Vertebrate community disassembly in human-induced Amazonian forest islands

© Adriano Gambarini

Anderson Saldanha Bueno

A thesis submitted to the University of East Anglia
for the degree of Doctor of Philosophy

July 2019

Dedicatória

À minha esposa, Carla, e aos meus filhos, Melina e Vicente (*in memoriam*).

Agradecimentos

Após tanta dedicação ao trabalho, é surpreendente se dar conta de que o mais fácil é justamente trabalhar. Sem condições e apoio, trabalhar é apenas um desejo. A realização do meu doutorado só foi possível graças às pessoas que estiveram comigo, contribuindo para que eu pudesse me dedicar integralmente à ciência.

Sou especialmente grato à minha esposa, Carla, por sempre ter dado todo o apoio que precisei para realizar o meu doutorado. Um apoio repleto de amor e companheirismo. Nessa nossa jornada, mudamos de país, vivemos em meio à floresta Amazônica e tivemos nossa amada filha, Melina. De coração, por tudo que tu fizeste por mim e pela nossa família, eu te agradeço, meu amor!

Ao meu orientador, Carlos, pela oportunidade, confiança, incentivo, parceria e amizade. Foi uma grande satisfação ter sido teu pupilo! Levarei comigo tua eterna empolgação, as ideias infinitas e o prazer em fazer ciência.

Aos meus pais, Heloisa e Emilson, pelo estímulo incondicional ao estudo. As portas que vocês me abriram, mesmo sem saber onde dariam, me trouxeram até aqui!

Às minhas irmãs, Thays e Laura, pelo carinho e inspiração. Enquanto eu trilhei um caminho linear, vocês me mostraram que reviravoltas também fazem parte do sucesso.

À Soninha (*in memoriam*), por ter cuidado sempre com amor, dedicação e lealdade da minha família.

Aos meus sogros, Marta e Almir, pela vinda até Norwich para nos dar apoio e suporte com a chegada da Melina, permitindo-me trabalhar com tranquilidade.

Ao meu amigo de longa data e irmão de coração, Vitor, que trouxe meu HD externo, esquecido no Brasil, com os dados do quinto capítulo.

À Maíra, pioneira em Balbina, por compartilhar sua experiência comigo, auxiliando-me nos momentos de dúvida.

Aos meus colegas de *office*, pelo agradável ambiente de trabalho: Agatha, Valentina, Rob e Claire.

*"Que ninguém se engane,
só se consegue a simplicidade através de muito trabalho".*

Clarice Lispector

Abstract

The extent to which land-use change impairs tropical biodiversity is controversial. This is because while some species may thrive in human-modified landscapes, others are locally extirpated – the so-called "few winners and many losers" paradigm. Furthermore, reliable environmental impact assessments imply that the correct drivers of biodiversity change are recognised, and suitable reference conditions are available. Herein, I examine vertebrate responses to habitat fragmentation induced by two mega hydroelectric reservoirs in Brazilian Amazonia. In the Tucuruí Hydroelectric Reservoir, bird point count surveys were carried out on 36 forest islands in 2006 and 2007. In the Balbina Hydroelectric Reservoir, I carried out bird mist-net surveys on 33 forest islands and five continuous forest sites in 2015 and 2016; I also used autonomous recording units to survey anuran species on 74 forest islands and four continuous forest sites in 2015. At Tucuruí, bird species traits associated with vulnerability to forest fragmentation included rarity and forest dependency. At Balbina, species richness of understorey insectivorous birds was more affected by island size than the amount of habitat surrounding mist-net lines, indicating that a patch-centric approach is most appropriate to measure species responses to habitat loss if species exhibit low dispersal ability and the surrounding habitat matrix is hostile. Regarding environmental impact assessments, the higher the intactness of the reference condition, the greatest the estimated impact of habitat fragmentation on bird species richness. In other words, studies of biodiversity responses to land-use change will likely yield over-optimistic results if they are masked by the low conservation status of either degraded or insufficiently large habitat patches regarded as the reference condition. Finally, island size *per se* played a decisive role in explaining anuran species richness on Amazonian forest islands, yet the inferential power of island species-area relationships is only reliable if derived from an appropriate study design.

Contents

Chapter 1

Introduction

Every introduction of a scientific paper is supposed to present the reasons why such research was carried out. That is an opportunity to convince readers of the relevance of the research to engage their attention to the end of the paper. Although sometimes engaging, the introduction of a scientific paper leaves behind its most interesting reasons. "It presents a logical argument for why you should have done the research, not the real reason, or even the true sequence of events that led to the results" (Medawar 1964 *apud* Magnusson 2015).

When I presented my MSc project in 2008, I was enquired by one of the examiners on "why birds?". I had a few seconds to decide to give him either the scientific or the real reason. The scientific reason was birds are a functionally diverse species-rich group which often responds predictably to environmental change. The real reason was that, having been raised in a large metropolis, I did not know how to identify any species apart from birds in the hand – along with a previous study of the expected species list, a couple of field guides and photographs to double-check some identifications with hardened ornithologists.

I was tempted to begin this introduction with something like *Land-use change is the major threat to terrestrial biodiversity* (Sala et al. 2000) as a rationale for my thesis. Instead, I will tell you *the true sequence of events* that led to the study, since the scientific reasons are already provided in the following data chapters. In this section, I aim to inform the origin and context of this thesis work as well as what lies behind the scene.

1.1   The thesis topic

1.1.1   What is this thesis *not* about?

In 2012, when I talked to my supervisor, Carlos, for the very first time, he suggested I should work on avifaunal recovery trajectories in insular vegetation regrowth patches following deforestation by oil and gas enterprises at tropical forest seismic clearings. The project was fantastic given the benign and unique surrounding landscape context dominated by vast areas of undisturbed primary forest: equal-sized clearings at different stages of regeneration embedded within Amazonian primary forest (Figure 1.1).

Figure 1.1: Forest clearing surrounded by primary Amazonian forest. Image from Google Earth.

Over the next two years, I went to two places in Amazonia to conduct preliminary fieldwork where my study areas would presumably become established. Discussions with Carlos, readings and field campaigns led me to apply for a PhD studentship and I eventually got the seismic clearings project. However, our plans changed abruptly as the oil company went bankrupt following two helicopter crashes and that project would no longer be viable.

Luckily, Carlos promptly came up with a Plan B. He invited me to join a large-scale project on forest fragmentation induced by the Balbina hydroelectric dam in Brazilian Amazonia. Although a bit reluctant at the beginning, I was even more sure about the risks involved in the clearings project. In 2014, I travelled to Balbina to meet Carlos and familiarise myself with the landscape. Fortunately, I left Balbina enthusiastic about all the logistical facilities and the research prospects.

1.1.2   What is this thesis about?

This thesis is about responses by forest vertebrates – birds and frogs – to habitat fragmentation, a process that involves both habitat loss and the breaking apart of remaining habitat (Fahrig 2003; Figure 1.2). Compared to continuous habitat, a fragmented landscape is characterised by a number of smaller patches of smaller total area, isolated from each other by an intervening matrix that is structurally distinct from habitat remnants (Figure 1.2). For example, the conversion of forests (habitat) into pastureland (matrix) where some forest patches remain (fragment; Figure 1.3). Likewise, flooding of lowland forest induced by river damming where former hilltops form forest islands is also a type of habitat fragmentation, which could be regarded as habitat insularization since the intervening matrix is water (Figure 1.3).

Habitat fragmentation research has largely focused on habitat patches surrounded by terrestrial matrices (Fahrig 2017). However, the theoretical backbone of habitat fragmentation research is rooted in the Theory of Island Biogeography (IBT; MacArthur & Wilson 1967), which was developed to explain biodiversity patterns on oceanic islands. Ironically, oceanic island formation ("creation" of habitat) is the opposite of habitat fragmentation ("destruction" of habitat; Figure 1.3). Furthermore, terrestrial matrix habitats are far less hostile than open water (Mendenhall et al. 2014) and may (1) favour some species by increasing habitat heterogeneity, (2) allow species to move among habitat patches, and (3) provide resources for patch-dependent species (Driscoll et al. 2013). Therefore, IBT has been shown to be of limited application to understand biodiversity patterns in habitat patches within terrestrial landscapes (Laurance 2008).

Habitat insularization consists of an intermediate scenario (Figure 1.3). On the one hand, it is habitat fragmentation. On the other hand, the intervening matrix is hostile. The extent to which IBT applies to reservoir islands is, therefore, a topic yet to reach a consensus.

Both the origin and hostility of the matrix determine the structure of species assemblages in habitat patches – here represented by habitat fragments,
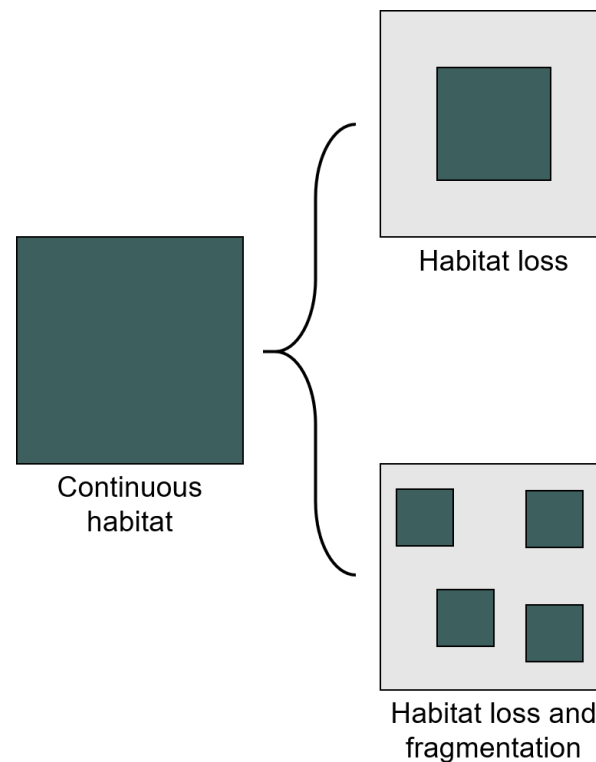
Figure 1.2: Illustration of habitat loss, and habitat loss combined with fragmentation. Habitat is represented in green and the intervening matrix in grey.

reservoir islands, and oceanic islands. According to Watson (2002), the patch biota is divided into three species groups based on their origin: relict species (i.e. present before fragmentation/insularization), matrix-derived species, and inter-patch dispersers. Different from habitat fragments where all three species groups are present, and oceanic islands where relict species are non-existent, reservoir islands are mostly comprised of relict species because both matrix-derived species and inter-patch dispersers are rare. However, only a fraction of the relict species will persist in the long-run (Jones et al. 2016) typically resulting in impoverished species assemblages.

Main aim

Broadly speaking, I aim to depict biodiversity patterns on forest islands induced by hydroelectric reservoirs as guided by the following questions:

- How many species remain on forest islands?

- Which are the remaining species?

- What are the characteristics of extinction-prone species?

- How do both island and landscape attribute shape species assemblages?
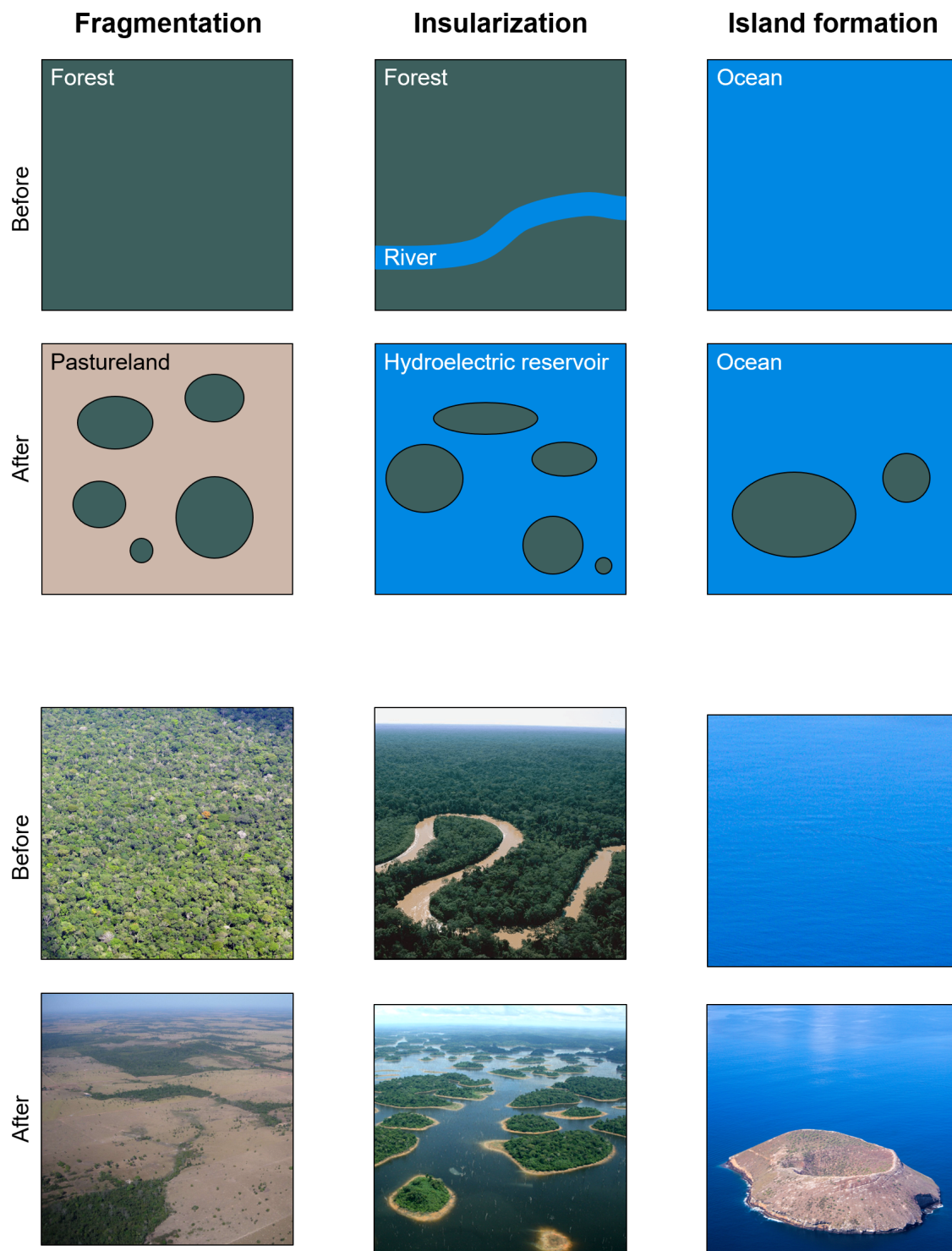
Figure 1.3: Examples of drivers of patchy systems: **Fragmentation** represented by forest fragments surrounded by pastureland; **insularization,** which is a form of fragmentation where the intervening matrix is water; **island formation** from volcanic eruption. Photo credits: (a) © Neil Palmer (CIAT), (b) Haag & Henriques (2016), (c, f) © Sam LaRussa, (d) Lees & Peres (2008), (e) Benchimol & Venticinque (2014). All photos were cropped.

1.2   Behind the scene

This thesis is structured as four data chapters. Although a thesis is a neat document, the process of getting from A to B is often nonlinear, and each data chapter has its own and sometimes independent conjectures. As a record of my PhD journey, I share here *the true sequence of events* that led me to each data chapter (Chapters 2 to 5).

1.2.1   Chapter 2

In early 2016, between the first and the second field campaigns, I did not have much to do apart from data entry, project reports and a literature review. Considering that I could not write a paper with only half of my data, Carlos suggested contacting a friend of him who had surveyed birds with another researcher on forest islands within the Tucuruí Hydroelectric Reservoir, eastern Brazilian Amazonia (Figure 2.1). The deal was quite simple: they would unearth the then defunct data buried in an old hard-drive, and I would write the paper.

For several reasons, Chapter 2 was the one from which I learned the most. First, I was in constant contact with Magalli Henriques and Sidnei Dantas to understand their data, learning how to conduct collaborative research. Second, I realised the importance of tedious data cleaning and repetitive checking; if you seek for mistakes, you will find them. Third, I extracted environmental variables from satellite images in a GIS environment having to start from scratch; the first time I opened a raster or vector file was during my PhD! Finally, Carlos told me I would have to cope with "the issue of phylogenetic non-independence among species" to perform the analysis. Again, I had to start from scratch by reading textbooks, spending numerous days reading tutorials and doing exercises in the R software.

This toolkit allowed me to investigate patterns of avian species loss and the role of morpho-ecological traits in explaining species vulnerability to forest fragmentation in an anthropogenic island system.

1.2.2   Chapter 3

I participated in the Brazilian Congress of Ornithology in 2012, where I had the opportunity to attend a talk by Lenore Fahrig, an eminent landscape ecologist. She presented the Habitat Amount Hypothesis (HAH), which was published a year later (Fahrig 2013). The HAH predicts that the number of species at sample sites is independent of the area of the particular patches in which samples are located, if the amount of habitat surrounding sample sites is held constant (figure 7

in Fahrig 2013). The HAH intrigued me because of a long-standing debate in protected area design – whether it is most optimal to have one Single Large Or Several Small (SLOSS) reserves of equivalent size (Diamond 1975) – would no longer be relevant, although she had never mentioned the SLOSS debate. Then, I looked for Lenore after her talk, and she kindly sat down with me to show me some extra slides on her laptop, confirming that I was right: according to the HAH, the SLOSS debate was presumably irrelevant.

The Island Biogeography Theory predicts that the number of species on islands is affected by island size and isolation (MacArthur & Wilson 1967). In fragmented landscapes, the HAH predicts that these two variables are simply driven by the sample area effect, so that the main variable affecting the number of species in equally-sized sample sites within habitat fragments would be the total amount of surrounding habitat (Fahrig 2013). The applicability of IBT for oceanic islands and of HAH for habitat fragments led me to the main question of Chapter 3: which of these two major concepts in fragmentation ecology (IBT or HAH) is the most appropriate theoretical framework for human-induced islands?

To address this question, I used mist nets to survey birds on 33 forest islands within the Balbina Hydroelectric Reservoir, central Brazilian Amazonia (Figure 3.1). Besides testing the HAH, I proposed a diagram, based on the degree of matrix permeability and species dispersal ability, to determine whether IBT or HAH is the most appropriate theoretical framework to understand patterns of species richness in reservoir islands and analogous habitat patches.

1.2.3   Chapter 4

Many reasons account for contrasting results in the ecological literature, such as landscape context, taxonomic group, sampling design and survey method. During my meetings with Carlos, when we discussed some "unexpected" result from the literature, he recurrently hypothesised that so-called control sites were often degraded, thereby masking the true effects of land-use change on biodiversity. The principle is simple: if the reference condition is degraded, the estimated impacts will be obscured. Although this issue in environmental impact assessments has been raised (Gardner et al. 2009), quantitative studies were still lacking.

To assess the effects of habitat fragmentation on biodiversity, it is common to use sites in continuous habitats (e.g. Almeida-Gomes et al. 2016) or in the largest available habitat fragment in the landscape (e.g. Daily et al. 2011) as a reference condition to contrast with smaller habitat fragments. However, the size of the largest available habitat fragment is rather variable. For example, the reference condition in a Costa Rican bird study by Daily et al. (2001) was a forest

fragment of *c.* 250 ha, whereas Uehara-Prado et al. (2007) used a forest fragment of *c.* 10,000 ha as a reference condition in a butterfly study in the Brazilian Atlantic Forest.

In Chapter 4, I assessed the extent to which the change in the size of forest islands used as a reference condition affects the estimated impact of forest fragmentation on species richness. To do so, I mist-netted birds on 33 forest islands and five continuous forest sites (Figure 4.1) and used different sets of sites as a reference condition to contrast with a set of small islands (Figure 4.3). Considering that species richness fails to account for species identity, I did the same comparative exercise using either only species recorded at references sites or the overall species assemblage. In doing so, I examined how including species not recorded at reference sites contributes to obscure the estimated effect of forest fragmentation on species richness.

1.2.4   Chapter 5

Initially, I had considered surveying only birds with mist nets, which is a labour-intensive method that limits spatial replication. However, autonomous recording units (ARUs) are cost-effective and allows one to survey multiple sites at the same time. Thus, during the first year of my fieldwork in 2015, I deployed ARUs at 151 plots located on 74 islands and in four continuous forest sites (Figure 5.1). Since ARUs record any vocalising species, I decided to dedicate a chapter of my thesis to frogs for several reasons. First, frogs comprise a less species-rich group than birds and also show stereotyped vocalisations, thereby facilitating species identification of the overall frog assemblage present at each plot. Second, I had never studied frogs before, and that would be an opportunity to open my mind to another taxonomic group, which is instrumental in developing ecological thinking. Third, there are much fewer studies on frogs in the fragmentation ecology literature compared to birds (Gardner et al. 2007); given the fact that journals always look for novelty, a fragmentation study on frogs *per se* was something appealing.

My lack of knowledge on frog species identification through vocalisation along with the huge amount of recordings obtained (over 3,500 hours) led me to establish a collaborative project with Ígor Kaefer and Gabriel Masseli, two Brazilian herpetologists, and to use a subset of recordings. Thus, 62 minutes of recordings per plot was listened to by Gabriel (over 150 hours in total), and the species identifications were validated by Ígor. My task was then to provide them with the recordings, analyse the data, and write the paper.

Chapter 5 is on species-area relationships. Although it is an over-studied topic, my study differs from all others because most of the studies fail to control

for several confounding factors before testing the effect of area *per se* on species richness (see Hill 1994). Furthermore, I assessed the role of sampling design – the number of islands and range in island size – in shaping species-area relationships aiming to provide guidelines to yield reliable estimates of area-driven species losses in habitat patches.

What follows next are four data-driven chapters which make up the empirical and theoretical body of this doctoral thesis work. Happy reading!

1.3   References

1. Almeida-Gomes M, Vieira MV, Rocha CFD, Metzger JP, Coster G De. 2016. Patch size matters for amphibians in tropical fragmented landscapes. Biological Conservation 195:89–96.

2. Benchimol M, Venticinque EM. 2014. Responses of primates to landscape change in Amazonian land-bridge islands – a multi-scale analysis. Biotropica 46:470–478.

3. Daily GC, Ehrlich PR, Sánchez-Azofeifa GA. 2001. Countryside biogeography: Use of human-dominated habitats by the avifauna of southern Costa Rica. Ecological Applications 11:1–13.

4. Diamond JM. 1975. The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. Biological Conservation 7:129–146.

5. Driscoll DA, Banks SC, Barton PS, Lindenmayer DB, Smith AL. 2013. Conceptual domain of the matrix in fragmented landscapes. Trends in Ecology & Evolution 28:605–613.

6. Fahrig L. 2003. Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, Evolution, and Systematics 34:487–515.

7. Fahrig L. 2013. Rethinking patch size and isolation effects: the habitat amount hypothesis. Journal of Biogeography 40:1649–1663.

8. Fahrig L. 2017. Ecological responses to habitat fragmentation *per se*. Annual Review of Ecology, Evolution, and Systematics 48:1–23.

9. Gardner TA, Barlow J, Chazdon R, Ewers RM, Harvey CA, Peres CA, Sodhi NS. 2009. Prospects for tropical forest biodiversity in a human-modified world. Ecology Letters 12:561–582.

10. Hill JL, Curran PJ, Foody GM. 1994. The effect of sampling on the species-area curve. Global Ecology and Biogeography Letters 4:97–106.

11. Gardner TA, Barlow J, Peres CA. 2007. Paradox, presumption and pitfalls in conservation biology: The importance of habitat change for amphibians and reptiles. Biological Conservation 138:166–179.

12. Haag NA, Henriques MH. 2016. The paleontological heritage of the Acre (Amazonia, Brazil): Contribution towards a national paleontological database. Geoheritage 8:381–391.

13. Jones IL, Bunnefeld N, Jump AS, Peres CA, Dent DH. 2016. Extinction debt on reservoir land-bridge islands. Biological Conservation 199:75–83.

14. Laurance WF. 2008. Theory meets reality: How habitat fragmentation research has transcended island biogeographic theory. Biological Conservation 141:1731–1744.

15. Lees AC, Peres CA. 2008. Conservation value of remnant riparian forest corridors of varying quality for Amazonian birds and mammals. Conservation Biology 22:439–449.

16. MacArthur RH, Wilson EO. 1967. The Theory of Island Biogeography. Princeton University Press, Princeton, New Jersey, USA.

17. Magnusson W (Bill) E. 2015. The Eye of the Crocodile. Open Science.

18. Sala OE et al. 2000. Global biodiversity scenarios for the year 2100. Science 287:1770–1774.

19. Mendenhall CD, Karp DS, Meyer CFJ, Hadly EA, Daily GC. 2014. Predicting biodiversity change and averting collapse in agricultural landscapes. Nature 509:213–217.

20. Uehara-Prado M, Brown KS, Freitas AVL. 2007. Species richness, composition and abundance of fruit-feeding butterflies in the Brazilian Atlantic Forest: comparison between a fragmented and a continuous landscape. Global Ecology and Biogeography 16:43–54.

21. Watson DM. 2002. A conceptual framework for studying species composition in fragments, islands and other patchy ecosystems. Journal of Biogeography 29:823–834.

Chapter 2

Ecological traits modulate bird species responses to forest fragmentation in an Amazonian anthropogenic archipelago
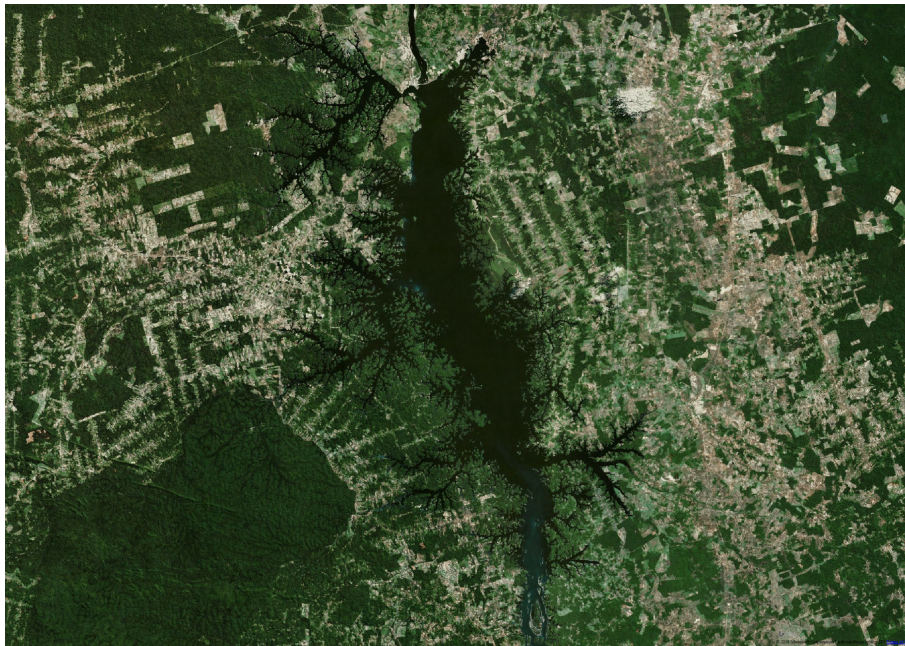
Image from Bing Maps

2.1   Abstract

Aim: We assessed patterns of avian species loss and the role of morpho-ecological traits in explaining species vulnerability to forest fragmentation in an anthropogenic island system. We also contrasted observed and detectability-corrected estimates of island occupancy, which are often used to infer species vulnerability.

Location: Tucuruí Hydroelectric Reservoir, eastern Brazilian Amazonia.

Methods: We surveyed forest birds within 36 islands (3.4–2,551.5 ha) after 22 years of post-isolation history. We applied species–area relationships to assess differential patterns of species loss among three data sets: all species, forest specialists and habitat generalists. After controlling for phylogenetic non-independence, we used observed and detectability-corrected estimates of island occupancy separately to build competing models as a function of species traits. The magnitude of the difference between these estimates of island occupancy was contrasted against species detectability.

Results: The rate of species loss as a function of island area reduction was higher for forest specialists than for habitat generalists. Accounting for the area effect, forest fragmentation did not affect the overall number of species regardless of the data set. Only the interactive model including natural abundance, habitat breadth and geographic range size was strongly supported for both estimates of island occupancy. For 30 species with detection probabilities below 30%, detectability-corrected estimates were at least tenfold higher than those observed. Conversely, differences between estimates were negligible or non-existent for all 31 species with detection probabilities exceeding 45.5%.

Main conclusions: Predicted decay of avian species richness induced by forest loss is affected by the degree of habitat specialisation of the species under consideration, and may be unrelated to forest fragmentation *per se*. Natural abundance was the main predictor of species island occupancy, although habitat breadth and geographic range size also played a role. We caution against using occupancy models for low-detectability species, because overestimates of island occupancy reduce the power of species-level predictions of vulnerability.

KEYWORDS
detectability, insularization, island biogeography, occupancy, rarity, species-area relationships

2.2   Introduction

Amazonian forests have been extensively converted to cattle pasture and crop-land inducing widespread loss and fragmentation of formerly continuous forests, especially in the eastern and southern portions of the basin (Laurance et al., 2001; Peres et al., 2010). This scenario is further exacerbated by a massive growth in hydroelectric dams, which invariably inundate large tracts of forest, creating archipelagic landscapes (Lees, Peres, Fearnside, Schneider, & Zuanon, 2016). Forest fragmentation is widely recognised as a pervasive and lasting threat to bio-diversity and ecosystem functioning as forest fragments are subject to the combined detrimental effects of core area reduction, edge proliferation, and isolation (Haddad et al., 2015). Nevertheless, the quality of the matrix surrounding forest fragments plays a major role in determining the severity of fragmentation (Kennedy, Marra, Fagan, & Neel, 2010). Old-growth forest fragments surrounded by secondary forests favour species that exploit matrix resources (Blake & Loiselle, 2001), are less affected by edge effects (Laurance et al., 2011), and are more permeable, ensuring species movements among forest fragments (Powell, Stouffer, & Johnson, 2013). Conversely, forest islands within hydroelectric reservoirs exhibit lower functional connectivity, are expected to be dominated by edge-mediated decay in forest structure (Benchimol & Peres, 2015), and harbour de-pauperate extinction-driven species assemblages (Wolfe, Stouffer, Mokross, Powell, & Anciães, 2015). The detrimental consequences of forest fragmentation are therefore amplified by a water matrix (Mendenhall, Karp, Meyer, Hadly, & Daily, 2014), rendering hydroelectric dams a more severe threat to forest biotas.

