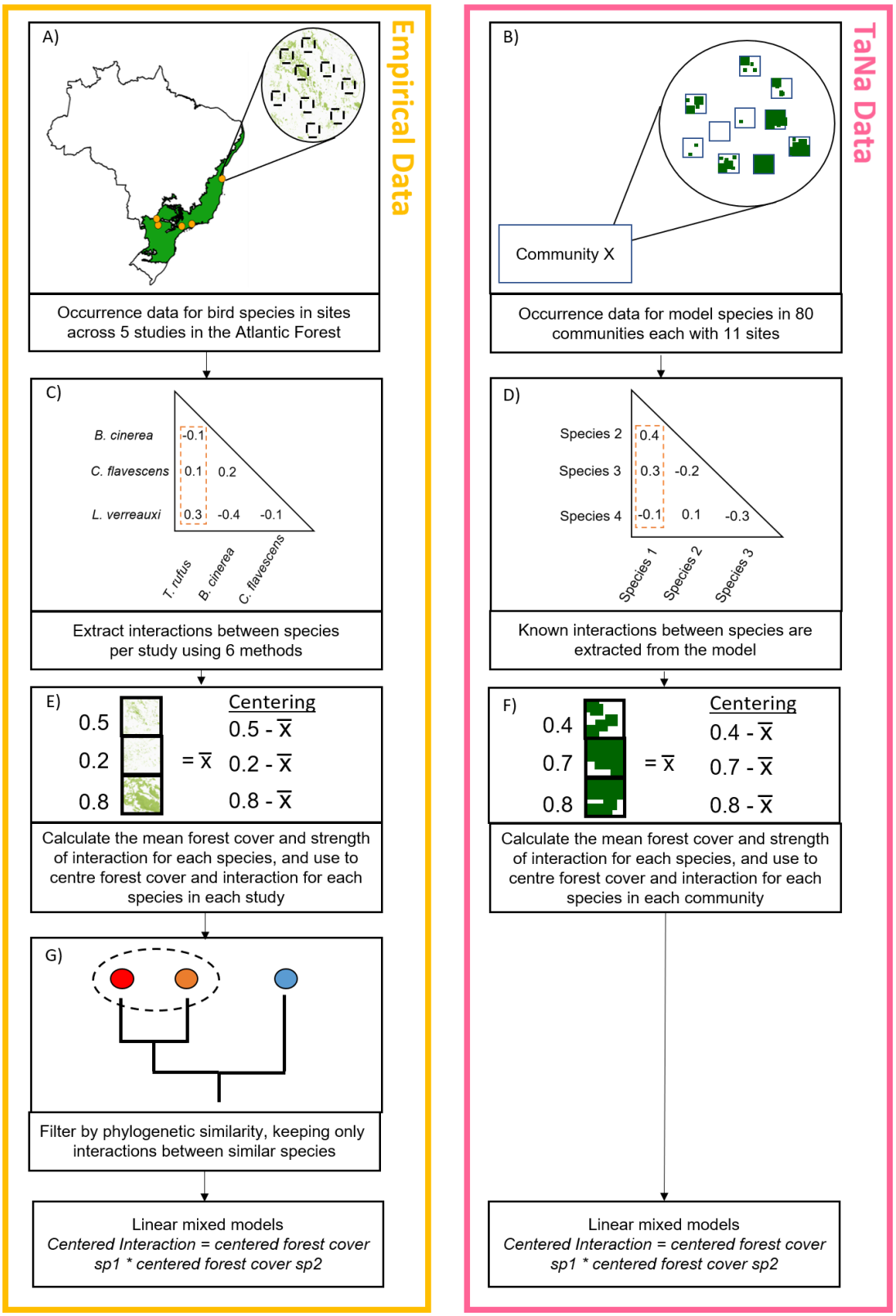


Figure 4.1. Methodological framework used. (A) Locations of the 5 studies from the Atlantic Forest. (B) Visualisation of the 11 landscape scenarios run in TaNa models for each community. (C and D) Triangular matrix of example interactions from the empirical and TaNa model data. (E and F) Calculate the mean interaction and forest cover for a species in the empirical and TaNa model data, then use this value to center the forest cover and interaction for each species and species pair by subtracting the mean (figure shows forest cover as an example). (G) Filtering of interactions to keep phylogenetically similar interactions (red and orange circle), and removing phylogenetically dissimilar interactions (blue circle).

Results

Analysis of empirical and TaNa model data

Both TaNa model and empirical data showed more positive interactions than average in landscapes and sites with higher forest cover, with a significant interaction term showing synergetic benefits when both species were in areas with higher than average forest cover (Table 4.1, Figure 4.2). Similarly, in both the TaNa model and empirical data, interactions were also more positive than average in areas with lower forest cover, a pattern consistent with the stress gradient hypothesis. This was particularly evident for TaNa model data where the predicted change to interactions was almost equal at comparably higher and lower forest cover.

Table 4.1. Coefficient estimates (and standard errors) obtained from linear mixed effect models predicting changes in species interactions as a function of relative forest cover. Results marked with a * are significant to $P < 0.05$.

Data	Response Variable	Centered Forest Cover Species 1	Centered Forest Cover Species 2	Interaction term
Empirical	Centered Interaction	0.118* (0.017)	0.109* (0.017)	0.935* (0.075)
TaNa model	Centered Interaction	0.088* (0.030)	-0.075* (0.029)	0.829* (0.046)

Empirical	Centered Pair Interaction	0.057* (0.018)	0.089* (0.017)	0.305* (0.061)

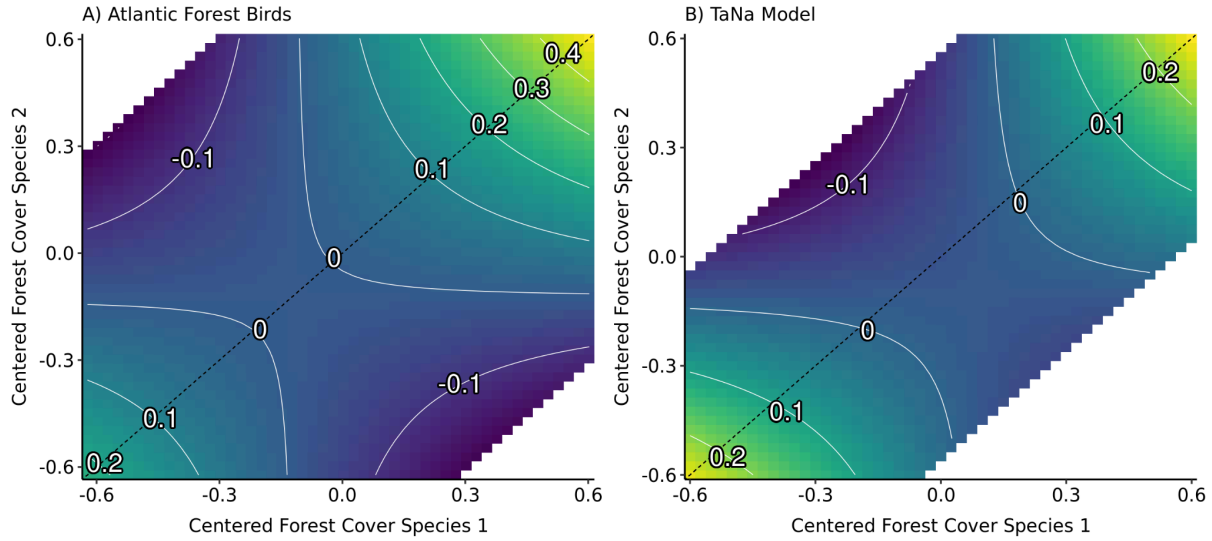


Figure 4.2. Predicted changes in species interaction values (colour gradients reflect predicted Centered Interaction) as a function of relative local forest cover for interacting species (Centered Forest Cover Species 1 and Centered Forest Cover Species 2) in the empirical (A) and complex system TaNa model (B) datasets. The lowest interaction values occur when forest cover is above average for one species but below average for the other species. Average forest cover varies considerably among species (minimum = 0.06, maximum = 0.99). Regression estimates in Table 4.1.

Changes in TaNa data were solely caused by species turnover because in this model pairwise interaction values were defined a priori and could not change over space or time. However, pairwise interaction values in the empirical dataset can also shift within a given species pair, such that two species would have a positive interaction in one region and a negative in another. To evaluate the potential role of interaction shifts within pairs of species across the gradient of deforestation we analysed only empirical data from pairs of species that appeared in at least two studies together (180 species, 8,179 interactions) and found

consistent results with more significant positive interactions when both species occurred in higher forest cover areas (Table 4.1, Figure 4.3).

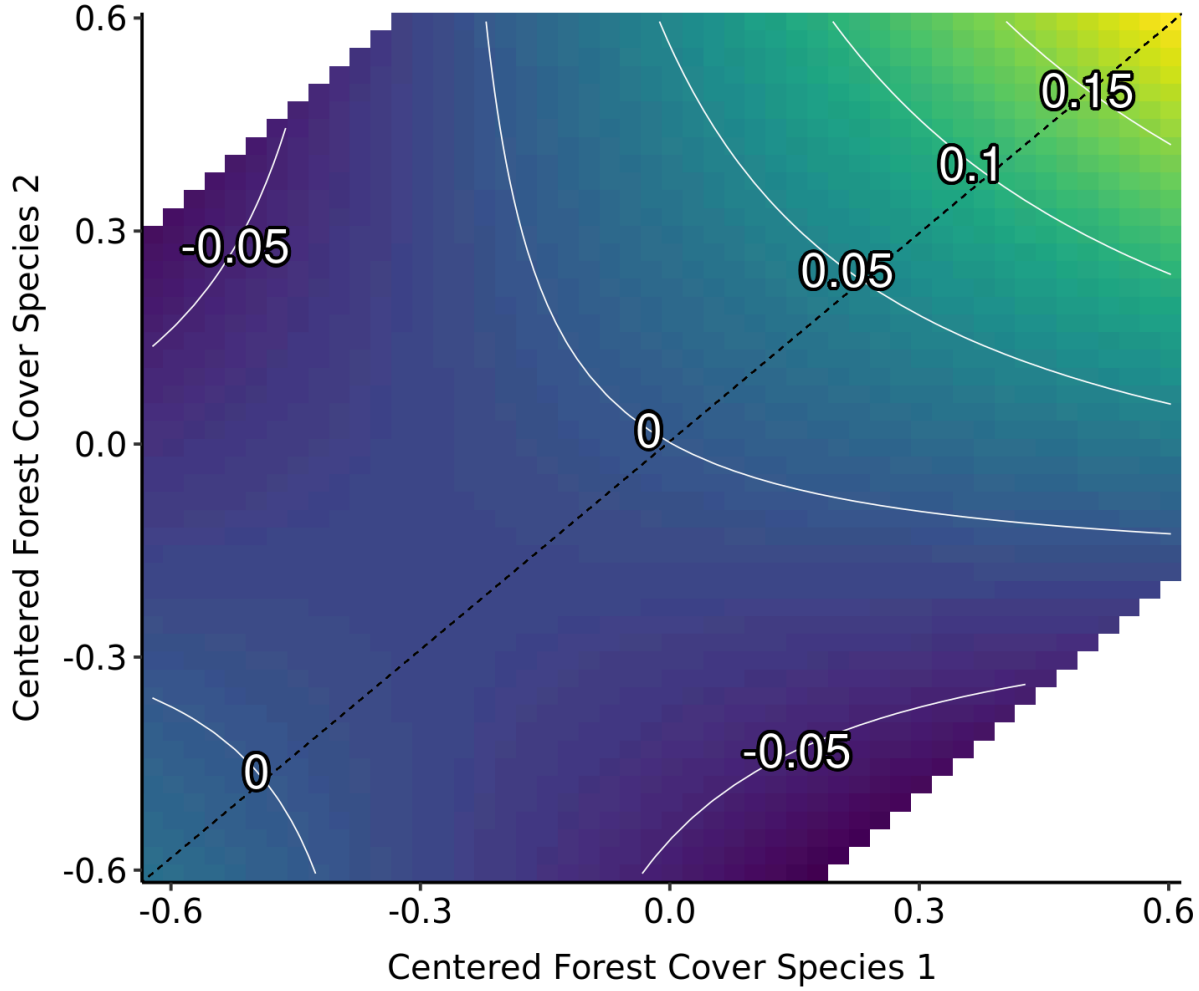


Figure 4.3. Predicted changes in pairwise interaction values (colour gradient reflects values of predicted Centered Pair Interaction) as a function of relative local forest cover for interacting species (Centered Forest Cover Species 1 and Species 2). Regression estimates in Table 4.1.

Discussion

Our results show that species interactions vary non-monotonically between ecological communities distributed along a gradient of forest cover. Both the empirical and TaNa model results showed more positive interactions than average occur when all species are present

in areas of higher forest cover. Interactions were also found to some extent to become more positive than average when species were at lower levels of forest cover. We found that these changes in interactions occurred both due to species turnover and shifts in interactions within pairs of species, which likely arises from changes in competition and behaviour. We suggest that ecological communities occurring in areas of higher forest cover are more stable, as they are predicted to have a larger proportion of facilitative interactions, as well as weaker antagonistic interactions. This increased stability will likely mean elevated resilience to ecological perturbations such as invasive species, removal of species, or changes to the abiotic environment (Allesina & Tang 2012; Landi *et al.* 2018).

Deforestation was found to alter species interactions via both changes in community composition as well as shifts in interaction strength within pairs of species, as shown in the analysis of the subset of empirical data. Shifts towards more positive interactions in highly forested areas may have occurred due to reduced competition for resources, especially amongst forest-specialists, both due to increased resource density, and reduced population density (Mönkkönen *et al.* 2004). Interaction shifts could also have occurred via benefits from facilitative behaviours such as mixed species flocks, which are expected to occur more in highly forested areas, in which individuals from different species gain fitness by increased foraging efficiency (e.g disturbing prey species) and decreased mortality from predators (e.g alarm calling) (Rubenstein *et al.* 2008; Sridhar *et al.* 2009).

Competition and flock forming will also be influenced by changes to species composition from turnover. Disturbance-adapted species may appear in deforested areas and outcompete forest specialists, and prevalence of behaviours such as mixed species flocks will depend on the presence of forest specialists, which are found in higher proportions and abundances, conferring functional redundancy, in areas of higher forest cover (Anjos *et al.* 2019; Hatfield *et al.* 2020; Kang *et al.* 2015; Rand & Tscharntke 2007; Rutt *et al.* 2019). Indeed, certain forest specialist species in the dataset analysed such as Habia rubica and Thamnomanes caesius are known to be vital to the formation and maintenance of

mixed-species flocks (Knowlton *et al.* 2017; Pagani-Núñez *et al.* 2018), and their absence from lower forested areas has previously led to the disintegration of such behaviours (Stouffer & Bierregaard 1995).

Species interactions became more positive at higher forest cover levels across all models. Recent studies on the stability of ecological communities have highlighted the importance of facilitative interactions, finding that network stability increases with higher proportions of facilitative interactions (Lurgi *et al.* 2016; Qian & Akçay 2020). Similarly, a decrease in the strength of negative interactions has been found to stabilise communities under most conditions (Loreau & de Mazancourt 2013). For example, Lurgi *et al.* (2016) found that food webs with a high proportion of mutualistic interactions between plants and animals were more stable than their antagonistic heavy counterparts, whilst Qian & Akçay (2020) showed that communities with more mutualistic interactions are more resistant to invasion. As such, it is expected that communities composed of species in areas of higher forest, which we found to have more positive interactions, will be more stable.

On the other hand, we found that more negative interactions occur when one species in the interacting pair is in higher levels of forest cover, and the other species in lower levels of forest cover. This may occur when the interacting pair is composed of species with contrasting forest cover requirements. This type of community is expected to exist in intermediately deforested areas, where disturbance adapted species have colonised, but many forest specialist species still persist due to previous populations (Stouffer *et al.* 2011). Based on studies by Loreau & de Mazancourt (2013) and Lurgi *et al* (2016) we would expect this shift towards more negative interactions in these communities to decrease their stability and resilience to further environmental change. Though note that some studies have found stabilising effects of antagonistic interactions, though these are usually focused on food webs as opposed to intra-trophic interactions (Coyte *et al.* 2015; Rooney *et al.* 2006; Stouffer & Bascompte 2011).

The results of the TaNa model dataset, and to a lesser extent, the full empirical data set suggest an increase in positive interactions at lower forest cover, however this was very weak in the model analysing interaction changes between pairs of species that occurred together in multiple studies. This suggests that this effect is primarily caused by the loss or addition of interactions via species turnover, and not changes to the existing species interactions per se. A decrease in forest cover is often associated with an influx of disturbance adapted species and a loss of forest specialist species (Anjos *et al.* 2019; Hatfield *et al.* 2020; Rand & Tscharntke 2007; Rutt *et al.* 2019). Disturbance adapted species are often associated with weak interactions and low levels of competition (Wootton & Stouffer 2016). Our findings suggest that these communities have more positive interactions, and are likely to be stable and relatively resilient to environmental changes (Landi *et al.* 2018; Lurgi *et al.* 2016; Qian & Akçay 2020).

Based on our results, it appears that communities vary in their stability along the gradient of deforestation, and the most stable communities occur in both highly forested and highly deforested areas. We hypothesise that communities in highly forested and highly deforested areas may be in contrasting stable states, whilst communities in intermediary forest cover may exist in long last transitory states (Hughes *et al.* 2013) (Figure 4.4). We propose that this occurs as the majority of species in these communities occur at either higher or lower forest cover than their average, and are therefore predicted to have more positive interactions which provide stability (Lurgi *et al.* 2016). On the other hand, our findings suggest more negative interactions in intermediately forested areas with mixed species composition (many species found in higher and lower forest cover) due to the influx of disturbance adapted species and the loss of forest specialist species, which are expected to be less stable (Loeuille 2010).

Stable communities in “degraded habitats” such as highly deforested areas may be seen as detrimental to conservation, as they are likely dominated by disturbance adapted species (Banks-Leite *et al.* 2014; Pardini *et al.* 2010; Prist *et al.* 2012). If these communities are

stable, they could be resistant to external perturbation, including restoration efforts, as found in previous studies (Lake 2013; Palmer *et al.* 2010; Suding *et al.* 2004) (Figure 4.4). This suggests that restoration efforts may be better placed in areas of intermediary deforestation which we suggest have less stable communities. Previous studies have already advocated for restoration efforts to prioritise areas of intermediary deforestation, due to a lack of source populations in highly deforested areas causing low recolonisation rates, and the need for the translocation of species (Banks-Leite *et al.* 2014; Santos Junior *et al.* 2016; Tambosi *et al.* 2014). Our study further suggests that restoration efforts coupled with species translocation to highly deforested areas may still not be enough, and that translocated species would be unlikely to persist due to the stability and resistance of the established community. Lastly, the observed shift in species interactions as a result of forest cover seen in our study may foreshadow and predict larger community changes in the future, as previous studies have found that interaction loss occurs at a faster rate than species loss (Valiente-Banuet *et al.* 2015).

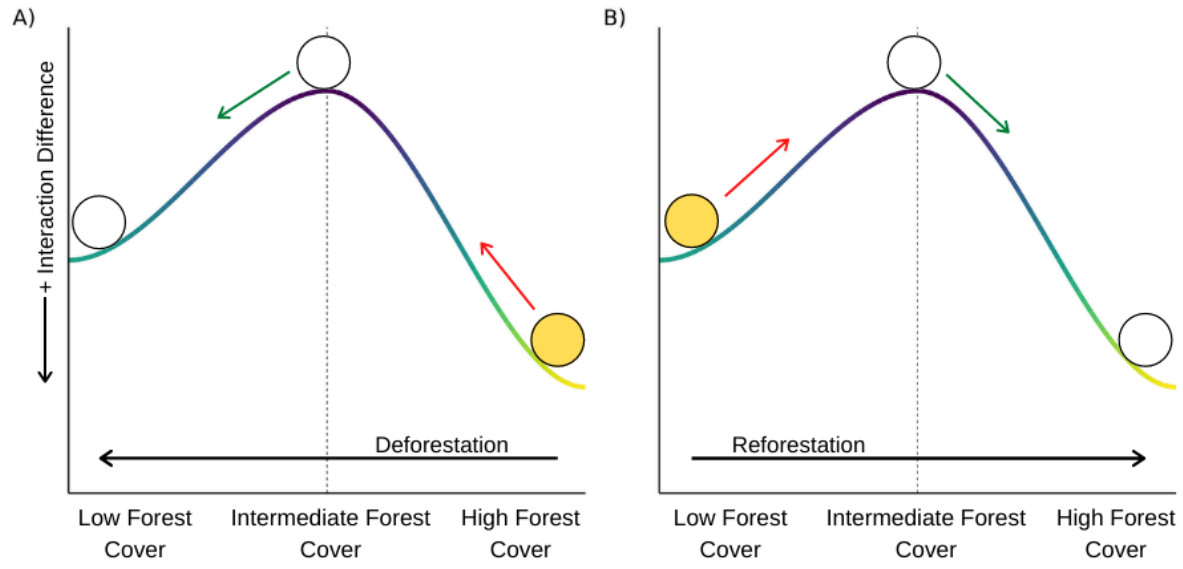


Figure 4.4. Cup and ball diagram of hypothesized stable states under deforestation (A) and reforestation (B) based on the observed changes in proportions of positive and negative interactions in communities at different relative forest covers shown in Figure 4.2A. In very highly forested areas nearly all species will be at, or above, their average forest cover, where

we expect an overall shift towards more positive interactions, with a similar, albeit weaker, result for highly deforested areas, where nearly all species will be at or below their average forest cover. On the other hand, we expect the highest proportion of species with contrasting forest cover requirements to occur in intermediately forested areas resulting in a shift towards more negative interactions. Arrow colours represent the hypothesized size of perturbation required to move the community to a new stable state (red = large, green = small). We expect that communities in highly forested and deforested areas to be relatively resilient to environmental changes (deforestation and restoration respectively), whilst those in intermediate forest areas to be comparatively sensitive to them.