Habitat loss and fragmentation are the twin processes associated with land-use change. From a species perspective, the former is defined as the conversion of a "habitat" into a "non-habitat" (i.e., habitat amount shrinkage), and the latter as the subdivision of a single large "habitat" into several smaller "habitat patches" separated from one another by an intervening "non-habitat" matrix (Lindenmayer & Fischer, 2007). While habitat loss has pervasive negative effects on native biodiversity, fragmentation affects species differently (Fahrig, 2003). For instance, habitat specialists are more consistently impaired by fragmentation than habitat generalists (Devictor, Julliard, & Jiguet, 2008). Predictions of species loss based on species–area relationships are therefore expected to be underestimated for habitat specialists if habitat generalists are included in the species pool (Matthews, Cottee-Jones, & Whittaker, 2014). Moreover, habitat fragmentation *per se* (i.e., accounting for habitat loss) may either decrease or increase the number of species that would be predicted by habitat loss alone (Yaacobi, Ziv, & Rosenzweig, 2007). Therefore, a proper assessment of species loss in variable-

sized habitat patches should focus on groups of target species (Matthews et al., 2014) and disentangle the effects of habitat loss from fragmentation (Fahrig, 2003; Yaacobi et al., 2007).

Species-level studies on responses to habitat fragmentation can further enhance our understanding of vulnerability-prone traits at both landscape (Feeley, Gillespie, Lebbin, & Walter, 2007) and global scales (Bregman, Sekercioglu, & Tobias, 2014), complementing assemblage-wide studies (Moura et al., 2016). Accordingly, low-density, large-bodied species at high trophic levels (Ewers & Didham, 2006), and those with restrict habitat breadth (Henle, Davies, Kleyer, Margules, & Settele, 2004) and narrow geographic range (Newbold et al., 2014) are expected to be at higher risk of extinction. The same holds true for bird species inhabiting the lower strata of closed-canopy forests (Sekercioglu et al., 2002), following ant-swarms and foraging in mixed-species flocks (Stouffer & Bierregaard, 1995). Understanding trait-based patterns of extinction proneness is therefore invaluable to anticipate species losses and tailor conservation programmes to vulnerable species. However, idiosyncratic species responses across different regions (Gage, Brooke, Symonds, & Wege, 2004; Moura et al., 2016), and the co-occurrence of confounding factors in human-modified landscapes, such as matrix type, may limit the extent to which clear patterns can be uncovered (Ewers & Didham, 2006), reinforcing the need for landscape-scale studies.

In fragmented landscapes, the area of remaining patches is the main driver of species patch occupancy (Keinath et al., 2017). Area-sensitive species can no longer occur in patches below a minimum spatial requirement, and are consequently relegated to fewer patches than species requiring smaller areas (Dardanelli, Nores, & Nores, 2006). Thus, the proportion of patches occupied in a landscape has often been used as a measure of species vulnerability to habitat fragmentation (e.g., Meyer, Fründ, Lizano, & Kalko, 2008; Thornton, Branch, & Sunquist, 2011; Wang, Thornton, Ge, Wang, & Ding, 2015). Due to inherent differences in species detectability and the fact that non-detections do not necessarily imply absences, observed estimates of patch occupancy can be underestimated. To overcome this bias, occupancy modelling has often been uncritically used as it can estimate patch occupancy while accounting for imperfect detectability (Banks-Leite et al., 2014). Unlike observed estimates, this analytical approach can overestimate patch occupancy for species with low detection probability ($< 30\%$; MacKenzie et al., 2002), which may degrade inferential power about species vulnerability.

Habitat fragmentation research has largely focused on terrestrial landscapes (Fahrig, 2017), where the remaining habitat (i.e., area of native vegetation) is typically termed as "remnant", "fragment", or "patch". Nonetheless, the term

"island" best describes remaining habitats in archipelagic landscapes. To avoid misleading terminology (Hall, Krausman, & Morrison, 1997), herein we refer to the remaining habitats in terrestrial and archipelagic landscapes as "fragments" and "islands", respectively. Meanwhile, the term "patch" is used to encompass both "fragment" and "island". In this study, we assessed bird species responses to forest fragmentation in a vast archipelagic landscape induced by a major hydroelectric dam in eastern Brazilian Amazonia, while addressing the four following questions. First, do habitat generalists show a less steep decline in species richness as a function of island area reduction compared to forest specialists? If so, we predict that assessments of overall species loss relying on species–area relationships also underestimate the loss of forest specialists in reservoir islands as previously shown for forest fragments (Matthews et al., 2014). Second, does forest fragmentation *per se* exacerbate or reduce the impact of forest loss on species richness for the overall species pool, forest specialists and habitat generalists? We predict a neutral fragmentation effect on the overall species pool due to a negative effect on forest specialists and a positive effect on habitat generalists. Third, which suite of morpho-ecological traits best explains species rates of island occupancy within the forest archipelago? This allowed us to determine which species are most or least prone to extirpation from anthropogenic island systems to anticipate species losses driven by existing and future hydroelectric impoundments in lowland tropical forests. Fourth, how divergent are observed and detectability-corrected estimates of island occupancy for rarely detected species? We show distinct responses to forest loss between forest specialists and habitat generalists, and that forest fragmentation *per se* may not affect the overall number of species in forest islands. We also emphasize the use of rarity metrics to assess species vulnerability to forest fragmentation, and caution against the use of occupancy models to infer island occupancy rates when species detection probability is low.

2.3 Methods

2.3.1 Study area

This study was carried out within the vast Tucuruí Hydroelectric Reservoir (hereafter, THR; 4°16′ S, 49°34′ W), located in the State of Pará, eastern Brazilian Amazonia (Figure 2.1). The reservoir was formed in 1984 when the Tocantins River was dammed, flooding over 250,000 ha of pristine lowland forests and creating some 2,200 islands on higher elevation terrain. In 2002, the entire archipelago and surrounding areas were set aside as a sustainable-use reserve (IUCN category VI), spanning 568,667 ha. This protected area—Tucuruí Lake Environmental

Protection Area (APA Lago de Tucuruí, in Portuguese)—is a multiple-use mosaic designated to meet both the interests of local communities and wildlife conservation.
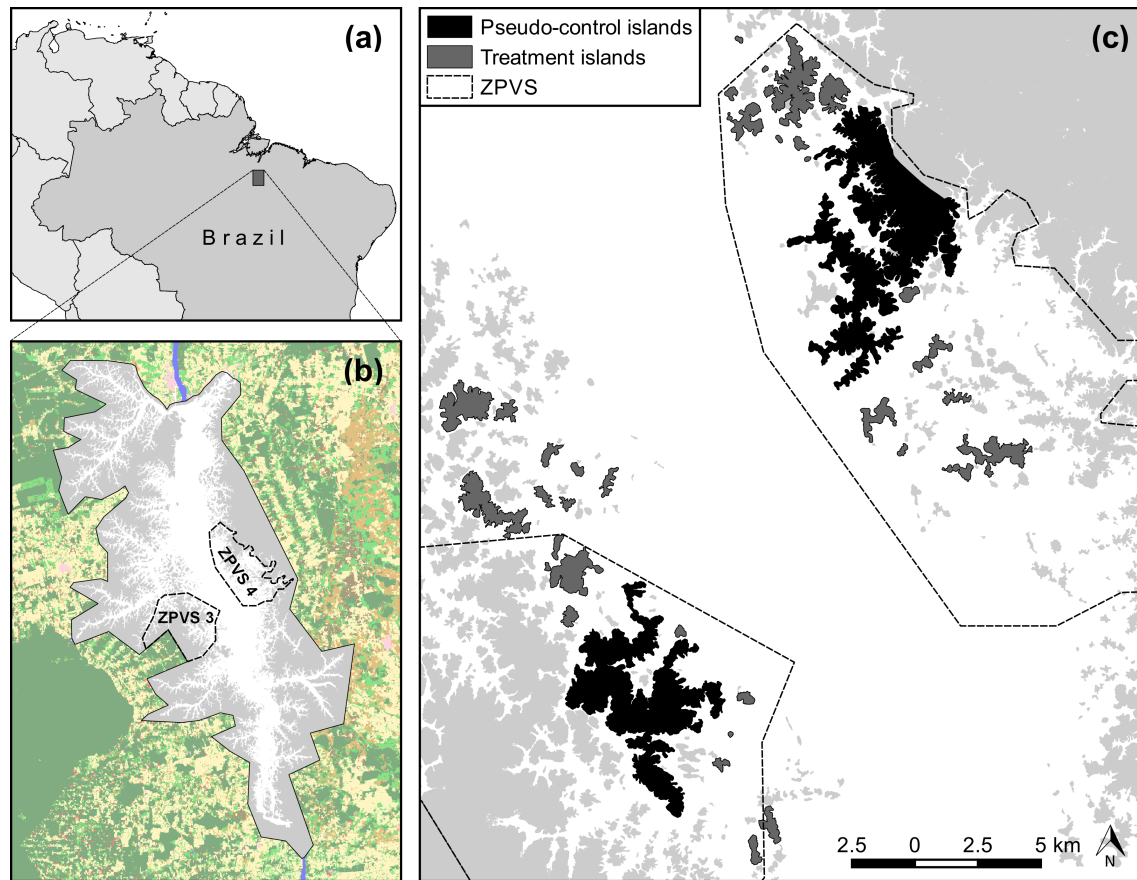
Figure 2.1: (a) Location of the study area in eastern Brazilian Amazonia; (b) Tucuruí Hydroelectric Reservoir (THR) within the Tucuruí Lake Environmental Protection Area (grey and white areas), showing the two Wildlife Conservation Zones (ZPVS 3 and 4, indicated by dotted lines) and heavily degraded areas surrounding the THR (yellow); and (c) distribution of the 36 surveyed islands (dark grey and black polygons) within or adjacent to the two ZPVS. The background image was extracted from the TerraClass project (de Almeida et al., 2016), available at http://www.inpe.br/cra/projetos_pesquisas/terraclass2008.php.

The vegetation is typical of Amazonian *terra firme* forests, containing 80%–90% forest cover and an understorey dominated by several palm species (Ferreira, Neckel-Oliveira, Galatti, Fáveri, & Parolin, 2012). The climate is equatorial monsoonal (Am), with a rainy season from December to May and a dry season from June to November (Alvares et al., 2013). Mean annual precipitation and temperature are 2,354 mm and 27.5 °C, respectively (Alvares et al., 2013).

The THR is located in the most deforested region of Brazilian Amazonia, known as the "Arc of Deforestation", and encompasses both the Xingú and Belém lowland areas of endemism, which are separated by the Tocantins River (da Silva,

Rylands, & Da Fonseca, 2005). To survey the forest avifauna of the reservoir, we selected an even number of islands across a comparable size range on each bank of the former river channel. Many islands and mainland sites surrounding the reservoir were heavily degraded, but we surveyed a set of 36 relatively undisturbed forest islands located within ($n = 26$) or adjacent to ($n = 10$) the two Wildlife Protection Zones (ZPVS): ZPVS 3 on the left bank and ZPVS 4 on the right bank (Figure 2.1b). The two largest islands ($> 1,800$ ha) were defined as "pseudo-controls", and 34 smaller islands as "treatments", which were selected to maximise the range of island sizes, shapes and degrees of connectivity (see Table S2.1). The pseudo-control island on the right bank is actually a mainland peninsula that was semi-isolated along a boundary of secondary forest.

2.3.2   Avian surveys

We conducted six field campaigns over a 15-month period: 6–25 August and 12–29 November in 2006, and 4–22 March, 12 April–1 May, 14–31 July and 22 September–10 October in 2007. During each field campaign, all 36 islands were surveyed once using 10-min point counts by experienced observers (S.M.D. or L.M.P.H.) accompanied by a field assistant, who simultaneously recorded bird vocal activity (using a Sony TCM-5000 recorder and a semi-directional microphone) as a voucher of species occurrences. To ensure that all birds recorded were within surveyed islands, we restricted all individuals seen or heard to within an estimated 50-m radius from the observer and discarded all flyovers. Given our focus on diurnal forest species, we also discarded all aquatic, nocturnal and aerial species, as well as every species that "does not normally occur in forest" (*sensu* BirdLife International, 2017). Surveys were usually carried out between 06:00 and 10:00 h avoiding rainy and windy weather. The number of point count stations (hereafter, PCs) surveyed per island, which ranged between 2 and 33 (see Table S2.1), was roughly proportional to island area on a *log-log* scale ($r^2_{adj} = 0.863$, $P < 0.001$). All 36 islands were surveyed along linear transects—three of each placed at the two pseudo-control islands (see Figure S2.1) and one at each of the 34 treatment islands—along which PCs were distributed at regular 200-m intervals. A total of 240 PCs were visited six times each, amounting to 1,388 samples.

Sampling sufficiency per island was represented by individual-based rarefaction curves produced with 1,000 bootstrap replications in the INEXT R package (Hsieh, Ma, & Chao, 2016; R Core Team, 2016). Sampling completeness per island was quantified as a percentage between the recorded and the estimated number of species based on the first-order Jackknife estimator (Willie, Petre, Tagg, & Lens, 2012) calculated using the VEGAN package (Oksanen et al., 2017).

2.3.3   Species traits

We classified the degree of habitat specialisation of each species into "forest specialist" or "habitat generalist" based on two attributes extracted from BirdLife International (2017), namely "forest dependency" and "habitats". Species had to meet two criteria to be classified as forest specialists: (1) "high" forest dependency, and (2) "Forest—Subtropical/Tropical Moist Lowland"—the equivalent to Amazonian lowland forest—listed as a habitat of "major" importance. Species that did not meet these criteria were classified as habitat generalists. Accordingly, a habitat generalist is a species that occurs in forest (i.e., "low", "medium" or "high" forest dependency) but does not have "Forest—Subtropical/Tropical Moist Lowland" listed as a habitat of "major" importance. For example, *Pitangus sulphuratus* has a "low" forest dependency and occurs throughout nine habitat types, including "Forest—Subtropical/Tropical Moist Lowland" which is listed as a habitat of "suitable" importance. Likewise, *Onychorhynchus coronatus* was classified as a habitat generalist, despite its "high" forest dependency, because this species is mostly associated with riparian habitats (Bueno, Bruno, Pimentel, Sanaiotti, & Magnusson, 2012) and this habitat type (i.e., "Forest—Subtropical/Tropical Swamp") was inundated by the THR floodwaters. Habitat generalist is then a species that may use the "Forest—Subtropical/Tropical Moist Lowland" habitat as an alternative habitat. Habitat specialisation was used to examine whether patterns of species loss differed between forest specialists and habitat generalists.

We also compiled data on seven additional traits associated with avian extinction risk (Sodhi, Liow, & Bazzaz, 2004): body mass, trophic level, vertical stratum, flocking behaviour, geographic range size, habitat breadth and natural abundance (herein defined as the total number of individuals recorded within pseudo-control islands; see Table S2.2 for variable descriptions and sources, and Table S2.3 for species traits). We log_{10}-transformed body mass (g), geographic range size (km^2) and natural abundance ($n + 1$) prior to analysis. Trophic level is a continuous variable estimated from proportional consumption of food items across five diet categories. Vertical stratum and flocking behaviour were converted from nominal to ordinal (rank) variables to produce a gradient from ground to canopy strata, and from low to high levels of gregariousness, respectively. Stotz, Fitzpatrick, Parker, and Moskovits (1996) classified 41 habitats for the Neotropical avifauna and assigned one to seven habitats used by each species. Habitat breadth was then defined as a count variable representing the number of habitats used, with lower and higher values indicating restricted and broad habitat breadth, respectively. These traits were used to assess patterns of species

occupancy across all 36 surveyed islands.

2.3.4   Island and landscape metrics

We used four RapidEye© imagery tiles (250,000 ha at 5-m resolution) covering all surveyed islands and an unsupervised classification performed in ESRI ARCMAP 10.2 to produce a categorical map with two land-cover classes: island and water (Figure 2.1c). We then extracted three spatial metrics for each island: island area in hectares (AREA), shape index (SHAPE), and proximity index (PROX). SHAPE is a measure of the deviation in the perimeter of a given island from the perimeter (m) of a perfect circle with the same area (m^2), and calculated as

$$\frac{\text{PERIMETER}}{2\sqrt{\pi \times \text{AREA}}},$$

with lower and higher values indicating simple and complex shapes, respectively (Burchell, Shake, Moorman, Riddle, & Burchell, 2012). PROX (*sensu* McGarigal, Cushman, & Ene, 2012) was used as a measure of connectivity, and considered the total area of any island ($\geqslant 1$ ha) that was partially or entirely within a 500-m external buffer (Benchimol & Peres, 2015), with smaller values indicating lower connectivity or higher isolation. We arbitrarily assigned a PROX value one order of magnitude greater than the most connected island for pseudo-control islands, and a value of 0.01 for the least connected island. Finally, we log_{10}-transformed both AREA and PROX prior to analysis.

2.3.5   Species–area relationships and forest fragmentation effect

The logarithmic form of the species–area relationship (type IV curve *sensu* Scheiner, 2003) was used to allow us to fit simple linear regression models (hereafter, SAR models; Rosenzweig, 1995) for three data sets—all species, only forest specialists, and only habitat generalists—according to the following equation:

$$log_{10}(S) = z \times log_{10}(A) + log_{10}(c),$$

where S = number of species, z = regression slope, A = island area (ha), c = regression intercept. As forest specialists were not recorded at one small surveyed island, S was standardised as $log_{10}(n+1)$. To test whether z-values for forest specialists and habitat generalists were significantly different ($P \leqslant 0.05$), we performed an ANCOVA model with habitat specialisation as the categorical independent variable.

We examined whether forest fragmentation *per se* either exacerbates or reduces species loss as a function of forest loss (i.e., island area reduction) following Yaacobi et al. (2007). Accordingly, after fitting SAR models for each data set, we extrapolated the number of species to a hypothetical island with the combined area of all 36 surveyed islands (6,502.6 ha). We then compared the overall number of species recorded across the whole set of surveyed islands (i.e., gamma diversity) with the extrapolated number of species to the hypothetical island (i.e., predicted alpha diversity) for each data set. If the gamma diversity is lower or higher than the predicted alpha diversity of the hypothetical island, forest fragmentation will have either exacerbated or reduced species loss, respectively. In other words, additional factors other than forest loss operate in explaining the gamma diversity, which we attributed to forest fragmentation. Values were considered significantly different if the overall number of species recorded was outside the 95% confidence interval of the extrapolated number of species. As the accuracy of this method relies on SAR model fits, we deemed the method as appropriate if the z-value was significant and the r^2_{adj} was $\geqslant 0.5$ (Matthews, Triantis et al., 2016).

Previous studies have shown that departures in the overall number of species recorded from that predicted by extrapolating SAR models are related to the nested structure of species assemblages (Matthews, Triantis et al., 2016; Santos et al., 2010). To examine how the degree of nestedness relates to the fragmentation effect on species richness, we quantified the nested structure of the three data sets using the nestedness metric based on overlap and decreasing fill (NODF) as this metric is statically robust to overestimating nestedness (i.e., type I statistical errors; Almeida-Neto, Guimarães, Guimaraes, Loyola, & Ulrich, 2008). We used the nodf-program, version 2.0 (Almeida-Neto & Ulrich, 2011), to calculate NODF values for all three data sets and for 1,000 simulated assemblages generated with the proportional-row and proportional-column (PP) null model algorithm (Ulrich & Gotelli, 2012). NODF Z-transformed scores (hereafter, Z-scores) were then used to determine whether the nested (positive Z-scores) or antinested (negative Z-scores) structure of species assemblages were significantly different from those of simulated assemblages (Matthews, Cottee-Jones, & Whittaker, 2015).

2.3.6   Species vulnerability to forest fragmentation

Species vulnerability to forest fragmentation corresponds to the risk of a species to become locally extinct across the whole set of forest patches remaining in the landscape. Hence, species occurring in a few patches would be more extinction-prone than those occurring in many patches, particularly if an extinction debt

has yet to be paid and patch colonization rates are low, which is likely the case of forest archipelagos within hydroelectric reservoirs (Jones, Bunnefeld, Jump, Peres, & Dent, 2016). In this case, patch occupancy is inversely related to vulnerability to forest fragmentation. However, species absences from a patch does not necessarily imply that local extinctions had occurred because such species could be initially absent from the patch at the time of its creation (Bolger, Alberts, & Soule, 1991), meaning that patch occupancy may not always indicate vulnerability to forest fragmentation (Keinath et al., 2017). Therefore, we first examined whether local extinction had actually occurred across surveyed islands by comparing the SARs for birds in pseudo-control islands with that in much smaller treatment islands (Bolger et al., 1991; Brown, 1971; Wang, Zhang, Feeley, Jiang, & Ding, 2009). To accomplish this, we used the number of bird species as a function of surveyed area (number of PCs × point count area), and performed an ANCOVA model with island type as the categorical independent variable (see Appendix S2.1 for further details). We found that the predicted line derived from treatment islands was well below and had a steeper slope than that of pseudo-control islands (see Figure S2.2a), indicating that local extinctions had occurred in the former. Subsequently, we estimated the number of local extinctions that had occurred in each treatment island by subtracting the predicted number of species in pseudo-control islands from the recorded number of species in treatment islands (Bolger et al., 1991). Accordingly, we estimated a total of 788 local extinctions across all 34 treatment islands over 22–23 years of post-isolation history at the THR landscape (see Figure S2.2b). Given these results, we used estimates of island occupancy (i.e., proportion of islands occupied—PIO) as a measure of species vulnerability to forest fragmentation based on species occurrence across all 36 surveyed islands.

Due to potential biases introduced by imperfect detectability, we calculated both the observed and detectability-corrected PIO for each species (Thornton et al., 2011; Wang et al., 2015). The latter was calculated using single-season occupancy models (MacKenzie et al., 2002) implemented in the UNMARKED package (Fiske & Chandler, 2011). As some species can occasionally disperse across islands by traversing the water matrix and our bird surveys were conducted over six discrete field campaigns, we relaxed the closure assumption of single-season models, which is defensible as long as (1) changes in island occupancy status occur at random—which is likely the case—and (2) "occupancy" is interpreted as "use" (Mackenzie & Royle, 2005).

We modelled species occupancy probability (ψ) as a function of island AREA, SHAPE and PROX, assuming an interactive effect between AREA and SHAPE due to their combined effects in determining the severity of edge effects. As

sampling effort increases the chances of detecting any given species, we modelled the detection probability (p) as a function of the number of PCs per island (EFFORT). We also considered both ψ and p as constants across islands. Accordingly, we built 16 competitive occupancy models for each species (Table 1). We then used the Akaike information criterion (AIC) to rank models and to calculate Akaike weights to indicate the best-fit models (Burnham & Anderson, 2002). From model-averaging based on all models with high support (ΔAIC $\leqslant$ 2), we summed the occupancy probability at each island and divided this by the total number of surveyed islands to obtain the detectability-corrected PIO for each species (Thornton et al., 2011; Wang et al., 2015). We also summed the detection probability for each visit per island and divided by 216 (36 islands $\times$ 6 surveys) to obtain the overall detection probability for each species.

Table 2.1: Structure of the 16 occupancy models used to estimate detectability-corrected proportions of islands occupied for 207 bird species surveyed across 36 islands at the Tucuruí Hydroelectric Reservoir landscape. Probability of occupancy (ψ) was modelled as a function of log_{10} island area in hectares (AREA), shape index (SHAPE), and log_{10} proximity index (PROX). The probability of detection (p) was modelled as a function of the number of point count stations surveyed per island (EFFORT).

Model description
psi(.) p(.)
psi(AREA) p(.)
psi(SHAPE) p(.)
psi(PROX) p(.)
psi(AREA $\times$ SHAPE) p(.)
psi(AREA + PROX) p(.)
psi(SHAPE + PROX) p(.)
psi(AREA $\times$ SHAPE + PROX) p(.)
psi(.) p(EFFORT)
psi(AREA) p(EFFORT)
psi(SHAPE) p(EFFORT)
psi(PROX) p(EFFORT)
psi(AREA $\times$ SHAPE) p(EFFORT)
psi(AREA + PROX) p(EFFORT)
psi(SHAPE + PROX) p(EFFORT)
psi(AREA $\times$ SHAPE + PROX) p(EFFORT)

2.3.7   Species traits and vulnerability to forest fragmentation

It is widely assumed that closely related species share more traits than distantly related species (Webb, Ackerly, McPeek, & Donoghue, 2002). Thus, analyses

involving species as sampling units should be corrected for phylogenetic non-independence among traits (Freckleton, Harvey, & Pagel, 2002). To account for this, we built a majority-rule consensus tree based on 1,000 trees obtained from birdtree.org ("Hackett All Species"; Jetz, Thomas, Joy, Hartmann, & Mooers, 2012) using the APE package (Paradis, Claude, & Strimmer, 2004). As a consensus tree does not include branch lengths, we set all branch lengths equal to one. We then performed Phylogenetic Generalised Least Squares (PGLS) models using the CAPER package (Orme et al., 2013) and Pagel's lambda branch length transformation optimised by maximum likelihood (Freckleton et al., 2002). We examined both observed and detectability-corrected PIO separately as response variables and species traits as explanatory variables.

To assess the role of morpho-ecological traits in explaining species vulnerability to forest fragmentation, we built 13 competing PGLS models: a univariate model for each of the seven traits, three additive models and three interactive models. Additive and interactive models were built under the same combination of traits. The first included natural abundance, habitat breadth and geographic range size, and is referred to as "rarity model" as it combines all three dimensions of rarity (*sensu* Rabinowitz, 1981). The second included natural abundance, body mass and trophic level, and is referred to as "population size model", following Meyer et al. (2008). The third included trophic level, vertical stratum and flocking behaviour, and is referred to as "foraging model".

2.4   Results

Considering all 36 surveyed islands, we recorded 10,575 individuals representing 207 bird species, 150 genera and 31 families (see Table S2.3). The number of individuals recorded per island ranged from 28 to 1,997 (mean $\pm SD = 293.8 \pm 359.1$), and the number of species from 7 to 128 (46.3 ± 26.8). The number of individuals recorded per species ranged widely from 1 to 1,385 (51.1 ± 124.2).

Despite our large sampling effort, individual-based rarefaction curves indicate that further surveys would be necessary to reach sampling sufficiency (i.e., to approach the asymptote of the curves; see Figure S2.3). Completeness of the inventories per island ranged from 64% to 89% ($73.6 \pm 5.1\%$; see Figure S2.3). As near-exhaustive inventories ($> 80\%$ completeness) were only obtained at four islands (see Figure S2.3), the number of species in most surveyed islands should be regarded as conservative.

2.4.1   Species–area relationships and forest fragmentation effect

Island area had a significant positive effect on the number of species for all species, forest specialists and habitat generalists (Figure 2.2). The z-value for habitat generalists was significantly lower than for forest specialists ($P = 0.028$; Figure 2.2), indicating that the rate of species loss as a function of island area reduction was higher for forest specialists.
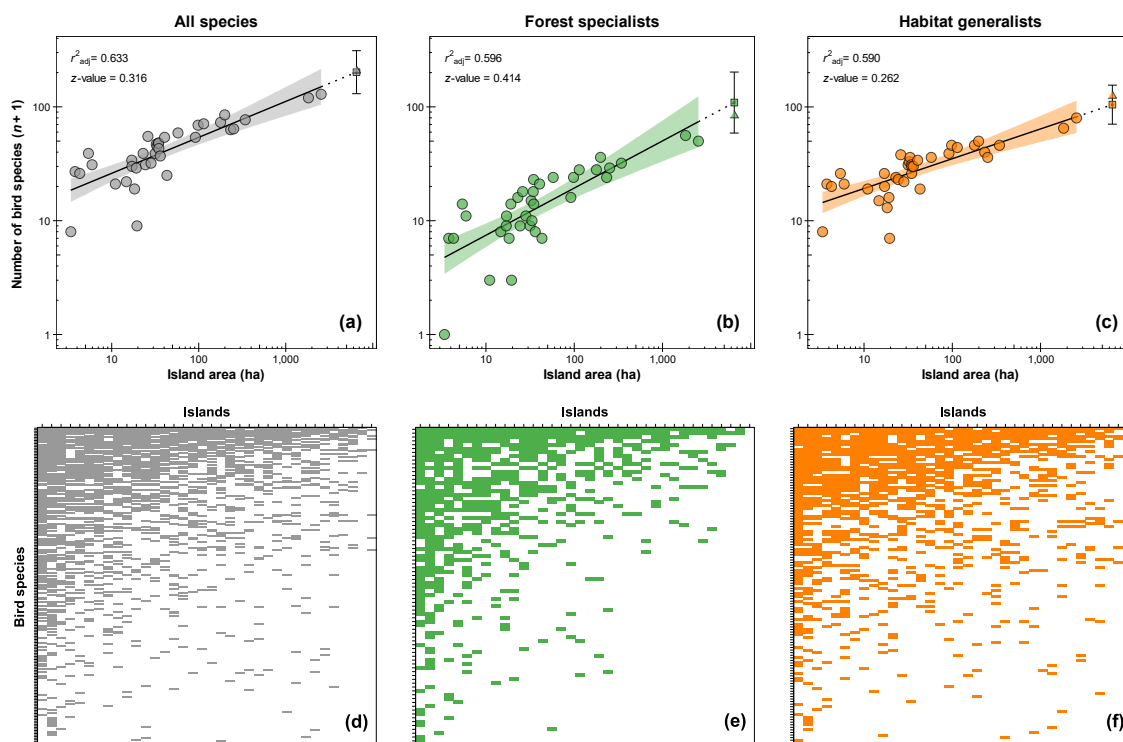
Figure 2.2: Plots at the top show the species–area relationships, and their r^2_{adj}- and z-values for (a) all species; (b) forest specialists; and (c) habitat generalists surveyed across 36 islands at the Tucuruí Hydroelectric Reservoir landscape ($P < 0.001$ in all instances). Dotted lines indicate null predicted numbers of species if forest fragmentation had no effect. Circles, squares and triangles correspond to the recorded, extrapolated, and overall number of species, respectively. Coloured regions and error bars show the 95% confidence intervals of predicted lines and extrapolated values, respectively. Note the base 10 logarithmic scales along both axes. Plots at the bottom show the maximally packed matrices for (d) all species; (e) forest specialists; and (f) habitat generalists based on the NODF nestedness metric (Almeida-Neto et al., 2008). Coloured bars indicate the islands (x-axis) where each species (y-axis) was recorded. None of the data sets was either significantly nested or antinested.

The SAR models were deemed as appropriate to assess the forest fragmentation effect on avian species richness as the z-value was significant and the r^2_{adj} was $\geqslant 0.5$ for all three data sets (Figure 2.2). We recorded a higher overall number of species in surveyed islands than that extrapolated to an unfragmented forest

area of 6,502.6 ha, the aggregate size of all 36 surveyed islands, considering both all species (207 + 1 > 201.6; Figure 2.2a) and only habitat generalists (124 + 1 > 104.7; Figure 2.2c). In contrast, this trend was reversed for forest specialists (83 + 1 < 109.1; Figure 2.2b). However, the difference between the recorded and extrapolated number of species was not significant for all three data sets.

The Z-scores for all species (0.34), forest specialists (0.24) and habitat generalists (0.64) were not statistically significant, indicating that the structure of all three data sets could not be described as either nested or antinested (Figure 2.2).

2.4.2   Trait-based vulnerability to forest fragmentation

Considering the observed PIO as a response variable, only the interactive PGLS "rarity model" including natural abundance, habitat breadth and geographic range size was highly supported based on AIC values (Table 2). This model explained most of the variance in observed PIO ($R^2_{adj} = 0.649$), outperforming the univariate PGLS models of natural abundance ($r^2_{adj} = 0.554$), habitat breadth ($r^2_{adj} = 0.031$) and geographic range size ($r^2_{adj} = 0.017$). Accordingly, species with higher natural abundance (Figure 2.3), broader habitat breadth and wider geographic range tended to have higher values of observed PIO (see Figure S2.4).

Considering the detectability-corrected PIO as a response variable, only the interactive PGLS "rarity model" was highly supported based on AIC values (Table 2). This model explained a fifth of the variance in detectability-corrected PIO ($R^2_{adj} = 0.199$), outperforming the univariate models of natural abundance ($r^2_{adj} = 0.113$), habitat breadth ($r^2_{adj} = 0.047$) and geographic range size ($r^2_{adj} = 0.018$). Accordingly, species with higher natural abundance, broader habitat breadth and wider geographic range tended to have higher values of detectability-corrected PIO (see Figure S2.4).

Table 2.2: Performance of 13 Phylogenetic Generalised Least Squares (PGLS) models relating either observed or detectability-corrected estimates of island occupancy to seven morpho-ecological traits, and combinations thereof, for 207 bird species surveyed across 36 islands at the Tucuruí Hydroelectric Reservoir landscape.

Model description	df	AIC	ΔAIC	ωi	R^2_{adj}
Response variable: observed proportion of islands occupied					
Univariate models					
Natural abundance	2	1,697.151	43.918	$2.901 \times 10{-}10$	0.554
Habitat breadth	2	1,847.325	194.091	$7.126 \times 10{-}43$	0.031
Geographic range size	2	1,850.594	197.361	$1.390 \times 10{-}43$	0.017
Body mass	2	1,851.311	198.077	$9.713 \times 10{-}44$	0.015
Flocking behaviour	2	1,854.150	200.917	$2.349 \times 10{-}44$	−0.001
Vertical stratum	2	1,854.443	201.209	$2.029 \times 10{-}44$	−0.002
Trophic level	2	1,854.845	201.611	$1.659 \times 10{-}44$	−0.004
Additive models					
Rarity: natural abundance + habitat breadth + geographic range size	4	1,666.006	12.773	0.0016815	0.620
Population size: natural abundance + body mass + trophic level	4	1,700.690	47.457	$4.945 \times 10{-}11$	0.551
Foraging: trophic level + vertical stratum + flocking behaviour	4	1,857.672	204.438	$4.037 \times 10{-}45$	−0.007
Interactive models					
Rarity: natural abundance × habitat breadth × geographic range size	8	1,653.233	0	0.9983185	0.649
Population size: natural abundance × body mass × trophic level	8	1,703.471	50.238	$1.231 \times 10{-}11$	0.553
Foraging: trophic level × vertical stratum × flocking behaviour	8	1,862.848	209.614	$3.035 \times 10{-}46$	−0.015

Continued on next page

Table 2.2: *Continued from previous page*

Response variable: detectability-corrected proportion of islands occupied

Model description	df	AIC	ΔAIC	ωi	R^2_{adj}
Univariate models					
Natural abundance	2	1,877.104	15.115	$4.063 \times 10{-}04$	0.113
Habitat breadth	2	1,891.950	29.962	$2.427 \times 10{-}07$	0.047
Geographic range size	2	1,898.307	36.319	$1.011 \times 10{-}08$	0.018
Body mass	2	1,901.063	39.074	$2.549 \times 10{-}09$	0.005
Trophic level	2	1,902.062	40.074	$1.546 \times 10{-}09$	-0.001
Flocking behaviour	2	1,902.317	40.328	$1.361 \times 10{-}09$	-0.002
Vertical stratum	2	1,902.804	40.815	$1.067 \times 10{-}09$	-0.004
Additive models					
Rarity: natural abundance + habitat breadth + geographic range size	4	1,864.512	2.524	0.2203577	0.173
Population size: natural abundance + body mass + trophic level	4	1,875.745	13.756	0.0008016	0.128
Foraging: trophic level + vertical stratum + flocking behaviour	4	1,905.494	43.505	$2.780 \times 10{-}10$	-0.008
Interactive models					
Rarity: natural abundance $\times$ habitat breadth $\times$ geographic range size	8	1,861.988	0	0.7782926	0.199
Population size: natural abundance $\times$ body mass $\times$ trophic level	8	1,879.215	17.227	0.0001414	0.130
Foraging: trophic level $\times$ vertical stratum $\times$ flocking behaviour	8	1,909.816	47.828	$3.202 \times 10{-}11$	-0.010
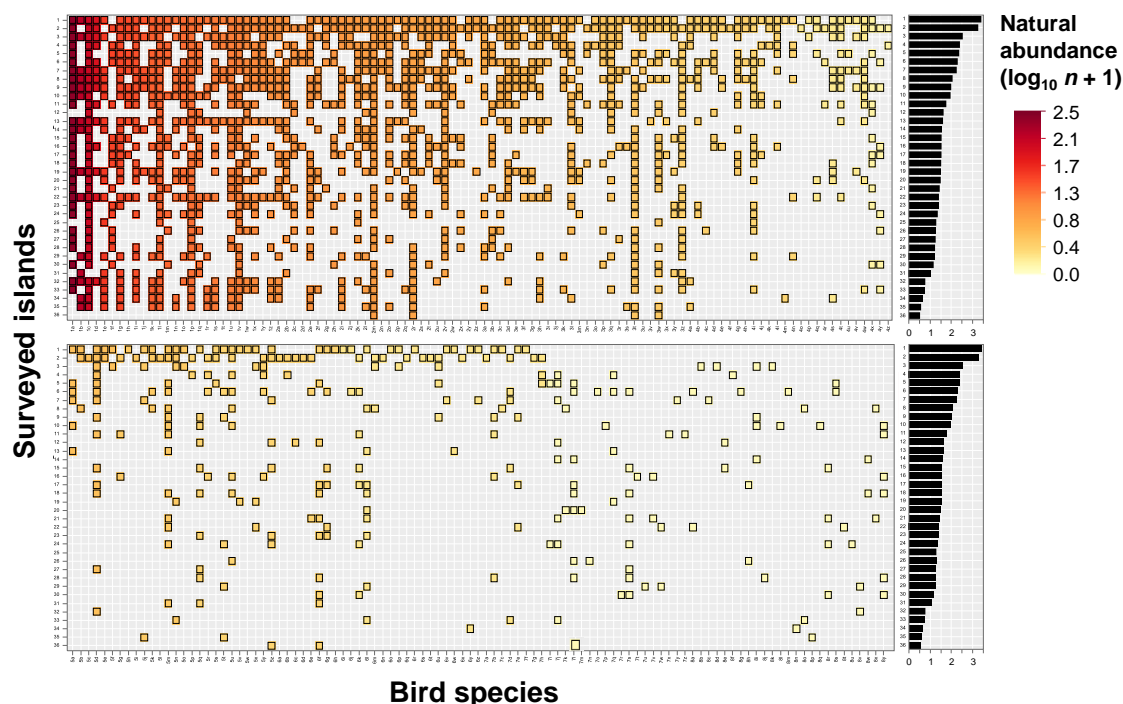
Figure 2.3: Site-by-species incidence matrix for 207 bird species surveyed across 36 islands at the Tucuruí Hydroelectric Reservoir landscape. Squares representing at least one individual recorded per island are coloured according to the natural abundance of each species, defined as the total number of individuals recorded within pseudo-control islands. Islands are ordered from the largest to the smallest (black bars on a log_{10} scale; see Table S2.1); species are ordered from the most to the least naturally abundant (see Table S2.3 for species codes).

2.4.3   Observed versus detectability-corrected estimates of island occupancy

Vulnerability to forest fragmentation was widely variable across the 207 species in terms of the proportion of islands occupied (PIO), regardless of whether we considered observed or detectability-corrected PIO (see Table S2.3). The variation in observed PIO ranged from 2.8% to 94.4% (22.4 $\pm$ 22.5%), whereas the variation in detectability-corrected PIO ranged from 5.6% to 96.4% (42.4 $\pm$ 24.4%). Estimates of island occupancy corrected for imperfect detectability were higher than those observed for 200 species, identical for five, and slightly lower for two (see Table S2.3). For 30 species with detection probabilities below 30%, the detectability-corrected PIO was at least tenfold higher than the observed PIO (16.0 $\pm$ 5.1%; Figure 2.4; see Table S2.3). Conversely, differences between these two estimates were negligible or non-existent (1.02 $\pm$ 0.02%) for all 31 species with detection probabilities exceeding 45.5% (Figure 2.4; see Table S2.3). Once phylogenetic non-independence was accounted for, detection probabilities was higher for more naturally abundant species (r^2_{adj} = 0.202, P < 0.001; Figure 2.4).
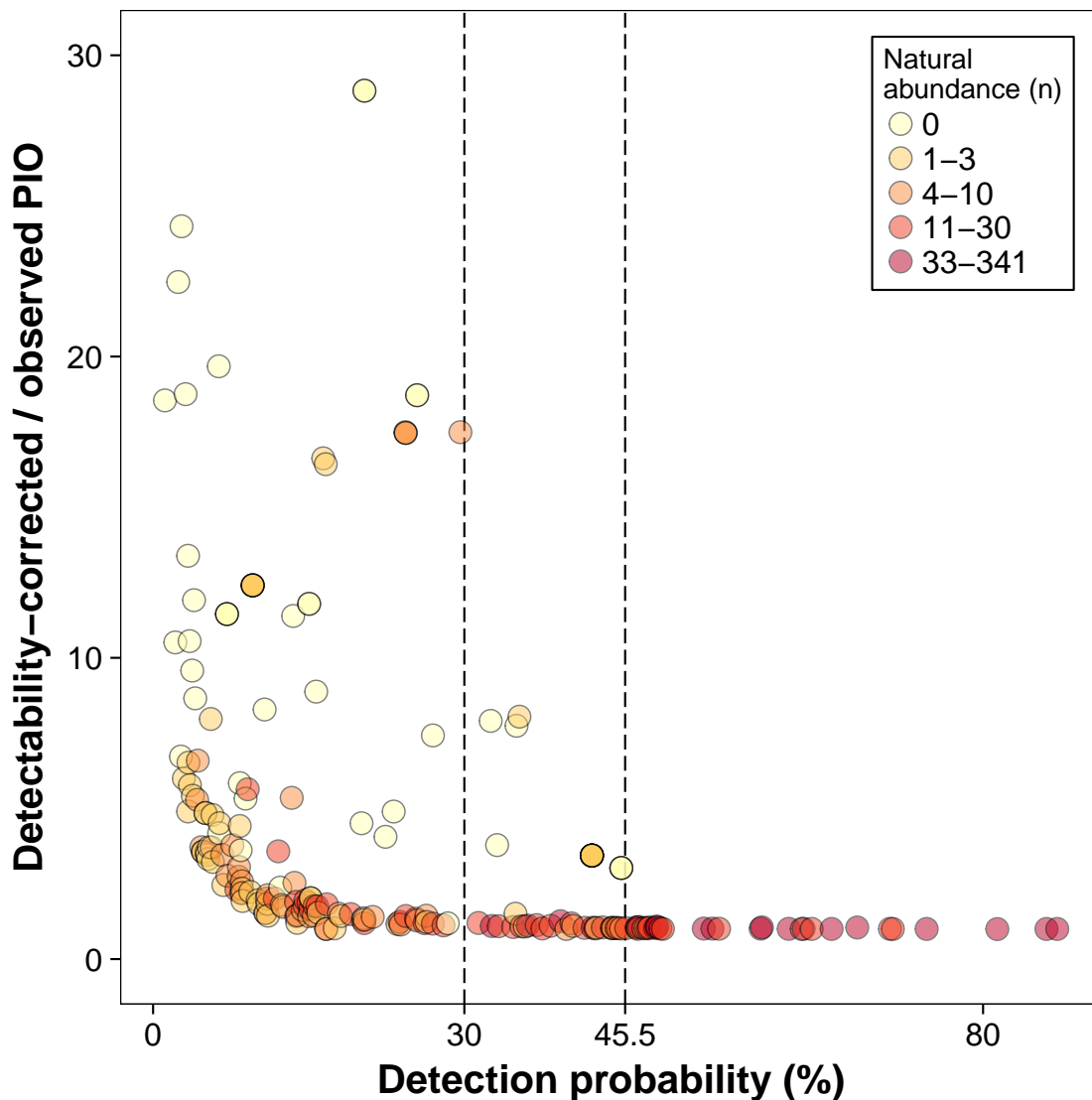
Figure 2.4: Ratio between detectability-corrected and observed estimates of proportion of islands occupied (PIO) as a function of species detectability for 207 bird species surveyed across 36 islands at the Tucuruí Hydroelectric Reservoir landscape; *y*-values indicate how many times detectability-corrected estimates are higher than observed estimates. Circles are coloured according to the natural abundance of each species, defined as the total number of individuals recorded within pseudo-control islands.

2.5   Discussion

Here, we present one of the most comprehensive landscape-scale efforts to date to assess the role of morpho-ecological traits in explaining species vulnerability to forest fragmentation, in terms of the number of surveyed islands (*n* = 36), range of island sizes (3.4–2,551.5 ha), overall sampling effort (*n* = 1,388 samples) and number of species surveyed (*n* = 207). This effort exploited a quasi-experimental

anthropogenic tropical forest archipelago, following an even-aged post-isolation history of 22–23 years, and allowed us to uncover which traits pose the greatest threats to bird species in forest islands within hydroelectric reservoirs. We also highlight potentially misleading applications of species occupancy models by contrasting observed and detectability-corrected estimates of island occupancy.

2.5.1   Species–area relationships and forest fragmentation effect

Although SARs are arguably the most ironclad relationship in ecology (Rosenzweig, 1995), rates of species loss induced by declining habitat areas are highly variable. Triantis, Guilhaumon, & Whittaker (2012) synthesised 449 data sets from *log-log* SAR applications to islands in inland, continental-shelf and oceanic systems, and reported z-values ranging from 0.064 to 1.312 (mean $\pm$ SD = 0.321 $\pm$ 0.164). Such variance was attributed to several factors, namely island type, taxonomic group and range of island areas (Triantis, Guilhaumon, & Whittaker, 2012). A reliable comparison of z-values among studies should therefore take these factors into account.

Z-values derived for forest islands have been shown to be higher than for forest fragments (Matthews, Guilhaumon, Triantis, Borregaard, & Whittaker, 2016), rendering forest fragmentation induced by hydroelectric dams (i.e., forest insularization) a more severe threat to biodiversity than that induced by agropastoral activities. We largely attribute such difference in z-values to the permeability of the intervening matrix, which may either preclude (increasing z-values; Moore, Robinson, Lovette, & Robinson, 2008) or allow species to disperse among forest patches, offsetting species losses through the rescue effect (decreasing z-values; Stouffer, Strong, & Naka, 2009). For instance, in a fragmented southern Amazonian landscape dominated by cattle pasture—where 338 bird species were surveyed across 30 forest fragments (1–14,476 ha)—Lees and Peres (2008) derived a z-value of 0.191, which is considerably lower than in this study (0.316). Although we do not have direct evidence on species dispersal in these two landscapes, both studies are comparable in most factors affecting z-values (Triantis, Guilhaumon, & Whittaker, 2012), except for the intervening matrix. Therefore, we predict that forest islands in existing and future hydroelectric reservoirs will experience a pronounced species richness decay, resulting in depauperate avian assemblages shaped by selective extinction (Mendenhall, Karp, Meyer, Hadly, & Daily, 2014; Si, Baselga, Leprieur, Song, & Ding, 2016; Wolfe et al., 2015).

Predictions of species losses based on the species–area relationship are affected by the degree of habitat specialisation of the species included in the analysis. In 16 of 23 data sets, avian species richness decreased at a greater rate as a

function of fragment area reduction for forest specialists than for habitat generalists (Matthews et al., 2014). Moreover, the inclusion of habitat generalist and edge species can even reverse the generally positive species–area relationship, whereby small patches will counter-intuitively harbour the most species-rich assemblages (Lövei, Magura, Tóthmérész, & Ködöböcz, 2006). In archipelagic landscapes, colonisation of habitat generalists into forest islands is expected to be hindered by the aquatic matrix, ultimately reducing their impact in reducing z-value estimates. In fact, our z-value derived for all species (0.316) approaches the mean value of island systems (0.321; Triantis, Guilhaumon, & Whittaker, 2012) rather than that of terrestrial landscapes (0.202; Watling & Donnelly, 2006). However, our z-value estimate for forest specialists (0.414) was significantly greater than that for habitat generalists (0.262). Including habitat generalists in the species pool therefore reduced our assemblage-wide rate of species loss, obscuring the more severe impact of habitat loss on forest specialists, which reinforces the notion that habitat patches must be defined from the perspective of target species (Lövei, Magura, Tóthmérész, & Ködöböcz 2006).

Forest fragmentation *per se* neither significantly decreased nor increased the overall number of species predicted by forest loss (i.e., island area reduction) regardless of the data set used, which corroborates our prediction regarding the fragmentation effect on all species (neutral) but not on both forest specialists (negative) and habitat generalists (positive). Likewise, species richness was unrelated to fragmentation in previous studies undertaken in different landscapes across a wide range of taxonomic groups. For instance, fragmentation effects on the overall number of species in forest fragments were not evident for perennial flowering plants and two beetle families in an agricultural landscape (Tenebrionidae and Carabidae; Yaacobi et al., 2007), and for butterflies in an urban landscape (Soga & Koike, 2012). Yet this failed to hold true for lizards in an archipelagic landscape, where the overall number of species in forest islands was significantly decreased by fragmentation (Wang et al., 2009). Hypothetically, terrestrial matrices can then buffer fragmentation effects as they are more permeable to species movements than water matrices (Soga & Koike, 2012), or even increase gamma diversity as shown for spider species in forest fragments of two agricultural landscapes in Israel (Gavish, Ziv, & Rosenzweig, 2012). To test this hypothesis, we reanalysed the bird data available from the Thousand Island Lake forest archipelago in China (Si, Baselga, & Ding, 2015) applying the same analysis carried out here (Yaacobi et al., 2007). We found no support for that hypothesis as forest fragmentation *per se* significantly increased the overall number of bird species in forest islands (60 recorded > 42.6 extrapolated; see Figure S2.5), which is partially explained by the low z-value (0.098; see Yu, Hu, Feeley, Wu, & Ding, 2012) and the antinested

structure (Si et al., 2015) of the avian assemblages in the Thousand Island Lake (Matthews, Triantis et al., 2016; Santos et al., 2010). Accordingly, antinested assemblages (i.e., species present at an island are not present at other islands) are expected to have a higher gamma diversity than nested assemblages (i.e., species present in smaller islands are subsets of larger islands; Santos et al., 2010), ultimately determining the direction (positive or negative) and magnitude of the fragmentation effect on species richness. In this study, the lack of fragmentation effects on species richness for the data sets including all species, forest specialists, and habitat generalists was thus unsurprisingly given the non-significant nested structure of these avian assemblages. Collectively, this indicates that the extrapolation of SAR models is an indirect method to infer the nested structure of species assemblages (this study; Santos et al., 2010; Matthews, Triantis et al., 2016).

In a recent SLOSS-type analysis, Fahrig (2017) uncovered a significantly higher overall number of species in several small patches compared to a single large patch based on 60 compiled data sets. This suggests that habitat fragmentation *per se* increases the overall number of species in habitat patches, but we caution against such assertion for three reasons. First, antinested assemblages are shaped by species turnover, which depends on landscape-dispersal processes determined by isolation (with lower isolation leading to greater antinested structure; Santos et al., 2010), matrix permeability (Stouffer, Johnson, Bierregaard, & Lovejoy, 2011) and species dispersal ability (Si, Pimm, Russell, & Ding, 2014). Second, methodological choices may lead to biased outcomes as exemplified by the widespread nested structure of species assemblages in fragmented landscapes (Watling & Donnelly, 2006; $n = 67$ data sets), which were recently deemed as an analytical artefact as most species assemblages were neither significantly nested nor antinested (Matthews et al., 2015; $n = 97$ data sets). As fragmentation effects on species richness are strictly related to the nested structure of species assemblages, we believe that a fragmentation effect on species richness would not be evident for most studies compiled by Matthews et al. (2015). This contradicts Fahrig's (2017) conclusions, which were largely grounded on the positive fragmentation effect on species richness when comparing species accumulation curves of sites ordered according to either increasing or decreasing patch area (Quinn & Harrison, 1988). Nevertheless, this method is biased towards detecting higher species richness in several small patches compared to a single large patch due to unequal sampling intensity (i.e., proportion of patch area that is surveyed) among surveyed patches (Gavish et al., 2012). Third, an assemblage-level approach may mask fragmentation effects on individual species, as measures of species richness completely disregard species identity.

2.5.2 Trait-based vulnerability to forest fragmentation

Rarity is an intrinsic property of certain species that results from variable cross-scale combinations of small local population size, restricted habitat breadth and narrow geographic range (Rabinowitz, 1981). Rare species are inherently predisposed to high extinction risk, which justifies the use of rarity as a measure of species vulnerability to a wide range of anthropogenic stressors (Kattan, 1992; Mace et al., 2008). Using a global-scale analysis, Newbold et al. (2014) reported that forest specialists and narrow-range bird species from tropical and subtropical forest biomes are more vulnerable to land-use change than habitat generalists and wide-range species. We corroborate this outcome at the scale of an archipelagic landscape, and endorse other comparative analyses incorporating field data (i.e., estimates of local population size; Feeley et al., 2007) and synergistic interactions among species traits that amplify the power of predictive models (Wang et al., 2015). Moreover, we identified rarity as a decisive factor exacerbating species vulnerability at all three spatial dimensions defined by Rabinowitz (1981), particularly because rarity is unrelated to several key traits, such as body mass and flocking behaviour (Thiollay, 1994; but see Kattan, 1992). As such, species with higher natural abundance, broader habitat breadth and wider geographic range were in general those with the highest rates of island occupancy at the THR landscape. Nevertheless, natural abundance played a disproportionately important role compared to habitat breadth and geographic range size, a pattern corroborated in another Amazonian fragmented landscape (Lees & Peres, 2008). A positive abundance-occupancy relationship, in which more abundant species occupy more sites, is widely considered a general rule in ecology (Hartley, 1998). Although many underlying mechanisms have been proposed to explain this relationship, there is no broadly accepted consensus as to why locally abundant species should be more ubiquitous (Gaston et al., 2000). We stress that our findings can be extended to other fragmented landscapes, including those dominated by variable-quality terrestrial matrices, in which non-random extirpations could also be predicted by metrics of rarity.

Based on our PGLS models, we failed to find support for some morpho-ecological traits that are often associated with avian extinction risk in human-modified tropical forest landscapes, namely body mass, trophic level, vertical foraging stratum and flocking behaviour (Sodhi et al., 2004). However, this does not imply that these traits are not meaningful (Hamer et al., 2015), although body mass, foraging specialisation and vertical stratum were unrelated to bird species vulnerability in a fragmented Atlantic Forest of southern Brazil (dos Anjos, 2006). In some instances, the role of species traits in predicting vulnerability to forest

fragmentation depends on the scale (global vs. landscape) and the response variable (e.g., population size vs. global extinction risk scores) used in the study (Keinath et al., 2017). For example, body mass has been often reported as a meaningful trait in broad-scale studies using global extinction risk scores (Keinath et al., 2017). Moreover, in model selection approaches, the best-fit models depend on the entire set of plausible competitive models (Aho, Derryberry, & Peterson, 2014). Had we considered only univariate models including each of those four traits separately, body mass (ΔAIC $\leqslant$ 2 in this instance) would have emerged as the most important trait in explaining observed island occupancy rates (Table 2), with small-bodied species occupying more islands than large-bodied species (r^2_{adj} = 0.015). Any given trait or combination of traits may therefore play a role in a comparative analysis, but collectively may operate as less meaningful variable (Keinath et al., 2017). Furthermore, the large number of species included in the analysis (n = 207) can obscure the role of ecological traits associated with only a few species (e.g., obligate ant-followers, n = 2), as the deviance of a few values may change the balance of strength in competing traits but not the main outcome.

It has been widely reported that insectivore species are particularly vulnerable to forest fragmentation (Bregman et al., 2014; Powell, Cordeiro, & Stratford, 2015), especially ground insectivores (Stratford & Stouffer, 1999) and obligate flocking species (i.e., mixed-species flock attendants and ant-followers; Van Houtan, Pimm, Bierregaard, Lovejoy, & Stouffer, 2006). Hence, species at higher trophic levels, using lower forest strata, and joining flocks were expected to exhibit lower rates of island occupancy. We failed to corroborate these expectations, which we largely attribute to differences in sampling design and analytical approaches among studies (Powell et al., 2015). For example, in an anthropogenic tropical forest archipelago in Malaysia, avian insectivores showed the steepest decline in the number of species with decreasing island area compared to either omnivores or frugivores (Yong et al., 2011). Had we applied the semi-log form of the species–area relationship [$S \sim log_{10}(A)$] to the same three avian foraging guilds, as the authors did, we would also have identified insectivores (*sensu* Wilman et al., 2014) as the most impaired foraging guild (see Figure S2.6). To provide further evidence of the impact of the analytical approach on the outcomes, we additionally applied the *log-log* form of the species–area relationship to both our data set and the data set available from the Malaysian archipelago (Yong et al., 2011). Although the outcomes converged between studies, at this time, frugivores emerged as the most impaired foraging guild, rather than insectivores (see Figure S2.6). Another noteworthy point is that species grouped into a foraging guild may span more than an entire trophic level (Hamer et al., 2015). As such, the trophic level of an insectivore species could overlap that of a carnivore

(Hamer et al., 2015), omnivore or granivore species (see Figure S2.7). In Bornean rainforests, insectivore species showed variable responses to selective logging, with species at higher trophic levels more adversely affected than those at lower trophic levels (Hamer et al., 2015). These authors used stable isotopes to quantify trophic levels, a more accurate approach than our energetic score, preventing a direct comparison between studies.

Ground insectivores were extirpated from small Amazonian forest remnants ($\leqslant$ 10 ha) following fragmentation (Stratford & Stouffer, 1999) as edge-dominated remnants could no longer sustain critical foraging microhabitats for these species (Stratford & Stouffer, 2013). Likewise, none of the five ground insectivores we recorded (*Conopophaga aurita*, *Conopophaga roberti*, *Formicarius analis*, *Formicarius colma* and *Hylopezus macularius*) was found in islands smaller than 30 ha (see Figure S2.8). Moreover, obligate flocking species were extirpated from small fragments ($\leqslant$ 10 ha) after isolation (Stouffer & Bierregaard, 1995), a pattern corroborated at the THR landscape, where smaller islands also harboured depauperate assemblage of these social species (see Figure S2.8). Although mixed-species flocks and obligate ant-followers can reassemble and recolonize small fragments following the regrowth of the intervening matrix (Stouffer & Bierregaard, 1995; Stouffer et al., 2011), these rebounds, by definition, cannot occur within hydroelectric reservoirs. Finally, the only comparable avian island biogeography study (Thousand Island Lake, China; Wang et al., 2015)—in terms of both the sampling design and analytical approach used here—is largely consistent with our findings, in which only natural abundance and habitat breadth had sufficiently high support in explaining species occupancy patterns in forest islands.

2.5.3 Observed versus detectability-corrected estimates of island occupancy

Occupancy modelling is assumed to derive more reliable estimates of patch occupancy as it accounts for potentially present species that go undetected in a given patch (MacKenzie et al., 2002). As a result, estimates of patch occupancy corrected for imperfect detectability are, as a general rule, equal to or higher than observed estimates (this study; Thornton et al., 2011; Wang et al., 2015). In an archipelagic landscape created by China's Thousand Island Lake, detectability-corrected proportions of islands occupied were up to sevenfold higher than that observed for a small raptor (*Accipiter soloensis*; Wang et al., 2015). At the THR landscape, those estimates were at least tenfold higher for 30 bird species, and almost 29-fold higher for two of them (Figure 2.4; see Table S2.3). These large

discrepancies can be explained by overestimates of patch occupancy for species with detection probabilities below 30% (MacKenzie et al., 2002). Overcoming this artefact to obtain more reliable estimates of patch occupancy would require increasing the number of samples per patch, but this is not always feasible due to logistical constraints (Mackenzie & Royle, 2005).

Estimates of patch occupancy for species with low detection probabilities ($< 30\%$) can be misleading and the large uncertainties they carry should be interpreted with caution (Welsh, Lindenmayer, & Donnelly, 2013). Such species may be defined as ubiquitous due to overestimates of patch occupancy, even though they have been recorded at only a few patches (Banks-Leite et al., 2014), which would invalidate species–specific predictions of vulnerability based on rates of patch occupancy. This was the case for *Myiopagis caniceps* and *Psarocolius bifasciatus*, which were recorded in only one island but were estimated to occupy 29. As species detectability tends to increase with increasing natural abundance, occupancy models yield far more reliable estimates of patch occupancy for common species than for those that are rare (Banks-Leite et al., 2014). Because over 200 species distributed across many lineages were considered in this study, identifying morpho-ecological characteristics that can best explain species vulnerability to forest fragmentation was largely unbiased. However, the same cannot be stated for species-poor assemblages in which most species have low detection probabilities. We argue that estimates of detectability-corrected proportions of patches occupied should always be reported and examined together with species detectability and observed estimates, to avoid misleading assessments of species vulnerability based on rates of patch occupancy.

2.6   Conclusions

On the basis of a comprehensive bird survey undertaken in forest islands within a major Amazonian hydroelectric reservoir, we addressed four questions: (1) Do habitat generalists show a less steep decline in species richness as a function of island area reduction compared to forest specialists? (2) Does forest fragmentation *per se* exacerbate or reduce the impact of forest loss on species richness for the overall species pool, forest specialists and habitat generalists? (3) Which suite of morpho-ecological traits best explains species rates of island occupancy within the forest archipelago? (4) How divergent are observed and detectability-corrected estimates of island occupancy for rarely detected species? Our findings show that (1) rates of species loss of forest specialists in land–bridge islands are underestimated if habitat generalists are included in the species pool because habitat generalists are less impacted by island area reduction than forest special-

ists; (2) fragmentation *per se* does not necessarily exacerbate the impact of forest loss on species richness; (3) rare species, especially those with low natural local abundance, are the most extinction-prone in fragmented landscapes; and (4) detectability-corrected estimates of island occupancy can be much higher than observed estimates for species with low detection probability, ultimately limiting the use of occupancy models for rare or elusive species. Finally, we conclude that forest islands within hydroelectric reservoirs are expected to typically harbour depauperate avian assemblages, mostly consisting of naturally abundant and habitat generalist species.