Recently, methods used to estimate interactions from co-occurrence data have come under severe criticism (Blanchet *et al.* 2020), however our own evaluation of these methods against known interactions in the TaNa model provided surprisingly accurate results; five of the six tested methods had R^2 above 0.4 (Table S4.2). Additionally, previous studies have found that abundance data provides more accurate approximations of interactions (Popovic *et al.* 2019), however, we did not find this in our evaluation (Table S4.2). Whilst the co-occurrence methods are unlikely to fully capture true interactions, the similarity between our results from these methods and the TaNa provide confidence that we are capturing a true pattern of interaction change. Similarly, whilst our empirical results are limited to bird species, we consider the results transferable across taxa, as the mechanisms expected to cause the observed interaction changes are universal. Despite this the study could be built upon in the future by including functional traits for every species, as this may uncover varying trends in how interactions change with forest cover dependent on species niche (Ulrich *et al.* 2018).

In conclusion, we provide the first theoretical and empirical evidence that interspecific interactions are altered in a predictable manner with varying forest cover. Interactions were consistently found to be more positive when species were in higher forest cover across all data sets and TaNa models, which we believe could confer greater stability and resilience to

external perturbations. Species in lower forest cover were similarly found to have more positive interactions, albeit to a lesser extent, and with less consistency across data sets. Future studies predicting community shifts as a result of environmental changes should account for this variation, as should conservation and restoration efforts, which should consider focusing more heavily on restoring areas of intermediate forest cover.

Chapter 5: Fragmentation increases species richness but disturbs community composition

Abstract

How fragmentation affects biodiversity remains a highly debated topic, with studies finding contrasting results. Instead of focusing on whether fragmentation is good or bad for biodiversity, it has been suggested that identifying general rules for when and how fragmentation affects communities may be more useful to inform conservation actions; allowing us to vary expectations based on the properties of the community and the conservation goal. In this study we use a spatially-explicit Complex Systems approach to identify how fragmentation affects communities with different dispersal and immigration properties when considering various biodiversity metrics. We find that landscape-level species richness nearly always increases with habitat fragmentation, whilst the similarity of the community composition to that of an undisturbed landscape often decreases with fragmentation. In general communities with higher levels of immigration experienced the strongest effects of fragmentation. Whilst communities with high dispersal ability and probability generally increased in species richness with fragmentation, but decreased community similarity. Our study highlights the diverse impacts of fragmentation on ecological communities, whose effects can vary both due to the community experiencing fragmentation and the metrics measured. These results suggest that fragmentation can be both a benefit and hindrance to conservation, dependent on the goals.

Introduction

Habitat loss caused by anthropogenic driven land use changes is well recognised as a key driver of global biodiversity loss (IPBES 2019; Mace et al. 2005). Habitat loss is usually accompanied by habitat fragmentation, whereby previously contiguous areas of natural habitat are split into smaller disconnected patches surrounded by a hostile matrix (unsuitable habitat). Whilst the effects of habitat loss on biodiversity are usually negative (Orme et al. 2019), the specific effects of fragmentation in the absence of habitat loss are debated (Arroyo-Rodríguez et al. 2020; Fahrig 2017; Fahrig et al. 2019; Fletcher et al. 2018). Indeed this debate forms the basis of the SLOSS (Single Large Or Several Small) question, which looks to determine whether a single large habitat patch (low fragmentation) has higher species richness than several small patches (high fragmentation) given the same total habitat area (Diamond 1975; Fahrig et al. 2022; Ovaskainen 2002; Tjørve 2010).

Many studies have assessed the impact of fragmentation on different taxa, with most finding that fragmentation has a significant effect on ecological communities (Bolgovics et al. 2019; Boulinier et al. 1998; Gutiérrez-Arellano & Mulligan 2020; Le Roux et al. 2015; McNeill & Fairweather 1993), though the direction of effect (positive or negative) are less consistent (Bolgovics et al. 2019; Le Roux et al. 2015; Lindenmayer et al. 2015; Tscharntke et al. 2002; Virolainen et al. 1998). Recent reviews of the fragmentation and SLOSS literature have similarly failed to find concordant responses of species and communities to fragmentation, with Fahrig (2017, 2020) finding mostly positive effects, and Fletcher et al. (2018), mostly negative. As such it has been recently proposed that fragmentation studies move away from simply identifying if a community responds positively or negatively to fragmentation, and instead focus on identifying why and when a particular community responds positively or negatively to fragmentation, as well as how it responds (Rybicki et al. 2020; Wilson et al. 2016).

The properties of ecological communities have previously been found, and would be expected, to determine whether fragmentation has positive or negative effects (Cushman 2006; Ewers & Didham 2006; Wilson et al. 2016). The best researched of these properties is the ability and propensity of species within the community to disperse between habitat patches (Ewers & Didham 2006; Niebuhr et al. 2015). Communities composed of species that are able to easily and efficiently move between the fragmented habitat patches are less negatively affected by habitat fragmentation than communities of poor dispersers (Ewers & Didham 2006; Niebuhr et al. 2015; Thomas 2000). This is because efficient dispersal across fragments can reduce isolation and local extinction risk (Cushman 2006; Thomas 2000), and because good dispersers spend less time in the matrix reducing individual mortality risks (E. Smith & O. Batzli 2006; Gruber & Henle 2008). While efficient dispersal reduces the impacts of fragmentation, species with greater dispersal propensity or frequency among patches can be more negatively affected by fragmentation. Frequent dispersal could increase migration (Cushman 2006; Fahrig et al. 2022), but it also increases the number of individuals moving across the matrix, where mortality risks are higher, potentially leading to reduced population sizes and ultimately local extinction (Homan et al. 2004). On the other hand, greater dispersal frequency from individuals in source populations of continuous habitat areas to other areas, immigration rates, can reduce the effects of fragmentation (Püttker et al. 2011). Communities with high levels of immigration from a source population can benefit from rescue effects that reduce the risk of local extinction (Pardini et al. 2010; Püttker et al. 2011).

Understanding the properties of communities to evaluate the impacts of fragmentation is key, but we also need to consider how we measure responses and ideally capture multiple aspects reflecting biodiversity changes. To date, most fragmentation and SLOSS studies have focused on a single response metric; species richness (Fahrig et al. 2019, 2022; Wilson et al. 2016). A majority of these studies have found that species richness increases with fragmentation due to increased habitat heterogeneity (Bolgovics et al. 2019; Fahrig 2020; Slancarova et al. 2014; Tscharntke et al. 2002). Species richness is however only one

aspect of biodiversity which can mask changes (turnover) in species composition and is therefore liable to miss impacts of fragmentation (Fleishman et al. 2006; Hillebrand 2004). As such studies which just measure species richness may miss important impacts of fragmentation, especially those changing the species composition of the community (Fleishman et al. 2006; Haas et al. 2020; Hillebrand 2004).

Measuring the effect of fragmentation across multiple biodiversity metrics such as community composition and species richness is key to inform conservation decisions (Aggemyr et al. 2018; Haas et al. 2020). Maximising areas for both species richness whilst maintaining the original community composition of the area may however be impractical, as the mechanisms by which fragmentation is expected to increase species richness may conversely be predicted to decrease similarity in species composition between the fragmented and pre-fragmented communities. The most obvious of these mechanisms are the increases in habitat heterogeneity caused by edge effects and increased spatial extent which are likely to create new niches for previously absent species to inhabit, thus reducing similarity in community composition (Fahrig 2020; McNeill & Fairweather 1993; Tscharntke et al. 2002; Virolainen et al. 1998). Additionally, the temporal stability of communities may provide another interesting biodiversity metric to measure when looking at the effects of fragmentation. Previous studies have found that communities in small patches are less temporally stable than those in larger patches, potentially reducing their conservation value (Boulinier et al. 1998). Temporally stable communities are expected to persist in their current state, and could therefore provide long-term refuge for conservationally important species, and thus be valuable from a conservation perspective; even if these communities have lower richness. On the other hand, temporally unstable communities are unlikely to be of high conservation value, as they are unlikely to persist in their current state in the long term.

Here we use a spatially-explicit Complex Systems modelling approach to understand when and how fragmentation affects communities. We test the effects of fragmentation on species richness, community composition and temporal stability across communities that differ in

their dispersal ability and frequency, as well as immigration rates. Notably, our model includes key mechanisms omitted from previous modelling studies including a fully interacting community with competitive, mutualistic and predator-prey interactions; species that range in their degree of habitat specialisation; and a novel means of including the impact of edge effects.

Methods

We used a modified version of the Tangled Nature Model (hereafter 'TaNa') presented in Brinck (2018) which is a short-term spatially explicit implementation of the original TaNa proposed by Christensen et al. (2002). TaNa models use a Complex Systems approach to modelling ecological communities whereby the properties of the community at the macroscopic scale emerge from the interactions and movement of individuals at the microscopic scale. TaNa models are characterised by their production of quasi-stable states in the community, which are stable states punctuated by short periods of drastic community composition and abundance changes (Christensen et al. 2002; Collobiano et al. 2003). TaNa models have been shown to accurately recreate realistic ecological results including species-area relationships, trophic networks, and species abundance distributions (Brinck & Jensen 2017; Christensen et al. 2002; Laird et al. 2008). An overview of the TaNa model and detailed descriptions of the modification implemented in our version are given below.

The TaNa model used here overcomes many of the limitations of previous modelling approaches, including all types of interactions between species, as well as a novel means of including the impact of habitat edge effects.

Landscape

The landscape used in our study is an 11 x 11 grid of uniformly sized cells, which are connected at the top and the bottom, and left and right, to create a torus. Within the landscape, cells can have two states: Habitat, or Non-Habitat. Landscapes in our study could hold one of four habitat configurations: Continuous Habitat; Single Large; Few

Medium; or Several Small (Figure 5.1). Under the Continuous Habitat configuration, all 121 cells were Habitat cells which represented an area of continuous habitat before fragmentation. The other three configurations represented fragmented landscapes with 25 Habitat cells and 96 Non-Habitat cells and differed in the spatial organisation of the Habitat cells (Figure 5.1).

An 11x11 cell landscape was chosen as it was large enough to allow for different habitat configurations to be created and for spatial heterogeneity, whilst being computationally viable.

Species

The model included a species pool δ of 100 species S ($S_{a,}S_{b}, S_{\ldots}$) which all had the potential to exist in the landscape. Each of these 100 species had two distinct traits: specialisation ζ and interactions J. A species' S_{a} specialisation value ζ_{a} was drawn from a random uniform distribution between 0 - 1 and represents how specialised a species is to the habitat, and therefore how much the fitness of an individual of a species is affected by the immediate habitat around it. A low specialisation value indicates that a species' fitness is only marginally affected by the state of the cell in which it resides (Habitat or Non-Habitat), whilst a high specialisation value indicates that a species' fitness is heavily affected by the state of the cell in which it resides. As such an individual of a species with a low specialisation value will have a similar fitness in a Habitat and Non-Habitat cell (all else equal) whilst a species with a high specialisation value will have a higher fitness in a Habitat cell than a Non-Habitat cell (all else equal).

Each species also had an interaction value with every other species which is stored in a matrix J, where $J(S_{a}, S_{b})$ stores the interaction value between species S_{a} and S_{b}. A set proportion θ (table 5.1) of interactions between species were randomly set to zero (no interaction) whilst the remaining interactions were set to a value randomly drawn from a

normal distribution with a mean of zero and standard deviation of 0.5, whilst interactions of individuals of the same species were set to zero. Interactions between species are not equally reciprocal, i.e. $J(S_a, S_b) \neq J(S_b, S_a)$ meaning species can have mutualistic (+/+), competitive (-/-), predator-prey (+/-), commensal (+/0), amensal (-/0) interactions with one another, or not interact at all (0/0). Interactions between species affect their fitness, whereby the fitness of an individual will increase when in proximity (defined by parameter γ) of individuals from a species with which it has a positive interaction, and decrease when in proximity to individuals from a species with which it interacts negatively.

Model Dynamics

The TaNa was run over 625,000 time steps. Within each time step cells were chosen in a random order without replacement, and four processes occurred:

1) **Death:** an individual was randomly chosen and removed from the cell with a constant probability P_Z; which was set to 0.15 as used in previous TaNa studies (Arthur *et al.* 2017; Hall *et al.* 2002)
2) **Dispersal:** an individual was randomly chosen from the cell and moved to another cell with probability P_D up to its maximum dispersal distance D.
3) **Immigration:** an individual from a random species within the pool of 100 species was introduced to the cell with probability P_I
4) **Reproduction:** an individual was randomly chosen from within the cell and reproduces with probability P_R, which is defined for species S_a at time t in cell f, as follows:

$$P_R(S_a, t, f) = \frac{exp(\Xi(S_a, t, f))}{1 + exp(\Xi(S_a, t, f))} \in (0, 1)$$

(eqn 5.1)

Where the reproductive capability Ξ is given by:

$$\Xi(S_a,\ t,\ f)\ = \frac{\sum_{f=C}\sum_{S_b=\delta} N_{b,f} J(S_a, S_b)}{C}\ +\ \mu\ -\ \frac{N_{\delta,f}}{R_f}$$

(eqn 5.2)

Where R_f is the carrying capacity of cell f(set to 10), C is the number of cells within proximity γ, $N_{b,f}$ is the number of individuals of species S_b in cell f, and $N_{\delta,f}$ is the total number of individuals in cell f, and the habitat effect μ is given by:

$$\mu(S_a,\ t,\ f)\ = \pm\ \zeta_a$$

(eqn 5.3)

Where μ is positive when species S_a is in a Habitat cell f, and negative when S_a is in a Non-Habitat cell f, which is dependent on time t.

Recreating Edge Effects

One of the key mechanisms by which fragmentation is known to impact species richness and community composition is via edge effects (Ewers & Didham 2006; Laurance et al. 2011; Pfeifer et al. 2017). Here we simulated edge effects by allowing the fitness of individuals within cells to be affected by their interactions with individuals in their surrounding neighbourhood of cells. This was controlled by the parameter γ which is a distance specifying which cells are included in the neighbourhood of cells; for instance, a γ value of 1.5 would include the cell the focus species is in, as well as the eight surrounding cells in the calculation of P_R. This means that the fitness of individuals residing in Habitat cells on the edge of habitat patches was influenced by the individuals residing in the Non-Habitat cell next to it, and vice versa. This approach emulated the different conditions experienced at the

edge of patches, which are not intrinsically either positive or negative, but do create a different environment for species to reside in.

Model Set-Up and Parameters

Each of the 121 cells of the landscape was initially filled with 100 individuals drawn at random from the species pool. For all configurations, the landscape was initially fully covered in habitat, with all 121 cells set as Habitat cells.

We varied the parameters of the model to test the influence of key community properties under different scenarios:

Dispersal: We considered three values for dispersal, each representing a different scenario. These were no dispersal ($D = 0$), low dispersal ability ($D = 1$) and high dispersal ability ($D = 2.5$). Under the no dispersal scenario individuals were confined to the cell in which they were placed, or immigrated to. Under the low dispersal scenario, individuals were able to move between cells and therefore patches. However, as all Habitat cells are at least two cells apart and individuals only move one cell at a time, movement required residing in the matrix for at minimum one time step, and this would decrease their fitness (hostile matrix), assuming the species was not a complete generalist ($\zeta_a = 0$). Under the high dispersal scenario individuals were capable of moving between patches without residing in the matrix, though since dispersal is random, many individuals still dispersed from Habitat cells to Non-Habitat cells.

Dispersal Probability: describes the density-dependent probability (P_D) that a species will disperse from its current cell to another within its dispersal ability (D), and regulates the frequency of dispersals (Equation 5.4).

$$P_D(S_a,\ t,\ f)\ = \left(\frac{N_{\delta,f} P_{D,max}}{R_f(4+log((1-P_Z)/P_Z)}\right)$$

(eqn 5.4)

Where $P_{D,max}$ is a user-defined value which describes the maximum dispersal probability, P_Z and R_f are set to the same constants as described above, $N_{\delta,f}$ is the number of individuals in cell f . We explored three $P_{D,max}$ scenarios: low dispersal probability ($P_{D,max} = 0.0001$), medium dispersal probability ($P_{D,max} = 0.001$), and high dispersal probability ($P_{D,max} = 0.01$). At cell densities $N_{\delta,f}$ of approximately 50 individuals, this equates to 1.2, 12, and 121 dispersal events from the 121 cells in the landscape per 100 time steps.

Immigration Rate: We considered three scenarios for immigration rate (P_I): no immigration ($P_I = 0$), low immigration ($P_I = 1.6529 \times 10^5$) which equates to one immigration event in the landscape per 500 time steps, and intermediate immigration ($P_I = 1.6529 \times 10^4$) which equates to one immigration event into the landscape per 50 time steps. Higher immigration rates of increasing magnitudes were tested in pilot studies, however, these were found to almost permanently maintain high abundances of all species, thus masking the effects of the other model dynamics.

Edge Effect: We considered two scenarios for edge effects (γ), one in which there were no edge effects ($\gamma = 0$), and one where edge effects were included ($\gamma = 1.5$); the fitness of individuals in a focal cell were affected by the presence of species in the eight cells surrounding them, as well as those in the focal cell.

Table 5.1. Model parameters, definitions, and values

Parameter	**Definition**	**Value**
$J(S_a, S_b)$	Interaction between species S_a and S_b	N(0,1)
ζ_a	Specialisation value of species S_a	U(0,1)
θ	Probability of interaction	0.5
R_f	Carrying capacity of a cell	10

P_Z	Probability of death	0.15
D	Dispersal ability	Varied (see table 5.2)
$P_{D,max}$	Maximum probability of dispersal	Varied (see table 5.2)
P_I	Probability of Immigration	Varied (see table 5.2)
γ	Interaction distance (edge effects)	Varied (see table 5.2)

Table 5.2. Values used in simulated experiment varying community properties

Parameter	**Low**	**Intermediate**	**High**
D	0	1	2.5
$P_{D,max}$	0.0001	0.001	0.01
P_I	0	1.6529×10^5	1.6529×10^5
γ	0	N/A	1.5

Simulation Experiments

We initially ran twenty-five models for each combination of the Dispersal (D), Dispersal Probability ($P_{D,max}$), Immigration Rate (P_I), and Edge Effect (γ) scenarios totalling 1,350 models (54 parameter combinations). Each of these models had a species pool which differed in their specialisation and interaction values specified by a unique seed.

After 500,000 time steps, the landscape of each of the 1,350 models was transformed into each of our four habitat configurations, three of which were fragmented; Single Large; Few Medium; or Several Small; and a Continuous Habitat configuration acting as a control, thus generating 5,400 models. Since each of the original 1,350 models transformed into all four configurations, each group of four configurations contained the same species pool with the same traits, and were independent from all of the other model groups. As such, results between these four configurations can be directly compared and is akin to visiting four

identical landscapes with identical species pools before and after fragmentation (Figure 5.1). After the configuration change, models were run for another 125,000 time steps. Initial runs of the model showed that communities usually assembled to stable states in less than 10,000 time steps, thus our choice to run models for 500,000 time steps before fragmentation was extremely conservative. Similarly, in pilot studies communities responded to fragmentation in a short time frame, and therefore 125,000 time steps is likewise a conservative choice. Fewer time steps were used after fragmentation than before simply because more models were run after fragmentation, and thus computational time had to be considered.

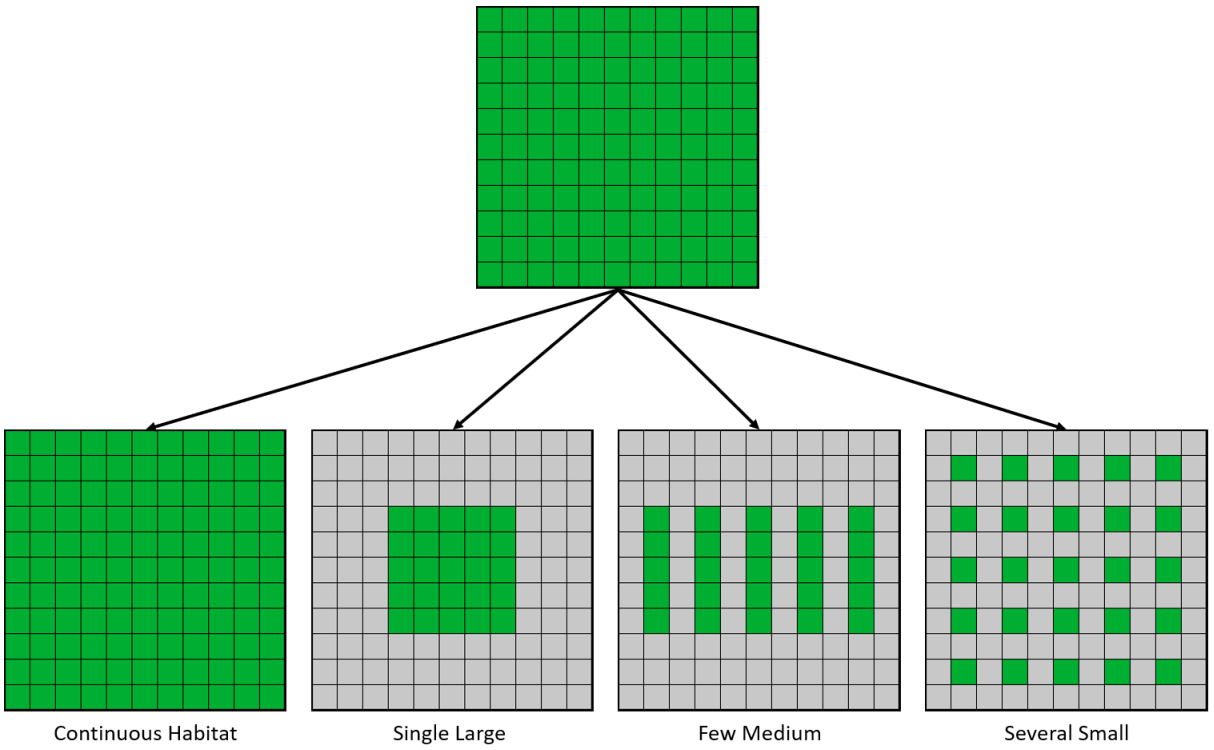


Figure 5.1. At 500,000 time steps, a single landscape of Habitat cells fragmented into all four of our habitat configurations, including the control Continuous Habitat Configuration and the three fragmented configurations (Single Large; Few Medium; Several Small). These configurations include all of the same species in the species pool, which means they have the same interactions and specialisation values, as well as the same community parameters.

We collected data on the presence and abundance of every species in each Habitat cell in the landscape every 1,250 time steps after fragmentation occurred, and at the final time step (500,000) before fragmentation. Data was only collected and analysed from Habitat cells as SLOSS and fragmentation studies focus on the impact of fragmentation on the populations within habitat patches, not from the surrounding matrix.

Analysis

We explored the effects of fragmentation on four metrics:

Total species Richness: The number of unique species across all Habitat cells in the landscape at the end of the 125,000 time steps.

Species Richness per Habitat Cell: The mean number of unique species per Habitat cell in the landscape at the end of the 125,000 time steps.

Community Composition: The Sorensen dissimilarity index (Equation S5.1) in community composition between the same Habitat Cells in the control Continuous Habitat configuration to those in each of the other three fragmented configurations at the end of the 125,000 time steps. To aid understanding, dissimilarity was converted to similarity by subtracting dissimilarity from one. We used a presence/absence based dissimilarity measure as most studies are interested in the presence of species as opposed to their abundance, but we also calculated Bray-Curtis dissimilarity based on abundance, and present results in the Supplementary (Figure S5.1). We also partitioned the drivers of community composition into nestedness (species loss) and turnover (species replacement) using the betapart package (Baselga 2010; Baselga & Orme 2012).

Temporal Stability: We used Weighted Coefficient of Variation (WPCV) as a measure of community temporal stability. WPCV is measured as the weighted coefficient of variation of the population of each individual species. Calculated using the abundance of each species in Habitat cells over the 125,000 post-fragmentation time steps at 1,250 time step intervals

Equation S2 in the Supplementary. A high WPCV suggests a low temporal stability, and vice versa.

For all metrics, the value for the Single Large patch configurations was used as a baseline, and the proportional change to the other fragmented configurations were calculated from this, e.g. if the Single Patch configurations had a species richness of five and the Several Small patch a species richness of six, after calculating proportional differences the Single Large configuration would have a value of one, and the Several Small configuration a value of 1.2. Since each group of four configurations is independent, this was done individually for each group of four configurations, and thus a total of 25 times per unique parameter combination.

We evaluated changes in these metrics in response to fragmentation. Our fragmentation configurations represent particular values within the continuum of possible number of habitat patches in the landscape (Single Large = 1, Few Medium = 5, and Several Small = 25) and thus, were modelled as a numerical predictor.

We used linear models as they were the most parsimonious after comparing to the AIC of models with a log-transformed predictor (number of habitat patches), which were tested as many of the relationships appeared to be logarithmic when initially plotted. To interpret model results we use a biologically significant approach as opposed to a statistically significant approach as p-values are dependent on the sample size which here could be manipulated by running more simulations (White et al. 2014). We decided on a biologically significant cut-off of 20% change, for example when testing for species richness, if a linear model predicted the Several Small configuration to have 20% more or fewer species than the Single Large configuration, it would be classified as a significant change. Coincidentally, a 20% cut-off also aligns well with statistical significance in our linear models, likely due to our relatively modest seventy-five data points per linear model.

Results

Fragmentation increased the total landscape level species richness under most scenarios, as long as individuals were able to disperse, but did not affect the average species richness of Habitat cells (Figure 5.2). Total species richness generally increased more when individuals' ability to disperse was low ($D = 1$) (individuals were exposed to the hostile matrix), immigration was included ($P_I \neq 0$), and dispersal probability was intermediate ($P_{D,max} = 0.001$) (Figure 5.2). Conversely, the average species richness of Habitat cells didn't change with fragmentation (Figure 5.2).

Fragmentation decreased similarity in community composition when edge effects were included so that species in habitat cells near matrix cells could be affected by those in the matrix ($\gamma = 1.5$) (Figure 5.3). With edge effects, dissimilarity was greatest when immigration, dispersal ability, and dispersal probability were at their highest ($P_I = 1.6529 \times 10^4$; $D = 2.5$; $P_{D,max} = 0.01$). Changes in similarity were driven by changes in community composition caused by turnover (Figure 5.3).

Lastly, the effects of fragmentation on temporal stability varied. On one hand temporal stability decreased with fragmentation when edge effects were included ($\gamma = 1.5$) and dispersal probability was intermediate ($P_{D,max} = 0.001$) or high ($P_{D,max} = 0.01$) (Figure 5.2). Contrastingly, when edge effects were not included ($\gamma = 0$), temporal stability often increased with fragmentation especially when dispersal probability was low ($P_{D,max} = 0.0001$) (Figure 5.2).

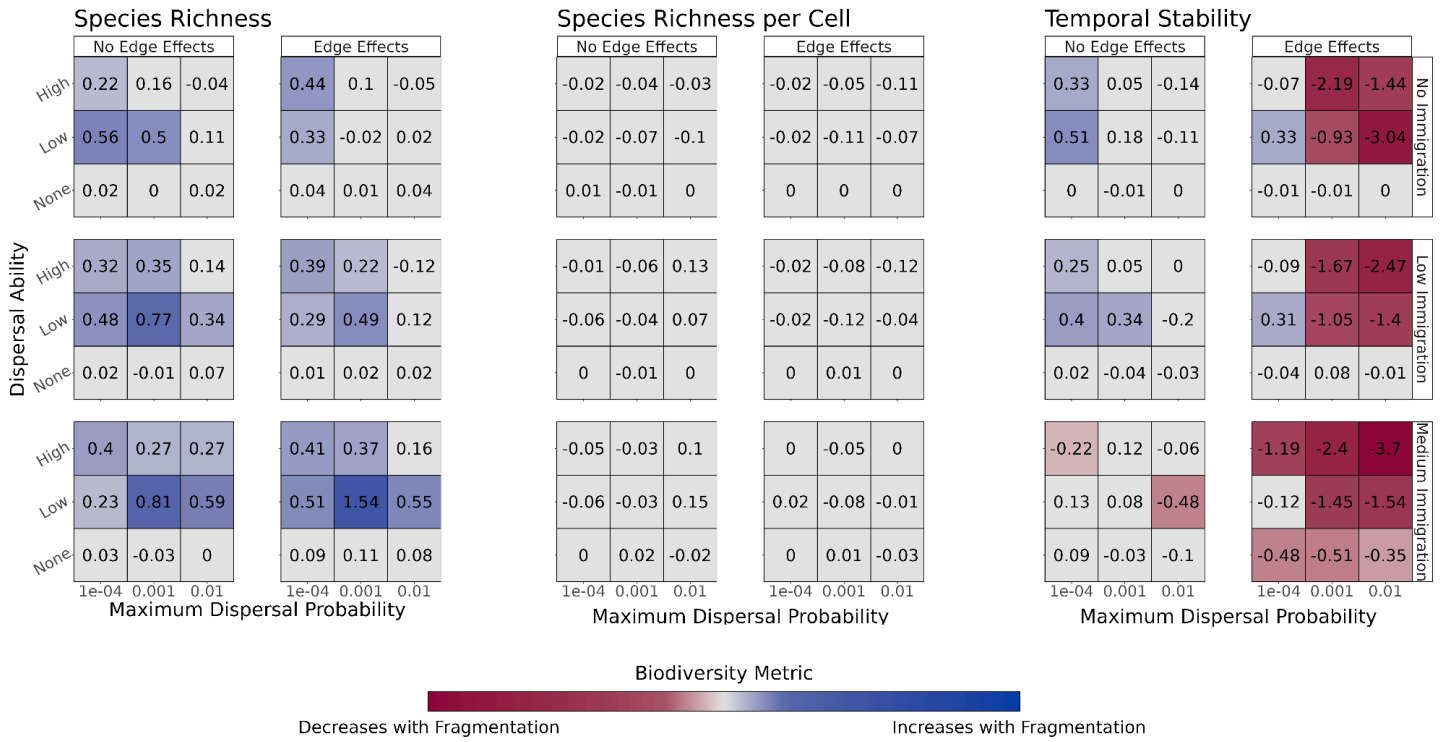


Figure 5.2. Proportional change in landscape scale species richness, species richness per forested cell, and Temporal Stability due to fragmentation, where the Single Large configuration is used as a baseline. Grey cells are below 0.2 change, and are deemed biologically insignificant. Red cells represent a biologically significant decrease as fragmentation increases, whilst Blue cells represent a biologically significant increase as fragmentation increases. Note that the value shown for temporal stability is the negative of the Weighted Population Coefficient of Variation (WPCV), flipping the sign, as temporal stability decreases when WPCV increases, and increases when WPCV decreases.

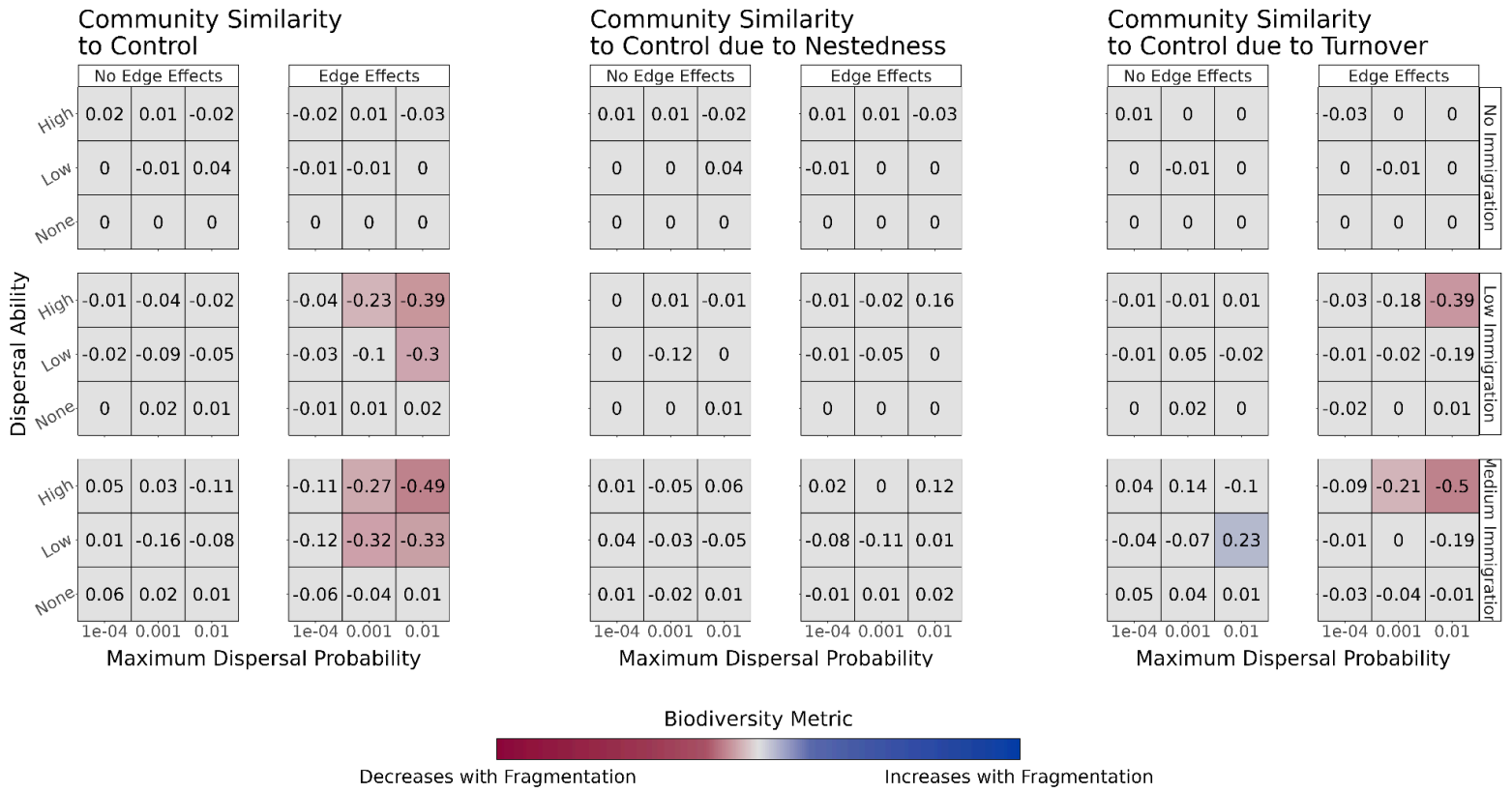


Figure 5.3. Proportional change in landscape scale community similarity to the comparable cells in the control Continuous Habitat configuration, as well as specific changes in community similarity caused by nestedness or turnover due to fragmentation, where the Single Large configuration is used as a baseline. Grey cells are below 0.2 change, and are deemed biologically insignificant. Red cells represent a biologically significant decrease as fragmentation increases, whilst Blue cells represent a biologically significant increase as fragmentation increases.

Discussion

Here we evaluated when and how fragmentation affects communities, focusing on communities with different properties and measuring changes in key biodiversity metrics often used in conservation assessments and decisions. Our results highlight the varied responses of even relatively simple communities to fragmentation, and the discordant relationship between fragmentation and various biodiversity metrics.

Our main conclusion can be summarised as follows. Firstly, as expected based on previous empirical and modelling approaches, habitat fragmentation is neither universally good nor universally bad (Fahrig et al. 2019; Fletcher et al. 2018; Rybicki et al. 2020), and varies both depending on the metric used to assess its impacts and the properties of the community studied. As such, the potential impacts of fragmentation will have to be assessed on a per community basis, and with a particular conservation goal in mind, linked to a specific biodiversity metric.

Secondly, fragmentation nearly always increased species richness, and never significantly decreased it; a finding which mirrors many empirical studies (Bolgovics et al. 2019; Fahrig et al. 2019, 2022; McNeill & Fairweather 1993; Tscharntke et al. 2002). This pattern was driven by increases in beta diversity between Habitat cells in the more fragmented configurations, as highlighted by the lack of increase in per cell species richness with fragmentation. Increases in species richness are particularly prevalent when dispersal ability is low, suggesting that the higher beta diversity between cells is driven by isolation as opposed to habitat heterogeneity. For conservation purposes our results suggest that fragmentation may be considered to have positive effects if the aim is solely to increase total landscape species richness.

Thirdly, Community similarity never increased with fragmentation, though it did frequently show no significant relationship. Community similarity was only found to significantly decrease with fragmentation for communities that included edge effects and had immigration

from a source population. Edge effects creating habitat heterogeneity and new niches, combined with the immigration of new species able to utilise these niches is most likely the driver of this change. This was further solidified by our finding that most of the change in community similarity with fragmentation was driven by turnover (species replacement) as opposed to nestedness (species loss), again suggesting that an influx of new species causes these changes, as opposed to loss of species alone. Since edge effects and their associated positive (Fahrig 2020; Terraube et al. 2016; Willmer et al. 2022) and negative (Fletcher et al. 2018; Laurance et al. 2011; Willmer et al. 2022) impacts are well established, and immigration of at least some species into the landscape that can utilise them is likely, our results would suggest that keeping a single large patch is the safest option when aiming to maintain a community composition as close as possible to that present before fragmentation occurred.

Fourthly, temporal stability exhibited varied responses to fragmentation, though it most commonly decreased with fragmentation. This negative response was most prevalent when edge effects were included and when species had a high propensity to disperse; likely disrupting the isolated communities formed by high levels of fragmentation and thus causing low stability. On the other hand, stability sometimes increased with fragmentation when dispersal probability was low, potentially as a result of limited movement between patches, which was likely a key cause of instability. Our results would suggest that under nearly all circumstances where individuals can move between patches in a landscape, temporal stability will decrease as fragmentation increases, and as such any benefits to species richness gained from fragmentation may be lost over time.

As with all modelling approaches our work has limitations. First and foremost, our model represents an abstract system rather than a particular environment. The majority of our parameters can only be assessed in relation to themselves and are not derived from empirical data on real-world communities. For example, our values and results when varying dispersal probability can be assessed against our other values, but do not have any

grounding in true dispersal probabilities of real ecological communities. Second, whilst our communities are fairly complex, they still represent an oversimplification of the complex nature of species and communities. Most notably, all of our species within each community have the same value for dispersal ability and dispersal probability, whereas in real communities some species are better or more frequent dispersers than others (Jenkins et al. 2007; Sheard et al. 2019; Tobias et al. 2022), likely leading to even more complex patterns in community responses to fragmentation. It is therefore even more surprising that we found such a diverse array of responses to fragmentation within our small, simplified set of communities.

Here we have limited our study to a small number of fragmentation configurations, and a minimal set of parameter changes for our community scenarios. Despite this, we have managed to uncover and illustrate the complexity of predicting and understanding community responses to fragmentation. Of particular note is the likely conservation trade-off between species richness, composition and long-term stability of communities, which emphasises the importance of having clear conservation aims and considering a diverse portfolio of metrics to monitor and assess conservation actions. Similarly, our unique Complex Systems modelling approach could be utilised in future work on similar topics. The most promising of these may be its ability to assess the underlying mechanisms driving the variation seen here in response to fragmentation. It is particularly well poised to assess the role competitive and predator-prey dynamics have in the outcomes of different fragmentation configurations; an as yet understudied, but expectedly important mechanism (Deane *et al.* 2020; Fahrig 2020; though see Rybicki *et al.* 2020). Ultimately, our study has uncovered the situations in which fragmentation can have positive and negative results that depend on community properties and conservation goals.

Chapter 6: Main Discussion

Overview

This thesis focused on uncovering variations in community responses to deforestation, highlighting which communities are most sensitive, as well as why. This thesis additionally found mechanisms which cause communities to change as deforestation occurs and showed the use of complex systems to study ecological communities. Whilst these thesis chapters have been created as stand-alone manuscripts for publication, this section focuses on their combined implications.

Firstly, I discuss the main output of the thesis; how results from the data chapters have combined to expand our understanding of variation in community responses to deforestation. I will then explore the implications of these results, as well as additional findings from the thesis for landscape-level conservation and restoration, before briefly outlining the use of complex systems modelling to study ecological communities and future research which could expand on my work.

Community Responses to Deforestation

Communities are expected to differ in their responses to deforestation, with some communities experiencing more negative impacts than others; more severe decreases in abundance and changes in community composition (Carrara *et al.* 2015; Halstead *et al.* 2019; Melo *et al.* 2018; Prist *et al.* 2012; Rutt *et al.* 2019). However, whilst lots of studies have identified determinants of sensitivity to deforestation at the species level, far less is known about what determines community sensitivity.

In this thesis, I demonstrate that communities most sensitive to deforestation are those at the equator (**Chapter 3**) and those in regions which have experienced anthropogenic disturbances (**Chapter 3**). Furthermore, I find that the proportion of forest specialist species (**Chapter 2**) as well as the dispersal ability of species within the community (**Chapter 5**) are the likely mechanisms which cause the observed variation in sensitivity.

In **Chapter 3** I found that ecological thresholds in the relationship between community integrity and forest cover were highest in undisturbed areas near the equator, and lowest in disturbed areas near the poles. Higher thresholds indicate higher sensitivity to deforestation since community composition begins to strongly alter with reductions in forest cover at an earlier point of deforestation. Previous studies have suggested that this would be the case, due to the poorer dispersal ability of species at the equator (Salisbury *et al.* 2012; Sheard *et al.* 2019; Shennan-Farpón *et al.* 2021) and more forest specialist species in undisturbed regions (Betts *et al.* 2019). Poorer dispersing species are expected to be strongly impacted by the fragmentation of forested areas which usually accompanies forest loss, since they are unable to move between fragmented forest patches, essentially reducing the overall habitat area available to them (Hatfield *et al.* 2018a). Forest specialist species are similarly expected to be more sensitive to forest loss since they require specific resources or environmental conditions that are only available in forested environments (Prist *et al.* 2012; Rutt *et al.* 2019). My results partially support these hypotheses; firstly finding that higher proportions of forest specialist species occur in highly forested equatorial regions in **Chapter 2**, and secondly showing that communities vary in their response to fragmentation based on community properties in **Chapter 5**.

In **Chapter 2** I found that there was a latitudinal gradient in the proportion of forest specialist species and that a higher proportion of forest specialist species occur at the equator. I similarly found that the proportion of forest specialist species varies with regional forest cover, with higher proportions of forest specialist species occurring in areas with high regional forest cover. This potentially explains the results of **Chapter 3**, since higher

proportions of forest specialist species were found at the equator in heavily forested regions; the same areas in which communities were found to be the most sensitive to deforestation in **Chapter 3**.