2.7   Acknowledgements

2.8 References

1. Aho, K., Derryberry, D., & Peterson, T. (2014). Model selection for ecologists: The worldviews of AIC and BIC. Ecology, 95, 631–636. https://doi.org/10.1890/13-1452.1

2. Almeida-Neto, M., Guimarães, P., Guimaraes, J. P. R., Loyola, R. D., & Ulrich, W. (2008). A consistent metric for nestedness analysis in ecological systems: Reconciling concept and measurement. Oikos, 117, 1227–1239. https://doi.org/10.1111/j.0030-1299.2008.16644.x

3. Almeida-Neto, M., & Ulrich, W. (2011). A straightforward computational approach for measuring nestedness using quantitative matrices. Environmental Modelling and Software, 26, 173–178. https://doi.org/10.1016/j.envsoft.2010.08.003

4. Alvares, C. A., Stape, J. L., Sentelhas, P. C., de Moraes, G., Leonardo, J., & Sparovek, G. (2013). Köppen's climate classification map for Brazil. Meteorologische Zeitschrift, 22, 711–728. https://doi.org/10.1127/0941-2948/2013/0507

5. Banks-Leite, C., Pardini, R., Boscolo, D., Cassano, C. R., Püttker, T., Barros, C. S., & Barlow, J. (2014). Assessing the utility of statistical adjustments for imperfect detection in tropical conservation science. Journal of Applied Ecology, 51, 849–859. https://doi.org/10.1111/1365-2664.12272

6. Benchimol, M., & Peres, C. A. (2015). Edge-mediated compositional and functional decay of tree assemblages in Amazonian forest islands after 26 years of isolation. Journal of Ecology, 103, 408–420. https://doi.org/10.1111/1365-2745.12371

7. BirdLife International (2017). IUCN Red List for birds. Retrieved from http://www.birdlife.org

8. Blake, J. G., & Loiselle, B. A. (2001). Bird assemblages in second-growth and old-growth forests, Costa Rica: Perspectives from mist nets and point counts. The Auk, 118, 304. https://doi.org/10.1642/0004-8038(2001)118[0304:BAISGA]2.0.CO;2

9. Bolger, D. T., Alberts, A. C., & Soule, M. E. (1991). Occurrence patterns of bird species in habitat fragments: Sampling, extinction, and nested species subsets. The American Naturalist, 137, 155–166. https://doi.org/10.1086/285151

10. 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. Biological Conservation, 169, 372–383. https://doi.org/10.1016/j.biocon.2013.11.024

11. Brown, J. H. (1971). Mammals on mountaintops: Nonequilibrium insular biogeography. The American Naturalist, 105, 467–478. https://doi.org/10.1086/282738

12. Bueno, A. S., Bruno, R. S., Pimentel, T. P., Sanaiotti, T. M., & Magnusson, W. E. (2012). The width of riparian habitats for understory birds in an Amazonian forest. Ecological Applications, 22, 722–734. https://doi.org/10.1890/11-0789.1

13. Burchell, M. R., Shake, C. S., Moorman, C. E., Riddle, J. D., & Burchell, M. R. II (2012). Influence of patch size and shape on occupancy by shrubland birds. Condor, 114, 268–278. https://doi.org/10.1525/cond.2012.110107

14. Burnham, K. P., & Anderson, D. R. (2002). Model selection and multi-model inference: A practical information-theoretic approach. New York, NY: Springer.

15. da Silva, J. M. C., Rylands, A. B., & Da Fonseca, G. A. B. (2005). The fate of the Amazonian areas of endemism. Conservation Biology, 19, 689–694. https://doi.org/10.1111/j.1523-1739.2005.00705.x

16. Dardanelli, S., Nores, M. L., & Nores, M. (2006). Minimum area requirements of breeding birds in fragmented woodland of Central Argentina. Diversity and Distributions, 12, 687–693. https://doi.org/10.1111/j.1472-4642.2006.00266.x

17. de Almeida, C. A., Coutinho, A. C., Esquerdo, J. C. D. M., Adami, M., Venturieri, A., Diniz, C. G., . . . Gomes, A. R. (2016). High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. Acta Amazonica, 46, 291–302. https://doi.org/10.1590/1809-4392201505504

18. Devictor, V., Julliard, R., & Jiguet, F. (2008). Distribution of specialist and generalist species along spatial gradients of habitat disturbance and fragmentation. Oikos, 117, 507–514. https://doi.org/10.1111/j.0030-1299.2008.16215.x

19. dos Anjos, L. (2006). Bird species sensitivity in a fragmented landscape of the Atlantic Forest in southern Brazil. Biotropica, 38, 229–234. https://doi.org/10.1111/j.1744-7429.2006.00122.x

20. Ewers, R. M., & Didham, R. K. (2006). Confounding factors in the detection of species responses to habitat fragmentation. Biological Reviews, 81, 117–142.

21. Fahrig, L. (2003). Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, Evolution, and Systematics, 34, 487–515. https://doi.org/10.1146/annurev.ecolsys.34.011802.132419

22. Fahrig, L. (2017). Ecological responses to habitat fragmentation *per se*. Annual Review of Ecology, Evolution, and Systematics, 48, 1–23. https://doi.org/10.1146/annurev-ecolsys-110316-022612

23. Feeley, K. J., Gillespie, T. W., Lebbin, D. J., & Walter, H. S. (2007). Species characteristics associated with extinction vulnerability and nestedness rankings of birds in tropical forest fragments. Animal Conservation, 10, 493–501. https://doi.org/10.1111/j.1469-1795.2007.00140.x

24. Ferreira, L. V., Neckel-Oliveira, S., Galatti, U., Fáveri, S. B., & Parolin, P. (2012). Forest structure of artificial islands in the Tucuruí dam reservoir in northern Brazil: A test core-area model. Acta Amazonica, 42, 221–226. https://doi.org/10.1590/S0044-59672012000200006

25. Fiske, I., & Chandler, R. (2011). unmarked: An R package for fitting hierarchical models of wildlife occurrence and abundance. Journal of Statistical Software, 43, 1–23.

26. Freckleton, R. P., Harvey, P. H., & Pagel, M. (2002). Phylogenetic analysis and comparative data: A test and review of evidence. The American Naturalist, 160, 712–726. https://doi.org/10.1086/343873

27. Gage, G. S., Brooke, M. D., Symonds, M. R., & Wege, D. (2004). Ecological correlates of the threat of extinction in Neotropical bird species. Animal Conservation, 7, 161–168. https://doi.org/10.1017/S1367943004001246

28. Gaston, K. J., Blackburn, T. M., Greenwood, J. J. D., Gregory, R. D., Quinn, R. M., & Lawton, J. H. (2000). Abundance-occupancy relationships. Journal of Applied Ecology, 37, 39–59. https://doi.org/10.1046/j.1365-2664.2000.00485.x

29. Gavish, Y., Ziv, Y., & Rosenzweig, M. L. (2012). Decoupling fragmentation from habitat loss for spiders in patchy agricultural landscapes. Conservation Biology, 26, 150–159. https://doi.org/10.1111/j.1523-1739.2011.01799.x

30. Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., ... Townshend, J. R. (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. Science Advances, 1, e1500052.

31. Hall, L. S., Krausman, P. R., & Morrison, M. L. (1997). The habitat concept and a plea for standard terminology. Wildlife Society Bulletin, 25, 173–182.

32. Hamer, K. C., Newton, R. J., Edwards, F. A., Benedick, S., Bottrell, S. H., & Edwards, D. P. (2015). Impacts of selective logging on insectivorous birds in Borneo: The importance of trophic position, body size and foraging height. Biological Conservation, 188, 82–88. https://doi.org/10.1016/j.biocon.2014.09.026

33. Hartley, S. (1998). A positive relationship between local abundance and regional occupancy is almost inevitable (but not all positive relationships are the same). Journal of Animal Ecology, 67, 992–994. https://doi.org/10.1046/j.1365-2656.1998.6760992.x

34. Henle, K., Davies, K. F., Kleyer, M., Margules, C., & Settele, J. (2004). Predictors of species sensitivity to fragmentation. Biodiversity and Conservation, 13, 207–251. https://doi.org/10.1023/B:BIOC.0000004319.91643.9e

35. Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: An R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods in Ecology and Evolution, 7, 1451–1456. https://doi.org/10.1111/2041-210X.12613

36. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of birds in space and time. Nature, 491, 444–448. https://doi.org/10.1038/nature11631

37. Jones, I. L., Bunnefeld, N., Jump, A. S., Peres, C. A., & Dent, D. H. (2016). Extinction debt on reservoir land-bridge islands. Biological Conservation, 199, 75–83. https://doi.org/10.1016/j.biocon.2016.04.036

38. Kattan, G. H. (1992). Rarity and vulnerability: The birds of the Cordillera Central of Colombia. Conservation Biology, 6, 64–70. https://doi.org/10.1046/j.1523-1739.1992.610064.x

39. Keinath, D. A., Doak, D. F., Hodges, K. E., Prugh, L. R., Fagan, W., Seker-cioglu, C. H., ... Kauffman, M. (2017). A global analysis of traits predicting species sensitivity to habitat fragmentation. Global Ecology and Biogeography, 26, 115–127. https://doi.org/10.1111/geb.12509

40. Kennedy, C. M., Marra, P. P., Fagan, W. F., & Neel, M. C. (2010). Landscape matrix and species traits mediate responses of Neotropical resident birds to forest fragmentation in Jamaica. Ecological Monographs, 80, 651–669. https://doi.org/10.1890/09-0904.1

41. Laurance, W. F., Camargo, J. L. C., Luizão, R. C. C., Laurance, S. G., Pimm, S. L., Bruna, E. M., ... Vasconcelos, H. L. (2011). The fate of Amazonian forest fragments: A 32-year investigation. Biological Conservation, 144, 56–67. https://doi.org/10.1016/j.biocon.2010.09.021

42. Laurance, W. F., Cochrane, M. A., Bergen, S., Fearnside, P. M., Delamônica, P., Barber, C., ... Fernandes, T. (2001). The future of the Brazilian Amazon. Science, 291, 438–439. https://doi.org/10.1126/science.291.5503.438

43. Lees, A. C., & Peres, C. A. (2008). Avian life-history determinants of local extinction risk in a hyper-fragmented neotropical forest landscape. Animal Conservation, 11, 128–137. https://doi.org/10.1111/j.1469-1795.2008.00162.x

44. Lees, A. C., Peres, C. A., Fearnside, P. M., Schneider, M., & Zuanon, J. A. S. (2016). Hydropower and the future of Amazonian biodiversity. Biodiversity and Conservation, 25, 451–466. https://doi.org/10.1007/s10531-016-1072-3

45. Lindenmayer, D. B., & Fischer, J. (2007). Tackling the habitat fragmentation panchreston. Trends in Ecology & Evolution, 22, 127–132. https://doi.org/10.1016/j.tree.2006.11.006

46. Lövei, G. L., Magura, T., Tóthmérész, B., & Ködöböcz, V. (2006). The influence of matrix and edges on species richness patterns of ground beetles (Coleoptera: Carabidae) in habitat islands. Global Ecology and Biogeography, 15, 283–289. https://doi.org/10.1111/j.1466-8238.2005.00221.x

47. Mace, G. M., Collar, N. J., Gaston, K. J., Hilton-Taylor, C., Akçakaya, H. R., Leader-Williams, N., ... Stuart, S. N. (2008). Quantification of extinction risk: IUCN's system for classifying threatened species. Conservation Biology, 22, 1424–1442. https://doi.org/10.1111/j.1523-1739.2008.01044.x

48. MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Andrew, Royle. J., & Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. Ecology, 83, 2248–2255. https://doi.org/10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2

49. Mackenzie, D. I., & Royle, J. A. (2005). Designing occupancy studies: General advice and allocating survey effort. Journal of Applied Ecology, 42, 1105–1114. https://doi.org/10.1111/j.1365-2664.2005.01098.x

50. 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. Diversity and Distributions, 20, 1136–1146. https://doi.org/10.1111/ddi.12227

51. Matthews, T. J., Cottee-Jones, H. E. W., & Whittaker, R. J. (2015). Quantifying and interpreting nestedness in habitat islands: A synthetic analysis of multiple datasets. Diversity and Distributions, 21, 392–404. https://doi.org/10.1111/ddi.12298

52. Matthews, T. J., Guilhaumon, F., Triantis, K. A., Borregaard, M. K., & Whittaker, R. J. (2016). On the form of species-area relationships in habitat islands and true islands. Global Ecology and Biogeography, 25, 847–858. https://doi.org/10.1111/geb.12269

53. Matthews, T. J., Triantis, K. A., Rigal, F., Borregaard, M. K., Guilhaumon, F., & Whittaker, R. J. (2016). Island species-area relationships and species accumulation curves are not equivalent: An analysis of habitat island datasets. Global Ecology and Biogeography, 25, 607–618. https://doi.org/10.1111/geb.12439

54. McGarigal, K., Cushman, S. A., & Ene, E. (2012). FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst, USA. Retrieved from http://www.umass.edu/landeco/research/fragstats/fragstats.html

55. Mendenhall, C. D., Karp, D. S., Meyer, C. F. J., Hadly, E. A., & Daily, G. C. (2014). Predicting biodiversity change and averting collapse in agricultural landscapes. Nature, 509, 213–217. https://doi.org/10.1038/nature13139

56. Meyer, C. F. J., Fründ, J., Lizano, W. P., & Kalko, E. K. V. (2008). Ecological correlates of vulnerability to fragmentation in Neotropical bats. Journal of Applied Ecology, 45, 381–391.

57. Moore, R. P., Robinson, W. D., Lovette, I. J., & Robinson, T. R. (2008). Experimental evidence for extreme dispersal limitation in tropical forest birds. Ecology Letters, 11, 960–968. https://doi.org/10.1111/j.1461-0248.2008.01196.x

58. Moura, N. G., Lees, A. C., Aleixo, A., Barlow, J., Berenguer, E., Ferreira, J., … Gardner, T. A. (2016). Idiosyncratic responses of Amazonian birds to primary forest disturbance. Oecologia, 180, 903–916. https://doi.org/10.1007/s00442-015-3495-z

59. Newbold, T., Hudson, L. N., Phillips, H. R. P., Hill, S. L. L., Contu, S., Lysenko, I., … Purvis, A. (2014). A global model of the response of tropical and sub-tropical forest biodiversity to anthropogenic pressures. Proceedings of the Royal Society B: Biological Sciences, 281, 20141371. https://doi.org/10.1098/rspb.2014.1371

60. Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Wagner, H. (2017). vegan: Community Ecology Package. R package version 2.4-3. Retrieved from https://CRAN.R-project.org/package=vegan

61. Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N., & Pears, W. (2013). caper: Comparative Analyses of Phylogenetics and Evolution in R. R package version 0.5.2. Retrieved from https://CRAN.R-project.org/package=caper

62. Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics, 20, 289–290. https://doi.org/10.1093/bioinformatics/btg412

63. Peres, C. A., Gardner, T. A., Barlow, J., Zuanon, J., Michalski, F., Lees, A. C., … Feeley, K. J. (2010). Biodiversity conservation in human-modified Amazonian forest landscapes. Biological Conservation, 143, 2314–2327. https://doi.org/10.1016/j.biocon.2010.01.021

64. Powell, L. L., Cordeiro, N. J., & Stratford, J. A. (2015). Ecology and conservation of avian insectivores of the rainforest understory: A pantropical perspective. Biological Conservation, 188, 1–10. https://doi.org/10.1016/j.biocon.2015.03.025

65. Powell, L. L., Stouffer, P. C., & Johnson, E. I. (2013). Recovery of understory bird movement across the interface of primary and secondary Amazon rainforest. The Auk, 130, 459–468. https://doi.org/10.1525/auk.2013.12202

66. Quinn, J. F., & Harrison, S. P. (1988). Effects of habitat fragmentation and isolation on species richness: Evidence from biogeographic patterns. Oecologia, 75, 132–140. https://doi.org/10.1007/BF00378826

67. R Core Team (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

68. Rabinowitz, D. (1981). Seven forms of rarity. In H. Synge (Ed.), The biological aspects of rare plant conservation (pp. 205–217). Chichester, UK: John Wiley & Sons.

69. Rosenzweig, M. L. (1995). Species diversity in space and time. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511623387

70. Santos, A. M. C., Whittaker, R. J., Triantis, K. A., Borges, P. A. V., Jones, O. R., Quicke, D. L. J., & Hortal, J. (2010). Are species-area relationships from entire archipelagos congruent with those of their constituent islands? Global Ecology and Biogeography, 19, 527–540.

71. Scheiner, S. M. (2003). Six types of species-area curves. Global Ecology and Biogeography, 12, 441–447. https://doi.org/10.1046/j.1466-822X.2003.00061.x

72. Sekercioglu, C. H., Ehrlich, P. R., Daily, G. C., Aygen, D., Goehring, D., & Sandi, R. F. (2002). Disappearance of insectivorous birds from tropical forest fragments. Proceedings of the National Academy of Sciences of the United States of America, 99, 263–267. https://doi.org/10.1073/pnas.012616199

73. Si, X., Baselga, A., & Ding, P. (2015). Revealing beta-diversity patterns of breeding bird and lizard communities on inundated land-bridge islands by separating the turnover and nestedness components. PLoS ONE, 10, e0127692. https://doi.org/10.1371/journal.pone.0127692

74. Si, X., Baselga, A., Leprieur, F., Song, X., & Ding, P. (2016). Selective extinction drives taxonomic and functional alpha and beta diversities in island bird assemblages. Journal of Animal Ecology, 85, 409–418. https://doi.org/10.1111/1365-2656.12478

75. Si, X., Pimm, S. L., Russell, G. J., & Ding, P. (2014). Turnover of breeding bird communities on islands in an inundated lake. Journal of Biogeography, 41, 2283–2292. https://doi.org/10.1111/jbi.12379

76. Sodhi, N. S., Liow, L. H., & Bazzaz, F. A. (2004). Avian extinctions from tropical and subtropical forests. Annual Review of Ecology, Evolution, and Systematics, 35, 323–345. https://doi.org/10.1146/annurev.ecolsys.35.112202.130209

77. Soga, M., & Koike, S. (2012). Relative importance of quantity, quality and isolation of patches for butterfly diversity in fragmented urban forests. Ecological Research, 27, 265–271. https://doi.org/10.1007/s11284-011-0896-2

78. Stotz, D. F., Fitzpatrick, J. W., Parker, T. A. III, & Moskovits, D. K. (1996). Neotropical birds: Ecology and conservation. Chicago, IL: The University of Chicago Press.

79. Stouffer, P. C., & Bierregaard, R. O. (1995). Use of Amazonian forest fragments by understory insectivorous birds. Ecology, 76, 2429–2445. https://doi.org/10.2307/2265818

80. Stouffer, P. C., Johnson, E. I., Bierregaard, R. O., & 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. https://doi.org/10.1371/journal.pone.0020543

81. Stouffer, P. C., Strong, C., & Naka, L. N. (2009). Twenty years of understorey bird extinctions from Amazonian rain forest fragments: Consistent trends and landscape-mediated dynamics. Diversity and Distributions, 15, 88–97. https://doi.org/10.1111/j.1472-4642.2008.00497.x

82. Stratford, J. A., & Stouffer, P. C. (1999). Local extinctions of terrestrial Insectivorous birds in a fragmented landscape near Manaus, Brazil. Conservation Biology, 13, 1416–1423. https://doi.org/10.1046/j.1523-1739.1999.98494.x

83. Stratford, J. A., & Stouffer, P. C. (2013). Microhabitat associations of terrestrial insectivorous birds in Amazonian rainforest and second-growth forests. Journal of Field Ornithology, 84, 1–12. https://doi.org/10.1111/jofo.12000

84. Thiollay, J.-M. (1994). Structure, density and rarity in an Amazonian rainforest bird community. Journal of Tropical Ecology, 10, 449. https://doi.org/10.1017/S0266467400008154

85. Thornton, D., Branch, L., & Sunquist, M. (2011). Passive sampling effects and landscape location alter associations between species traits

and response to fragmentation. Ecological Applications, 21, 817–829. https://doi.org/10.1890/10-0549.1

86. Triantis, K. A., Guilhaumon, F., & Whittaker, R. J. (2012). The island species-area relationship: Biology and statistics. Journal of Biogeography, 39, 215–231. https://doi.org/10.1111/j.1365-2699.2011.02652.x

87. Ulrich, W., & Gotelli, N. J. (2012). A null model algorithm for presence-absence matrices based on proportional resampling. Ecological Modelling, 244, 20–27. https://doi.org/10.1016/j.ecolmodel.2012.06.030

88. Van Houtan, K. S., Pimm, S. L., Bierregaard, R. O., Lovejoy, T. E., & Stouffer, P. C. (2006). Local extinctions in flocking birds in Amazonian forest fragments. Evolutionary Ecology Research, 8, 129–148.

89. Wang, Y., Thornton, D. H., Ge, D., Wang, S., & Ding, P. (2015). Ecological correlates of vulnerability to fragmentation in forest birds on inundated subtropical land-bridge islands. Biological Conservation, 191, 251–257. https://doi.org/10.1016/j.biocon.2015.06.041

90. Wang, Y., Zhang, J., Feeley, K. J., Jiang, P., & Ding, P. (2009). Life-history traits associated with fragmentation vulnerability of lizards in the Thousand Island Lake, China. Animal Conservation, 12, 329–337. https://doi.org/10.1111/j.1469-1795.2009.00257.x

91. Watling, J. I., & Donnelly, M. A. (2006). Fragments as islands: A synthesis of faunal responses to habitat patchiness. Conservation Biology, 20, 1016–1025. https://doi.org/10.1111/j.1523-1739.2006.00482.x

92. Webb, C. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies and community ecology. Annual Review of Ecology and Systematics, 33, 475–505. https://doi.org/10.1146/annurev.ecolsys.33.010802.150448

93. Welsh, A. H., Lindenmayer, D. B., & Donnelly, C. F. (2013). Fitting and interpreting occupancy models. PLoS ONE, 8, e52015. https://doi.org/10.1371/journal.pone.0052015

94. Willie, J., Petre, C.-A., Tagg, N., & Lens, L. (2012). Evaluation of species richness estimators based on quantitative performance measures and sensitivity to patchiness and sample grain size. Acta Oecologica, 45, 31–41. https://doi.org/10.1016/j.actao.2012.08.004

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

96. Wolfe, J. D., Stouffer, P. C., Mokross, K., Powell, L. L., & Anciães, M. M. (2015). Island vs. countryside biogeography: An examination of how Amazonian birds respond to forest clearing and fragmentation. Ecosphere, 6, art295. https://doi.org/10.1890/ES15-00322.1

97. Yaacobi, G., Ziv, Y., & Rosenzweig, M. L. (2007). Habitat fragmentation may not matter to species diversity. Proceedings of the Royal Society B: Biological Sciences, 274, 2409–2412. https://doi.org/10.1098/rspb.2007.0674

98. Yong, D. L., Qie, L., Sodhi, N. S., Koh, L. P., Peh, K. S.-H., Lee, T. M., . . . Lim, S. L.-H. (2011). Do insectivorous bird communities decline on land-bridge forest islands in Peninsular Malaysia? Journal of Tropical Ecology, 27, 1–14. https://doi.org/10.1017/S0266467410000520

99. Yu, M., Hu, G., Feeley, K. J., Wu, J., & Ding, P. (2012). Richness and composition of plants and birds on land-bridge islands: Effects of island attributes and differential responses of species groups. Journal of Biogeography, 39, 1124–1133. https://doi.org/10.1111/j.1365-2699.2011.02676.x

2.9   Supporting Information

This supporting information contains:

- Appendix S2.1

- Figure S2.1

- Figure S2.2

- Figure S2.3

- Figure S2.4

- Figure S2.5

- Figure S2.6

- Figure S2.7

- Figure S2.8

- Table S2.1

- Table S2.2

- Table S2.3

- R code

    Click here to download the R code in .Rmd file format

Appendix S2.1

Details on estimates of local extinctions

The use of estimates of island occupancy (i.e., proportion of islands occupied –
PIO) as a measure of species vulnerability to forest fragmentation is only mean-
ingful if local extinctions have occurred at the study islands (Bolger et al., 1991;
Keinath et al., 2017). Ideally, bird surveys would be carried out at the time of is-
land creation, which could be compared with present-day species distributions to
determine the occurrence of local extinctions (Bolger et al., 1991). In the absence
of historical data, which is typically the case in ecological studies, the comparison
between species-area relationships (SARs) for birds in pseudo-control islands and
in treatment islands can be alternatively used to infer the occurrence of local ex-
tinctions across study islands (Brown, 1971; Bolger et al., 1991; Wang et al., 2009).
As such, bird surveys were carried out in pseudo-control islands within plots of
similar size to those in treatment islands to represent species distributions in an
unfragmented habitat (Bolger et al., 1991; Wang et al., 2009).

   To produce the species-area curves for both pseudo-control islands and
treatment islands, we used the number of bird species recorded as a function of
surveyed area rather than total island area to make the spatial extent of radial
surveys around point count stations comparable in both pseudo-control islands
and treatment islands (Wang et al., 2009). Only species that had been recorded
in pseudo-control islands (n = 164) were considered for this comparison. For
example, one of the seven species recorded at our smallest site (3.39 ha), Island
Caua (Table S2.1), was not recorded in pseudo-control islands, so the number of
species in that island was restricted to six. Surveyed areas (expressed in hectares)
were calculated as the survey area of a point count station ($\pi \times 50^2$; hereafter,
PCs) times the number of PCs sampled. For instance, the surveyed area within
Island Caua was 1.57 ha as we deployed two PCs in that island, each of which
covering an area of 0.785 ha.

   For treatment islands, each island was considered as a data point (n =
34; blue circles in Figure S2.2a) with the total number of species recorded de-
fined as the dependent variable and total surveyed area as the independent vari-
able. In contrast, for pseudo-control islands, we used the mean number of species
recorded in each of the six survey transects placed therein (Figure S2.1) as the de-
pendent variable, which was calculated from the aggregated number of species
recorded across all possible combinations of adjacent PCs (Figure S2.1) for any
given area surveyed (i.e., independent variable) in treatment islands. As such,
the total survey area was held constant in both pseudo-control and treatment is-
lands. For example, to calculate the mean number of species along the transect

Pedral-B (which contained five PCs; Figure S2.1) for an area equivalent to two adjacent PCs (1.57 ha), we used the number of species recorded by pairing PCs 1-2, 2-3, 3-4 and 4-5 (Figure S2.1). In this case, the number of species recorded was 38, 40, 49 and 45, respectively, with a mean of 43 species. We followed the same procedure for all transects and combinations of 2, 3, 4, 6, 8, 10, 11 and 13 adjacent PCs, amounting to 33 data points (red circles in Figure S2.2a). Subsequently, we performed an ANCOVA model with island type as the categorical independent variable to determine whether the intercept (c-value) and the slope (z-value) of the predicted lines (red and blue lines in Figure S2.2a) were statistically different.

Predicted lines derived from pseudo-control islands ($r^2_{adj} = 0.950$, $P = 0.001$) and from treatment islands ($r^2_{adj} = 0.614$, $P = 0.001$) were different in both the c-values ($P < 0.001$) and z-values ($P = 0.008$). For treatment islands, the c-value was 1.125 and the z-value was 0.745. For pseudo-control islands, the c-value was 1.533 and the z-value was 0.450. As such, treatment islands supported depauperate species assemblages and had experienced a much higher rate of species loss as a function of surveyed area in relation to equivalent-sized survey areas within pseudo-control islands, which indicates that local extinction had indeed occurred in treatment islands.