Variation in the dispersal ability of species within communities is similarly expected to cause differences in responses to deforestation. In **Chapter 5** I assessed the impact of fragmentation on communities with different properties, focusing on differing responses of communities with varying dispersal ability and frequency to fragmentation. I found that communities with well-dispersing species were generally less impacted by fragmentation than those with poorly dispersing species; weaker change in species richness and community composition. Previous studies have found that a latitudinal gradient in dispersal ability exists (Salisbury *et al.* 2012; Sheard *et al.* 2019), and therefore this may explain the variation in community responses to deforestation seen in **Chapter 3**. Furthermore, in **Chapter 2** I conduct a preliminary analysis showing that forest specialist species generally have poorer dispersal ability than non-forest specialist species, indicating that communities with high proportions of forest specialist species will also be poor dispersers, and thus sensitive to deforestation.

Implications for Landscape-Level Conservation and Restoration

Area-based conservation techniques are perhaps the most promising conservation action to mitigate biodiversity loss due to anthropogenic habitat loss (Hannah *et al.* 2020; Maxwell *et al.* 2020; Tambosi *et al.* 2014). Policies to implement such techniques are becoming more popular (e.g. 30x30), however the question of where to focus such efforts is still debated (Hannah *et al.* 2020; Maxwell *et al.* 2020; Tambosi *et al.* 2014). This thesis suggests that at the global level, a larger proportion of forests should be conserved in equatorial regions with few anthropogenic disturbances since these contain ecological communities which are most sensitive to deforestation.

As well as uncovering variation in community sensitivity to deforestation at large regional scales, much of the research in this thesis could be used to directly mitigate the impacts of deforestation on communities and to identify suitable areas for restoration at the local level. Specifically, findings from **Chapter 3** and **Chapter 4** both indicate that reforestation of intermediately forested areas is more beneficial for biodiversity than highly deforested areas, whilst **Chapter 5** highlights how different spatial configurations of forest restoration should be used depending on the intended biodiversity outcome.

In **Chapter 4** I found that interspecific interactions were most positive in areas of higher forest cover, relatively positive in areas of very low forest cover, and most negative in areas of intermediate forest cover. This indicated that communities in intermediately forested areas are less stable than those in areas of high or low forest cover since positive interactions have been found to stabilise communities (Lurgi *et al.* 2016; Qian & Akçay 2019). The instability of communities in intermediately forested areas and the stability of communities in poorly forested areas suggest that reforestation efforts should primarily focus on areas of intermediate forest cover, since these communities should have little resistance to changing to a conservationally-positive alternative stable state. This finding is corroborated by the theoretical framework proposed from the results of **Chapter 3**. Specifically, in **Chapter 3** I found support for a non-linear relationship between community integrity and forest cover, whereby community composition markedly changes with initial forest loss, then remains fairly stable, before once again significantly changing as forest cover becomes minimal (Figure 3.3). This finding similarly suggests that areas of intermediate forest cover are most suitable for restoration since their communities still maintain many of the species found in undisturbed locations.

Lastly, in **Chapter 5** I focused on how the spatial configuration of habitat can influence biodiversity, which could be used to inform restoration plans; is it better to have a small number of larger habitat patches, or many small habitat patches (Fahrig *et al.* 2022; Le Roux *et al.* 2015; Lindenmayer *et al.* 2015). Unlike **Chapters 3** and **4**, this chapter specifically

aimed to identify optimal configurations of habitat patches for conservation planning, dependent on the intended biodiversity target. I found that many small patches usually increase species richness, but a few large patches more closely maintain the community composition of a natural landscape. The strength of these results, but generally not the direction, varied based on the properties of species within the community, with communities with poorly dispersing species generally more impacted than those with well-dispersing species; as predicted from previous literature (Cushman 2006; Thomas 2000).

Combining results from these three chapters can give insights into which local areas should be prioritised for restoration and conservation.

1) Intermediately forested landscapes should be prioritised for restoration over highly deforested areas, as previously proposed (Banks-Leite *et al.* 2014; Lake 2013; Tambosi *et al.* 2014), though uniquely I show that this is due to the instability of communities in intermediately forested areas; driven by less positive interactions and the mixing of forest specialist and general species
2) Restoration should prioritise connecting patches if the conservation target is to maintain the ecological community present before anthropogenic disturbance
3) Restoration should prioritise creating as many individual habitat patches if the conservation target is to maximise species richness. Though the resulting communities will likely contain invasive species and could lose larger, often rarer species, which have high minimum patch-area requirements (Henle *et al.* 2004)

Complex Systems for Ecological Research

Complex systems modelling approaches are a regular feature of many disciplines, from physics to economics (Farmer *et al.* 2012; Holovatch *et al.* 2017). Despite ecological communities and ecosystems being quintessential complex systems, their use in ecological research has been minimal (Riva *et al.* 2023). In this thesis I have used a complex systems

modelling approach, the Tangled Nature Model in **Chapter 4** and **Chapter 5**, demonstrating its use in modelling the impact of deforestation on ecological communities.

Whilst many other modelling approaches are used in ecology (Ferrier & Guisan 2006; Gandu *et al.* 2004; He & Legendre 2002) complex systems approaches excel at modelling interconnected parts within larger systems (Anand *et al.* 2010; Riva *et al.* 2023). I utilised this unique aspect of complex systems to study how deforestation impacts species interactions in **Chapter 4**, finding that interactions vary with deforestation. However, as with many theoretical modelling approaches, showing that results are empirically feasible is difficult (DiRenzo *et al.* 2023). Impressively, in **Chapter 4** I show that the results from the complex systems modelling approach very closely match those from an empirical dataset where interactions were estimated using empirical methods (Clark *et al.* 2018; Friedman *et al.* 2019; Ovaskainen *et al.* 2010). This strongly indicates that the results obtained from the complex systems modelling approach are correct, and adds weight to any future results obtained from a similar modelling approach, when empirical comparisons may be impossible.

I demonstrate the Tangled Nature Models use in a scenario where empirical data is a limiting factor in **Chapter 5,** where I explore the effects of varying fragmentation *per se* on communities with different properties, using a variety of biodiversity metrics. Such a study could only be carried out using a theoretical approach since empirical studies looking at fragmentation *per se* have been suggested to have multiple limitations, and of course, the specific properties of real-world communities can never be fully known which limits our ability to uncover causal mechanisms. Using such an approach allows for the extraction of the underlying mechanisms causing community responses to fragmentation *per se*. In summary, complex systems are a promising avenue for ecological research, specifically for use in understanding interactions between species, and to model entire community responses to perturbations. Suggestions for future work using such approaches are outlined in the ‘Future Work’ section below.

Future Work

Here I outline future work that could be conducted to fill knowledge gaps left in the thesis, or studies which could utilise methods or analysis pipelines developed within the thesis. Below I break down the suggestions into the four chapters of my thesis, suggesting what work could be conducted, why it should be done, and the implications.

In **Chapter 2** I uncover a latitudinal gradient in the proportion of forest species, as well as identifying key factors which explain the global distribution of forest species. Future work could utilise the analysis pipeline created within this chapter to assess the global distribution of species associated with other ecotypes (grasslands, deserts), and therefore identify whether the latitudinal gradient in the proportion of forest species is unique to the habitat, or is a part of a larger latitudinal trend in specialisation. Results from such an analysis would have two main outputs. Firstly, identifying general trends in where specialist species occur can aid in the designation of protected areas, since habitat specialist species are usually more sensitive to habitat loss than generalist species (Betts *et al.* 2017, 2019). Secondly, results from such an analysis would aid in the identification of factors driving the well-known latitudinal biodiversity gradient (Allen & Gillooly 2006; Condamine *et al.* 2012; Fine & Ree 2006). For example, analysis in **Chapter 2** showed that forest specialist species are generally poor dispersers, suggesting that a higher incidence of forest species at the equator could be due to a lack of dispersal to higher latitudes after the last glacial maxima, however open-area specialists, such as grassland species, are generally good dispersers (Sheard *et al.* 2020), and therefore any latitudinal gradient in their distribution is unlikely to be due to dispersal limitations, ruling out the hypothesis as a driver of the latitudinal biodiversity gradient.

In **Chapter 3** I assessed how the position of ecological thresholds varied with latitude and regional deforestation. Whilst I found a clear relationship between the position of the threshold and latitude as well as regional deforestation, other factors are likely to, and have previously been found to, alter the position of the threshold (Boesing *et al.* 2018; Hatfield *et*

al. 2020). The permeability of the matrix, as well as the connectivity of the forest patches (fragmentation) at both the regional and landscape scale are known to alter the effect of deforestation on species and communities (Bellotto-Trigo *et al.* 2023; Boesing *et al.* 2018; Hatfield *et al.* 2020; Pardini *et al.* 2010). Future analysis should include these factors, which may have an impact on threshold position, and therefore influence conservation actions based on results. In **Chapter 3** I also propose a theoretical framework for a general relationship between community integrity and forest cover (Figure 3.3). In the analysis conducted, I was only able to find limited support for this, which largely seemed to occur due to data and methodological limitations; many studies had few sites with low sampling effort, and forest cover data was taken from a relatively crude source (Hansen *et al.* 2013). Future work should utilise a likely smaller set of high-quality studies which have well-sampled biodiversity data from across the gradient of deforestation, coupled with high-quality land cover data, to test whether this theoretical framework is globally consistent.

In **Chapter 4** and **Chapter 5** I utilise a complex systems modelling approach (Tangled Nature Model) to assess the influence of deforestation on species interactions, and identify the effect of fragmentation *per se* on ecological communities with varying properties. This approach could be used to answer a myriad of ecological questions for which empirical data is currently lacking. Firstly, it could be used to identify optimal restoration practices, both in terms of when to restore an area, and the optimal configuration. In my thesis both **Chapter 3** and **Chapter 4** suggest that reforesting intermediately forested areas is optimal for conservation, however using the Tangled Nature Model this could be tested. Furthermore, as in **Chapter 5** different spatial configurations of reforestation could be assessed, alongside communities which varied in their properties, ultimately creating a theoretical guide of when and how to reforest based on the properties of the ecological community in question. Secondly, the Tangled Nature Model could be parameterised based on ecological theory to make outputs more realistic and therefore useful to furthering our understanding of ecosystems and conservation practices. One potential way to do this is to implement scaling

laws from the metabolic theory of ecology (Brown *et al.* 2004; Gillooly *et al.* 2001; Pawar *et al.* 2012), which could be used to infer lifespan, fecundity, dispersal, and most importantly interactions between species (Pawar *et al.* 2019; Rall *et al.* 2012; Savage *et al.* 2004).

Summary

In conclusion, I have shown that communities vary in their sensitivity to deforestation, both at large scales (**Chapters 2 and 3**)**,** as well as at local scales (**Chapters 4 and 5**). This variation is driven by differences in the species present within communities, which themselves differ in their sensitivity to deforestation, primarily due to habitat specialisation and dispersal ability. My findings suggest that the protection of forested areas should be prioritised in undisturbed equatorial regions where communities are most sensitive, and that reforestation of intermediately deforested areas should be prioritised over completely deforested areas. Lastly, I demonstrated the use of complex systems modelling approaches to answer ecological questions for which empirical data is scarce, jointly highlighting how interactions change with deforestation, and that results are comparable to empirical observations, and then using a similar approach to tackle the long-standing SLOSS debate.

References

Aggemyr, E., Auffret, A.G., Jädergård, L. & Cousins, S.A.O. (2018). Species richness and composition differ in response to landscape and biogeography. *Landsc. Ecol.*, 33, 2273–2284.

Allan, J.R., Possingham, H.P., Atkinson, S.C., Waldron, A., Di Marco, M., Butchart, S.H.M., *et al.* (2022). The minimum land area requiring conservation attention to safeguard biodiversity. *Science*, 376, 1094–1101.

Allen, A.P. & Gillooly, J.F. (2006). Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. *Ecol. Lett.*, 9, 947–954.

Allesina, S. & Tang, S. (2012). Stability criteria for complex ecosystems. *Nature*, 483, 205–208.

Anand, M., Gonzalez, A., Guichard, F., Kolasa, J. & Parrott, L. (2010). Ecological Systems as Complex Systems: Challenges for an Emerging Science. *Diversity*, 2, 395–410.

Anderson, P.E. & Jensen, H.J. (2005). Network properties, species abundance and evolution in a model of evolutionary ecology. *J. Theor. Biol.*, 232, 551–558.

Anjos, L. dos, Bochio, G.M., Medeiros, H.R., Almeida, B. de A., Lindsey, B.R.A., Calsavara, L.C., *et al.* (2019). Insights on the functional composition of specialist and generalist birds throughout continuous and fragmented forests. *Ecol. Evol.*, 9, 6318–6328.

Arroyo-Rodríguez, V., Fahrig, L., Tabarelli, M., Watling, J.I., Tischendorf, L., Benchimol, M., *et al.* (2020). Designing optimal human-modified landscapes for forest biodiversity conservation. *Ecol. Lett.*, 23, 1404–1420.

Arthur, R., Nicholson, A., Sibani, P. & Christensen, M. (2017). The Tangled Nature Model for organizational ecology. *Comput. Math. Organ. Theory*, 23, 1–31.

Ávila-Gómez, E.S., Moreno, C.E., García-Morales, R., Zuria, I., Sánchez-Rojas, G. & Briones-Salas, M. (2015). Deforestation Thresholds for Phyllostomid Bat Populations in Tropical Landscapes in the Huasteca Region, Mexico. *Trop. Conserv. Sci.*, 8, 646–661.

Banks-Leite, C., Betts, M.G., Ewers, R.M., Orme, C.D.L. & Pigot, A.L. (2022). The macroecology of landscape ecology. *Trends Ecol. Evol.*, 0.

Banks-Leite, C., Larrosa, C., Carrasco, L.R., Tambosi, L.R. & Milner-Gulland, E.J. (2021). The suggestion that landscapes should contain 40% of forest cover lacks evidence and is problematic. *Ecol. Lett.*, 24, 1112–1113.

Banks-Leite, C., Pardini, R., Tambosi, L.R., Pearse, W.D., Bueno, A.A., Bruscagin, R.T., *et al.* (2014). Using ecological thresholds to evaluate the costs and benefits of set-asides in a biodiversity hotspot. *Science*, 345, 1041–1045.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.*, 67, 1–48.

Bellotto-Trigo, F.C., Uezu, A., Hatfield, J.H., Morante-Filho, J.C., dos Anjos, L., Develey, P.F., *et al.* (2023). Intraspecific variation in sensitivity to habitat fragmentation is influenced by forest cover and distance to the range edge. *Biol. Conserv.*, 284, 110167.

Benstead, J.P. & Pringle, C.M. (2004). Deforestation alters the resource base and biomass of endemic stream insects in eastern Madagascar. *Freshw. Biol.*, 49, 490–501.

Bentham, R. (2011). Macroevolution and Climate Change: A Mathematical Study with the Tangled Nature Model. Imperial College London.

Berlow, E.L. (1999). Strong effects of weak interactions in ecological communities. *Nature*, 398, 330–334.

Bertness, M.D. & Callaway, R. (1994). Positive interactions in communities. *Trends Ecol. Evol.*, 9, 191–193.

Betts, M.G., Fahrig, L., Hadley, A.S., Halstead, K.E., Bowman, J., Robinson, W.D., *et al.* (2014). A species-centered approach for uncovering generalities in organism responses to habitat loss and fragmentation. *Ecography*, 37, 517–527.

Betts, M.G., Forbes, G.J. & Diamond, A.W. (2007). Thresholds in Songbird Occurrence in

Relation to Landscape Structure. *Conserv. Biol.*, 21, 1046–1058.
Betts, M.G., Hagar, J.C., Rivers, J.W., Alexander, J.D., McGarigal, K. & McComb, B.C. (2010). Thresholds in forest bird occurrence as a function of the amount of early-seral broadleaf forest at landscape scales. *Ecol. Appl.*, 20, 2116–2130.
Betts, M.G., Wolf, C., Pfeifer, M., Banks-Leite, C., Arroyo-Rodríguez, V., Ribeiro, D.B., *et al.* (2019). Extinction filters mediate the global effects of habitat fragmentation on animals. *Science*, 366, 1236–1239.
Betts, M.G., Wolf, C., Ripple, W.J., Phalan, B., Millers, K.A., Duarte, A., *et al.* (2017). Global forest loss disproportionately erodes biodiversity in intact landscapes. *Nature*, 547, 441–444.
Billick, I. & Tonkel, K. (2003). The relative imporance of spatial vs temporal variability in generating a conditional mutualism. *Ecology*, 84, 289–295.
BirdLife Data Zone. (2021). . Available at: http://datazone.birdlife.org/species/search. Last accessed 2 June 2021.
Blanchet, F.G., Cazelles, K. & Gravel, D. (2020). Co‐occurrence is not evidence of ecological interactions. *Ecol. Lett.*, 23, 1050–1063.
Boesing, A.L., Nichols, E. & Metzger, J.P. (2018). Biodiversity extinction thresholds are modulated by matrix type. *Ecography*, 41, 1520–1533.
Bolgovics, Á., B-Béres, V., Várbíró, G., Krasznai-K, E.Á., Ács, É., Kiss, K.T., *et al.* (2019). Groups of small lakes maintain larger microalgal diversity than large ones. *Sci. Total Environ.*, 678, 162–172.
Boscolo, D. & Metzger, J.P. (2009). Is bird incidence in Atlantic forest fragments influenced by landscape patterns at multiple scales? *Landsc. Ecol.*, 24, 907–918.
Boulinier, T., Nichols, J.D., Hines, J.E., Sauer, J.R., Flather, C.H. & Pollock, K.H. (1998). Higher temporal variability of forest breeding bird communities in fragmented landscapes. *Proc. Natl. Acad. Sci.*, 95, 7497–7501.
Bracewell, S.A., Clark, G.F. & Johnston, E.L. (2018). Habitat complexity effects on diversity and abundance differ with latitude: an experimental study over 20 degrees. *Ecology*, 99, 1964–1974.
Bregman, T.P., Sekercioglu, C.H. & Tobias, J.A. (2014). Global patterns and predictors of bird species responses to forest fragmentation: Implications for ecosystem function and conservation. *Biol. Conserv.*, 169, 372–383.
Brinck, K. (2018). Dynamics of life: Self organisation, co-adaptation & the evolution of resilience in complex ecosystems. Imperial College London.
Brinck, K. & Jensen, H.J. (2017). The evolution of ecosystem ascendency in a complex systems based model. *J. Theor. Biol.*, 428, 18–25.
Brooks, E, M., Kristensen, Kasper, van Benthem, J, K., *et al.* (2017). Modeling Zero-Inflated Count Data With glmmTMB. *BioRxiv Prepr. BioRxiv132753*.
Brown, J.H. (2014). Why are there so many species in the tropics? *J. Biogeogr.*, 41, 8–22.
Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M. & West, G.B. (2004). Toward a Metabolic Theory of Ecology. *Ecology*, 85, 1771–1789.
Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L. & Smets, B. (2020). Copernicus Global Land Cover Layers—Collection 2. *Remote Sens.*, 12, 1044.
Buchmann, C.M., Schurr, F.M., Nathan, R. & Jeltsch, F. (2013). Habitat loss and fragmentation affecting mammal and bird communities—The role of interspecific competition and individual space use. *Ecol. Inform.*, The analysis and application of spatial ecological data to support the conservation of biodiversity, 14, 90–98.
Bunker, D.E., DeClerck, F., Bradford, J.C., Colwell, R.K., Perfecto, I., Phillips, O.L., *et al.* (2005). Species Loss and Aboveground Carbon Storage in a Tropical Forest. *Science*, 310, 1029–1031.
Cairoli, A., Piovani, D. & Jensen, H.J. (2014). Forecasting Transitions in Systems with High-Dimensional Stochastic Complex Dynamics: A Linear Stability Analysis of the Tangled Nature Model. *Phys. Rev. Lett.*, 113, 264102.
Callaway, R.M. (2007). *Positive interactions and interdependence in plant communities*. Springer, Dordrecht.

Callaway, R.M., Brooker, R.W., Choler, P., Kikvidze, Z., Lortie, C.J., Michalet, R., *et al.* (2002). Positive interactions among alpine plants increase with stress. *Nature*, 417, 844–848.

Cao, K., Condit, R., Mi, X., Chen, L., Ren, H., Xu, W., *et al.* (2021). Species packing and the latitudinal gradient in beta-diversity. *Proc. R. Soc. B Biol. Sci.*, 288, 20203045.

Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P., *et al.* (2012). Biodiversity loss and its impact on humanity. *Nature*, 486, 59–67.

Carlson, A. (2000). The effect of habitat loss on a deciduous forest specialist species: the White-backed Woodpecker (Dendrocopos leucotos). *For. Ecol. Manag.*, 131, 215–221.

Carrara, E., Arroyo-Rodríguez, V., Vega-Rivera, J.H., Schondube, J.E., de Freitas, S.M. & Fahrig, L. (2015). Impact of landscape composition and configuration on forest specialist and generalist bird species in the fragmented Lacandona rainforest, Mexico. *Biol. Conserv.*, 184, 117–126.

Case, T.J., Holt, R.D., McPeek, M.A. & Keitt, T.H. (2005). The community context of species' borders: ecological and evolutionary perspectives. *Oikos*, 108, 28–46.

Case, T.J. & Taper, M.L. (2000). Interspecific Competition, Environmental Gradients, Gene Flow, and the Coevolution of Species' Borders. *Am. Nat.*, 155, 583–605.

Chamberlain, S.A., Bronstein, J.L. & Rudgers, J.A. (2014). How context dependent are species interactions? *Ecol. Lett.*, 17, 881–890.