To estimate the number of extinctions that had occurred in treatment islands, we rounded down to the nearest integer the predicted number of species in pseudo-control islands (red line in Figure S2.2a) which was subtracted from the number of species recorded in treatment islands (blue circles in Figure S2.2a; Bolger et al., 1991). For example, the estimated number of local extinctions at the small Island Caua was 35, since the predicted number of species in pseudo-control islands for an equivalent survey area of 1.57 ha was 41.85 (41), whereas the recorded number of species in that island was only 6 (Figure S2.2b). Overall, we estimated that a total of 788 local extinctions had occurred across all 34 treatment islands over 22-23 years of post-isolation history at the Tucuruí Hydroelectric Reservoir landscape.
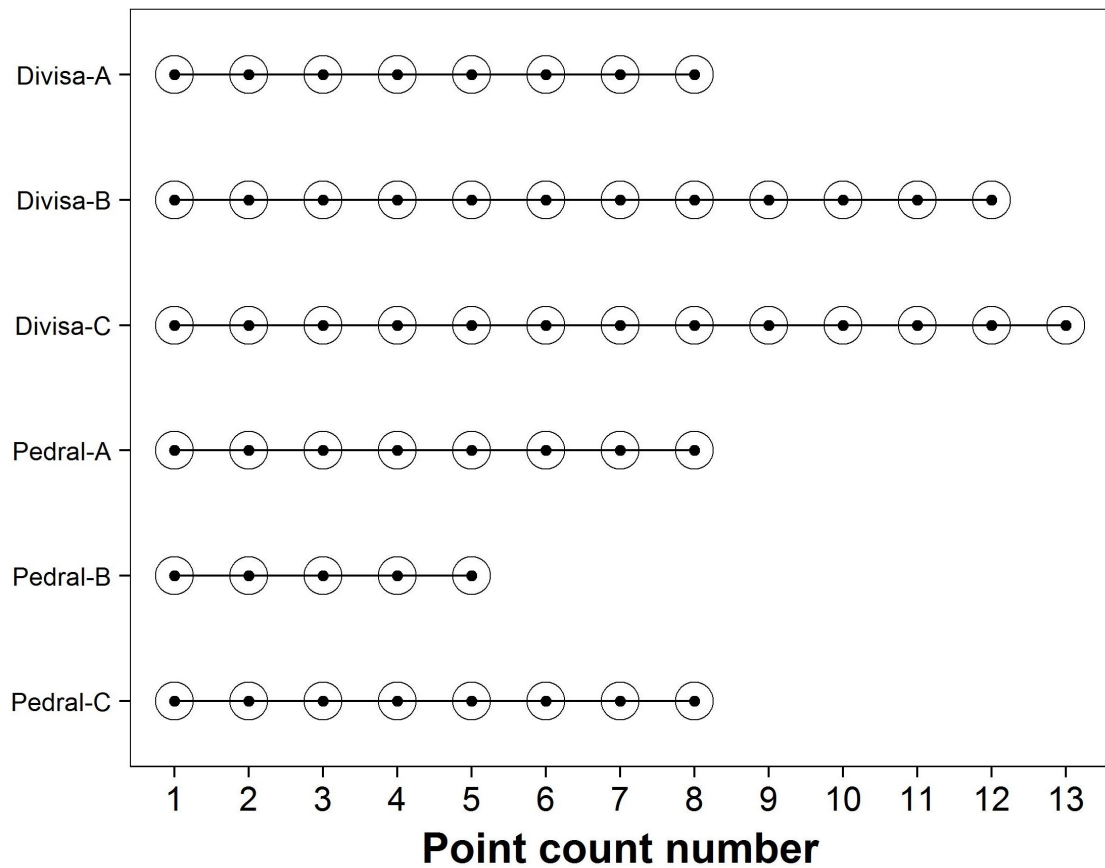
Figure S2.1: Schematic representation of all six transects and 54 point count stations (PCs) deployed within the two pseudo-control islands (Divisa and Pedral) at the Tucuruí Hydroelectric Reservoir landscape. PCs (black dots) and their 50-m fixed-radius survey areas (circles) were distributed at regular 200-m intervals along transects (black lines).
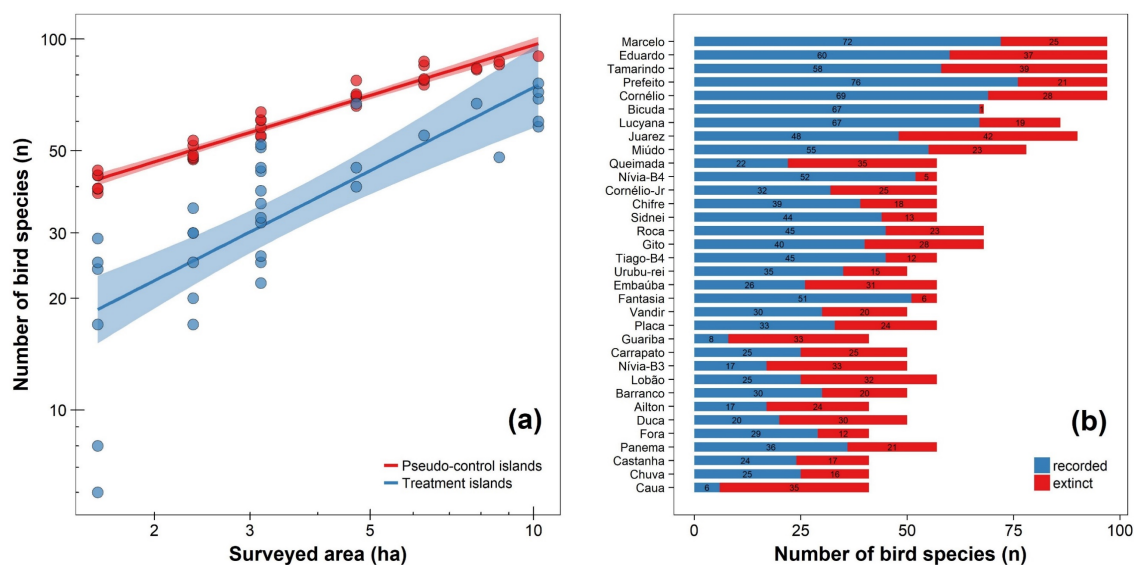
Figure S2.2: (a) Species-area relationships for birds in pseudo-control islands (red lines and circles) and in treatment islands (blue lines and circles) as a function of surveyed area around point count stations. Coloured regions show the 95% confidence intervals of predicted lines. Note the base 10 logarithmic scales along both axes. (b) Numbers of bird species within pseudo-control islands ($n = 164$) that were either recorded (blue horizontal bars) or estimated to have been locally extinct (red horizontal bars) following isolation across 34 treatment islands at the Tucuruí Hydroelectric Reservoir landscape. Islands are ordered top to bottom from the largest to the smallest (Table S2.1).
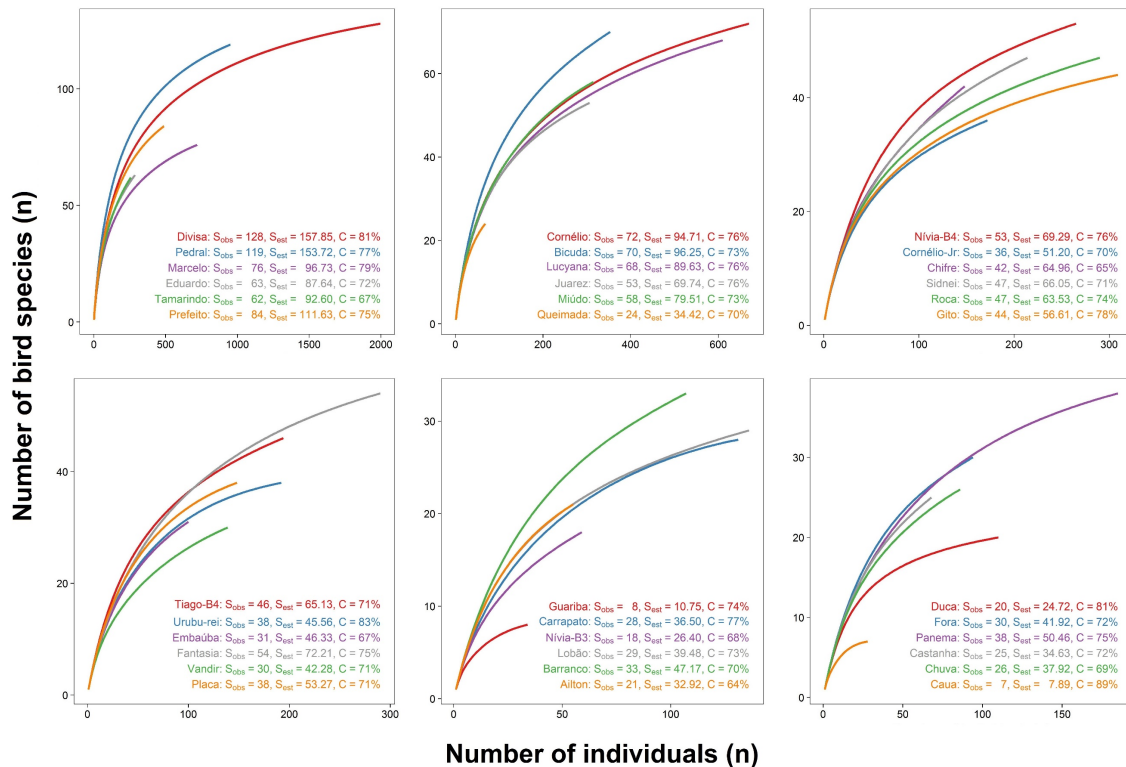
Figure S2.3: Individual-based rarefaction curves of the number of bird species recorded per surveyed island at the Tucuruí Hydroelectric Reservoir landscape. Each line represents one island coloured according to its sampling completeness, which was quantified as a percentage between the recorded and the estimated number of species based on the first-order Jackknife estimator. Islands are ordered by decreasing size as in Table S2.1. Note the different scales on both the x and y axes.
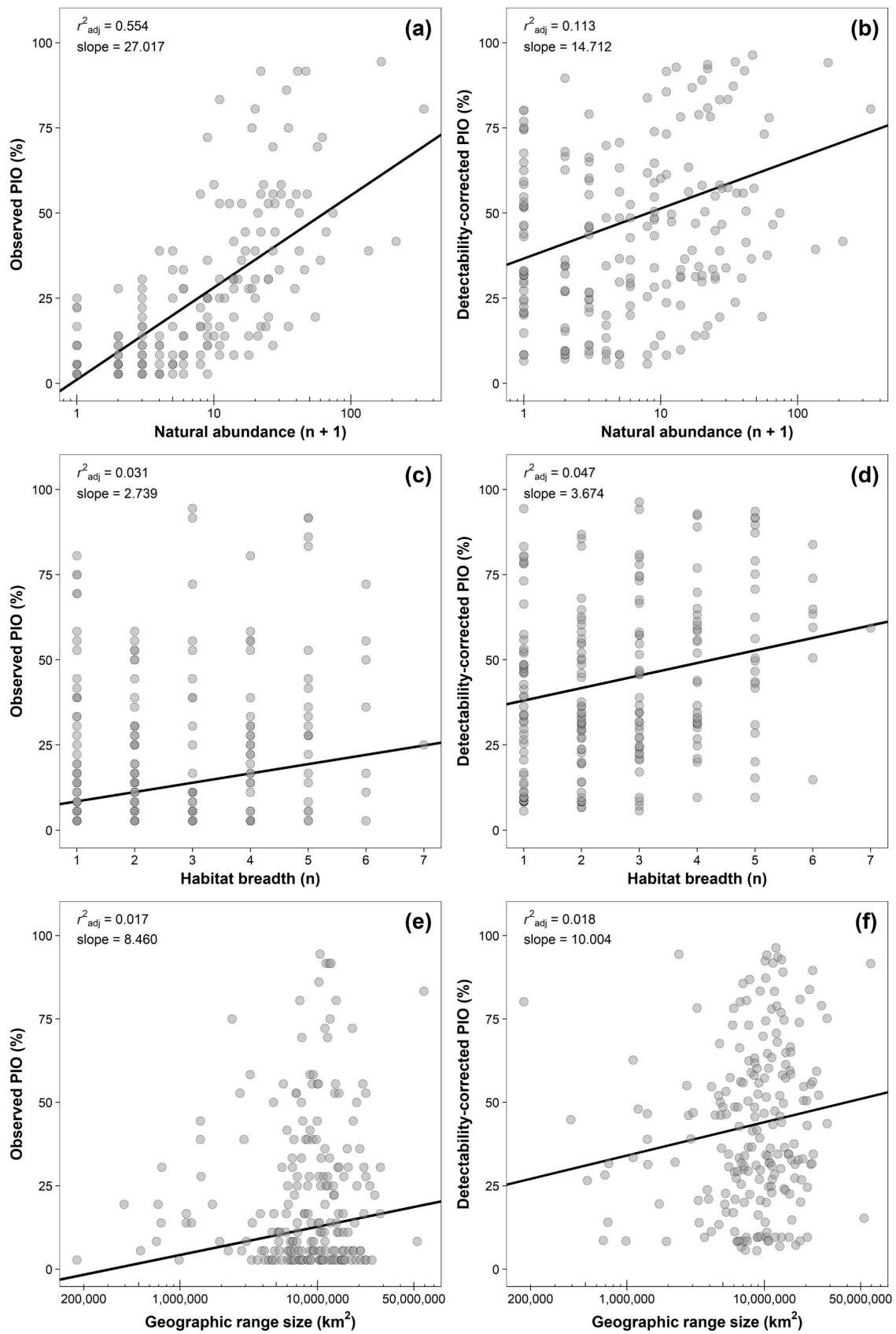
Figure S2.4: *Caption on the next page*

Figure S2.4: Relationships between the proportion of islands occupied (PIO) and species traits included in Phylogenetic Generalised Least Squares (PGLS) models with high support (ΔAIC $\leqslant$ 2), namely natural abundance (a and b), habitat breadth (c and d) and geographic range size (e and f). Grey circles represent the 207 bird species surveyed across 36 islands at the Tucuruí Hydroelectric Reservoir landscape. Observed PIO was quantified as a percentage between the number of islands where a species was recorded divided by the total number of surveyed islands, whereas detectability-corrected PIO was quantified from single-season occupancy models (MacKenzie et al., 2002). See Table S2.3 for a description of species traits. Note the base 10 logarithmic scales of the x-axes in (a), (b), (e) and (f).
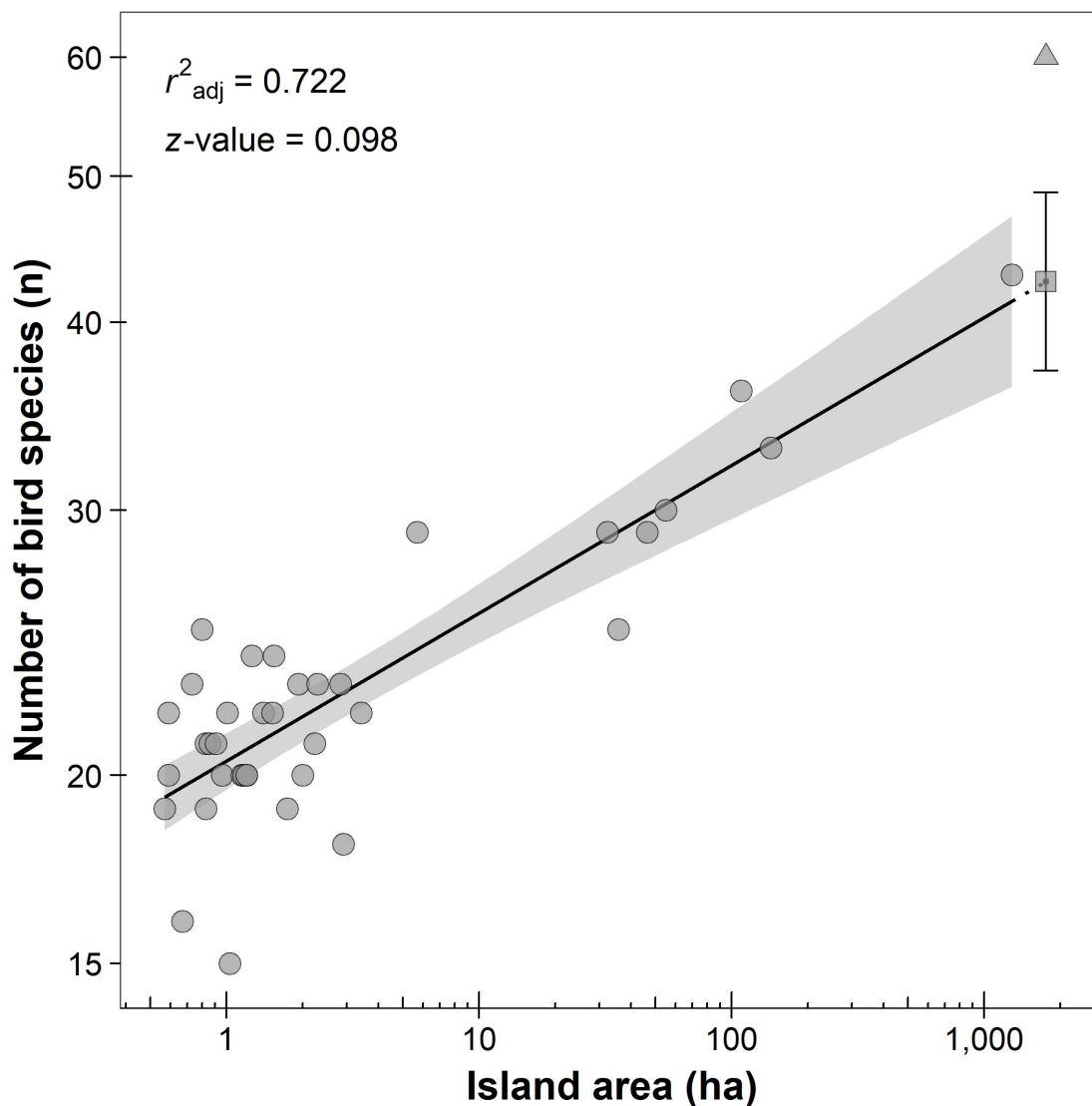
Figure S2.5: Species-area relationship for birds surveyed across 37 islands at the Thousand Island Lake in China (data reanalysed from Si et al., 2015) ($P < 0.001$). The dotted line indicates null predicted numbers of species if forest fragmentation had no effect. Circles, squares, and triangles correspond to the recorded, extrapolated, and overall number of species, respectively. Grey region and error bars show the 95% confidence intervals of the predicted line and the extrapolated value, respectively. See Yu et al. (2012) for an explanation for the low z-value for bird species at the Thousand Island Lake. Note the base 10 logarithmic scales along both axes.
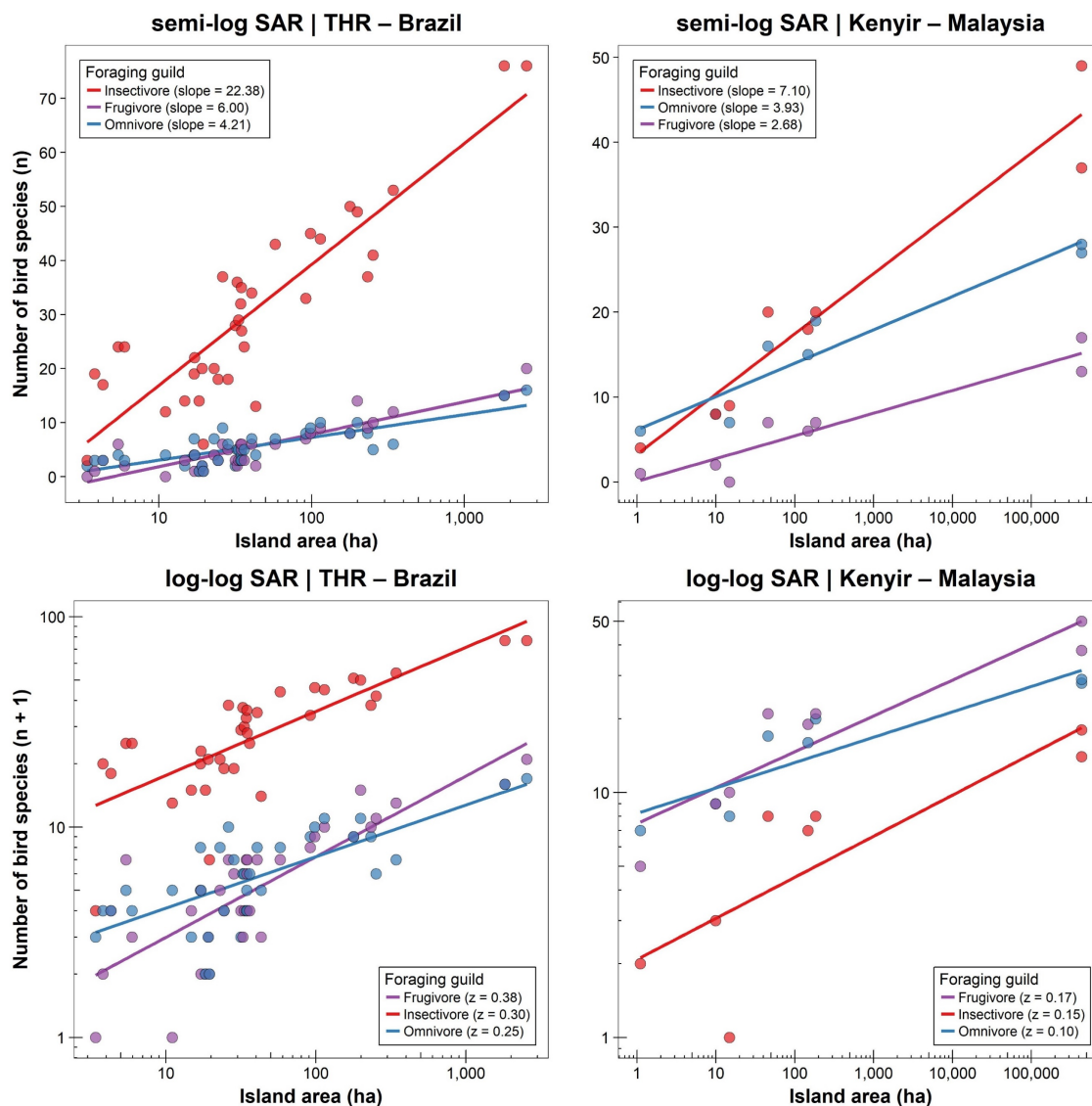
Figure S2.6: Comparison of the species-area relationships (SARs) for three avian foraging guilds at the Tucuruí Hydroelectric Reservoir landscape (see Figure 2.S7) and Lake Kenyir in Malaysia (data from Yong et al., 2011) ($P \leqslant 0.01$ in all 12 instances). Plots at the top show the semi-log form of the SARs and their respective slope values, and those at the bottom show the *log-log* form of the SARs and their respective *z*-values (*z*). Slope and *z*-values in each plot are sorted by decreasing order, indicating the most (higher values) to the least (lower values) impaired foraging guilds in terms of species losses as a function of island area reduction.
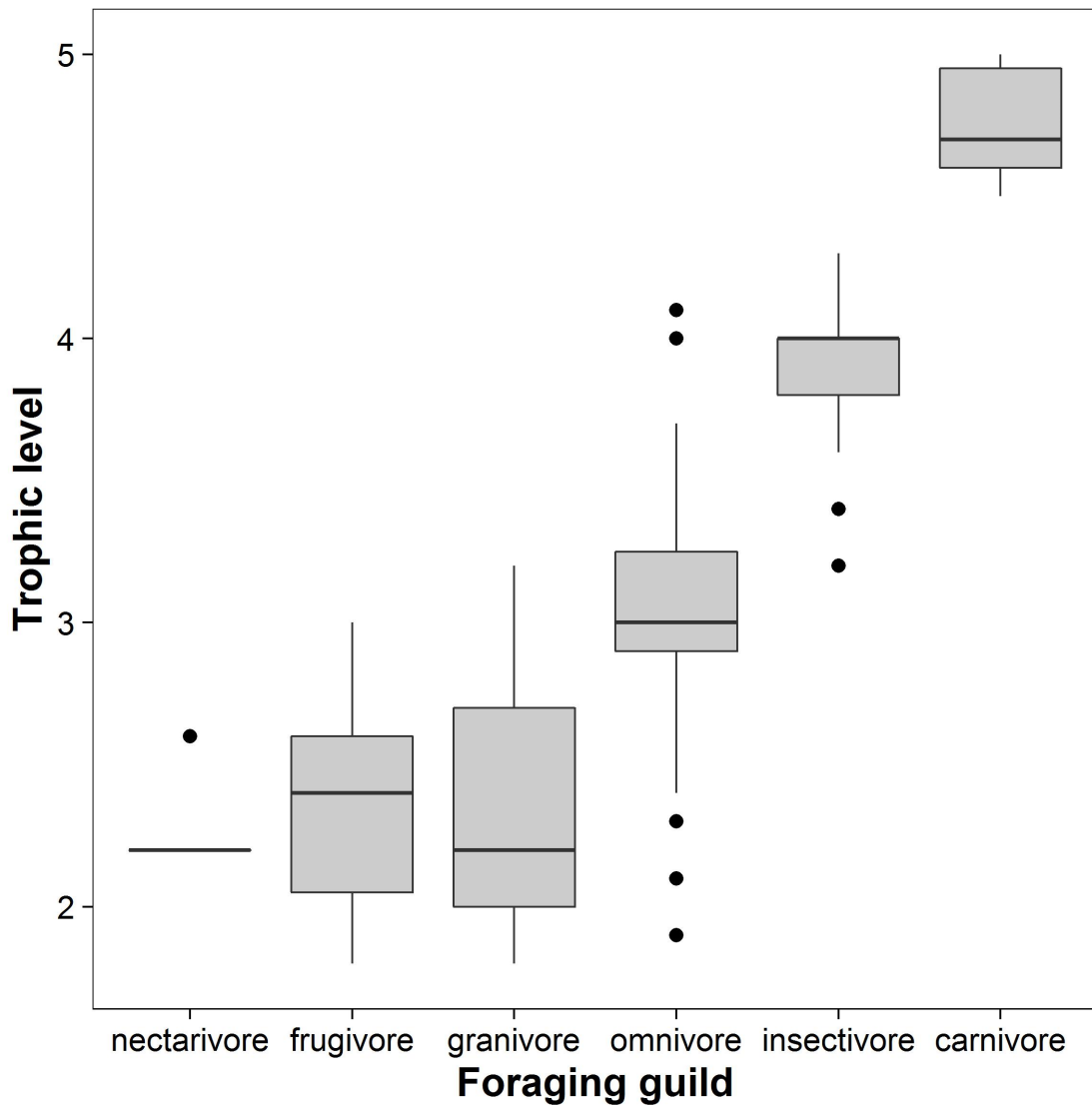
Figure S2.7: Distribution of trophic level scores in six foraging guilds for 207 bird species surveyed across 36 islands at the Tucuruí Hydroelectric Reservoir landscape. Species guilds were assigned according to Wilman et al. (2014), expect for nectarivore-frugivores, which were split into nectarivores (Trochilidae and Coerebidae) and frugivores. Note that the trophic level of some species belonging to a foraging guild may overlap those of another foraging guild.
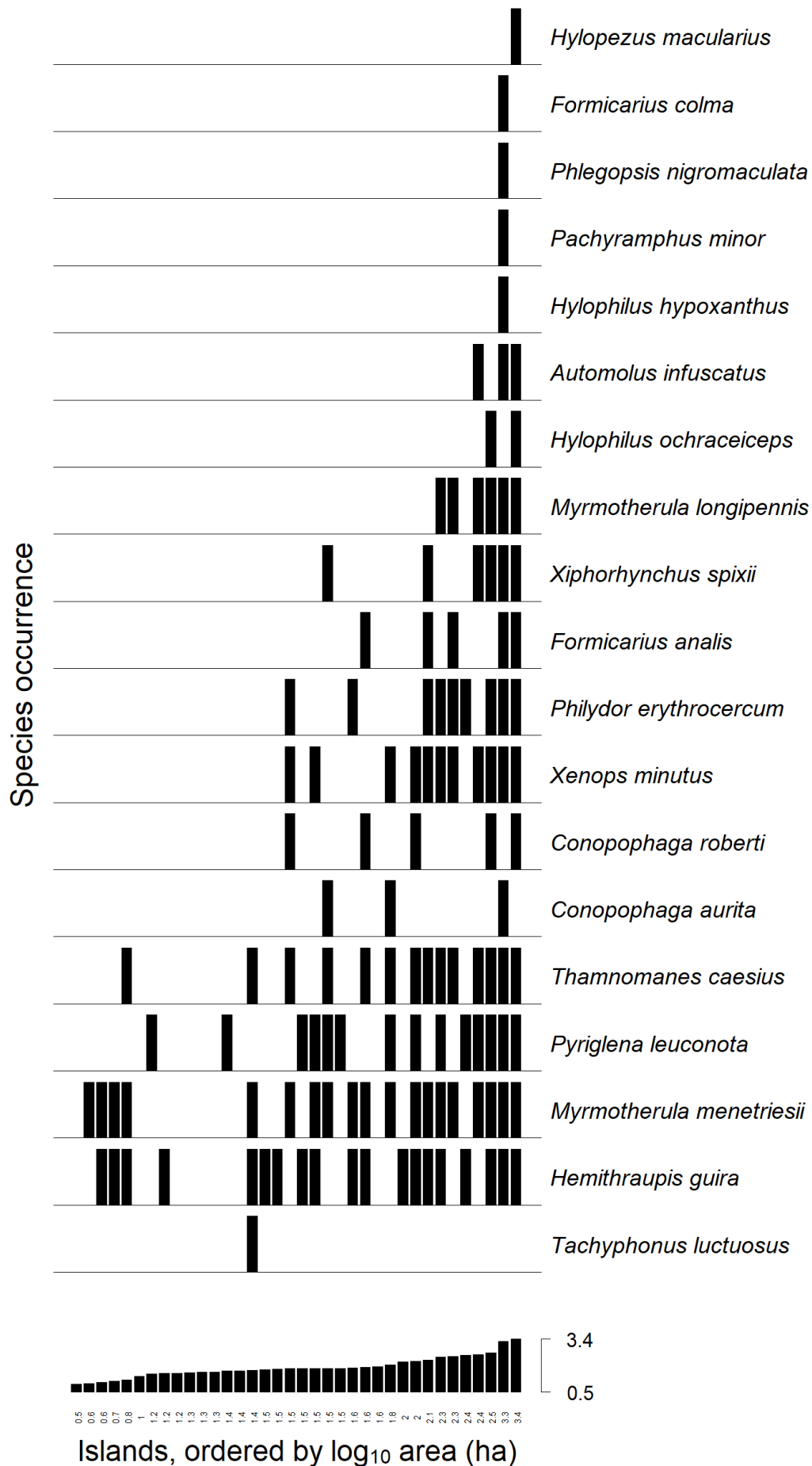
Figure S2.8: *Caption on the next page*

Figure S2.8: Species occurrence of ground insectivores, obligate ant-followers, and obligate mixed-species flock attendants along the island area gradient. Bars indicate species occurrence on islands, ordered left to right from the smallest to the largest. Bars at the bottom indicate the area of each island, which ranged from 3.4 to 2551.5 ha. Ground insectivores: *C. aurita, C. roberti, F. analis, F. colma, H. macularius*; obligate ant-followers: *P. nigromaculata, P. leuconota*; obligate mixed-species flock attendants: *A. infuscatus, H. guira, H. hypoxanthus, H. ochraceiceps, M. longipennis, M. menetriesii, P. minor, P. erythrocercum, T. luctuosus, T. caesius, X. minutus, X. spixii*.

Table S2.1: Description of the 36 islands surveyed at the Tucuruí Hydroelectric Reservoir landscape, and number of bird species occurring therein. 'Effort' indicates the number of point count stations (PCs) and 'Samples' indicates the number of PCs times the number of survey visits per PCs.

Island name	Latitude (S)	Longitude (W)	Area (ha)	Shape index	Proximity index	Effort (PCs)	Samples	S_{fs}[a]	S_{hg}[b]	S_{all}[c]
Divisa	4°12'22"	49°30'04"	2551.45	8.42	40111.28	33	195	49	79	128
Pedral	4°22'44"	49°35'31"	1823.35	5.63	40111.28	21	124	55	64	119
Marcelo	4°09'17"	49°32'45"	342.43	3.97	99.15	13	78	31	45	76
Eduardo	4°18'32"	49°39'17"	251.93	3.59	699.83	13	70	28	35	63
Tamarindo	4°16'09"	49°39'49"	232.74	2.39	1.07	13	78	23	39	62
Prefeito	4°19'51"	49°37'35"	198.66	2.29	813.48	13	76	35	49	84
Cornélio	4°17'15"	49°28'05"	178.27	3.52	64.99	13	78	27	45	72
Bicuda	4°09'30"	49°32'01"	113.99	2.43	4011.13	6	36	27	43	70
Lucyana	4°10'09"	49°33'47"	98.23	2.42	72.43	10	59	23	45	68
Juarez	4°16'29"	49°30'56"	91.55	2.43	9.83	11	65	15	38	53
Miúdo	4°25'15"	49°33'16"	57.75	1.91	1131.61	8	45	23	35	58
Queimada	4°16'23"	49°38'55"	43.09	2.02	450.06	4	19	6	18	24
Nívia-B4	4°14'58"	49°29'41"	40.52	1.91	584.42	4	24	20	33	53
Cornélio-Jr	4°17'42"	49°29'05"	36.18	2.01	86.66	4	20	7	29	36
Chifre	4°17'16"	49°38'04"	34.79	1.80	5.68	4	23	13	29	42
Sidnei	4°20'47"	49°37'32"	34.62	1.54	42.16	4	21	22	25	47
Roca	4°25'41"	49°33'38"	34.35	1.58	18.53	6	36	17	30	47
Gito	4°17'52"	49°36'47"	33.19	2.26	0.42	6	33	9	35	44
Tiago-B4	4°13'52"	49°30'19"	32.53	1.32	1.74	4	23	14	32	46
Urubu-rei	4°16'05"	49°29'25"	31.61	2.70	0.80	3	18	8	30	38
Embaúba	4°18'08"	49°38'12"	28.32	1.97	25.75	4	24	10	21	31
Fantasia	4°15'24"	49°30'07"	26.06	1.51	913.16	4	24	17	37	54
Vandir	4°22'31"	49°33'47"	24.39	1.36	85.16	3	18	8	22	30
Placa	4°19'12"	49°37'51"	22.95	1.56	128.96	4	22	15	23	38
Guariba	4°10'39"	49°32'40"	19.50	1.37	186.97	2	12	2	6	8
Carrapato	4°23'56"	49°34'22"	19.17	1.65	218.36	3	18	13	15	28
Nívia-B3	4°17'38"	49°37'25"	18.30	1.23	9.98	3	15	6	12	18
Lobão	4°18'22"	49°37'38"	17.09	1.89	32.06	4	21	10	19	29
Barranco	4°09'34"	49°34'07"	16.99	1.29	66.34	3	18	8	25	33
Ailton	4°21'06"	49°35'13"	14.74	1.20	2136.92	2	12	7	14	21
Duca	4°17'11"	49°30'25"	11.02	1.44	8.53	3	18	2	18	20
Fora	4°08'26"	49°33'57"	5.95	1.15	20.74	2	12	10	20	30
Panema	4°10'33"	49°33'05"	5.40	1.11	19.78	4	24	13	25	38
Castanha	4°09'03"	49°33'24"	4.29	1.09	52.51	2	8	6	19	25
Chuva	4°10'59"	49°32'29"	3.79	1.09	2200.61	2	12	6	20	26
Caua	4°23'18"	49°34'08"	3.39	1.06	0.01	2	9	0	7	7

[a] S_{fs}: number of forest specialist species

[b] S_{hg}: number of habitat generalist species

[c] S_{all}: overall number of species

Table S2.2: Description and sources of seven morpho-ecological traits for bird species considered in this study.