Chen, Y. & Peng, S. (2017). Evidence and mapping of extinction debts for global forest-dwelling reptiles, amphibians and mammals. *Sci. Rep.*, 7, 44305.

Christensen, K., Di Collobiano, S.A., Hall, M. & Jensen, H.J. (2002). Tangled Nature: A Model of Evolutionary Ecology. *J. Theor. Biol.*, 216, 73–84.

Clark, N.J., Wells, K. & Lindberg, O. (2018). Unravelling changing interspecific interactions across environmental gradients using Markov random fields. *Ecology*, 99, 1277–1283.

Collobiano, S.A. di, Christensen, K. & Jensen, H.J. (2003). The tangled nature model as an evolving quasi-species model. *J. Phys. Math. Gen.*, 36, 883–891.

Condamine, F.L., Sperling, F.A.H., Wahlberg, N., Rasplus, J.-Y. & Kergoat, G.J. (2012). What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. *Ecol. Lett.*, 15, 267–277.

Coyte, K.Z., Schluter, J. & Foster, K.R. (2015). The ecology of the microbiome: Networks, competition, and stability. *Science*, 350, 663–666.

Cullen, R. (2012). Biodiversity protection prioritisation: a 25-year review. *Wildl. Res.*, 40, 108–116.

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

Cushman, S.A. (2006). Effects of habitat loss and fragmentation on amphibians: A review and prospectus. *Biol. Conserv.*, 128, 231–240.

Cushman, S.A., Shirk, A. & Landguth, E.L. (2012). Separating the effects of habitat area, fragmentation and matrix resistance on genetic differentiation in complex landscapes. *Landsc. Ecol.*, 27, 369–380.

Dangles, O., Herrera, M. & Anthelme, F. (2013). Experimental support of the stress-gradient hypothesis in herbivore-herbivore interactions. *New Phytol.*, 197, 405–408.

Darwin, C. (1859). *The Origin of Species: By Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. 6th edn. Cambridge University Press.

Daskin, J.H. & Alford, R.A. (2012). Context-dependent symbioses and their potential roles in wildlife diseases. *Proc. R. Soc. B Biol. Sci.*, 279, 1457–1465.

Deane, D.C., Nozohourmehrabad, P., Boyce, S.S.D. & He, F. (2020). Quantifying factors for understanding why several small patches host more species than a single large patch. *Biol. Conserv.*, 249, 108711.

Dell, A.I., Pawar, S. & Savage, V.M. (2014). Temperature dependence of trophic interactions are driven by asymmetry of species responses and foraging strategy. *J. Anim. Ecol.*, 83, 70–84.

Diamond, J.M. (1975). The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. *Biol. Conserv.*, 7, 129–146.
Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N.D., Wikramanayake, E., *et al.* (2017). An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm. *BioScience*, 67, 534–545.
DiRenzo, G.V., Hanks, E. & Miller, D.A.W. (2023). A practical guide to understanding and validating complex models using data simulations. *Methods Ecol. Evol.*, 14, 203–217.
Dornelas, M., Antão, L.H., Moyes, F., Bates, A.E., Magurran, A.E., Adam, D., *et al.* (2018). BioTIME: A database of biodiversity time series for the Anthropocene. *Glob. Ecol. Biogeogr.*, 27, 760–786.
E. Smith, J. & O. Batzli, G. (2006). Dispersal and mortality of prairie voles (Microtus ochrogaster) in fragmented landscapes: a field experiment. *Oikos*, 112, 209–217.
Ehbrecht, M., Seidel, D., Annighöfer, P., Kreft, H., Köhler, M., Zemp, D.C., *et al.* (2021). Global patterns and climatic controls of forest structural complexity. *Nat. Commun.*, 12, 519.
Estavillo, C., Pardini, R. & Rocha, P.L.B. da. (2013). Forest Loss and the Biodiversity Threshold: An Evaluation Considering Species Habitat Requirements and the Use of Matrix Habitats. *PLOS ONE*, 8, e82369.
Ewers, R.M. & Didham, R.K. (2006). Confounding factors in the detection of species responses to habitat fragmentation. *Biol. Rev. Camb. Philos. Soc.*, 81, 117–142.
Ewers, R.M., Kapos, V., Coomes, D.A., Lafortezza, R. & Didham, R.K. (2009). Mapping community change in modified landscapes. *Biol. Conserv.*, 142, 2872–2880.
Fahrig, L. (2001). How much habitat is enough? *Biol. Conserv.*, 100, 65–74.
Fahrig, L. (2017). Ecological Responses to Habitat Fragmentation Per Se. *Annu. Rev. Ecol. Evol. Syst.*, 48, 1–23.
Fahrig, L. (2020). Why do several small patches hold more species than few large patches? *Glob. Ecol. Biogeogr.*, 29, 615–628.
Fahrig, L. (2021). What the habitat amount hypothesis does and does not predict: A reply to Saura. *J. Biogeogr.*, 48, 1530–1535.
Fahrig, L., Arroyo-Rodríguez, V., Bennett, J.R., Boucher-Lalonde, V., Cazetta, E., Currie, D.J., *et al.* (2019). Is habitat fragmentation bad for biodiversity? *Biol. Conserv.*, 230, 179–186.
Fahrig, L., Watling, J.I., Arnillas, C.A., Arroyo-Rodríguez, V., Jörger-Hickfang, T., Müller, J., *et al.* (2022). Resolving the SLOSS dilemma for biodiversity conservation: a research agenda. *Biol. Rev.*, 97, 99–114.
FAO (Ed.). (2020). *Forest, biodiversity and people*. State of the world's forests. FAO, Rome.
Farmer, J.D., Gallegati, M., Hommes, C., Kirman, A., Ormerod, P., Cincotti, S., *et al.* (2012). A complex systems approach to constructing better models for managing financial markets and the economy. *Eur. Phys. J. Spec. Top.*, 214, 295–324.
Ferger, S.W., Schleuning, M., Hemp, A., Howell, K.M. & Böhning-Gaese, K. (2014). Food resources and vegetation structure mediate climatic effects on species richness of birds. *Glob. Ecol. Biogeogr.*, 23, 541–549.
Ferrier, S. & Guisan, A. (2006). Spatial modelling of biodiversity at the community level. *J. Appl. Ecol.*, 43, 393–404.
Fine, P.V.A. (2015). Ecological and Evolutionary Drivers of Geographic Variation in Species Diversity. *Annu. Rev. Ecol. Evol. Syst.*, 46, 369–392.
Fine, P.V.A. & Ree, R.H. (2006). Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. *Am. Nat.*, 168, 796–804.
Fleishman, E., Noss, R.F. & Noon, B.R. (2006). Utility and limitations of species richness metrics for conservation planning. *Ecol. Indic.*, 6, 543–553.
Fletcher, R.J., Didham, R.K., Banks-Leite, C., Barlow, J., Ewers, R.M., Rosindell, J., *et al.* (2018). Is habitat fragmentation good for biodiversity? *Biol. Conserv.*, 226, 9–15.
Fox, J. & Morin, P. (2001). Effects of intra- and interspecific interactions on species responses to environmental change. *J. Anim. Ecol.*, 70, 80–90.
Friedl, Mark & Sulla-Menashe, Damien. (2022). MODIS/Terra+Aqua Land Cover Type

Yearly L3 Global 500m SIN Grid V061.
Friedman, J., Hastie, T. & Tibshirania, R. (2019). glasso: Graphical Lasso: Estimation of Gaussian Graphical Models.
Furlani, D., Ficetola, G.F., Colombo, G., Ugurlucan, M. & Bernardi, F.D. (2009). Deforestation and the Structure of Frog Communities in the Humedale Terraba-Sierpe, Costa Rica. *Zoolog. Sci.*, 26, 197–202.
Gandu, A.W., Cohen, J.C.P. & de Souza, J.R.S. (2004). Simulation of deforestation in eastern Amazonia using a high-resolution model. *Theor. Appl. Climatol.*, 78.
García, F.C., Clegg, T., O'Neill, D.B., Warfield, R., Pawar, S. & Yvon-Durocher, G. (2023). The temperature dependence of microbial community respiration is amplified by changes in species interactions. *Nat. Microbiol.*, 8, 272–283.
Gaston, K.J. (2007). Latitudinal gradient in species richness. *Curr. Biol.*, 17, R574.
Giam, X. (2017). Global biodiversity loss from tropical deforestation. *Proc. Natl. Acad. Sci.*, 114, 5775–5777.
Gillooly, J.F., Brown, J.H., West, G.B., Savage, V.M. & Charnov, E.L. (2001). Effects of Size and Temperature on Metabolic Rate. *Science*, 293, 2248–2251.
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.*, Big Remotely Sensed Data: tools, applications and experiences, 202, 18–27.
Grigaltchik, V.S., Ward, A.J.W. & Seebacher, F. (2012). Thermal acclimation of interactions: differential responses to temperature change alter predator–prey relationship. *Proc. R. Soc. B Biol. Sci.*, 279, 4058–4064.
Gruber, B. & Henle, K. (2008). Analysing the effect of movement on local survival: a new method with an application to a spatially structured population of the arboreal gecko Gehyra variegata. *Oecologia*, 154, 679–690.
Gutiérrez-Arellano, C. & Mulligan, M. (2020). Small-sized protected areas contribute more per unit area to tropical crop pollination than large protected areas. *Ecosyst. Serv.*, 44, 101137.
Haas, A.R., Kross, S.M. & Kneitel, J.M. (2020). Avian community composition, but not richness, differs between urban and exurban parks. *J. Urban Ecol.*, 6, juaa028.
Haddad, N.M., Brudvig, L.A., Clobert, J., Davies, K.F., Gonzalez, A., Holt, R.D., *et al.* (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. *Sci. Adv.*, 1, e1500052.
Hagan, J.M., Haegen, W.M.V. & McKinley, P.S. (1996). The Early Development of Forest Fragmentation Effects on Birds. *Conserv. Biol.*, 10, 188–202.
Hagen, M., Kissling, W.D., Rasmussen, C., De Aguiar, M.A.M., Brown, L.E., Carstensen, D.W., *et al.* (2012). Biodiversity, Species Interactions and Ecological Networks in a Fragmented World. In: *Advances in Ecological Research*. Elsevier, pp. 89–210.
Hall, M., Christensen, K., di Collobiano, S.A. & Jeldtoft Jensen, H. (2002). Time-dependent extinction rate and species abundance in a tangled-nature model of biological evolution. *Phys. Rev. E*, 66, 011904.
Halstead, K.E., Alexander, J.D., Hadley, A.S., Stephens, J.L., Yang, Z. & Betts, M.G. (2019). Using a species-centered approach to predict bird community responses to habitat fragmentation. *Landsc. Ecol.*, 34, 1919–1935.
Haney, A., Apfelbaum, S. & Burris, J.M. (2008). Thirty Years of Post-fire Succession in a Southern Boreal Forest Bird Community. *Am. Midl. Nat.*, 159, 421–433.
Hannah, L., Roehrdanz, P.R., Marquet, P.A., Enquist, B.J., Midgley, G., Foden, W., *et al.* (2020). 30% land conservation and climate action reduces tropical extinction risk by more than 50%. *Ecography*, 43, 943–953.
Hansen, A.J. & Urban, D.L. (1992). Avian response to landscape pattern: The role of species' life histories. *Landsc. Ecol.*, 7, 163–180.
Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., *et al.* (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 342, 850–853.

Harris, D.J. (2016). Inferring species interactions from co-occurrence data with Markov networks. *Ecology*, 97, 3308–3314.
Hartig, F. (2022). *DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models*.
Hatfield, J.H., Barlow, J., Joly, C.A., Lees, A.C., Parruco, C.H. de F., Tobias, J.A., *et al.* (2020). Mediation of area and edge effects in forest fragments by adjacent land use. *Conserv. Biol.*, 34, 395–404.
Hatfield, J.H., Orme, C.D.L. & Banks-Leite, C. (2018a). Using functional connectivity to predict potential meta-population sizes in the Brazilian Atlantic Forest. *Perspect. Ecol. Conserv.*, 16, 215–220.
Hatfield, J.H., Orme, C.D.L., Tobias, J.A. & Banks‐Leite, C. (2018b). Trait-based indicators of bird species sensitivity to habitat loss are effective within but not across data sets. *Ecol. Appl.*, 28, 28–34.
He, F. & Legendre, P. (2002). Species Diversity Patterns Derived from Species–Area Models. *Ecology*, 83, 1185–1198.
He, Q. & Bertness, M.D. (2014). Extreme stresses, niches, and positive species interactions along stress gradients. *Ecology*, 95, 1437–1443.
He, Q., Bertness, M.D. & Altieri, A.H. (2013). Global shifts towards positive species interactions with increasing environmental stress. *Ecol. Lett.*, 16, 695–706.
Henle, K., Davies, K.F., Kleyer, M., Margules, C. & Settele, J. (2004). Predictors of Species Sensitivity to Fragmentation. *Biodivers. Conserv.*, 13, 207–251.
Hesselbarth, M.H.K., Sciaini, M., With, K.A., Wiegand, K. & Nowosad, J. (2019). landscapemetrics: an open-source R tool to calculate landscape metrics. *Ecography*, 42, 1648–1657.
Hillebrand, H. (2004). On the Generality of the Latitudinal Diversity Gradient. *Am. Nat.*, 163, 192–211.
Holland, J.D., Bert, D.G. & Fahrig, L. (2004). Determining the Spatial Scale of Species' Response to Habitat. *BioScience*, 54, 227–233.
Holland, J.D., Fahrig, L. & Cappuccino, N. (2005). Body size affects the spatial scale of habitat–beetle interactions. *Oikos*, 110, 101–108.
Holovatch, Y., Kenna, R. & Thurner, S. (2017). Complex systems: physics beyond physics. *Eur. J. Phys.*, 38, 023002.
Homan, R.N., Windmiller, B.S. & Reed, J.M. (2004). Critical Thresholds Associated with Habitat Loss for Two Vernal Pool-Breeding Amphibians. *Ecol. Appl.*, 14, 1547–1553.
Hudson, L.N., Newbold, T., Contu, S., Hill, S.L.L., Lysenko, I., De Palma, A., *et al.* (2017). The database of the PREDICTS (Projecting Responses of Ecological Diversity In Changing Terrestrial Systems) project. *Ecol. Evol.*, 7, 145–188.
Hughes, T.P., Linares, C., Dakos, V., van de Leemput, I.A. & van Nes, E.H. (2013). Living dangerously on borrowed time during slow, unrecognized regime shifts. *Trends Ecol. Evol.*, 28, 149–155.
Humboldt, A.V. (1849). *Ansichten der Natur : mit wissenschaftlichen Erläuterungen /*. J.G. Cotta, Stuttgart :
IPBES, I. (2019). *Summary for policymakers of the global assessment report on biodiversity and ecosystem services*.
Jablonski, D., Roy, K. & Valentine, J.W. (2006). Out of the Tropics: Evolutionary Dynamics of the Latitudinal Diversity Gradient. *Science*, 314, 102–106.
Jankowski, J.E., Merkord, C.L., Rios, W.F., Cabrera, K.G., Revilla, N.S. & Silman, M.R. (2013). The relationship of tropical bird communities to tree species composition and vegetation structure along an Andean elevational gradient. *J. Biogeogr.*, 40, 950–962.
Janzen, D.H. (1967). Why Mountain Passes are Higher in the Tropics. *Am. Nat.*, 101, 233–249.
Jenkins, D.G., Brescacin, C.R., Duxbury, C.V., Elliott, J.A., Evans, J.A., Grablow, K.R., *et al.* (2007). Does size matter for dispersal distance? *Glob. Ecol. Biogeogr.*, 16, 415–425.
Jetz, W. & Fine, P.V.A. (2012). Global Gradients in Vertebrate Diversity Predicted by

Historical Area-Productivity Dynamics and Contemporary Environment. *PLOS Biol.*, 10, e1001292.
Kang, S., Ma, W., Li, F.Y., Zhang, Q., Niu, J., Ding, Y., *et al.* (2015). Functional Redundancy Instead of Species Redundancy Determines Community Stability in a Typical Steppe of Inner Mongolia. *PLOS ONE*, 10, e0145605.
Keinath, D.A., Doak, D.F., Hodges, K.E., Prugh, L.R., Fagan, W., Sekercioglu, C.H., *et al.* (2017). A global analysis of traits predicting species sensitivity to habitat fragmentation: Species sensitivity. *Glob. Ecol. Biogeogr.*, 26, 115–127.
Knowlton, J.L., Phifer, C.C., Cerqueira, P.V., Barro, F. de C., Oliveira, S.L., Fiser, C.M., *et al.* (2017). Oil Palm Plantations Affect Movement Behavior of a Key Member of Mixed-Species Flocks of Forest Birds in Amazonia, Brazil. *Trop. Conserv. Sci.*, 10, 1940082917692800.
Kofron, C.P. & Chapman, A. (1995). Deforestation and bird species composition in Liberia, West Africa. *Trop. Zool.*, 8, 239–256.
Kotliar, N.B., Kennedy, P.L. & Ferree, K. (2007). Avifaunal Responses to Fire in Southwestern Montane Forests Along a Burn Severity Gradient. *Ecol. Appl.*, 17, 491–507.
Laird, S. & Jensen, H.J. (2006). The Tangled Nature model with inheritance and constraint: Evolutionary ecology restricted by a conserved resource. *Ecol. Complex.*, 3, 253–262.
Laird, S., Lawson, D. & Jensen, H.J. (2008). The Tangled Nature Model of Evolutionary Ecology: An Overview. In: *Mathematical Modeling of Biological Systems, Volume II* (eds. Deutsch, A., Parra, R.B. de la, Boer, R.J. de, Diekmann, O., Jagers, P., Kisdi, E., et al.). Birkhäuser Boston, Boston, MA, pp. 49–62.
Lake, P.S. (2013). Resistance, Resilience and Restoration. *Ecol. Manag. Restor.*, 14, 20–24.
Landi, P., Minoarivelo, H.O., Brännström, Å., Hui, C. & Dieckmann, U. (2018). Complexity and stability of ecological networks: a review of the theory. *Popul. Ecol.*, 60, 319–345.
Laurance, W.F., Camargo, J.L.C., Luizão, R.C.C., Laurance, S.G., Pimm, S.L., Bruna, E.M., *et al.* (2011). The fate of Amazonian forest fragments: A 32-year investigation. *Biol. Conserv.*, 144, 56–67.
Lavorel, S., Flannigan, M.D., Lambin, E.F. & Scholes, M.C. (2007). Vulnerability of land systems to fire: Interactions among humans, climate, the atmosphere, and ecosystems. *Mitig. Adapt. Strateg. Glob. Change*, 12, 33–53.
Le Roux, D.S., Ikin, K., Lindenmayer, D.B., Manning, A.D. & Gibbons, P. (2015). Single large or several small? Applying biogeographic principles to tree-level conservation and biodiversity offsets. *Biol. Conserv.*, 191, 558–566.
Lechón-Alonso, P., Clegg, T., Cook, J., Smith, T.P. & Pawar, S. (2021). The role of competition versus cooperation in microbial community coalescence. *PLOS Comput. Biol.*, 17, e1009584.
Legendre, P. (1993). Spatial Autocorrelation: Trouble or New Paradigm? *Ecology*, 74, 1659–1673.
Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., *et al.* (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecol. Lett.*, 7, 601–613.
Lenth, R. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means*.
Liao, Z., Peng, S. & Chen, Y. (2022). Half-millennium evidence suggests that extinction debts of global vertebrates started in the Second Industrial Revolution. *Commun. Biol.*, 5, 1–8.
Lindenmayer, D.B., Wood, J., McBurney, L., Blair, D. & Banks, S.C. (2015). Single large versus several small: The SLOSS debate in the context of bird responses to a variable retention logging experiment. *For. Ecol. Manag.*, 339, 1–10.
Location of tropical cyclones. (2024). *Met Off.* Available at: https://www.metoffice.gov.uk/weather/learn-about/weather/types-of-weather/hurricanes/location. Last accessed 3 January 2024.
Loeuille, N. (2010). Influence of evolution on the stability of ecological communities. *Ecol.*

Lett., 13, 1536–1545.
Looijen, R.C. & van Andel, J. (1999). Ecological communities: conceptual problems and definitions. *Perspect. Plant Ecol. Evol. Syst.*, 2, 210–222.
Loreau, M. & de Mazancourt, C. (2013). Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecol. Lett.*, 16, 106–115.
Lurgi, M., Montoya, D. & Montoya, J.M. (2016). The effects of space and diversity of interaction types on the stability of complex ecological networks. *Theor. Ecol.*, 9, 3–13.
Macarthur, R.H. (1965). Patterns of Species Diversity. *Biol. Rev.*, 40, 510–533.
Mace, G., Masundire, H. & Baillie, J. (2005). Biodiversity. In: *Ecosystem and Human Well-Being Scenarios*.
Magrach, A., Laurance, W.F., Larrinaga, A.R. & Santamaria, L. (2014). Meta-Analysis of the Effects of Forest Fragmentation on Interspecific Interactions. *Conserv. Biol.*, 28, 1342–1348.
Malhi, Y., Roberts, J.T., Betts, R.A., Killeen, T.J., Li, W. & Nobre, C.A. (2008). Climate Change, Deforestation, and the Fate of the Amazon. *Science*, 319, 169–172.
Matthews, T.J., Cottee-Jones, H.E. & Whittaker, R.J. (2014). Habitat fragmentation and the species-area relationship: a focus on total species richness obscures the impact of habitat loss on habitat specialists. *Divers. Distrib.*, 20, 1136–1146.
Maxwell, S.L., Cazalis, V., Dudley, N., Hoffmann, M., Rodrigues, A.S.L., Stolton, S., *et al.* (2020). Area-based conservation in the twenty-first century. *Nature*, 586, 217–227.
McNeill, S.E. & Fairweather, P.G. (1993). Single Large or Several Small Marine Reserves? An Experimental Approach with Seagrass Fauna. *J. Biogeogr.*, 20, 429–440.
Melo, I., Ochoa‐Quintero, J.M., Roque, F. de O. & Dalsgaard, B. (2018). A review of threshold responses of birds to landscape changes across the world. *J. Field Ornithol.*, 89, 303–314.
Menegotto, A. & Rangel, T.F. (2018). Mapping knowledge gaps in marine diversity reveals a latitudinal gradient of missing species richness. *Nat. Commun.*, 9, 4713.
Michonneau, F., Brown, J.W. & Winter, D.J. (2016). rotl: an R package to interact with the Open Tree of Life data. *Methods Ecol. Evol.*, 7, 1476–1481.
Mittelbach, G.G., Schemske, D.W., Cornell, H.V., Allen, A.P., Brown, J.M., Bush, M.B., *et al.* (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol. Lett.*, 10, 315–331.
Mogg, S., Fastre, C., Jung, M. & Visconti, P. (2019). *Targeted expansion of Protected Areas to maximise the persistence of terrestrial mammals* (preprint). Ecology.
Mönkkönen, M., Forsman, J.T. & Thomson, R.L. (2004). Qualitative geographical variation in interspecific interactions. *Ecography*, 27, 112–118.
Morante-Filho, J.C., Faria, D., Mariano-Neto, E. & Rhodes, J. (2015). Birds in Anthropogenic Landscapes: The Responses of Ecological Groups to Forest Loss in the Brazilian Atlantic Forest. *PLOS ONE*, 10, e0128923.
Morueta-Holme, N., Blonder, B., Sandel, B., McGill, B.J., Peet, R.K., Ott, J.E., *et al.* (2016). A network approach for inferring species associations from co-occurrence data. *Ecography*, 39, 1139–1150.
Muggeo, V. (2016). *Segmented mixed models with random changepoints in R*.
Muggeo, V.M. (2008). Segmented: an R package to fit regression models with broken-line relationships. *R News*, 8, 20–25.
Muggeo, V.M., Atkins, D.C., Gallop, R.J. & Dimidjian, S. (2014). Segmented mixed models with random changepoints: a maximum likelihood approach with application to treatment for depression study. *Stat. Model.*, 14, 293–313.
Niebuhr, B.B.S., Wosniack, M.E., Santos, M.C., Raposo, E.P., Viswanathan, G.M., da Luz, M.G.E., *et al.* (2015). Survival in patchy landscapes: the interplay between dispersal, habitat loss and fragmentation. *Sci. Rep.*, 5, 11898.
NOAA National Centers for Environmental Information. (2022). ETOPO 2022 15 Arc-Second Global Relief Model.
Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., *et al.* (2020).

vegan: Community Ecology Package.
Oliver, T.H., Heard, M.S., Isaac, N.J.B., Roy, D.B., Procter, D., Eigenbrod, F., *et al.* (2015). Biodiversity and Resilience of Ecosystem Functions. *Trends Ecol. Evol.*, 30, 673–684.
Orme, C.D.L., Mayor, S., dos Anjos, L., Develey, P.F., Hatfield, J.H., Morante-Filho, J.C., *et al.* (2019). Distance to range edge determines sensitivity to deforestation. *Nat. Ecol. Evol.*, 3, 886–891.
Ovaskainen, O. (2002). Long-Term Persistence of Species and the SLOSS Problem. *J. Theor. Biol.*, 218, 419–433.
Ovaskainen, O., Hottola, J. & Siitonen, J. (2010). Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. *Ecology*, 91, 2514–2521.
Pagani-Núñez, E., Xia, X., Beauchamp, G., He, R., Husson, J.H.D., Liang, D., *et al.* (2018). Are vocal characteristics related to leadership patterns in mixed-species bird flocks? *J. Avian Biol.*, 49, jav-01674.
Paiva, P.F.P.R., de Lourdes Pinheiro Ruivo, M., da Silva Júnior, O.M., de Nazaré Martins Maciel, M., Braga, T.G.M., de Andrade, M.M.N., *et al.* (2019). Deforestation in protect areas in the Amazon: a threat to biodiversity. *Biodivers. Conserv.*
Palmer, M.A., Menninger, H.L. & Bernhardt, E. (2010). River restoration, habitat heterogeneity and biodiversity: a failure of theory or practice? *Freshw. Biol.*, 55, 205–222.
Paquette, A. & Messier, C. (2011). The effect of biodiversity on tree productivity: from temperate to boreal forests. *Glob. Ecol. Biogeogr.*, 20, 170–180.
Paradis, E. & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35, 526–528.
Pardini, R., Bueno, A. de A., Gardner, T.A., Prado, P.I. & Metzger, J.P. (2010). Beyond the Fragmentation Threshold Hypothesis: Regime Shifts in Biodiversity Across Fragmented Landscapes. *PLoS ONE*, 5, e13666.
Pawar, S., Dell, A.I., Lin, T., Wieczynski, D.J. & Savage, V.M. (2019). Interaction Dimensionality Scales Up to Generate Bimodal Consumer-Resource Size-Ratio Distributions in Ecological Communities. *Front. Ecol. Evol.*, 7.
Pawar, S., Dell, A.I., & Van M. Savage. (2012). Dimensionality of consumer search space drives trophic interaction strengths. *Nature*, 486, 485–489.
Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *R J.*, 10, 439.
Pfeifer, M., Lefebvre, V., Gardner, T.A., Arroyo-Rodriguez, V., Baeten, L., Banks-Leite, C., *et al.* (2014). BIOFRAG – a new database for analyzing BIOdiversity responses to forest FRAGmentation. *Ecol. Evol.*, 4, 1524–1537.
Pfeifer, M., Lefebvre, V., Peres, C.A., Banks-Leite, C., Wearn, O.R., Marsh, C.J., *et al.* (2017). Creation of forest edges has a global impact on forest vertebrates. *Nature*, 551, 187–191.
Pillay, R., Venter, M., Aragon-Osejo, J., González-del-Pliego, P., Hansen, A.J., Watson, J.E., *et al.* (2022). Tropical forests are home to over half of the world's vertebrate species. *Front. Ecol. Environ.*, 20, 10–15.
Pimm, S.L. (1984). The complexity and stability of ecosystems. *Nature*, 307, 321–326.
Pimm, S.L. & Askins, R.A. (1995). Forest losses predict bird extinctions in eastern North America. *Proc. Natl. Acad. Sci.*, 92, 9343–9347.
Pontarp, M., Bunnefeld, L., Cabral, J.S., Etienne, R.S., Fritz, S.A., Gillespie, R., *et al.* (2019). The Latitudinal Diversity Gradient: Novel Understanding through Mechanistic Eco-evolutionary Models. *Trends Ecol. Evol.*, 34, 211–223.
Popovic, G.C., Warton, D.I., Thomson, F.J., Hui, F.K.C. & Moles, A.T. (2019). Untangling direct species associations from indirect mediator species effects with graphical models. *Methods Ecol. Evol.*, 10, 1571–1583.
Prist, P.R., Michalski, F. & Metzger, J.P. (2012). How deforestation pattern in the Amazon influences vertebrate richness and community composition. *Landsc. Ecol.*, 27,

799–812.
Pruitt, J.N. & Ferrari, M.C.O. (2011). Intraspecific trait variants determine the nature of interspecific interactions in a habitat-forming species. *Ecology*, 92, 1902–1908.
Püttker, T., Bueno, A.A., dos Santos de Barros, C., Sommer, S. & Pardini, R. (2011). Immigration Rates in Fragmented Landscapes – Empirical Evidence for the Importance of Habitat Amount for Species Persistence. *PLoS ONE*, 6, e27963.
Qian, J.J. & Akçay, E. (2019). The balance of interaction types determines the assembly and stability of ecological communities. *bioRxiv*, 643478.
Qian, J.J. & Akçay, E. (2020). The balance of interaction types determines the assembly and stability of ecological communities. *Nat. Ecol. Evol.*, 4, 356–365.
Quesnelle, P.E., Lindsay, K.E. & Fahrig, L. (2014). Low Reproductive Rate Predicts Species Sensitivity to Habitat Loss: A Meta-Analysis of Wetland Vertebrates. *PLOS ONE*, 9, e90926.
R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
Rall, B.C., Brose, U., Hartvig, M., Kalinkat, G., Schwarzmüller, F., Vucic-Pestic, O., *et al.* (2012). Universal temperature and body-mass scaling of feeding rates. *Philos. Trans. R. Soc. B Biol. Sci.*, 367, 2923–2934.
Rand, T.A. & Tscharntke, T. (2007). Contrasting effects of natural habitat loss on generalist and specialist aphid natural enemies. *Oikos*, 116, 1353–1362.
Ratzke, C., Barrere, J. & Gore, J. (2019). Strength of species interactions determines biodiversity and stability in microbial communities. *bioRxiv*, 671008.
Rees, J. & Cranston, K. (2017). Automated assembly of a reference taxonomy for phylogenetic data synthesis. *Biodivers. Data J.*, 5, e12581.
Ritchie, H. & Roser, M. (2023a). Deforestation and Forest Loss. *Our World Data*.
Ritchie, H. & Roser, M. (2023b). The world has lost one-third of its forest, but an end of deforestation is possible. *Our World Data*.
Riva, F., Graco-Roza, C., Daskalova, G.N., Hudgins, E.J., Lewthwaite, J.M.M., Newman, E.A., *et al.* (2023). Toward a cohesive understanding of ecological complexity. *Sci. Adv.*, 9, eabq4207.
Rivas-Salvador, J., Hořák, D. & Reif, J. (2019). Spatial patterns in habitat specialization of European bird communities. *Ecol. Indic.*, 105, 57–69.
Rooney, N., McCann, K., Gellner, G. & Moore, J.C. (2006). Structural asymmetry and the stability of diverse food webs. *Nature*, 442, 265–269.
Rubenstein, D.I., Barnett, R.J., Ridgely, R.S. & Klopfer, P.H. (2008). Adaptive advantages of mixed-species feeding flocks among seed-eating finches in costa rica. *Ibis*, 119, 10–21.
Rutt, C.L., Jirinec, V., Cohn‐Haft, M., Laurance, W.F. & Stouffer, P.C. (2019). Avian ecological succession in the Amazon: A long‐term case study following experimental deforestation. *Ecol. Evol.*, 9, 13850–13861.
Rybicki, J., Abrego, N. & Ovaskainen, O. (2020). Habitat fragmentation and species diversity in competitive communities. *Ecol. Lett.*, 23, 506–517.
Sabatini, F.M., Jiménez-Alfaro, B., Jandt, U., Chytrý, M., Field, R., Kessler, M., *et al.* (2022). Global patterns of vascular plant alpha diversity. *Nat. Commun.*, 13, 4683.
Salisbury, C.L., Seddon, N., Cooney, C.R. & Tobias, J.A. (2012). The latitudinal gradient in dispersal constraints: ecological specialisation drives diversification in tropical birds. *Ecol. Lett.*, 15, 847–855.
Sánchez González, I., Hopper, G.W., Bucholz, J.R., Kubala, M.E., Lozier, J.D. & Atkinson, C.L. (2023). Niche specialization and community niche space increase with species richness in filter-feeder assemblages. *Ecosphere*, 14, e4495.
Santos Junior, P.C.A., Marques, F.C., Lima, M.R. & dos Anjos, L. (2016). The importance of restoration areas to conserve bird species in a highly fragmented Atlantic forest landscape. *Nat. Conserv.*, 14, 1–7.

Saupe, E.E. (2023). Explanations for latitudinal diversity gradients must invoke rate variation. *Proc. Natl. Acad. Sci.*, 120, e2306220120.
Savage, V.M., Gillooly, J.F., Brown, J.H., West, G.B. & Charnov, E.L. (2004). Effects of Body Size and Temperature on Population Growth. *Am. Nat.*, 163, 429–441.
Schädler, M., Brandl, R. & Haase, J. (2007). Antagonistic interactions between plant competition and insect herbivory. *Ecology*, 88, 1490–1498.
Schafer, J., Opgen-Rhein, R., Zuber, V., Ahdesmaki, M., Duarte Silva, P. & Strimmer, K. (2017). corpcor: Efficient Estimation of Covariance and (Partial) Correlation.
Schumm, M., Edie, S.M., Collins, K.S., Gómez-Bahamón, V., Supriya, K., White, A.E., *et al.* (2019). Common latitudinal gradients in functional richness and functional evenness across marine and terrestrial systems. *Proc. R. Soc. B Biol. Sci.*, 286, 20190745.
Şekercioğlu, Ç.H., Daily, G.C. & Ehrlich, P.R. (2004). Ecosystem consequences of bird declines. *Proc. Natl. Acad. Sci.*, 101, 18042–18047.
Sheard, C., Neate-Clegg, M.H.C., Alioravainen, N., Jones, S.E.I., Vincent, C., MacGregor, H.E.A., *et al.* (2019). The latitudinal gradient in hand-wing-index: global patterns and predictors of wing morphology in birds.
Sheard, C., Neate-Clegg, M.H.C., Alioravainen, N., Jones, S.E.I., Vincent, C., MacGregor, H.E.A., *et al.* (2020). Ecological drivers of global gradients in avian dispersal inferred from wing morphology. *Nat. Commun.*, 11, 2463.
Shennan-Farpón, Y., Visconti, P. & Norris, K. (2021). Detecting ecological thresholds for biodiversity in tropical forests: Knowledge gaps and future directions. *Biotropica*, 53, 1276–1289.
Simpson, G.L. (2018). Modelling Palaeoecological Time Series Using Generalised Additive Models. *Front. Ecol. Evol.*, 6.
Slancarova, J., Benes, J., Kristynek, M., Kepka, P. & Konvicka, M. (2014). Does the surrounding landscape heterogeneity affect the butterflies of insular grassland reserves? A contrast between composition and configuration. *J. Insect Conserv.*, 18, 1–12.
Smith, B.T., Seeholzer, G.F., Harvey, M.G., Cuervo, A.M. & Brumfield, R.T. (2017). A latitudinal phylogeographic diversity gradient in birds. *PLOS Biol.*, 15, e2001073.
Sommer, R.S. & Nadachowski, A. (2006). Glacial refugia of mammals in Europe: evidence from fossil records. *Mammal Rev.*, 36, 251–265.
Soulé, M.E. & Sanjayan, M.A. (1998). ECOLOGY: Conservation Targets: Do They Help? *Science*, 279, 2060–2061.
Souza, C.M., Z. Shimbo, J., Rosa, M.R., Parente, L.L., A. Alencar, A., Rudorff, B.F.T., *et al.* (2020). Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. *Remote Sens.*, 12, 2735.
Sridhar, H., Beauchamp, G. & Shanker, K. (2009). Why do birds participate in mixed-species foraging flocks? A large-scale synthesis. *Anim. Behav.*, 78, 337–347.
Stouffer, D.B. & Bascompte, J. (2011). Compartmentalization increases food-web persistence. *Proc. Natl. Acad. Sci.*, 108, 3648–3652.
Stouffer, P.C. & Bierregaard, R.O. (1995). Use of Amazonian Forest Fragments by Understory Insectivorous Birds. *Ecology*, 76, 2429–2445.
Stouffer, P.C., Johnson, E.I., Jr, R.O.B. & Lovejoy, T.E. (2011). Understory Bird Communities in Amazonian Rainforest Fragments: Species Turnover through 25 Years Post-Isolation in Recovering Landscapes. *PLOS ONE*, 6, e20543.
Suding, K.N., Gross, K.L. & Houseman, G.R. (2004). Alternative states and positive feedbacks in restoration ecology. *Trends Ecol. Evol.*
Sunahara, T. & Mogi, M. (2002). Variability of intra- and interspecific competitions of bamboo stump mosquito larvae over small and large spatial scales. *Oikos*, 97, 87–96.
Sverdrup-Thygeson, A., Skarpaas, O., Blumentrath, S., Birkemoe, T. & Evju, M. (2017). Habitat connectivity affects specialist species richness more than generalists in veteran trees. *For. Ecol. Manag.*, 403, 96–102.
Swift, T.L. & Hannon, S.J. (2010). Critical thresholds associated with habitat loss: a review of the concepts, evidence, and applications. *Biol. Rev.*, 85, 35–53.

Tabarelli, M., Aguiar, A.V., Ribeiro, M.C., Metzger, J.P. & Peres, C.A. (2010). Prospects for biodiversity conservation in the Atlantic Forest: Lessons from aging human-modified landscapes. *Biol. Conserv.*, Conserving complexity: Global change and community-scale interactions, 143, 2328–2340.

Tabarelli, M., Pinto, L.P., Silva, J.M.C., Hirota, M. & Bedê, L. (2005). Challenges and Opportunities for Biodiversity Conservation in the Brazilian Atlantic Forest. *Conserv. Biol.*, 19, 695–700.

Tabor, K., Hewson, J., Tien, H., González-Roglich, M., Hole, D. & Williams, J.W. (2018). Tropical Protected Areas Under Increasing Threats from Climate Change and Deforestation. *Land*, 7, 90.

Tambosi, L.R., Martensen, A.C., Ribeiro, M.C. & Metzger, J.P. (2014). A Framework to Optimize Biodiversity Restoration Efforts Based on Habitat Amount and Landscape Connectivity. *Restor. Ecol.*, 22, 169–177.

Terraube, J., Archaux, F., Deconchat, M., van Halder, I., Jactel, H. & Barbaro, L. (2016). Forest edges have high conservation value for bird communities in mosaic landscapes. *Ecol. Evol.*, 6, 5178–5189.

Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M.C., Schwager, M., *et al.* (2004). Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. *J. Biogeogr.*, 31, 79–92.

The IUCN Red List of Threatened Species. (2021a). *IUCN Red List Threat. Species*. Available at: https://www.iucnredlist.org/en. Last accessed 2 June 2021.

The IUCN Red List of Threatened Species. (2023b). *IUCN Red List Threat. Species*. Available at: https://www.iucnredlist.org/en. Last accessed 6 April 2023.

Thebault, E. & Fontaine, C. (2010). Stability of Ecological Communities and the Architecture of Mutualistic and Trophic Networks. *Science*, 329, 853–856.

Thomas, C.D. (2000). Dispersal and extinction in fragmented landscapes. *Proc. R. Soc. Lond. B Biol. Sci.*, 267, 139–145.

Thompson, J.N. & Cunningham, B.M. (2002). Geographic structure and dynamics of coevolutionary selection. *Nature*, 417, 735–738.

Thornton, D.H. & Fletcher Jr, R.J. (2014). Body size and spatial scales in avian response to landscapes: a meta-analysis. *Ecography*, 37, 454–463.

Thuiller, W., Broennimann, O., Hughes, G., Alkemade, J.R.M., Midgley, Guy.F. & Corsi, F. (2006). Vulnerability of African mammals to anthropogenic climate change under conservative land transformation assumptions. *Glob. Change Biol.*, 12, 424–440.

Tilman, D., May, R.M., Lehman, C.L. & Nowak, M.A. (1994). Habitat destruction and the extinction debt. *Nature*, 371, 65–66.

Tjørve, E. (2010). How to resolve the SLOSS debate: Lessons from species-diversity models. *J. Theor. Biol.*, 264, 604–612.

Tobias, J.A., Sheard, C., Pigot, A.L., Devenish, A.J.M., Yang, J., Sayol, F., *et al.* (2022). AVONET: morphological, ecological and geographical data for all birds. *Ecol. Lett.*, 25, 581–597.

Tscharntke, T., Steffan-Dewenter, I., Kruess, A. & Thies, C. (2002). Contribution of Small Habitat Fragments to Conservation of Insect Communities of Grassland–Cropland Landscapes*. *Ecol. Appl.*, 12, 354–363.

Tucker, C.M., Davies, T.J., Cadotte, M.W. & Pearse, W.D. (2018). On the relationship between phylogenetic diversity and trait diversity. *Ecology*, 99, 1473–1479.

Ulrich, W., Banks‐Leite, C., Coster, G.D., Habel, J.C., Matheve, H., Newmark, W.D., *et al.* (2018). Environmentally and behaviourally mediated co‐occurrence of functional traits in bird communities of tropical forest fragments. *Oikos*, 127, 274–284.

University of Geneva, Switzerland, Ray, N. & Adams, J.M. (2001). A GIS-based Vegetation Map of the World at the Last Glacial Maximum (25,000-15,000 BP). *Internet Archaeol.*

Valiente‐Banuet, A., Aizen, M.A., Alcántara, J.M., Arroyo, J., Cocucci, A., Galetti, M., *et al.* (2015). Beyond species loss: the extinction of ecological interactions in a changing world. *Funct. Ecol.*, 29, 299–307.

Vargas-Cárdenas, F., Arroyo-Rodríguez, V., Morante-Filho, J.C., Schondube, J.E., Auliz-Ortiz, D.M. & Ceccon, E. (2022). Landscape forest loss decreases bird diversity with strong negative impacts on forest species in a mountain region. *Perspect. Ecol. Conserv.*, 20, 386–393.

Vásquez-Grandón, A., Donoso, P.J. & Gerding, V. (2018). Forest Degradation: When Is a Forest Degraded? *Forests*, 9, 726.

Virolainen, K.M., Suomi, T., Suhonen, J. & Kuitunen, M. (1998). Conservation of vascular plants in single large and several small mires: species richness, rarity and taxonomic diversity. *J. Appl. Ecol.*, 35, 700–707.