Species trait	Description	Source
Body mass	Species mean body mass (*g*)	Wilman et al., 2014
Trophic level	Sum of the proportional food consumption in each diet category weighted by an energetic score: (1) foliage and other plant material, (2) fruit and nectar, (3) seed, (4) invertebrate, (5) vertebrate, including carrion. For example, a species relying entirely on invertebrates is assigned a value 4, and a species relying on 50% fruits and 50% invertebrates is assigned a value 3	Wilman et al., 2014
Vertical stratum	Preferred foraging stratum classified into four categories: (1) ground, (2) understorey, (3) midstorey, (4) canopy	Stotz et al., 1996; Henriques et al., 2003; Wilman et al., 2014; personal observation
Flocking behaviour	Degree of gregariousness classified into six categories: (1) solitary or pairs, (2) monospecific flocks, (3) facultative ant-follower, (4) facultative mixed-species flock attendant, (5) obligate ant-follower, and (6) obligate mixed-species flock attendant	Willis & Oniki, 1978; Munn & Terborgh, 1979; Jullien & Thiollay, 1998; Thiollay & Jullien, 1998; Jullien & Clobert, 2000; Willson, 2004; Martínez et al., 2016; personal observation
Geographic range size	Breeding/resident extent of occurrence (*km²*)	BirdLife International, 2017
Habitat breadth	Number of habitats used	Stotz et al., 1996
Natural abundance	Total number of individuals recorded within pseudo-control islands	Field surveys

Table S2.3: Morpho-ecological traits and measures of vulnerability to forest fragmentation for 207 bird species surveyed across 36 islands at the Tucuruí Hydroelectric Reservoir landscape. Taxonomy follows Jetz et al. (2012).

| Species, by family | Habitat specialisation | Morpho-ecological traits | | | | | | | Measures of vulnerability | | | Species code |
		Body mass (g)	Trophic Level	Vertical stratum[a]	Flocking behaviour[b]	Geographic range size (km²)	Habitat breadth[c]	Natural abundance (n)	Observed PIO[d] (%)	Detectability-corrected PIO[e] (%)	Detectability (%)	
Tinamidae												
Crypturellus cinereus	habitat generalist	506.63	2	1	1	6630000	3	2	8.33	8.51	17.5	5s
Crypturellus soui	habitat generalist	216.16	2.4	1	1	15200000	3	0	5.56	22.57	22.4	7w
Crypturellus strigulosus	forest specialist	430.58	2.9	1	1	4650000	1	0	2.78	52.03	25.45	7x
Crypturellus variegatus	forest specialist	378	2	1	1	9010000	1	1	8.33	16.25	10.08	6m
Tinamus guttatus	forest specialist	686.18	3	1	1	5250000	1	0	5.56	53.25	3.77	8x
Tinamus tao	forest specialist	1600.1	2.3	1	1	8520000	2	7	16.67	61.76	5.64	4n
Cracidae												
Crax fasciolata	forest specialist	2600	2	1	1	4720000	3	0	2.78	67.62	2.74	7t
Mitu tuberosum	forest specialist	2769.46	2.2	1	1	4810000	1	7	16.67	46.12	7.16	4j
Penelope pileata	forest specialist	1249.79	1.8	4	2	1210000	1	8	13.89	48.01	6.66	4e
Pipile cujubi	forest specialist	1195.82	2.3	4	1	2810000	2	0	5.56	46.06	10.75	8o
Accipitridae												
Buteo magnirostris	habitat generalist	269	4.6	4	1	23900000	7	0	25	59.32	12.25	7l
Buteo nitidus	habitat generalist	519.04	4.8	4	1	15500000	4	2	27.78	65.05	8.57	5q
Buteogallus urubitinga	habitat generalist	1152.87	4.5	1	1	22100000	4	0	2.78	31.66	13.5	7m
Harpagus bidentatus	habitat generalist	215.25	4.2	4	2	17600000	2	0	2.78	32.78	15.04	8c
Columbidae												
Claravis pretiosa	habitat generalist	68.2	3.2	1	1	18900000	4	0	2.78	33.12	3.94	7p
Geotrygon montana	habitat generalist	133.86	2.9	1	1	22600000	3	1	2.78	34.49	9.6	6o
Leptotila rufaxilla	habitat generalist	157	3	1	1	14300000	4	10	30.56	61.35	10.96	3p
Leptotila verreauxi	habitat generalist	146.88	3	1	1	26000000	5	2	22.22	79.04	4.82	5z
Patagioenas speciosa	habitat generalist	258.47	2	4	1	16900000	3	1	5.56	7.22	20.19	6y
Patagioenas subvinacea	forest specialist	162.48	2	4	1	9170000	3	1	5.56	26.96	05.05	6z
Cuculidae												
Crotophaga ani	habitat generalist	110.09	4.1	2	2	24800000	2	0	2.78	52.13	3.13	7u
Crotophaga major	habitat generalist	148.25	3.5	2	2	15500000	4	0	5.56	58.44	2.14	7v
Piaya cayana	habitat generalist	101.98	4	4	1	20200000	5	8	36.11	50.51	21.17	4g
Trochilidae												
Anthracothorax nigricollis	habitat generalist	7	2.2	4	1	14900000	4	1	2.78	34.49	9.6	6i
Glaucis hirsutus	habitat generalist	6.76	2.2	2	1	13000000	3	2	25	45.84	11.01	5u
Heliothryx auritus	forest specialist	5.4	2.2	4	1	10700000	1	0	2.78	8.41	45.13	8d
Hylocharis sapphirina	habitat generalist	4.4	2.2	2	1	10800000	2	0	2.78	31.84	7.12	8g
Phaethornis ruber	habitat generalist	2.4	2.2	2	1	11300000	3	61	72.22	78	48.54	1g
Phaethornis superciliosus	forest specialist	6.3	2.2	2	1	3030000	3	17	8.33	46.95	9.12	2z
Thalurania furcata	habitat generalist	4.19	2.2	2	1	12500000	3	1	2.78	34.49	9.6	7f

Table S2.3: *Continued from previous page*

| | | Morpho-ecological traits | | | | | | | Measures of vulnerability | | | |
Species, by family	Habitat specialisation	Body mass (g)	Trophic Level	Vertical stratum[a]	Flocking behaviour[b]	Geographic range size (km²)	Habitat breadth[c]	Natural abundance (n)	Observed PIO[d] (%)	Detectability-corrected PIO[e] (%)	Detectability (%)	Species code
Trogonidae												
Trogon melanurus	habitat generalist	114	2.6	3	1	8520000	3	2	2.78	45.67	16.64	6h
Trogon rufus	forest specialist	53.8	3.2	3	1	14200000	2	10	16.67	23.69	15.03	3u
Trogon viridis	habitat generalist	89.69	2.9	3	1	11800000	1	26	69.44	83.29	31.36	2e
Momotidae												
Momotus momota	habitat generalist	114.96	3.6	3	1	11300000	6	15	36.11	63.45	15.55	3c
Galbulidae												
Galbula cyanicollis	forest specialist	23.23	4	2	1	3350000	1	17	13.89	13.95	49.14	2y
Galbula dea	habitat generalist	27.4	4	4	1	6560000	2	13	30.56	33.35	33.32	3g
Jacamerops aureus	forest specialist	62.9	4.1	3	1	8930000	1	1	2.78	9.56	42.3	6q
Bucconidae												
Bucco capensis	habitat generalist	54	4.7	3	1	6080000	1	2	5.56	11.24	15.19	5p
Bucco tamatia	habitat generalist	35.5	4.1	3	1	6240000	2	0	5.56	27.23	23.19	7k
Malacoptila rufa	habitat generalist	47.52	4	2	1	3660000	1	1	2.78	9.56	42.3	6s
Monasa morphoeus	habitat generalist	87.9	3.9	3	2	10300000	2	2	2.78	9.56	42.3	6d
Monasa nigrifrons	habitat generalist	80.7	4.3	3	2	8880000	4	3	2.78	9.56	42.3	5l
Notharchus hyperrhynchus	habitat generalist	95.89	3.8	4	1	11400000	2	1	5.56	26.96	05.05	6v
Notharchus tectus	habitat generalist	26.8	4	4	1	6720000	2	4	38.89	56.24	15.47	5d
Nystalus striolatus	forest specialist	47	4.2	4	1	4870000	2	1	5.56	36.26	3.41	6w
Ramphastidae												
Pteroglossus aracari	habitat generalist	250.16	2.4	4	2	6980000	2	10	52.78	85.62	14.42	3q
Pteroglossus bitorquatus	forest specialist	142	2.5	4	2	1110000	2	1	13.89	62.67	6.4	7d
Pteroglossus inscriptus	habitat generalist	125.82	2.6	4	2	3920000	3	0	5.56	21.04	33.16	8s
Ramphastos tucanus	habitat generalist	659.58	2.7	4	1	3240000	1	22	58.33	78.29	25.34	2k
Ramphastos vitellinus	forest specialist	360.36	3	4	1	5990000	1	13	41.67	78.2	14.98	3h
Picidae												
Campephilus melanoleucos	habitat generalist	256	3.6	3	1	13300000	4	12	52.78	92.79	14.6	3l
Campephilus rubricollis	forest specialist	210.71	4	3	1	7810000	2	16	44.44	86.83	14.68	3a
Celeus flavus	forest specialist	147.33	3.7	4	1	10300000	3	0	2.78	51.55	1.14	7n
Celeus undatus	forest specialist	64.5	3.8	4	1	2240000	1	1	5.56	32.1	3.55	6j
Colaptes melanochloros	habitat generalist	127.27	3.6	3	1	6390000	3	0	2.78	29.34	3.52	7r
Dryocopus lineatus	habitat generalist	183.19	3.7	4	1	21300000	6	7	55.56	83.83	17.89	4i
Melanerpes cruentatus	forest specialist	58.1	3	4	2	7810000	3	2	2.78	22.37	35.31	6a
Piculus chrysochloros	habitat generalist	88	4	4	1	13300000	3	1	2.78	9.56	42.3	7a
Piculus flavigula	habitat generalist	55	4	3	4	10300000	2	3	13.89	20.01	11.07	5n
Veniliornis affinis	forest specialist	34.83	3.4	4	4	9290000	2	30	58.33	83.4	24.36	1x
Falconidae												
Daptrius ater	habitat generalist	351.75	4.6	4	2	8090000	3	4	16.67	36.62	8.48	5a
Falco rufigularis	forest specialist	163.64	4.7	4	1	21700000	4	2	16.67	55.41	5.3	5t
Herpetotheres cachinnans	habitat generalist	623.58	5	4	1	20400000	5	0	5.56	43.06	35.03	8e
Ibycter americanus	habitat generalist	624	4	4	2	10700000	2	2	11.11	60.25	3.83	5x

Table S2.3: *Continued from previous page*

| Species, by family | Morpho-ecological traits | | | | | | | Measures of vulnerability | | | | |
	Habitat specialisation	Body mass (g)	Trophic Level	Vertical stratum[a]	Flocking behaviour[b]	Geographic range size (km²)	Habitat breadth[c]	Natural abundance (n)	Observed PIO[d] (%)	Detectability-corrected PIO[e] (%)	Detectability (%)	Species code
Micrastur mintoni	forest specialist	209.5	4.5	3	1	4110000	1	2	5.56	11.24	15.19	6b
Micrastur ruficollis	habitat generalist	177.64	5	3	1	19300000	2	8	25	54.73	8.58	4b
Micrastur semitorquatus	habitat generalist	621.68	5	4	1	20500000	3	2	5.56	24.56	8.35	6c
Psittacidae												
Amazona amazonica	habitat generalist	370	2	4	2	12200000	5	4	33.33	70.74	11.03	4x
Amazona farinosa	habitat generalist	625.99	2.1	4	2	11900000	1	11	25	47.45	13.77	3m
Amazona ochrocephala	habitat generalist	476.94	2	4	2	8010000	4	8	16.67	58.96	5.15	3x
Ara chloropterus	habitat generalist	1214	2	4	2	10500000	2	8	25	64.64	8.56	3y
Ara macao	forest specialist	1015	2	4	2	10200000	3	0	5.56	74.43	3.37	7i
Ara severus	forest specialist	343	2.3	4	2	8470000	4	4	25	63.26	13.63	4y
Aratinga leucophthalma	habitat generalist	158	2.2	4	2	13200000	4	0	22.22	76.94	5.2	7j
Brotogeris chrysoptera	habitat generalist	54.5	2.3	4	2	3840000	2	34	16.67	23.78	13.79	1t
Deroptyus accipitrinus	forest specialist	246	1.8	4	2	5470000	2	3	2.78	34.49	9.6	5h
Guaruba guarouba	forest specialist	194	1.8	4	2	516000	1	2	5.56	26.6	5.72	5v
Pionus menstruus	habitat generalist	251	2.1	4	2	10100000	4	21	55.56	92.37	15.15	2o
Pyrrhura picta	habitat generalist	62.1	1.9	4	2	392000	2	24	19.44	44.82	08.03	2g
Thamnophilidae												
Cercomacra cinerascens	forest specialist	14.3	4	3	1	7420000	1	341	80.56	80.56	86.11	1a
Cercomacra laeta	habitat generalist	15.96	4	2	1	1720000	2	54	19.44	19.57	62.7	1j
Cercomacra nigrescens	habitat generalist	16.5	4	2	1	6120000	4	2	19.44	20.98	35.83	6k
Cymbilaimus lineatus	habitat generalist	35.8	4	3	4	9350000	1	4	5.56	29.76	13.36	4z
Dysithamnus mentalis	habitat generalist	14.87	3.8	2	4	16700000	2	24	30.56	31.46	47.48	2f
Epinecrophylla ornata	forest specialist	9.4	4	3	4	1420000	2	17	27.78	31.37	36.98	2x
Formicivora grisea	habitat generalist	10.36	4	2	1	8830000	4	9	58.33	60.14	44.67	3w
Herpsilochmus rufimarginatus	habitat generalist	10.58	3.8	4	4	11800000	5	5	27.78	28.47	44.32	4s
Hylophylax naevius	forest specialist	14.2	4	2	1	6790000	2	0	5.56	6.6	28.4	8h
Hypocnemis striata	forest specialist	12.29	4	2	1	1410000	3	26	38.89	38.91	71.3	2c
Myrmoborus myotherinus	forest specialist	18.8	4	2	3	6590000	1	2	8.33	66.39	5.56	6e
Myrmotherula axillaris	forest specialist	08.09	4	2	4	10400000	3	166	94.44	94.15	81.37	1c
Myrmotherula brachyura	forest specialist	6.4	4	2	4	6790000	3	0	2.78	31.84	7.12	8m
Myrmotherula hauxwelli	forest specialist	10.7	4.1	2	4	5640000	1	30	55.56	57.42	44.79	1w
Myrmotherula longipennis	forest specialist	9.4	4	2	6	5930000	1	21	16.67	16.93	37.52	2n
Myrmotherula menetriesii	habitat generalist	8.6	4	3	6	6980000	1	35	52.78	55.88	48.59	1s
Phlegopsis nigromaculata	forest specialist	45.24	4.2	2	5	6080000	2	8	2.78	48.62	29.62	4f
Pyriglena leuconota	forest specialist	32.3	4.2	2	5	8450000	3	19	38.89	58.1	19.06	2t
Taraba major	habitat generalist	59.2	4	2	1	18300000	5	0	5.56	20.09	8.44	8v
Thamnomanes caesius	forest specialist	15.7	4	2	6	9240000	1	134	38.89	39.35	58.56	1d
Thamnophilus aethiops	forest specialist	25.7	4	2	4	8180000	1	41	38.89	41.42	46.81	1n
Thamnophilus amazonicus	habitat generalist	18.68	4	3	4	7120000	4	5	13.89	20.02	26.34	4v
Thamnophilus palliatus	habitat generalist	23.3	4	3	1	7260000	3	7	5.56	5.71	44.22	4m
Thamnophilus schistaceus	forest specialist	20.3	4	2	4	5160000	2	26	11.11	19.47	15.82	2d

Table S2.3: Continued from previous page

Species, by family	Habitat specialisation	Morpho-ecological traits							Measures of vulnerability			Species code
		Body mass (g)	Trophic Level	Vertical stratum[a]	Flocking behaviour[b]	Geographic range size (km²)	Habitat breadth[c]	Natural abundance (n)	Observed PIO[d] (%)	Detectability-corrected PIO[e] (%)	Detectability (%)	
Thamnophilus stictocephalus	forest specialist	21.5	4	2	1	1410000	5	65	44.44	46.61	67.87	1f
Willisornis poecilinotus	forest specialist	18.4	4.1	2	3	738000	1	19	30.56	31.69	41.58	2u
Conopophagidae												
Conopophaga aurita	forest specialist	26.3	4	1	1	674000	1	3	8.33	8.63	42.71	5g
Conopophaga roberti	forest specialist	20.8	4	1	1	730000	2	20	13.89	14.12	48.9	2p
Formicariidae												
Formicarius analis	forest specialist	62.19	3.9	1	1	10100000	2	9	13.89	14.09	45.1	3v
Formicarius colma	forest specialist	47	4	1	1	12000000	1	1	2.78	9.56	42.3	6n
Hylopezus macularius	forest specialist	44.2	4	1	1	4380000	1	2	2.78	46.21	16.4	5w
Dendrocolaptidae												
Dendrexetastes rufigula	habitat generalist	69.6	3.8	4	1	6340000	2	4	8.33	8.33	16.67	5b
Dendrocincla fuliginosa	habitat generalist	38.7	4.2	2	3	12300000	1	18	75	78.88	47.83	2v
Dendrocolaptes certhia	habitat generalist	68.7	4.2	3	3	8970000	1	5	13.89	52.48	7.64	4q
Dendroplex picus	habitat generalist	41.34	4.1	3	4	12500000	5	21	91.67	93.6	47.71	2m
Glyphorynchus spirurus	habitat generalist	14.6	4	2	4	13500000	2	26	55.56	58.21	46.6	2b
Hylexetastes brigidai	forest specialist	117	4.1	3	3	3310000	1	0	2.78	20.64	26.97	8f
Lepidocolaptes albolineatus	habitat generalist	20.3	4	4	4	2390000	1	34	75	94.39	39.26	1u
Sittasomus griseicapillus	habitat generalist	13.12	3.7	3	4	13700000	5	11	13.89	49.69	12.07	3n
Xiphorhynchus guttatus	habitat generalist	59.69	4.2	3	4	7680000	1	56	69.44	73.14	46.65	1i
Xiphorhynchus spixii	forest specialist	31.2	4	2	6	1110000	1	23	16.67	33.49	11.73	2j
Furnariidae												
Automolus infuscatus	forest specialist	32.9	4.1	2	6	1940000	1	10	8.33	8.33	39.83	3o
Philydor erythrocercum	forest specialist	25.27	4	3	6	7100000	2	19	25	29.87	20.35	2s
Synallaxis rutilans	forest specialist	16.7	4	2	1	6770000	1	7	22.22	25.47	23.86	4l
Xenops minutus	forest specialist	10.6	4	3	6	14000000	2	14	30.56	36.49	26.17	3e
Tyrannidae												
Attila cinnamomeus	habitat generalist	38.8	3.7	3	1	6400000	1	5	2.78	48.55	24.35	4o
Attila spadiceus	habitat generalist	39.1	4	3	1	18100000	3	21	44.44	80.89	16.74	2l
Camptostoma obsoletum	habitat generalist	8.1	3.4	4	1	17900000	6	8	72.22	73.94	47.26	3z
Cnemotriccus fuscatus	habitat generalist	13.6	4	2	1	14300000	6	0	11.11	64.88	8.36	7q
Empidonomus varius	habitat generalist	27.1	3.2	4	1	12900000	4	0	2.78	52.03	25.45	7z
Hemitriccus griseipectus	forest specialist	8.9	4	3	1	5220000	2	5	8.33	22.74	8.25	4r
Hemitriccus minor	forest specialist	7.4	4	2	1	2920000	1	16	38.89	39.04	71.02	3b
Inezia subflava	habitat generalist	8.4	4	2	1	4130000	3	0	2.78	54.7	6.33	8j
Lathrotriccus euleri	habitat generalist	11.33	4	2	1	15000000	3	1	2.78	34.49	9.6	6r
Legatus leucophaius	habitat generalist	22.2	2.6	4	1	18600000	4	2	13.89	26.79	8.62	5y
Lophotriccus galeatus	habitat generalist	6.6	4	3	1	4780000	2	73	50	50	87.15	1e
Megarynchus pitangua	habitat generalist	69.91	3.8	4	1	20500000	5	1	2.78	9.56	42.3	6t
Mionectes oleagineus	habitat generalist	11.17	2.4	2	4	16600000	3	1	11.11	27.31	6.77	6u
Myiarchus ferox	habitat generalist	27.5	3	3	1	13600000	4	19	80.56	89.05	38.32	2r
Myiarchus tuberculifer	habitat generalist	17.7	3.7	4	1	22400000	5	15	52.78	56.24	34.71	3d

Table S2.3: *Continued from previous page*

		Morpho-ecological traits							Measures of vulnerability			
Species, by family	Habitat specialisation	Body mass (g)	Trophic Level	Vertical stratum[a]	Flocking behaviour[b]	Geographic range size (km^2)	Habitat breadth[c]	Natural abundance (n)	Observed PIO[d] (%)	Detectability-corrected PIO[e] (%)	Detectability (%)	Species code
Myiodynastes maculatus	habitat generalist	43.2	3.7	4	1	9770000	4	3	38.89	69.82	12.39	5m
Myiopagis caniceps	forest specialist	10.5	3.6	4	4	6650000	1	0	2.78	80.13	20.37	8l
Myiopagis gaimardii	forest specialist	12.02	3.6	4	4	12100000	3	46	91.67	96.37	58.68	1l
Myiornis ecaudatus	habitat generalist	5.25	4	3	4	7600000	2	20	50	50.38	54.55	2q
Onychorhynchus coronatus	habitat generalist	14	4	3	4	7160000	1	8	11.11	33.97	8.28	4c
Ornithion inerme	habitat generalist	7	4	2	4	10200000	2	18	27.78	34.3	23.77	2w
Pitangus lictor	habitat generalist	25.5	4	2	1	12400000	2	1	13.89	68.06	3.37	7b
Pitangus sulphuratus	habitat generalist	62.85	3.7	4	1	28600000	5	2	30.56	43.63	18.11	6f
Platyrinchus platyrhynchos	forest specialist	12	4	3	1	6410000	1	1	5.56	8.38	34.93	7c
Poecilotriccus sylvia	habitat generalist	7.1	4	2	1	7120000	5	38	27.78	30.9	36.11	1r
Rhynchocyclus olivaceus	forest specialist	21.3	4	2	4	6050000	2	10	25	29.12	23.55	3r
Rhytipterna simplex	forest specialist	31.8	3.6	4	1	11300000	1	13	19.44	22.9	26.96	3i
Todirostrum chrysocrotaphum	habitat generalist	7	4	4	1	5860000	3	10	11.11	73.19	4.3	3s
Tolmomyias flaviventris	habitat generalist	12.2	3.6	3	4	10200000	5	33	86.11	87.24	53.08	1v
Tolmomyias poliocephalus	habitat generalist	10.8	4	4	4	10000000	3	47	55.56	57.28	48.39	1k
Tolmomyias sulphurescens	habitat generalist	14.3	3.8	3	1	19200000	6	41	50	50.56	53.91	1o
Tyrannulus elatus	habitat generalist	7	3.2	4	1	8700000	4	13	27.78	31.1	27.99	3k
Tyrannus melancholicus	habitat generalist	37.4	4	4	1	28500000	5	0	16.67	75.15	20.09	8y
Zimmerius acer	forest specialist	7.13	3.4	4	1	2730000	2	24	52.78	55.02	44.27	2i
Cotingidae												
Gymnoderus foetidus	forest specialist	275.49	2.4	4	1	7480000	2	0	2.78	8.41	45.13	8b
Lipaugus vociferans	forest specialist	75.42	3	3	1	10600000	1	29	33.33	33.93	42.62	1y
Pachyramphus castaneus	habitat generalist	19.5	3.8	4	1	11500000	4	8	25	43.35	12.5	4d
Pachyramphus marginatus	forest specialist	18.4	3	4	4	11000000	1	5	27.78	36.26	20.38	4t
Pachyramphus minor	forest specialist	36.6	3	4	6	7380000	1	1	2.78	9.56	42.3	6x
Querula purpurata	forest specialist	107.35	3	4	2	8790000	1	4	5.56	5.56	16.7	5e
Schiffornis turdina	forest specialist	31.7	3	2	1	10400000	2	8	11.11	11.1	62.52	4h
Tityra inquisitor	habitat generalist	43.1	2.4	4	1	18100000	2	4	5.56	8.51	10.83	5f
Pipridae												
Lepidothrix iris	forest specialist	8	2.4	2	1	963000	1	0	2.78	8.41	45.13	8k
Pipra fasciicauda	forest specialist	15.9	2.2	2	1	7620000	4	45	33.33	36.58	32.65	1m
Pipra pipra	habitat generalist	11.11	2.8	2	1	11300000	2	5	8.33	30.93	4.65	4u
Pipra rubrocapilla	forest specialist	12	2.4	2	1	7060000	1	7	8.33	8.36	46.72	4k
Piprites chloris	forest specialist	16	3.8	4	4	12900000	2	3	5.56	6.85	26.35	5o
Tyranneutes stolzmanni	forest specialist	7.2	2.4	3	1	6540000	1	5	33.33	42.78	25.41	4w
Vireonidae												
Cyclarhis gujanensis	habitat generalist	28.8	4	3	1	21700000	4	13	30.56	31.43	45.6	3f
Hylophilus hypoxanthus	forest specialist	17	4	4	6	498000	1	3	2.78	48.62	24.32	5i
Hylophilus ochraceiceps	habitat generalist	11.6	3.6	2	6	10600000	1	1	5.56	8.51	10.83	6p
Hylophilus semicinereus	habitat generalist	13	4	3	4	5270000	3	3	11.11	17.1	15.91	5j
Vireo olivaceus	forest specialist	16.06	3.4	4	1	52900000	5	1	8.33	15.29	10.17	7h

Table S2.3: *Continued from previous page*

Species, by family	Habitat specialisation	Body mass (g)	Trophic Level	Vertical stratum[a]	Flocking behaviour[b]	Geographic range size (km²)	Habitat breadth[c]	Natural abundance (n)	Observed PIO[d] (%)	Detectability-corrected PIO[e] (%)	Detectability (%)	Species code
Troglodytidae												
Campylorhynchus turdinus	forest specialist	32.6	3.6	4	1	9110000	4	39	55.56	55.82	61.26	1q
Microcerculus marginatus	forest specialist	18.22	4	2	1	8480000	1	4	2.78	48.54	24.35	5c
Thryothorus coraya	habitat generalist	17.2	3.8	2	1	5500000	3	24	30.56	30.63	63.48	2h
Thryothorus genibarbis	habitat generalist	19.2	4	2	1	8790000	5	213	41.67	41.7	74.56	1b
Troglodytes aedon	habitat generalist	10.85	3.4	2	1	59200000	5	10	83.33	91.62	40.47	3t
Polioptilidae												
Polioptila plumbea	habitat generalist	6	4	4	4	15800000	6	0	2.78	14.81	8.91	8p
Ramphocaenus melanurus	forest specialist	9.7	4	3	4	16400000	2	59	36.11	37.6	42.37	1h
Turdidae												
Turdus albicollis	habitat generalist	54	3.2	3	1	15100000	1	1	2.78	9.56	42.3	7g
Thraupidae												
Cissopis leverianus	habitat generalist	76	2.6	4	2	11000000	3	0	2.78	31.84	7.12	7o
Conirostrum speciosum	habitat generalist	8.8	3.4	4	1	13100000	4	0	22.22	30.15	20.3	7s
Dacnis cayana	habitat generalist	13	2.9	4	4	15700000	4	0	2.78	32.78	15.04	7y
Euphonia violacea	habitat generalist	15	2	3	4	9350000	3	0	5.56	20.64	5.39	8a
Hemithraupis guira	habitat generalist	12	3.7	4	6	14300000	3	27	52.78	57.28	35.49	2a
Lamprospiza melanoleuca	forest specialist	34	2.9	4	2	4670000	1	3	11.11	13.37	13.89	5k
Ramphocelus carbo	habitat generalist	25.92	3	2	2	11500000	5	40	91.67	91.77	65.4	1p
Tachyphonus luctuosus	habitat generalist	13	3.6	3	6	10700000	3	0	2.78	24.68	15.72	8t
Tangara mexicana	habitat generalist	20.5	3	4	2	7710000	5	0	2.78	62.48	2.42	8u
Thraupis episcopus	habitat generalist	35	2.7	4	2	13300000	4	0	5.56	43.96	32.54	8w
Thraupis palmarum	habitat generalist	39	2.4	4	4	15300000	6	2	16.67	59.5	4.77	6g
Coerebidae												
Coereba flaveola	habitat generalist	10.01	2.6	4	4	22400000	5	1	27.78	89.57	5.76	6l
Emberizidae												
Arremon taciturnus	forest specialist	24.8	2.9	1	1	9910000	1	27	44.44	46.7	43.34	1z
Oryzoborus angolensis	habitat generalist	13	2.4	2	1	13900000	3	0	2.78	24.06	04.06	8n
Cardinalidae												
Cyanocompsa cyanoides	habitat generalist	32.5	2.5	2	1	7600000	2	5	5.56	29.4	4.24	4p
Granatellus pelzelni	habitat generalist	11.2	4	3	4	693000	1	8	19.44	28.23	13.89	4a
Saltator grossus	habitat generalist	44.2	3.2	3	1	9340000	1	13	8.33	9.87	40.32	3j
Saltator maximus	habitat generalist	47.62	3.2	3	1	15400000	3	1	11.11	66.62	2.96	7e
Icteridae												
Cacicus cela	habitat generalist	85.45	2.8	4	2	11200000	4	2	11.11	24.82	9.36	5r
Icterus cayanensis	habitat generalist	35.44	3.4	4	2	5290000	3	0	11.11	46.44	6.37	8i
Psarocolius bifasciatus	forest specialist	335.7	3	4	2	178000	3	0	2.78	80.13	20.37	8q
Psarocolius decumanus	habitat generalist	206.3	2.8	4	2	13900000	3	0	11.11	74.69	2.68	8r

[a] **Vertical stratum:** (1) ground, (2) understorey, (3) midstorey, (4) canopy.