Waldron, A., Adams, V., Allan, J., Arnell, A., Asner, G., Atkinson, S., *et al.* (2020). Protecting 30% of the planet for nature: costs, benefits and economic implications.

Wallace, A.R. (1878). *Tropical Nature, and Other Essays*. Macmillan and Company.

Waltari, E., Hijmans, R.J., Peterson, A.T., Nyári, Á.S., Perkins, S.L. & Guralnick, R.P. (2007). Locating Pleistocene Refugia: Comparing Phylogeographic and Ecological Niche Model Predictions. *PLOS ONE*, 2, e563.

Watson, J.E.M., Whittaker, R.J. & Freudenberger, D. (2005). Bird community responses to habitat fragmentation: how consistent are they across landscapes?: Avifaunal responses to habitat fragmentation. *J. Biogeogr.*, 32, 1353–1370.

Weeks, T.L., Betts, M.G., Pfeifer, M., Wolf, C., Banks-Leite, C., Barbaro, L., *et al.* (2023). Climate-driven variation in dispersal ability predicts responses to forest fragmentation in birds. *Nat. Ecol. Evol.*, 7, 1079–1091.

Wiens, J.J. & Graham, C.H. (2005). Niche Conservatism: Integrating Evolution, Ecology, and Conservation Biology. *Annu. Rev. Ecol. Evol. Syst.*, 36, 519–539.

Williams, D.R., Rondinini, C. & Tilman, D. (2022). Global protected areas seem insufficient to safeguard half of the world's mammals from human-induced extinction. *Proc. Natl. Acad. Sci.*, 119, e2200118119.

Williams, J.J. & Newbold, T. (2021). Vertebrate responses to human land use are influenced by their proximity to climatic tolerance limits. *Divers. Distrib.*, n/a.

Willig, M.R., Kaufman, D.M. & Stevens, R.D. (2003). Latitudinal Gradients of Biodiversity: Pattern, Process, Scale, and Synthesis. *Annu. Rev. Ecol. Evol. Syst.*, 34, 273–309.

Willmer, J.N.G., Püttker, T. & Prevedello, J.A. (2022). Global impacts of edge effects on species richness. *Biol. Conserv.*, 272, 109654.

Wilson, J.D., Whittingham, M.J. & Bradbury, R.B. (2005). The management of crop structure: a general approach to reversing the impacts of agricultural intensification on birds? *Ibis*, 147, 453–463.

Wilson, M.C., Chen, X.-Y., Corlett, R.T., Didham, R.K., Ding, P., Holt, R.D., *et al.* (2016). Habitat fragmentation and biodiversity conservation: key findings and future challenges. *Landsc. Ecol.*, 31, 219–227.

Wood, S.N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. B*, 73, 3–36.

Wootton, K.L. & Stouffer, D.B. (2016). Many weak interactions and few strong; food-web feasibility depends on the combination of the strength of species' interactions and their correct arrangement. *Theor. Ecol.*, 9, 185–195.

WorldClim. (2024). . Available at: https://www.worldclim.org/. Last accessed 3 January 2024.

Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S.M., *et al.* (2018). A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. *ISPRS J. Photogramm. Remote Sens.*, 146, 108–123.

Zakaria, M., Rajpar, M.N., Ozdemir, I. & Rosli, Z. (2016). Fauna Diversity in Tropical Rainforest: Threats from Land-Use Change. In: *Tropical Forests - The Challenges of Maintaining Ecosystem Services while Managing the Landscape* (eds. Blanco, J.A., Chang, S.-C. & Lo, Y.-H.). InTech.

Zheng, F.-L. (2006). Effect of Vegetation Changes on Soil Erosion on the Loess Plateau. *Pedosphere*, 16, 420–427.

Zuur, A.F. (Ed.). (2009). *Mixed effects models and extensions in ecology with R*. Statistics for

biology and health. Springer, New York, NY.

Supplementary Material

Chapter 2: A global latitudinal gradient in the proportion of terrestrial vertebrate forest species

Model Structure

All predictors in models were scaled and centered to allow for direct comparison of effect sizes.

Latitude Model Per Taxa

bam(proportion_of_forest_species ~ latitude + s(x, y, bs = “gaussian process”) + s(ecoregion, bs = “random”), weights = total_species, family = binomial)

Latitude Model All Taxa

bam(proportion_of_forest_species ~ latitude + s(x, y, bs = “gaussian process”) + s(ecoregion, bs = “random”) + s(taxa, bs = “random”), weights = total_species, family = binomial)

Full Model Per Taxa

bam(proportion_of_forest_species ~ current_forest_cover + historical_deforestation + natural_disturbances + geological_forest_time + geologica_forest_stability + plant_alpha_diversity + latitude + s(x, y, bs = “gaussian process”) + s(ecoregion, bs = “random”), weights = total_species, family = binomial)

Full Model All Taxa

bam(proportion_of_forest_species ~ current_forest_cover + historical_deforestation + natural_disturbances + geological_forest_time + geologica_forest_stability + plant_alpha_diversity + latitude + s(x, y, bs = “gaussian process”) + s(ecoregion, bs = “random”) + s(taxa, bs = “random”), weights = total_species, family = binomial)

Hand Wing Index

Hand wing index data was obtained from Avonet (Tobias *et al.* 2022) and paired with our categorisation of bird species using IUCN habitat classification into forest and non-forest. Hand wing index is a well-studied proxy for dispersal ability (Sheard *et al.* 2020; Tobias *et al.* 2022; Weeks *et al.* 2023), and is calculated as the ratio of the Kipp's distance (the distance between the tip of the first secondary feather and the tip of the longest primary feather) to the total wing chord (Sheard *et al.* 2019). A linear regression was run using habitat classification (forest or non-forest) as a categorical explanatory variable, and hand wing index as the response. Forest species were found to have significantly lower hand wing indexes than non-forest species (Non-forest = 24.94, 95% CI (24.62, 25.26), Forest = 21.32 95% CI (20.84, 21.81)) suggesting forest species are generally poorer dispersers than non-forest species, at least in birds.

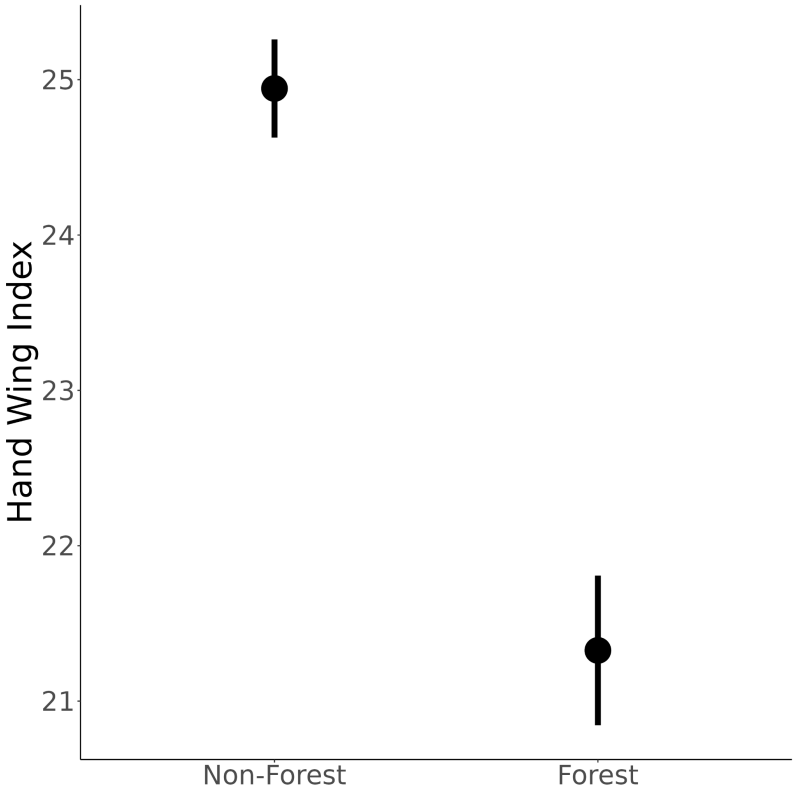


Figure S2.1. Forest species have on average a significantly smaller hand wing index than non-forest species, indicating that they are generally poorer dispersers. Estimates and 95%

confidence intervals of a linear model where habitat classification (forest or non-forest) was used as a predictor variable, and hand wing index as the response.

Figure S2.2 Model diagnostics for all models including models not accounting for spatial autocorrelation for comparison

https://doi.org/10.6084/m9.figshare.25054940.v1

Table S2.1 Table of coefficients for models

https://doi.org/10.6084/m9.figshare.25054955.v1

Chapter 3: Global Variation in Forest Thresholds for Community Persistence

Sensitivity Analysis

Forest Range of Studies

In the main text we used a cut-off value of 30% for the forest range of studies. So any study that did not have sites covering at least a 30% range of forest cover at the optimal scale of effect was excluded. We carried out analysis on how the position of the threshold varies with regional deforestation and latitude, as well as an analysis on how this changes with threshold type (Accelerating or Decelerating) using a 0% and a 60% cut-off to determine the impact of this choice. We present the model with the lowest AIC and informative parameters at 85% CI.

In all cases for both models, the optimal model was exactly the same, or very similar. Variables generally retained significance using 95% confidence intervals across cut-offs, and they always retained the same direction (positive or negative).

Table S3.1. Best model based on AIC and the corresponding coefficients (95% CI) for model predicting the position of thresholds. RD = Regional Deforestation, L = Latitude. Only variables that appeared in at least one model are included in columns, and main effects are only included for terms significant using 95% confidence intervals.

Forest range cut-off	**Data Points**	**Best Model**	**RD**	**L**	**RD:L**
0%	69	RD*L	-2.35 (-2/96, -1.74)	0.02 (0.02, 0.04)	-0.05 (-0.07, -0.03)
30%	53	RD*L	-1.81 (-2.50, -1.13)	0.05 (0.03,0.06)	-0.07 (-0.11, -0.04)
60%	25	RD*L	-1.75 (-3.41, -0.09)	(-0.00, 0.08)	-0.07 (-0.13, -0.00)

Table S3.2. Best model based on AIC and the corresponding coefficients (95% CI) for model predicting the position of thresholds including threshold type as a variable. RD = Regional Deforestation, L = Latitude, and T = Threshold Type (Accelerating or Decelerating). In all models Accelerating Thresholds were used as the reference type, and therefore estimates for RD, L, and RD:L all pertain to Accelerating Thresholds, whilst T, RD:T, L:T, and RD:L:T are all tests of whether these variables effect Decelerating Thresholds significantly differently from Accelerating Thresholds. Only variables that appeared in at least one model are included in columns, and main effects are only included for terms significant using 95% confidence intervals.

Forest range cut-off	**Data Points**	**Best Model**	**RD**	**L**	**T**	**RD:L**	**RD:T**	**L:T**	**RD:L:T**
0%	69	RD*L*T	-3.20 (-3.95, -2.46)	(-0.02, 0.01)	-0.60 (-1.03, -0.16)	(-0.03, 0.03)	1.96 (0.82, 3.10)	0.05 (0.03, 0.07)	-0.07 (-0.11, 0.03)
30%	53	RD*L*T	-3.06 (-3.95, -2.16)	(-0.00, , 0.06)	(-0.80, 0.16)	(-0.09, 0.01)	2.57 (1.35, 3.79)	(-0.03, 0.04)	(-0.08, 0.05)
60%	25	RD*L + T + T:RD	-3.42 (-5.20, -1.63)	(-0.00, 0.05)	(-1.53, 0.51)	(-0.09, 0.02)	2.49 (0.93, 4.04)	NA	NA

Table S3.3 Overview of each study, including data on number of sites, study taxa, optimal scale of effect, threshold position, threshold type, and all regional explanatory variables

https://figshare.com/s/64b2bea284d15ef883d4

Table S3.4 Number of control sites at each scale in each study

https://figshare.com/s/dd2a1ef08ebbc8d75ee6

Table S3.5 R2 for piecewise regression models fit at scales between 100 - 5000m for each study

https://figshare.com/s/b8029f0ac464cea6e67c

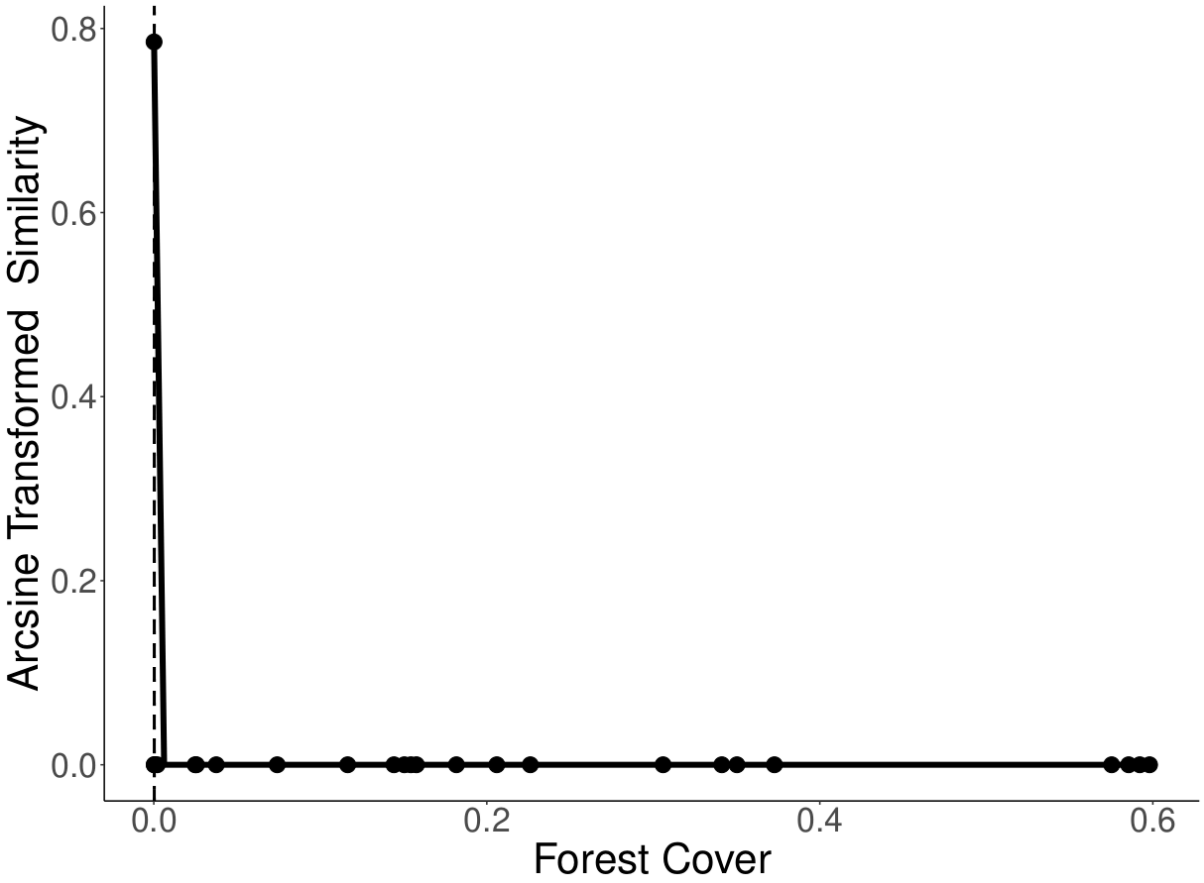


Figure S3.1. Piecewise regression plot for PID0073 which was removed for analysis due to erroneous model fit. This became clearer upon inspection of the model itself, which had an R^2 of one.

https://figshare.com/s/b67819ec0f485f0db430

Figure S3.2. Visualisation of piecewise regressions for all studies for which piecewise regression was found to have a lower AIC than its linear counterpart.

Chapter 4: Deforestation Alters Species Interactions

Table S4.1. Information on the five studies providing bird data. Coordinates are approximately the centre of the area surveyed.

Study	Coordinates	No. sites	No. bird species	Elevation
Anjos	23° 3’ S 51° 9’ W	14	171	~500m
Develey	23° 5’ S 47° 2’ W	32	154	~700m
Jack	23° 1’ S 45° 2’ W	24	190	~900m
Morante-Filho	15º 2’ S 39º 2’ W	36	182	~100m
Uezu	22° 2’ S 52° 2’ W	28	98	~400m

Evaluating Methods for Estimating Interactions

Six methods for estimating interactions from co-occurrence and abundance data were evaluated for accuracy against the Tangled Nature Model. The Tangled Nature Model was set up similarly to that used in the main text, however a species pool of 20 species was used, all species had the same specialism value (value of 1), and all cells remained forest cells (no deforestation event).

A total of 159 models were run, with the occurrence and abundance of each species in each cell recorded after the last time step of the Tangled Nature model. Six methods for estimating interactions were then used: correlation, partial correlation, GLMs, graphical lasso on log transformed counts (ggmlog), gaussian copula graphical models (ecoCopula). Each of these methods were applied as per their use in Harris (2016 and Popovic *et al.* (2019), changing family to binomial, poisson, or negative binomial as appropriate given the use of occurrence or abundance data.

The estimated interactions between each species using the aforementioned methods was then compared against the actual interactions as specified in the Tangled Nature Model

using linear models. Twelve linear models were created in R 4.0.3 (R Core Team 2020), two for each of the interaction estimation methods, with one using occurrence data, and the other using abundance data. The interaction as specified in the Tangled Nature Model was used as the response variable with the estimated interaction by each method used as the explanatory variable.

The adjusted R^2 for each model was reported as a measure of how well the estimated interaction method explains the variability seen in the interactions specified in the Tangled Nature Model (Table S4.2 and Table S4.3).

Table S4.2. Adjusted R2, intercept, and slope for each model for each interaction estimation method using occurrence data.

Method	**Adjusted R^2 Co-occurrence (2dp)**	**Intercept (2dp)**	**Slope (2dp)**
GLM	0.42	0.13	0.70
gllvm	0.22	0.17	0.39
ggmlog	0.44	0.12	0.54
ecoCopula	0.42	0.12	0.59
Correlation	0.44	0.12	0.54
Partial correlation	0.41	0.10	0.76

Table S4.3. Adjusted R2, intercept, and slope for each model for each interaction estimation method using abundance data.

Method	**Adjusted R^2 Abundance (2dp)**	**Intercept (2dp)**	**Slope (2dp)**
GLM	0.04	0.33	0.33
gllvm	0.34	0.23	0.22
ggmlog	0.43	0.25	0.24
ecoCopula	0.42	0.25	0.25

Correlation	0.43	0.26	0.25
Partial correlation	0.16	0.35	0.30

Since gllvm performed noticeably worse than all other methods when using co-occurrence data, this method was not used in the analysis.

Difference in Methods for Estimating Interactions - Empirical Data Set

The interaction estimates from the six methods that were used to estimate interactions from the empirical data set: correlation, partial correlation, GLMs, graphical lasso on log transformed counts (ggmlog), gaussian copula graphical models (ecoCopula), and gaussian copula graphics models with forest cover as an environmental variable (ecoCopulawith Env Variable) were compared against one another using generalised linear models and emmeans for post hoc tests in R 4.0.3 (Lenth 2020; R Core Team 2020).

All methods were found to be significantly different from one another ($P < 0.001$), with the exception of correlation and ggmlog, and partial correlation and GLM (see figure S4.1).

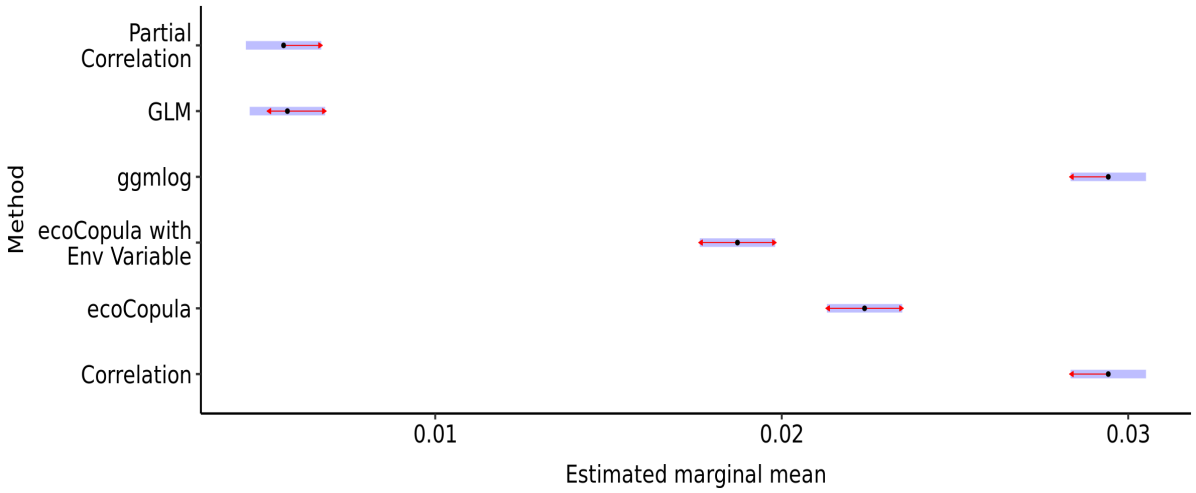


Figure S4.1. Pairwise comparisons of methods, if red arrows overlap then the methods are not significantly different from one another.

Each interaction estimation method was then regressed against each other using linear models, with the adjusted R^2 reported as a measure of similarity between methods. These methods had an average R2 of 0.42 (range = 0.41-0.44), with correlation and graphical lasso on log transformed counts methods performing the best. Interestingly, abundance methods performed worse than co-occurrence methods with an average R2 of 0.30 (range = 0.04-0.43) (Table S4.3).

Table S4.4. Adjusted R2 of linear models between all six methods for estimating interactions.

Method 1	**Method 2**	**Adjusted R^2 (2dp)**
GLM	ggmlog	0.55
GLM	ecoCopula	0.57
GLM	ecoCopula with Env Variable	0.58
GLM	Correlation	0.55
GLM	Partial Correlation	0.85
ggmlog	ecoCopula	0.92
ggmlog	ecoCopula with Env Variable	0.74
ggmlog	Correlation	1.00
ggmlog	Partial Correlation	0.55
ecoCopula	ecoCopula with Env Variable	0.81
ecoCopula	Correlation	0.92
ecoCopula	Partial Correlation	0.48
ecoCopula with Env Variable	Correlation	0.74
ecoCopula with Env Variable	Partial Correlation	0.48
Correlation	Partial Correlation	0.55

Linear mixed effect models

Full equations for the mixed effect linear models used to assess changes in interaction strength with forest cover can be seen below. Each model was chosen using the methods outlined in Zuur (2009).

Centered_interaction = centered_forest_cover_species1*centered_forest_cover_species2 + (centered_forest_cover_species1*centered_forest_cover_species2|species1) + (centered_forest_cover_species1*centered_forest_cover_species2|species2)

Centered_pair_interaction = centered_forest_cover_species1*centered_forest_cover_species2 + (centered_forest_cover_species1*centered_forest_cover_species2|species1) + (centered_forest_cover_species1*centered_forest_cover_species2|species2)

Table S4.5. Results of mixed effect models without phylogenetic filtering

Data	Response Variable	Centered Forest Cover Species 1	Centered Forest Cover Species 2	Interaction term
Empirical	Centered Interaction	0.109* (0.015)	0.097* (0.014)	0.825* (0.065)
Empirical	Centered Pair Interaction	0.073* (0.013)	0.081* (0.013)	0.423* (0.043)

Example Centering Forest Cover and Interactions

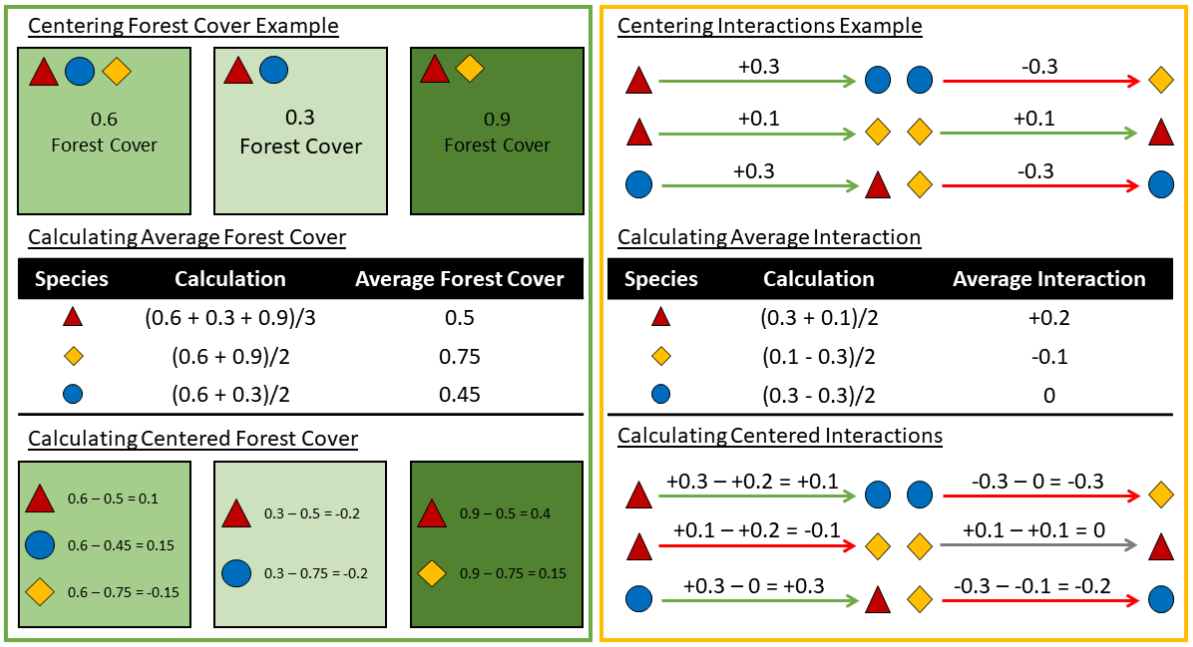


Figure S4.2. Example of centering forest cover and interactions on a per species basis.

Average Interactions per Species

Table S4.6. Average interaction used for centering for each species in the empirical dataset

Species	Centered Interaction	Species	Centered Interaction
Accipiter_striatus	-0.003	Myiarchus_ferox	-0.016
Amazilia_fimbriata	0.007	Myiarchus_swainsoni	-0.001
Amazilia_lactea	-0.011	Myiarchus_tuberculifer	0.019
Amazilia_versicolor	0.02	Myiarchus_tyrannulus	0.011
Amazona_aestiva	0.055	Myiobius_atricaudus	0
Amazona_amazonica	-0.003	Myiobius_barbatus	0.037
Amazona_rhodocorytha	-0.01	Myiodynastes_solitarius	0.014
Anabacerthia_amaurotis	0.024	Myiopagis_caniceps	0.018
Anabacerthia_lichtensteini	0.068	Myiopagis_viridicata	-0.006
Anabazenops_fuscus	0.003	Myiophobus_fasciatus	0.054

Anthracothorax_nigricollis	-0.022	Myiornis_auricularis	0.023
Antrostomus_rufus	0.032	Myiothlypis_flaveola	-0.015
Ara_chloropterus	0.072	Myiothlypis_leucoblephara	0.015
Aramides_saracura	-0.049	Myiothlypis_rivularis	0.015
Aratinga_auricapillus	0.032	Myiozetetes_similis	0.022
Arremon_flavirostris	-0.049	Myrmoderus_squamosus	0.029
Arremon_semitorquatus	0.005	Myrmotherula_luctuosa	0.02
Asemospiza_fuliginosa	0.016	Myrmotherula_minor	0.028
Attila_phoenicurus	0.071	Myrmotherula_unicolor	0.008
Attila_rufus	0.03	Myrmotherula_urosticta	0.011
Attila_spadiceus	0.014	Nemosia_pileata	0
Automolus_leucophthalmus	0.024	Neopelma_chrysolophum	0.01
Baryphthengus_ruficapillus	0.083	Nonnula_rubecula	0.052
Basileuterus_auricapilla	0.13	Notharchus_swainsoni	0.068
Batara_cinerea	0.002	Nyctidromus_albicollis	0.017
Brotogeris_tirica	0.021	Nyctiphrynus_ocellatus	-0.007
Cacicus_cela	0.018	Nystalus_chacuru	0.002
Cacicus_chrysopterus	0.024	Nystalus_maculatus	0.006
Cacicus_haemorrhous	0.04	Odontophorus_capueira	0.043
Campephilus_robustus	0.049	Orthogonys_chloricterus	0.108
Camptostoma_obsoletum	-0.027	Oxyruncus_cristatus	0.054
Campylorhamphus_falcularius	0.031	Pachyramphus_castaneus	0.047
Campylorhynchus_turdinus	0.017	Pachyramphus_marginatus	0.032
Capsiempis_flaveola	0.028	Pachyramphus_polychopterus	-0.017
Caracara_plancus	0.003	Pachyramphus_validus	-0.001
Carpornis_cucullata	0.064	Pachyramphus_viridis	-0.015
Carpornis_melanocephala	0.065	Patagioenas_cayennensis	-0.044

Caryothraustes_canadensis	0.021	Patagioenas_picazuro	-0.027
Casiornis_rufus	-0.064	Patagioenas_plumbea	0.068
Cathartes_aura	-0.001	Patagioenas_speciosa	0.025
Celeus_flavescens	0.003	Penelope_obscura	-0.036
Celeus_tinnunculus	-0.001	Penelope_superciliaris	-0.097
Ceratopipra_rubrocapilla	0.013	Phacellodomus_erythrophthalmus	0.002
Chamaeza_campanisona	0.073	Phaethornis_eurynome	-0.001
Chamaeza_meruloides	0.055	Phaethornis_pretrei	-0.024
Chelidoptera_tenebrosa	0.004	Phaethornis_ruber	0.021
Chiroxiphia_caudata	0.065	Pheugopedius_genibarbis	0.018
Chiroxiphia_pareola	-0.03	Philydor_atricapillus	0.037
Chlorophanes_spiza	0.009	Philydor_rufum	0.036
Chlorostilbon_lucidus	-0.004	Phyllomyias_burmeisteri	-0.014
Ciccaba_virgata	0.029	Phyllomyias_fasciatus	0.033
Cichlocolaptes_holti	0.072	Phyllomyias_griseocapilla	0.017
Cissopis_leverianus	0.053	Phyllomyias_virescens	0.024
Claravis_pretiosa	0.025	Phylloscartes_oustaleti	0.023
Clytolaema_rubricauda	0.038	Phylloscartes_paulista	0.07
Cnemotriccus_fuscatus	-0.012	Phylloscartes_ventralis	0.003
Coccyzus_americanus	0.093	Piaya_cayana	-0.013
Coccyzus_euleri	-0.002	Piculus_aurulentus	0.004
Coccyzus_melacoryphus	0.008	Piculus_flavigula	-0.006
Coereba_flaveola	0.011	Picumnus_albosquamatus	-0.01
Colaptes_campestris	0.003	Picumnus_cirratus	0.009
Colaptes_melanochloros	-0.007	Picumnus_exilis	0.012
Colonia_colonus	0.033	Pionopsitta_pileata	0.059
Columbina_talpacoti	-0.013	Pionus_maximiliani	0.04

Conirostrum_speciosum	-0.009	Pionus_reichenowi	0.001
Conopophaga_lineata	-0.032	Pipra_fasciicauda	-0.013
Conopophaga_melanops	0.054	Pipraeidea_bonariensis	-0.018
Coragyps_atratus	0	Pipraeidea_melanonota	0.021
Corythopis_delalandi	0.001	Piprites_chloris	0.05
Cranioleuca_obsoleta	0.059	Pitangus_sulphuratus	-0.033
Cranioleuca_pallida	0.003	Platyrinchus_leucoryphus	0.065
Crotophaga_major	0.028	Platyrinchus_mystaceus	0.028
Crypturellus_obsoletus	0.039	Poecilotriccus_plumbeiceps	0.011
Crypturellus_parvirostris	-0.042	Primolius_maracana	0.043
Crypturellus_soui	0.021	Procnias_nudicollis	0.057
Crypturellus_tataupa	-0.003	Psarocolius_decumanus	-0.003
Crypturellus_undulatus	0.06	Pseudastur_polionotus	0.038
Crypturellus_variegatus	0.009	Pseudopipra_pipra	-0.001
Cyanerpes_cyaneus	-0.009	Psilorhamphus_guttatus	0.082
Cyanocorax_chrysops	0.004	Psittacara_leucophthalmus	0.018
Cyanocorax_cristatellus	-0.018	Pteroglossus_aracari	0.032
Cyclarhis_gujanensis	-0.02	Pteroglossus_bailloni	0.059
Dacnis_cayana	0.006	Pteroglossus_castanotis	0.07
Dendrocincla_turdina	0.028	Pulsatrix_koeniswaldiana	0.052
Dendrocolaptes_platyrostris	0.057	Pygochelidon_cyanoleuca	-0.011
Dendroplex_picus	0.011	Pyriglena_leucoptera	0.055
Diopsittaca_cumanensis	0.005	Pyroderus_scutatus	0.033
Dromococcyx_pavoninus	0.048	Pyrrhura_cruentata	0.023
Drymophila_ferruginea	-0.015	Pyrrhura_frontalis	0.058
Drymophila_malura	0.01	Ramphastos_ariel	0.055
Drymophila_ochropyga	0.015	Ramphastos_dicolorus	0.045
Drymophila_rubricollis	0.06	Ramphastos_toco	0.001

Drymophila_squamata	0.053	Ramphocaenus_melanurus	0.044
Dysithamnus_mentalis	0.047	Ramphocelus_bresilius	0.006
Dysithamnus_stictothorax	0.084	Ramphodon_naevius	0.108
Dysithamnus_xanthopterus	0.034	Rhopias_gularis	0.055
Elaenia_flavogaster	-0.002	Rhynchocyclus_olivaceus	0.022
Elaenia_mesoleuca	-0.033	Rhytipterna_simplex	0.016
Eleoscytalopus_indigoticus	0.098	Rupornis_magnirostris	-0.032
Eleoscytalopus_psychopompus	0.029	Saltator_fuliginosus	0.043
Empidonomus_varius	0.016	Saltator_maximus	0.012
Eupetomena_macroura	0.031	Saltator_similis	-0.053
Euphonia_chalybea	-0.002	Sarcoramphus_papa	-0.005
Euphonia_chlorotica	0.002	Schiffornis_turdina	0.033
Euphonia_cyanocephala	0.003	Schiffornis_virescens	0.067
Euphonia_pectoralis	0.051	Sclerurus_mexicanus	-0.006
Euphonia_violacea	0.034	Sclerurus_scansor	0.038
Euphonia_xanthogaster	0.009	Scytalopus_speluncae	0
Eupsittula_aurea	0.011	Selenidera_maculirostris	0.048
Florisuga_fusca	0.013	Serpophaga_subcristata	-0.012
Formicarius_colma	0.058	Setophaga_pitiayumi	0.011
Formicivora_grisea	-0.005	Sicalis_flaveola	-0.019
Forpus_xanthopterygius	-0.006	Sirystes_sibilator	0.064
Furnarius_figulus	-0.001	Sittasomus_griseicapillus	0.047
Furnarius_rufus	0.002	Spinus_magellanicus	-0.007
Galbula_ruficauda	0.043	Spizaetus_tyrannus	0.016
Geotrygon_montana	0.018	Sporophila_falcirostris	-0.013
Geranospiza_caerulescens	0.001	Sporophila_frontalis	0.022
Glaucidium_brasilianum	0.016	Stelgidopteryx_ruficollis	-0.009

Glaucidium_minutissimum	0.007	Synallaxis_cinerascens	-0.002
Glaucis_hirsutus	-0.003	Synallaxis_frontalis	-0.008
Glyphorynchus_spirurus	0.03	Synallaxis_ruficapilla	-0.012
Gnorimopsar_chopi	0.007	Synallaxis_spixi	0.025
Grallaria_varia	0.04	Syndactyla_rufosuperciliata	-0.002
Habia_rubica	0.057	Tachyphonus_coronatus	0.018
Haplospiza_unicolor	0.04	Tachyphonus_rufus	0.002
Harpagus_diodon	0.012	Tangara_brasiliensis	-0.006
Heliobletus_contaminatus	0.001	Tangara_cyanocephala	0.097
Hemithraupis_flavicollis	0.035	Tangara_cyanomelas	0.015
Hemithraupis_guira	0.13	Tangara_cyanoptera	0.067
Hemithraupis_ruficapilla	0.02	Tangara_cyanoventris	-0.008
Hemitriccus_diops	0.018	Tangara_desmaresti	-0.001
Hemitriccus_nidipendulus	-0.005	Tangara_flava	-0.017
Hemitriccus_obsoletus	0.037	Tangara_ornata	-0.009
Hemitriccus_orbitatus	0.042	Tangara_palmarum	0.002
Herpetotheres_cachinnans	0.022	Tangara_sayaca	-0.015
Herpsilochmus_atricapillus	0.146	Tangara_seledon	0.049
Herpsilochmus_pileatus	-0.005	Taraba_major	-0.008
Herpsilochmus_rufimarginatus	0.032	Terenura_maculata	0.048
Herpsilochmus_scapularis	0.01	Tersina_viridis	0.009
Hylatomus_lineatus	0.007	Thalurania_glaucopis	0.006
Hylocharis_chrysura	0.01	Thamnomanes_caesius	0.03
Hylocharis_cyanus	0.008	Thamnophilus_ambiguus	-0.002
Hylopezus_nattereri	0.033	Thamnophilus_caerulescens	-0.035
Hylophilus_poicilotis	0.014	Thamnophilus_doliatus	-0.056
Hypoedaleus_guttatus	0.046	Thamnophilus_palliatus	0.013

Icterus_jamacaii	-0.007	Thamnophilus_torquatus	-0.027
Icterus_pyrrhopterus	-0.071	Thlypopsis_pyrrhocoma	0.064
Ictinia_plumbea	0.011	Tinamus_solitarius	0.082
Ilicura_militaris	0.04	Tityra_braziliensis	0.054
Islerothraupis_cristata	0.041	Tityra_inquisitor	-0.014
Laniisoma_elegans	0.085	Todirostrum_cinereum	-0.004
Lathrotriccus_euleri	-0.007	Todirostrum_poliocephalum	0.007
Legatus_leucophaius	0.02	Tolmomyias_flaviventris	0.007
Lepidocolaptes_angustirostris	0.019	Tolmomyias_poliocephalus	0.022
Lepidocolaptes_falcinellus	0.018	Tolmomyias_sulphurescens	0.009
Leptodon_cayanensis	0.01	Trichothraupis_melanops	0
Leptopogon_amaurocephalus	0.033	Triclaria_malachitacea	0.052
Leptotila_rufaxilla	0.004	Troglodytes_aedon	-0.014
Leptotila_verreauxi	-0.012	Trogon_rufus	0.05
Leucochloris_albicollis	0.013	Trogon_surrucura	0.067
Lipaugus_vociferans	0.012	Trogon_viridis	0.057
Lochmias_nematura	-0.023	Turdus_albicollis	0.017
Lurocalis_semitorquatus	0.101	Turdus_amaurochalinus	0.005
Machaeropterus_regulus	0.038	Turdus_flavipes	0.033
Mackenziaena_leachii	-0.027	Turdus_leucomelas	0.01
Malacoptila_striata	0.033	Turdus_rufiventris	-0.019
Manacus_manacus	0.019	Turdus_subalaris	0.022
Megarynchus_pitangua	-0.022	Tyrannus_melancholicus	-0.013
Megascops_choliba	-0.058	Vanellus_chilensis	-0.002
Melanerpes_flavifrons	0.044	Veniliornis_affinis	-0.002
Merulaxis_ater	0.029	Veniliornis_spilogaster	0.028
Micrastur_ruficollis	0.026	Vireo_olivaceus	0.02

Micrastur_semitorquatus	0.047	Xenops_minutus	0.034
Milvago_chimachima	-0.009	Xenops_rutilus	0.022
Mimus_saturninus	0.006	Xiphocolaptes_albicollis	0.077
Mionectes_oleagineus	0.029	Xipholena_atropurpurea	-0.016
Mionectes_rufiventris	0.053	Xiphorhynchus_fuscus	0.042
Momotus_momota	0.093	Xiphorhynchus_guttatus	0.035
Monasa_morphoeus	0.016	Zonotrichia_capensis	0.009

Chapter 5: Fragmentation increases species richness but disturbs community composition

Equations

The Sørensen dissimilarity index was calculated as follows:

$$S = \frac{2c}{a+b}$$

(eqn S5.1)

Where c is the number of species in common between the communities in the Habitat cells in the fragmented configurations and the community in the comparable Habitat cells in the Continuous Habitat configuration. Whilst a is the total number of species in the Habitat cells in the fragmented configurations and b is the number of species in the comparable Habitat cells in the fragmented configurations.

The Weighted Population Coefficient of Variation for community X made up of populations of species $X_1 \ldots X_n$ where n is the total species in the species pool, which is always 100, is calculated as follows:

$$WPCV(X) = \frac{\sum_{I=1}^{n} \left(\frac{\sigma(X_i)}{\mu(X_i)} x \frac{\mu(X_i)}{\mu(X)}\right)}{n}$$

(eqn S5.2)

Where $\sigma(X_i)$ is the standard deviation of the abundance of the population of species i and $\mu(X_i)$ is the mean abundance of species i, over the 125,000 post-fragmentation time steps.

Results

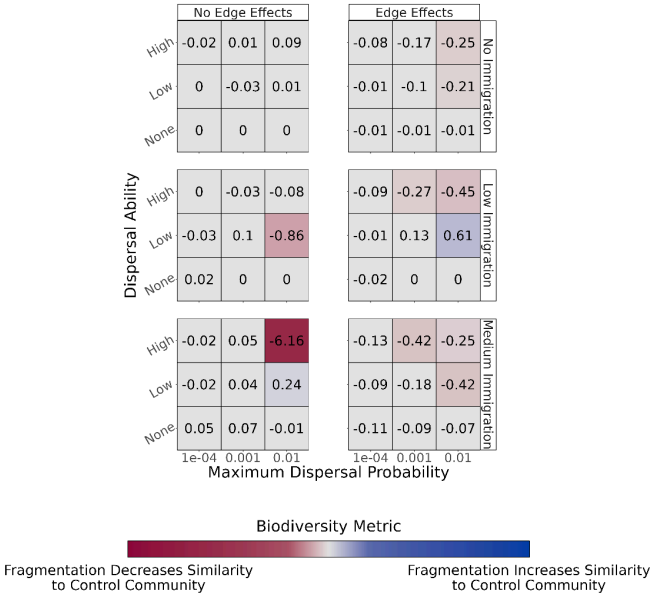


Figure S5.1. Proportional change in community similarity based on bray-curtis similarity compared to comparable cells in the Continuous Habitat control configuration due to fragmentation, where the Single Large patch configuration is used as a baseline. Grey cells are below 0.2 change, and are deemed biologically insignificant. Red cells represent a significant decrease as fragmentation increases, whilst Blue cells represent a significant increase as fragmentation increases.