[b] **Flocking behaviour:** (1) solitary or pairs, (2) monospecific flocks, (3) facultative ant-follower, (4) facultative mixed-species flock attendant, (5) obligate ant-follower, (6) obligate mixed-species flock attendant.

[c] **Habitat breadth:** number of habitats used.

[d] **Observed PIO:** proportion of islands occupied not corrected for imperfect detectability.

[e] **Detectability-corrected PIO:** proportion of islands occupied corrected for imperfect detectability.

References for the Supporting Information

1. BirdLife International (2017) IUCN Red List for birds. Downloaded from http://www.birdlife.org in February 2017.

2. Bolger D.T., Alberts A.C., & Soule M.E. (1991) Occurrence patterns of bird species in habitat fragments: Sampling, extinction, and nested species subsets. The American Naturalist, 137, 155–166.

3. Brown J.H. (1971) Mammals on mountaintops: nonequilibrium insular biogeography. The American Naturalist, 105, 467–478.

4. Henriques L.M.P., Wunderle Jr. J.M., & Willig M.R. (2003) Birds of the Tapajós National Forest, Brazilian Amazon: A preliminary assessment. Ornitología Neotropical, 14, 307–338.

5. Jetz W., Thomas G.H., Joy J.B., Hartmann K., & Mooers A.O. (2012) The global diversity of birds in space and time. Nature, 491, 444–448.

6. Jullien M. & Clobert J. (2000) The survival value of flocking in Neotropical birds: Reality or fiction? Ecology, 81, 3416–3430.

7. Jullien M. & Thiollay J.-M. (1998) Multi-species territoriality and dynamic of neotropical forest understorey bird flocks. Journal of Animal Ecology, 67, 227–252.

8. Keinath D.A., Doak D.F., Hodges K.E., Prugh L.R., Fagan W., Sekercioglu C.H., Buchart S.H.M., & Kauffman M. (2017) A global analysis of traits predicting species sensitivity to habitat fragmentation. Global Ecology and Biogeography, 26, 115–127.

9. MacKenzie D.I., Nichols J.D., Lachman G.B., Droege S., Andrew Royle J., & Langtimm C.A. (2002) Estimating site occupancy rates when detection probabilities are less than one. Ecology, 83, 2248–2255.

10. Martínez A.E., Gomez J.P., Ponciano J.M., & Robinson S.K. (2016) Functional traits, flocking propensity, and perceived predation risk in an Amazonian understory bird community. The American Naturalist, 187, 607–619.

11. Munn C.A. & Terborgh J.W. (1979) Multi-species territoriality in Neotropical foraging flocks. The Condor, 81, 338.

12. Si X., Baselga A., & Ding P. (2015) Revealing beta-diversity patterns of breeding bird and lizard communities on inundated land-bridge islands

by separating the turnover and nestedness components. PLoS ONE, 10, e0127692.

13. Stotz D.F., Fitzpatrick J.W., Parker III T.A., & Moskovits D.K. (1996) Neotropical Birds: Ecology and Conservation. The University of Chicago Press, Chicago, USA.

14. Thiollay J.-M. & Jullien M. (1998) Flocking behaviour of foraging birds in a neotropical rain forest and the antipredator defence hypothesis. Ibis, 140, 382–394.

15. Wang Y., Zhang J., Feeley K.J., Jiang P., & Ding P. (2009) Life-history traits associated with fragmentation vulnerability of lizards in the Thousand Island Lake, China. Animal Conservation, 12, 329–337.

16. Willis E.O. & Oniki Y. (1978) Birds and army ants. Annual Review of Ecology and Systematics, 9, 243–263.

17. Willson S.K. (2004) Obligate army-ant-following birds: A study of ecology, spatial movement patterns, and behavior in Amazonian Peru. Ornithological Monographs, 55, 1–67.

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

19. Yong D.L., Qie L., Sodhi N.S., Koh L.P., Peh K.S.-H., Lee T.M., Lim H.C., & Lim S.L.-H. (2011) Do insectivorous bird communities decline on land-bridge forest islands in Peninsular Malaysia? Journal of Tropical Ecology, 27, 1–14.

20. Yu M., Hu G., Feeley K.J., Wu J., & Ding P. (2012) Richness and composition of plants and birds on land-bridge islands: effects of island attributes and differential responses of species groups. Journal of Biogeography, 39, 1124–1133.

Supporting Information

Ecological traits modulate bird species responses to forest fragmentation in an Amazonian anthropogenic archipelago

```r
# Clear workspace and disable scientific notation
remove(list = ls()); options(scipen = 999)
```

Packages

```r
library(caper)
library(qpcR)
```

Species-area relationships (SAR models)

```r
tables1 = read.table("https://ndownloader.figshare.com/files/15158654", header = T, row.names = 1)
```

```r
sar.all = lm(log10(Sall + 1) ~ log10(area), data = tables1)
sar.fs = lm(log10(Sfs + 1) ~ log10(area), data = tables1)
sar.hg = lm(log10(Shg + 1) ~ log10(area), data = tables1)
```

```r
summary(sar.all)
```

```
##
## Call:
## lm(formula = log10(Sall + 1) ~ log10(area), data = tables1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.55351 -0.05124  0.02985  0.10188  0.25938
##
## Coefficients:
##             Estimate Std. Error t value           Pr(>|t|)
## (Intercept)  1.10043    0.06963    15.80 < 0.0000000000000002 ***
## log10(area)  0.31574    0.04028     7.84       0.00000000398 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.158 on 34 degrees of freedom
## Multiple R-squared:  0.6438, Adjusted R-squared:  0.6334
## F-statistic: 61.46 on 1 and 34 DF,  p-value: 0.000000003985
```

```r
summary(sar.fs)
```

```
##
## Call:
## lm(formula = log10(Sfs + 1) ~ log10(area), data = tables1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.67741 -0.08069  0.02970  0.14995  0.38494
```

```
##
## Coefficients:
##              Estimate Std. Error t value    Pr(>|t|)
## (Intercept)  0.45772    0.09878   4.634 0.0000510189 ***
## log10(area)  0.41435    0.05714   7.252 0.0000000215 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2242 on 34 degrees of freedom
## Multiple R-squared:  0.6073, Adjusted R-squared:  0.5958
## F-statistic: 52.59 on 1 and 34 DF,  p-value: 0.00000002148
```

```
summary(sar.hg)
```

```
##
## Call:
## lm(formula = log10(Shg + 1) ~ log10(area), data = tables1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.51392 -0.04700  0.04992  0.08502  0.20200
##
## Coefficients:
##              Estimate Std. Error t value          Pr(>|t|)
## (Intercept)  1.02117    0.06324   16.15 < 0.0000000000000002 ***
## log10(area)  0.26190    0.03658    7.16          0.0000000281 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1435 on 34 degrees of freedom
## Multiple R-squared:  0.6012, Adjusted R-squared:  0.5895
## F-statistic: 51.26 on 1 and 34 DF,  p-value: 0.00000002805
```

Forest specialists vs. habitat generalists

```
sar.comparison = as.data.frame(cbind(c(tables1$area, tables1$area),
                                     c(tables1$Sfs, tables1$Shg)))
names(sar.comparison) = c("area", "richness")
sar.comparison$group = "forest_specialists"
sar.comparison[37:72,]$group = "habitat_generalists"
```

```
summary(lm(log10(richness + 1) ~ log10(area) * group, data = sar.comparison))
```

```
##
## Call:
## lm(formula = log10(richness + 1) ~ log10(area) * group, data = sar.comparison)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.67741 -0.06330  0.04875  0.11563  0.38494
##
## Coefficients:
##                                 Estimate Std. Error t value
## (Intercept)                      0.45772    0.08294   5.519
```

```
## log10(area)                              0.41435    0.04797   8.637
## grouphabitat_generalists                 0.56344    0.11729   4.804
## log10(area):grouphabitat_generalists    -0.15246    0.06784  -2.247
##                                           Pr(>|t|)
## (Intercept)                         0.00000057627628 ***
## log10(area)                         0.00000000000153 ***
## grouphabitat_generalists            0.00000893061520 ***
## log10(area):grouphabitat_generalists         0.0279 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1882 on 68 degrees of freedom
## Multiple R-squared:  0.6968, Adjusted R-squared:  0.6834
## F-statistic: 52.09 on 3 and 68 DF,  p-value: < 0.00000000000000022
```

Fragmentation effect

```
combined.area = data.frame(area = sum(tables1$area))

# All species
10^(predict(sar.all, combined.area, interval = "confidence"))
```

```
##        fit      lwr      upr
## 1 201.5509 130.4784 311.3373
```

```
# Forest specialists
10^(predict(sar.fs, combined.area, interval = "confidence"))
```

```
##        fit      lwr      upr
## 1 109.0632 58.85262 202.1113
```

```
# Habitat generalists
10^(predict(sar.hg, combined.area, interval = "confidence"))
```

```
##        fit      lwr      upr
## 1 104.6634 70.51507 155.3486
```

Phylogenetic Generalised Least Squares (PGLS)

```
tables3 = read.table("https://ndownloader.figshare.com/files/15158657", header = T, row.names = 1)

# Bird tree (1,000 phylogenetic download from birdtree.org - Hacket All Species)
tree.1000 = read.nexus("https://ndownloader.figshare.com/files/15158651")
tree.consensus = consensus(tree.1000, p = 0.5)
tree = compute.brlen(tree.consensus, 1)

dataset = comparative.data(tree, tables3, species, vcv = TRUE)
```

PGLS models

```
# Response variable: observed.PIO
```

```r
# Univariate models
obs.body.mass = pgls(observed.PIO ~
                     log10(body.mass),
               dataset, lambda = "ML")

obs.trophic.level = pgls(observed.PIO ~
                         trophic.level,
                  dataset, lambda = "ML")

obs.vertical.stratum = pgls(observed.PIO ~
                            vertical.stratum,
                      dataset, lambda = "ML")

obs.flocking.behaviour = pgls(observed.PIO ~
                              flocking.behaviour,
                        dataset, lambda = "ML")

obs.geographic.range.size = pgls(observed.PIO ~
                                 log10(geographic.range.size),
                           dataset, lambda = "ML")

obs.habitat.breadth = pgls(observed.PIO ~
                           habitat.breadth,
                     dataset, lambda = "ML")

obs.natural.abundance = pgls(observed.PIO ~
                             log10(natural.abundance + 1),
                       dataset, lambda = "ML")

# Additive models
obs.add.rarity = pgls(observed.PIO ~
                      log10(geographic.range.size) +
                      habitat.breadth +
                      log10(natural.abundance + 1),
                    dataset, lambda = "ML")

obs.add.pop.size = pgls(observed.PIO ~
                        log10(natural.abundance + 1) +
                        log10(body.mass) +
                        trophic.level,
                      dataset, lambda = "ML")

obs.add.foraging = pgls(observed.PIO ~
                        trophic.level +
                        vertical.stratum +
                        flocking.behaviour,
                      dataset, lambda = "ML")

# Interactive models
obs.int.rarity = pgls(observed.PIO ~
                      log10(geographic.range.size) *
                      habitat.breadth *
                      log10(natural.abundance + 1),
```

```r
                              dataset, lambda = "ML")

obs.int.pop.size = pgls(observed.PIO ~
                              log10(natural.abundance + 1) *
                              log10(body.mass) *
                              trophic.level,
                          dataset, lambda = "ML")


obs.int.foraging = pgls(observed.PIO ~
                              trophic.level *
                              vertical.stratum *
                              flocking.behaviour,
                          dataset, lambda = "ML")

# Model selection
obs.AIC = AIC(
  obs.body.mass,
  obs.trophic.level,
  obs.vertical.stratum,
  obs.flocking.behaviour,
  obs.geographic.range.size,
  obs.habitat.breadth,
  obs.natural.abundance,
  obs.add.rarity,
  obs.add.pop.size,
  obs.add.foraging,
  obs.int.rarity,
  obs.int.pop.size,
  obs.int.foraging)

obs.AIC$deltaAIC = akaike.weights(obs.AIC$AIC)$deltaAIC
obs.AIC$weights = akaike.weights(obs.AIC$AIC)$weights

obs.R2 = c(
  summary(obs.body.mass)[11],
  summary(obs.trophic.level)[11],
  summary(obs.vertical.stratum)[11],
  summary(obs.flocking.behaviour)[11],
  summary(obs.geographic.range.size)[11],
  summary(obs.habitat.breadth)[11],
  summary(obs.natural.abundance)[11],
  summary(obs.add.rarity)[11],
  summary(obs.add.pop.size)[11],
  summary(obs.add.foraging)[11],
  summary(obs.int.rarity)[11],
  summary(obs.int.pop.size)[11],
  summary(obs.int.foraging)[11])

obs.AIC$R2 = obs.R2
obs.AIC[order(obs.AIC$df, obs.AIC$AIC),]
```

```
##                        df      AIC  deltaAIC
## obs.natural.abundance   2 1697.151  43.91801
## obs.habitat.breadth     2 1847.325 194.09149
```

```
## obs.geographic.range.size  2 1850.594 197.36105
## obs.body.mass              2 1851.311 198.07724
## obs.flocking.behaviour     2 1854.150 200.91651
## obs.vertical.stratum       2 1854.443 201.20923
## obs.trophic.level          2 1854.845 201.61137
## obs.add.rarity             4 1666.006  12.77283
## obs.add.pop.size           4 1700.690  47.45681
## obs.add.foraging           4 1857.672 204.43835
## obs.int.rarity             8 1653.233   0.00000
## obs.int.pop.size           8 1703.471  50.23766
## obs.int.foraging           8 1862.848 209.61422
##                                                                         weights
## obs.natural.abundance        0.0000000002901317601347173152741276691557459344039671
## obs.habitat.breadth          0.00000000000000000000000000000000000007125875949
## obs.geographic.range.size    0.00000000000000000000000000000000000001389511267
## obs.body.mass                0.00000000000000000000000000000000000000971277103
## obs.flocking.behaviour       0.00000000000000000000000000000000000000234857139
## obs.vertical.stratum         0.00000000000000000000000000000000000000202880563
## obs.trophic.level            0.00000000000000000000000000000000000000165926934
## obs.add.rarity               0.0016814538813696659064200744637673778925091028213501
## obs.add.pop.size             0.0000000000494486035020742518628078521203406126005575
## obs.add.foraging             0.00000000000000000000000000000000000000040368805
## obs.int.rarity               0.9983185457667386186741964593238662928342819213867187
## obs.int.pop.size             0.0000000001231121540635354307349248748071346426651363
## obs.int.foraging             0.00000000000000000000000000000000000003034732
##                                        R2
## obs.natural.abundance         0.5540062
## obs.habitat.breadth           0.03140069
## obs.geographic.range.size     0.01672049
## obs.body.mass                 0.01524488
## obs.flocking.behaviour       -0.0007736182
## obs.vertical.stratum         -0.002445984
## obs.trophic.level            -0.004384838
## obs.add.rarity                0.6199413
## obs.add.pop.size              0.5506145
## obs.add.foraging             -0.007402409
## obs.int.rarity                0.6493197
## obs.int.pop.size              0.5529954
## obs.int.foraging             -0.01492667
```

```r
# Response variable: detectability.corrected.PIO

# Univariate models
cor.body.mass = pgls(detectability.corrected.PIO ~
                     log10(body.mass),
                   dataset, lambda = "ML")


cor.trophic.level = pgls(detectability.corrected.PIO ~
                         trophic.level,
                       dataset, lambda = "ML")


cor.vertical.stratum = pgls(detectability.corrected.PIO ~
                            vertical.stratum,
                          dataset, lambda = "ML")
```

```r
cor.flocking.behaviour = pgls(detectability.corrected.PIO ~
                              flocking.behaviour,
                         dataset, lambda = "ML")

cor.geographic.range.size = pgls(detectability.corrected.PIO ~
                                 log10(geographic.range.size),
                            dataset, lambda = "ML")

cor.habitat.breadth = pgls(detectability.corrected.PIO ~
                           habitat.breadth,
                      dataset, lambda = "ML")

cor.natural.abundance = pgls(detectability.corrected.PIO ~
                             log10(natural.abundance + 1),
                        dataset, lambda = "ML")

# Additive models
cor.add.rarity = pgls(detectability.corrected.PIO ~
                      log10(geographic.range.size) +
                      habitat.breadth +
                      log10(natural.abundance + 1),
                   dataset, lambda = "ML")

cor.add.pop.size = pgls(detectability.corrected.PIO ~
                        log10(natural.abundance + 1) +
                        log10(body.mass) +
                        trophic.level, dataset,
                      lambda = "ML")

cor.add.foraging = pgls(detectability.corrected.PIO ~
                        trophic.level +
                        vertical.stratum +
                        flocking.behaviour,
                      dataset, lambda = "ML")

# Interactive models
cor.int.rarity = pgls(detectability.corrected.PIO ~
                      log10(geographic.range.size) *
                      habitat.breadth *
                      log10(natural.abundance + 1),
                   dataset, lambda = "ML")

cor.int.pop.size = pgls(detectability.corrected.PIO ~
                        log10(natural.abundance + 1) *
                        log10(body.mass) *
                        trophic.level,
                      dataset, lambda = "ML")

cor.int.foraging = pgls(detectability.corrected.PIO ~
                        trophic.level *
                        vertical.stratum *
                        flocking.behaviour,
                      dataset, lambda = "ML")
```

```r
# Model selection
cor.AIC = AIC(
  cor.body.mass,
  cor.trophic.level,
  cor.vertical.stratum,
  cor.flocking.behaviour,
  cor.geographic.range.size,
  cor.habitat.breadth,
  cor.natural.abundance,
  cor.add.rarity,
  cor.add.pop.size,
  cor.add.foraging,
  cor.int.rarity,
  cor.int.pop.size,
  cor.int.foraging)

cor.AIC$deltaAIC = akaike.weights(cor.AIC$AIC)$deltaAIC
cor.AIC$weights = akaike.weights(cor.AIC$AIC)$weights

cor.R2 = c(
  summary(cor.body.mass)[11],
  summary(cor.trophic.level)[11],
  summary(cor.vertical.stratum)[11],
  summary(cor.flocking.behaviour)[11],
  summary(cor.geographic.range.size)[11],
  summary(cor.habitat.breadth)[11],
  summary(cor.natural.abundance)[11],
  summary(cor.add.rarity)[11],
  summary(cor.add.pop.size)[11],
  summary(cor.add.foraging)[11],
  summary(cor.int.rarity)[11],
  summary(cor.int.pop.size)[11],
  summary(cor.int.foraging)[11])

cor.AIC$R2 = cor.R2
cor.AIC[order(cor.AIC$df, cor.AIC$AIC),]
```

```
##                             df      AIC  deltaAIC                   weights
## cor.natural.abundance        2 1877.104 15.115382 0.00040633043973531
## cor.habitat.breadth          2 1891.950 29.961541 0.00000024270399151
## cor.geographic.range.size    2 1898.307 36.318896 0.00000001010636301
## cor.body.mass                2 1901.063 39.074212 0.00000000254850565
## cor.trophic.level            2 1902.062 40.073670 0.00000000154616540
## cor.flocking.behaviour       2 1902.317 40.328237 0.00000000136137392
## cor.vertical.stratum         2 1902.804 40.815409 0.00000000106706179
## cor.add.rarity               4 1864.512  2.523701 0.22035774683845807
## cor.add.pop.size             4 1875.745 13.756426 0.00080162814017317
## cor.add.foraging             4 1905.494 43.505318 0.00000000027802451
## cor.int.rarity               8 1861.988  0.000000 0.77829264307246515
## cor.int.pop.size             8 1879.215 17.226645 0.00014139186565967
## cor.int.foraging             8 1909.816 47.827843 0.00000000003202277
##                                     R2
## cor.natural.abundance        0.1129218
```

```
## cor.habitat.breadth          0.04737096
## cor.geographic.range.size    0.01786144
## cor.body.mass                0.004830527
## cor.trophic.level            -0.0006297909
## cor.flocking.behaviour       -0.0020095
## cor.vertical.stratum         -0.004370121
## cor.add.rarity               0.1731607
## cor.add.pop.size             0.1281231
## cor.add.foraging             -0.007763853
## cor.int.rarity               0.1985205
## cor.int.pop.size             0.129537
## cor.int.foraging             -0.0100225
```

Detectability as a function of natural abundance

```
summary(pgls(detectability ~ log10(natural.abundance + 1), dataset, lambda = "ML"))
```

```
##
## Call:
## pgls(formula = detectability ~ log10(natural.abundance + 1),
##     data = dataset, lambda = "ML")
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -14.5863  -3.3878   0.6103   3.3986  13.3980
##
## Branch length transformations:
##
## kappa  [Fix] : 1.000
## lambda [ ML] : 0.437
##    lower bound : 0.000, p = 0.0000010465
##    upper bound : 1.000, p = 0.000000000037904
##    95.0% CI   : (0.193, 0.691)
## delta  [Fix] : 1.000
##
## Coefficients:
##                             Estimate Std. Error t value        Pr(>|t|)
## (Intercept)                   7.2878     3.6098  2.0189          0.0448
## log10(natural.abundance + 1) 13.6458     1.8700  7.2972 0.000000000006332
##
## (Intercept)                  *
## log10(natural.abundance + 1) ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.033 on 205 degrees of freedom
## Multiple R-squared: 0.2062,  Adjusted R-squared: 0.2023
## F-statistic: 53.25 on 1 and 205 DF,  p-value: 0.000000000006332
```

Chapter 3

Patch-scale biodiversity retention in fragmented landscapes: Reconciling the habitat amount hypothesis with the island biogeography theory

3.1   Abstract

Aim: To test whether the species richness of understorey insectivorous birds on forest islands induced by a major hydroelectric dam is best explained by either the island biogeography theory (IBT) or the habitat amount hypothesis (HAH). Given the low dispersal ability of the focal species group and the hostile water matrix, we predict that the species richness will be predominantly driven by an island effect as posited by the IBT, rather than a sample area effect as posited by the HAH.

Location: Forest islands within the Balbina Hydroelectric Reservoir, central Brazilian Amazonia.

Taxon: Birds.

Methods: We mist-netted birds at 33 forest islands (0.63–1,699 ha), totalling 874 individuals of 59 species. The size of the local landscape used to calculate the habitat amount was determined by a multi-scale analysis in which buffers around mist-net lines ranged from 50 to 2,000 m. We applied four tests to examine whether the species richness on forest islands is predominantly driven by either an island effect (island size) or a sample area effect (habitat amount).

Results: From the four tests applied, one was consistent with an island effect, two were regarded as inappropriate to test the HAH, and one could not be adequately addressed due to island size being highly correlated with habitat amount in the local landscape (200-m buffer).

Main conclusions: Some of the proposed ways of testing the HAH may lead to misleading conclusions. The relative importance of island size in determining the species richness of understorey insectivorous birds on forest islands is higher than that of surrounding habitat amount, thereby providing stronger support for IBT. We propose a conceptual framework, based on the degree of matrix permeability and species dispersal ability, to determine to what extent a patch- or landscape-centric worldview in landscape ecology provides the most appropriate framework to assess the effects of habitat fragmentation on biodiversity.

KEYWORDS
Amazonia, habitat amount hypothesis, habitat fragmentation, habitat loss, insularization, island biogeography theory, SLOSS, species richness, Species–area relationship, tropical forest

3.2   Introduction

MacArthur and Wilson's (1967) island biogeography theory (hereafter, IBT) has been widely applied as a paradigmatic conceptual framework in habitat fragmentation ecology, implying that habitat patches are analogous to oceanic islands surrounded by a hostile matrix (Haila, 2002; Laurance, 2008). However, such analogy has been repeatedly challenged as IBT does not account for many factors operating in fragmented landscapes (Laurance, 2008; Wiens, 2008), which were later incorporated into a landscape ecology framework (Haila, 2002). For example, species move among suitable habitat patches as a function of varying degrees of terrestrial matrix permeability (Powell, Stouffer, & Johnson, 2013), indicating that habitat patches exert weaker boundaries to local populations and their derivative assemblages compared to oceanic islands. If habitat patches fail to behave as discrete spatial units, the universally celebrated species–area relationship (hereafter, SAR)—which is widely observed in fragmented landscapes (Matthews, Guilhaumon, Triantis, Borregaard, & Whittaker, 2016)—may be governed at spatial scales larger than that of island effects driven by habitat patch size.

With this in mind, Fahrig (2013) proposed the habitat amount hypothesis (hereafter, HAH), which posits that habitat patches are *not* discrete spatial units, and the habitat surrounding any given patch is the main source of immigrants. The underlying mechanism of SARs in fragmented landscapes is therefore predicted to be the sample area effect, rather than the island effect. Accordingly, sample sites within larger habitat patches harbour more species because they are also associated with a greater amount of surrounding habitat. Meanwhile, sample sites associated with the same amount of landscape-scale habitat should harbour the same number of species, regardless of patch size (fig. 7 in Fahrig, 2013). Such notion implies that conservation efforts should primarily focus on increasing the overall habitat amount (i.e. proportion of habitat in the landscape) without necessarily considering its spatial arrangement (i.e. size and isolation of individual habitat patches) (Seibold et al., 2017).

The generalisation of the HAH was initially criticised as its application was considered to be restricted to small-scale landscapes containing large habitat amounts (Hanski, 2015), although the HAH was yet to be tested (Fahrig, 2015). Recent empirical studies carried out in a variety of natural (e.g. forest fragments, fluvial islands, calcareous grasslands) and experimental fragmented landscapes (e.g. dead-wood microhabitats, moss fragments), across a wide range of taxonomic groups (e.g. small and arboreal mammals, birds, vascular plants, saproxylic beetles, micro-arthropods), have either supported (Melo, Sponchiado,

Cáceres, & Fahrig, 2017; Rabelo, Bicca-Marques, Aragón, & Nelson, 2017; Seibold et al., 2017) or refuted (Evju & Sverdrup-Thygeson, 2016; Haddad et al., 2016; Torrenta & Villard, 2017) the HAH. Therefore, further empirical studies are needed to appraise the degree to which the HAH can be generalised to different landscape scenarios and taxonomic groups (Rabelo et al., 2017).

The IBT and HAH were originally developed within a context of oceanic islands (MacArthur & Wilson, 1967) and habitat patches within terrestrial landscapes (Fahrig, 2013), respectively. These two landscape scenarios may be seen as extremes along a continuum. In a global synthesis, Matthews et al. (2016) showed that z-values of SARs are higher in true islands than in habitat patches. They also reported gradients in z-values ranging from inland water-body to oceanic islands, and from forest to mountaintop habitat patches. Hence, the magnitude of island effects is context-dependent regarding the type of matrix surrounding habitat patches (Prugh, Hodges, Sinclair, & Brashares, 2008). Patterns of species richness in intermediate landscape scenarios, such as inland water-body islands and mountaintops, could therefore be explained by either IBT or HAH.

The HAH was erected under the assumption that species perceive the wider macrohabitat mosaic as functionally connected (Fahrig, 2013). Matrix permeability, as measured by the structural similarity between habitat patches and any surrounding matrix (Prevedello & Vieira, 2010), along with inherent differences in species dispersal ability (Lees & Peres, 2009), would then determine whether species use their habitat primarily at the patch- or landscape-scale. Accordingly, we hypothesise that patterns of species richness in fragmented landscapes can be better explained under either the HAH, if species exhibit high levels of dispersal ability across a permeable matrix, or the IBT, if species exhibit low dispersal ability across a hostile matrix.

Here, we examined whether the HAH can be extended to anthropogenic archipelagic landscapes using the number of understorey insectivorous bird species on forest islands induced by a large hydroelectric dam in central Brazilian Amazonia. We focused on understorey insectivorous birds because they are particularly vulnerable to forest loss and fragmentation (Powell, Cordeiro, & Stratford, 2015), and exhibit low dispersal ability through non-forest matrix habitats (Laurance, Stouffer, & Laurance, 2004; Şekercioğlu et al., 2002). We show that the number of understorey insectivorous bird species on forest islands is best explained by an island effect, which is consistent with the IBT. Moreover, we propose a conceptual framework, based on the degree of matrix permeability and species dispersal ability, to determine which point along the continuum between a patch- and landscape-centric worldview in fragmentation ecology—represented here by either IBT or HAH—provides the most appropri-

ate guiding framework for biodiversity studies in fragmented landscapes.

3.3 Methods

3.3.1 Study area

This study was carried out within the Balbina Hydroelectric Reservoir (BHR) in central Brazilian Amazonia (1°40′ S, 59°40′ W; Figure 3.1). The BHR spans *c.* 300,000 ha and was formed by the damming of the Uatumã River in 1987 (Fearnside, 2016), creating over 3,500 land-bridge islands of variable size (range = 0.2–4,878 ha), which are surrounded by a vast water reservoir often containing dead tree snags rising above the water level (Benchimol & Peres, 2015a). To offset the environmental impacts of the dam, the left bank of the former Uatumã River, including all islands, became strictly protected by the 940,358-ha Uatumã Biological Reserve (IUCN category Ia), the largest of its category in Brazil (Figure 3.1).
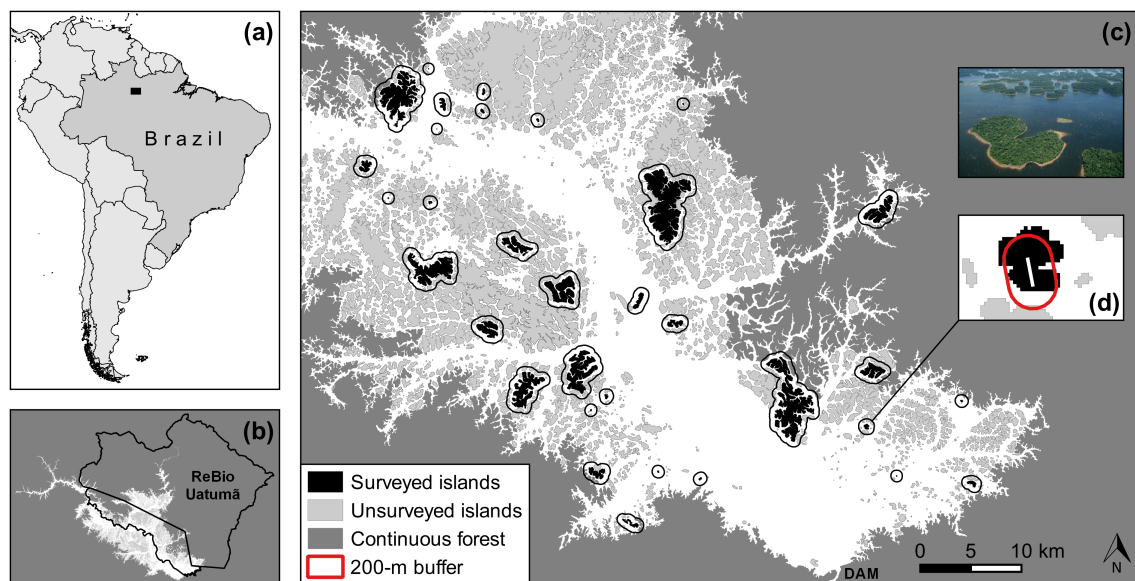
Figure 3.1: (a) Location of the study area in central Brazilian Amazonia, indicated by a solid rectangle containing (b) the Balbina Hydroelectric Reservoir (BHR) landscape, showing the boundaries of the Uatumã Biological Reserve, a strictly-protected area safeguarding most of this landscape; (c) larger inset map showing the spatial distribution of the 33 surveyed islands; and (d) the 200-m buffer area (red polygon) around a mist-net line (white line) representing the local landscapes derived for the understorey insectivorous birds examined here. Photo credit: Eduardo M. Venticinque.

The vegetation is comprised primarily of submontane dense ombrophilous (*terra firme*) forest, although *igapó* forest subjected to seasonal flooding

formerly occurred along the margins of the Uatumã River before damming. Forest structure varies among islands due to both island size and associated edge-mediated disturbance: smaller islands are species-poor and dominated by pioneer tree species, whereas larger islands are species-rich and contain a higher dominance of large-seeded canopy tree species (Benchimol & Peres, 2015a). According to the Köppen classification, the climate is equatorial fully humid (Af), with mean annual precipitation and temperature of 2,464 mm and 26.5°C, respectively (Alvares, Stape, Sentelhas, Gonçalves, & Sparovek, 2013).

3.3.2   Sampling design

We selected 33 forest islands within the BHR, ranging in size from 0.63 to 1,699 ha. Sixteen islands were on the left bank, whereas 17 islands were on the right bank (Figure 3.1). The combined study meta-landscape encompassed 175,583 ha where sample sites were spaced apart by an average distance of 27.9 km ($SD =$ 15.0 km; range = 2.0–68.5 km).

   We surveyed birds using mist nets (12 × 2.5 m, *Ecotone* 1016/12) from July to December in two consecutive years (2015 and 2016). We placed 16 mist nets end-to-end in the understorey along a continuous near-linear net-line (*c.* 200 m) whenever possible, but used a cross-shaped net-line design on islands smaller than 4 ha, thereby ensuring the same sampling effort across all 33 surveyed islands. Herein, each mist-net line corresponds to one sample site. Mist nets were operated from 06:00 to 15:00 h for 2 days at each site each year, resulting in a total sampling effort of 19,008 net-hours (16 mist nets × 9 hr × 2 days × 2 years × 33 sites). To avoid double-counting, we ringed birds with coded aluminium rings and excluded recaptures. Rings were provided by the Brazilian National Center for Bird Conservation and Research (CEMAVE) under research permits SISBIO 49068 and CEMAVE 3984.

3.3.3   Response variable and species group

We used the *number of species* of understorey insectivorous birds as the response variable, and limited our analysis to forest species because the habitat type (i.e. forest) must be appropriately defined for the focal species group (Fahrig, 2013). We defined forest species as those classified as having "medium" or "high" levels of forest dependency (*sensu* BirdLife International, 2018), and insectivorous species as those classified under the "invertebrate" dietary category (*sensu* Wilman et al., 2014). The only forest insectivorous species omitted from the analysis was the Amazonian Pygmy-Owl (*Glaucidium hardyi*) because surveys were diurnal, and this species is nocturnal (Wilman et al., 2014). As understorey mist

nets primarily capture understorey birds and occasionally those that walk on the ground or forage at forest strata higher than 2.5 m (Karr, 1981), we considered all species captured as understorey birds to avoid misinterpretation.

3.3.4   Predictor variables

We extracted data on *island size* and *habitat amount* for all 33 sample sites using a classified image (Collection 2, 2015, Amazon) derived from 30-m resolution LANDSAT imagery downloaded from the Brazilian Annual Land Use and Land Cover Mapping Project (available at http://mapbiomas.org). To do so, we used the QGIS software (QGIS Development Team, 2016) and the *LecoS* plugin (Jung, 2016). Island size corresponds to the total forest area (in hectares) within an island, and habitat amount corresponds to the percentage of forest cover within a given surrounding landscape at varying scales. In extracting the predictor variables, only "dense forest" (pixel value 3) was defined as forest, because other pixel values effectively represent either heavily degraded forests or non-forest land cover types.

3.3.5   Data analysis

Scale of effect

Species-landscape relationships are strongly affected by the scale at which landscape attributes are measured (Jackson & Fahrig, 2015). We therefore employed a multi-scale analysis to determine the "scale of effect"—the landscape-scale at which the relationship between the number of species and habitat amount peaks (Jackson & Fahrig, 2015). We defined the scale of effect as the "local landscape" for understorey insectivorous birds at the Balbina forest archipelago. Our multi-scale analysis examined 40 different buffer sizes around sample sites (i.e. mist-net lines), ranging from 50 to 2,000 m at 50-m intervals. The smallest landscape-scale (50 m) corresponds to the average between the reluctance of Amazonian understorey birds to cross forest clearings as narrow as 30 m (Laurance et al., 2004) and an assemblage-wide avian gap-crossing ability of up to 70 m (Lees & Peres, 2009). The largest landscape-scale (2,000 m) includes those frequently used in avian fragmentation studies (Aurélio-Silva, Anciães, Henriques, Benchimol, & Peres, 2016; Jackson & Fahrig, 2015; Morante-Filho, Faria, Mariano-Neto, & Rhodes, 2015). For this analysis, we included all 33 surveyed islands and log-transformed the response and predictor variables ($log_{10} x + 1$).

IBT versus HAH

The number of species in fragmented landscapes can be explained by either patch size (e.g. Torrenta & Villard, 2017) or habitat amount (e.g. Melo et al., 2017), which represents two worldviews in assessing the total area of suitable habitat. This means that the iconic SAR (Rosenzweig, 1995) holds true regardless of its spatial drivers (patch size or habitat amount), but that the underlying mechanism may be either the island effect driven by patch size as posited by the IBT (MacArthur & Wilson, 1967), or the sample area effect driven by habitat amount as posited by the HAH (Fahrig, 2013). We applied four tests to determine whether the IBT or the HAH is the most appropriate theoretical framework to explain the number of understorey insectivorous bird species on forest islands within the BHR.

Test 1: Multiple linear regression

We used multiple linear regression analysis to examine the independent effects of island size and habitat amount in the local landscape on species richness. This method allows one to estimate how much of the variation in the response variable (i.e. number of species) can be attributed solely to a predictor variable (e.g. island size), once the effects of another predictor (e.g. habitat amount) are controlled for (Legendre & Legendre, 1998).

An effect of island size, rather than one of habitat amount, would provide support for IBT, whereas the reverse would provide support for HAH (Figure 3.2). The response and predictor variables were log-transformed ($log_{10} x + 1$) prior to analysis. The predictor variables were also standardised (mean = 0, SD = 1) to allow comparison of regression slopes. Finally, we examined the strength of correlation values between island size and habitat amount across the entire spectrum of 40 landscape-scales (50–2,000 m) to assess the suitability of the multiple linear regression test.

Test 2: Z-values

We used the logarithmic form of the SAR (type IV curve *sensu* Scheiner, 2003) to fit simple linear regression models (Rosenzweig, 1995) for islands surrounded by either low habitat amounts (up to *c.* 50% of the landscape; Morante-Filho et al., 2015) or high habitat amounts, according to the following equation:

$$log_{10}(S + 1) = z \times log_{10}(A + 1) + log_{10}(c),$$

where S = number of species, z = regression slope, A = island size, c = regression intercept. To assess whether the z-values derived from either SARs were significantly different ($p < 0.05$), we performed an ANCOVA model with habitat amount (low or high) as an independent categorical variable. To support the IBT, the SAR for islands with low habitat amounts should have a higher z-value than those with high habitat amounts (Figure 3.2). Conversely, z-values should be statistically equivalent to support the HAH (Figure 3.2).

Test 3: Species accumulation curves

We compared the cumulative number of species across all 33 surveyed islands ordered according to either increasing (small-to-large) or decreasing (large-to-small) island sizes (Quinn & Harrison, 1988), which may lead to three possible outcomes. First, the small-to-large accumulation curve lies below the large-to-small, supporting IBT (Figure 3.2). Second, the curves overlap, supporting HAH (Figure 3.2). Third, the small-to-large accumulation curve lies above the large-to-small, supporting neither IBT nor HAH (Figure 3.2).

Test 4: Extrapolation of SAR

We fit a SAR model (see Test 2: Z-values) to all 33 surveyed islands. We further extrapolated the number of species to a hypothetical island containing the area (+1) of all 33 surveyed islands combined (7,874 ha), and compared the total number of species observed across surveyed islands to the extrapolated value (Yaacobi, Ziv, & Rosenzweig, 2007). Compared to the extrapolated value, a lower observed number of species would support IBT; a statistically equivalent number, HAH; a higher number, neither IBT nor HAH (Figure 3.2).
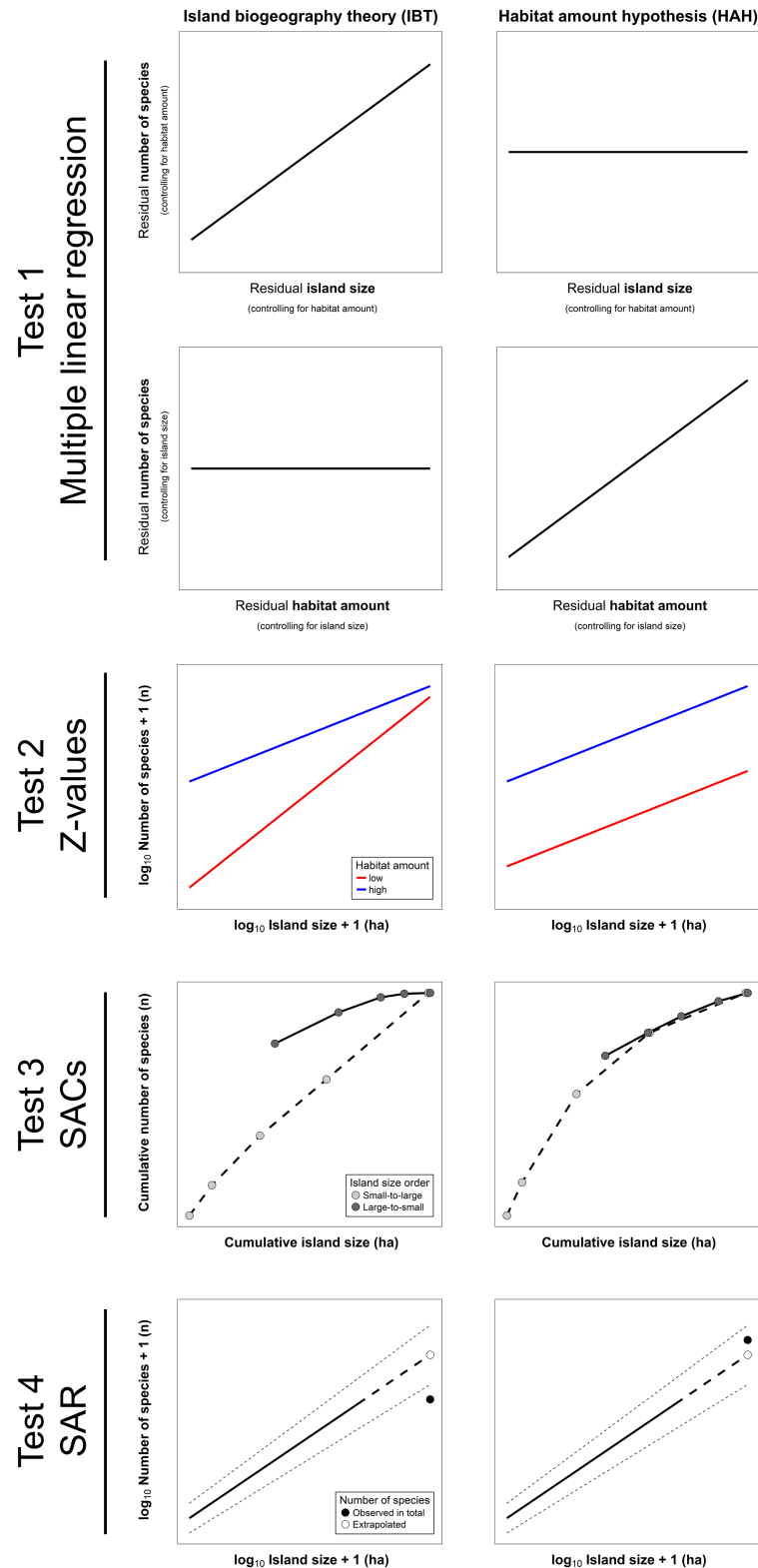
Figure 3.2: Possible conceptual relationships of the four empirical tests applied to determine whether either the island biogeography theory (IBT; graphs on the left) or the habitat amount hypothesis (HAH; graphs on the right) is the most appropriate framework to explain the number of understorey insectivorous bird species on forest islands within the Balbina Hydroelectric Reservoir in central Brazilian Amazonia.

3.4　Results

We captured a total of 874 individual understorey insectivorous birds representing 59 species, 49 genera, and 19 families across all 33 sample sites (see Table S3.1 in Supporting Information). The number of individuals per island ranged from 0 to 84 (26.5 ± 23.4), and the number of species from 0 to 27 (8.8 ± 7.5; see Table S3.2).

3.4.1　Scale of effect

The correlational peak between the number of species and habitat amount (i.e. the scale of effect) occurred at 200-m buffers around sample sites (r = 0.873; see Figure S3.1), thereby corresponding to the local landscape for understorey insectivorous birds at the Balbina forest archipelago. The fact that this local landscape is intermediate between the smallest and the largest landscape-scales examined here indicates that our multi-scale analysis included the true scale of effect (Jackson & Fahrig, 2015). Therefore, habitat amount is defined as the percentage of forest cover only within 200-m buffer local landscapes for all subsequent analyses.

3.4.2　IBT versus HAH

Test 1: Multiple linear regression

A multiple linear regression model including island size and habitat amount showed that both predictor variables were strongly and positively related to the number of species (R^2_{adj} = 0.804, $p < 0.001$). Partial regressions also showed that habitat amount had a slightly better fit and higher regression slope than island size (see Figure S3.2), which in itself would lend more support for HAH than IBT. However, island size and habitat amount were positively correlated across the entire range of 40 landscape-scales examined and peaked exactly at the scale of effect (200-m buffer; see Figure S3.3). Due to the high collinearity between predictors (r = 0.857), regression coefficients could change depending on the random component in the response variable (Legendre & Legendre, 1998), thereby precluding us from raising any conclusions derived from Test 1.

Test 2: Z-values

The SAR for islands surrounded by low habitat amounts (<55%) had a statistically higher z-value (0.747) than islands surrounded by high habitat amounts (>70%; 0.311), as shown by the ANCOVA test (p = 0.009; Figure 3.3). This outcome supports an island effect, rather than solely a sample area effect (Fahrig, 2013), thereby lending stronger support for IBT.
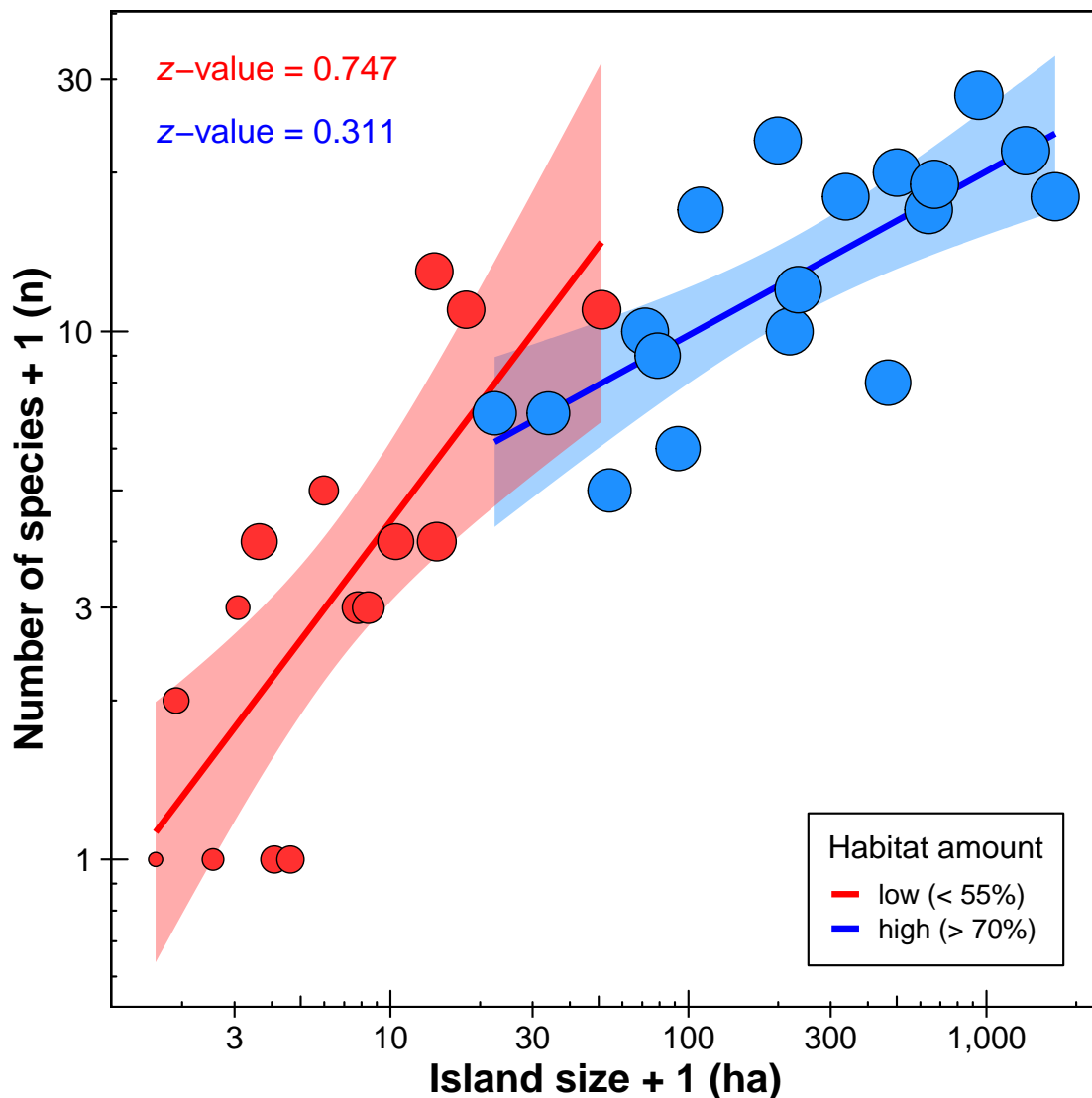
Figure 3.3: Divergent species–area relationships for understorey insectivorous birds surveyed at 33 forest islands surrounded by either low (r^2_{adj} = 0.598, $p <$ 0.001, n = 15) or high (r^2_{adj} adj = 0.547, $p <$ 0.001, n = 18) habitat amounts. Circle sizes scale to the landscape-scale habitat amount, with smaller and larger circles representing islands surrounded by either low or high habitat amounts, respectively. Note the different z-values for these two landscape scenarios (p = 0.009) and the base 10 logarithmic scales along both axes.

Test 3: Species accumulation curves

Species accumulation curves (SACs) did not overlap whether sampling sites were ordered according to either increasing (small-to-large) or decreasing (large-to-small) island sizes. The small-to-large curve lay above the large-to-small curve (Figure 3.4), which contradicts both IBT and HAH.
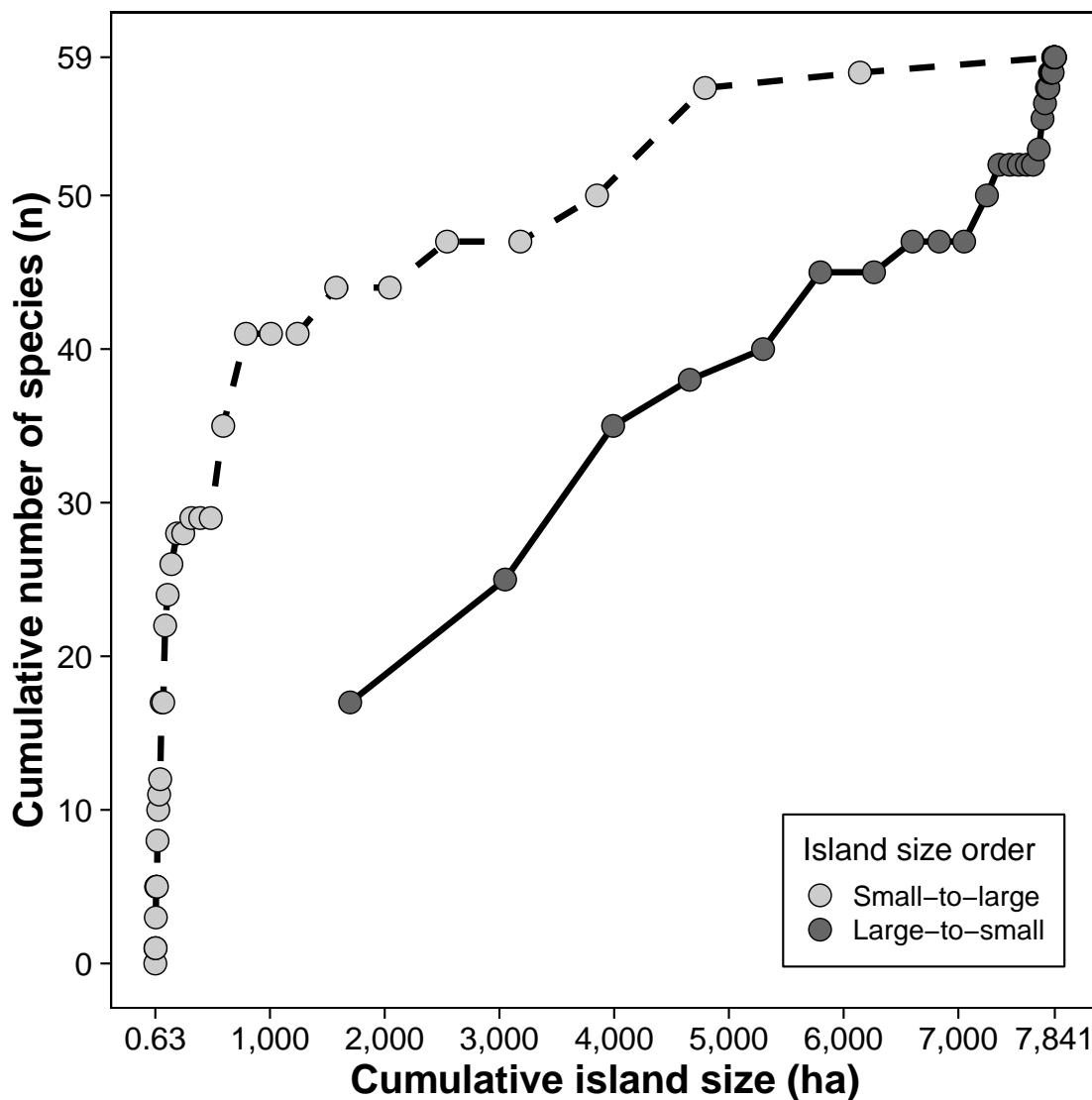
Figure 3.4: Species accumulation curves of understorey insectivorous birds for islands ordered according to either increasing (light grey circles, dashed line) or decreasing (dark grey circles, solid line) island sizes.

Test 4: Extrapolation of SAR

We observed a larger number of species (59 + 1) across all 33 surveyed islands than that extrapolated (55.2) to a hypothetical island containing the area (+1) of all surveyed islands combined (7,874 ha). However, the difference between the observed and extrapolated number of species was not significant (Figure 3.5), which lends support for HAH.
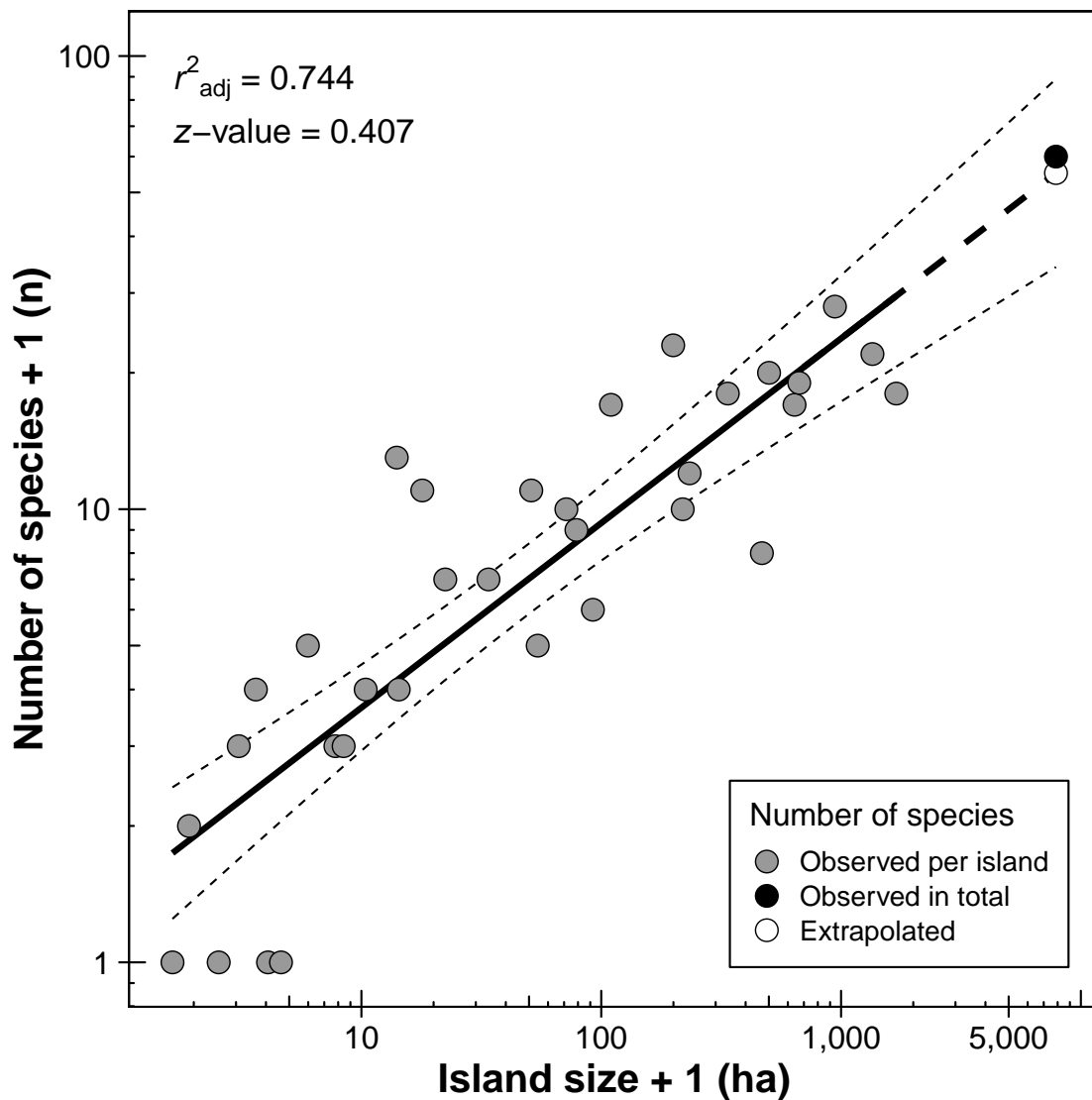
Figure 3.5: Species–area relationship for understorey insectivorous birds surveyed at 33 forest islands. The white circle shows the extrapolated number of species (55.2) to a hypothetical island containing the area (+1) of all 33 surveyed islands combined (7,874 ha), whereas the black circle shows the total number of species observed in this study (59 + 1). Dashed lines show the 95% confidence intervals of the predicted line. Note the base 10 logarithmic scales along both axes

3.5 Discussion

Compared to the HAH, there was more evidence giving support to the IBT in explaining the number of understorey insectivorous bird species within Amazonian forest islands in one of the largest hydroelectric reservoirs on Earth. Considering the four tests applied, test 2 (z-values) was consistent with an island effect as posited by the IBT, tests 3 (SACs) and 4 (extrapolation of SAR) were regarded as inappropriate to test the HAH (see below), and test 1 (multiple linear regression) could not be adequately addressed due to a prohibitively high correlation between island size and habitat amount at the local landscape. As both an island effect (e.g. Evju & Sverdrup-Thygeson, 2016) and a sample area effect (e.g. Rabelo et al., 2017) may explain patterns of species richness in fragmented landscapes, the key question becomes which of these two theoretical frameworks provides the best fit to different scenarios in "real-world" fragmented landscapes. This question has critical implications to biodiversity conservation strategies as empirical evidence primarily supporting IBT would imply a management focus on the spatial arrangement of remaining habitat patches, whereas support for HAH would imply a management strategy focused on retaining the maximum overall amount of habitat regardless of its configuration (Seibold et al., 2017).

The independent effects of predictor variables may be disentangled using statistical methods such as multiple regression analysis. However, as the degree of collinearity between predictor variables increases, the accuracy in determining their independent effects decreases, particularly above a high threshold ($r > 0.7$) from which parameter estimates begin to be severely distorted in regression-type analyses (Dormann et al., 2013). In our set of sample sites, the highest correlation between island size and habitat amount occurred exactly at the scale of effect (i.e. 200-m buffer; $r = 0.857$), which precluded us from directly testing the predictions of the HAH. Ideally, patch size and habitat amount should be either orthogonally independent or negatively correlated (fig. 7 in Fahrig, 2013). However, the pervasive positive correlation between habitat patch size and habitat amount in landscapes worldwide is the rule rather than the exception (Fahrig, 2003), and this correlation becomes even stronger and more ubiquitous for smaller local landscapes (Rabelo et al., 2017). For instance, island size and habitat amount were more likely to be independent in our study system at larger scales, well beyond a demographically realistic local landscape for our focal species group.

The scale of effect is indeed unlikely to be known before sampling design is established; thereby a multi-scale analysis is necessary to determine the local landscape (Fahrig, 2013). This implies that sample sites selected *a priori* to control for the positive correlation between patch size and habitat amount may fail

to achieve this goal if the size of the local landscape is different than initially thought. To illustrate this, consider a set of sample sites where the size of the focal habitat patches increases while the amount of habitat remains constant (Figure 3.6). If the size of the local landscape derived from a multi-scale analysis is found to be half of that defined *a priori*, patch size and habitat amount will be positively correlated (Figure 3.6). Therefore, directly testing the HAH under its main assumptions is expected to be less feasible if the scale of effect is small or not known *a priori*. Despite these shortcomings, there are alternative ways of testing the HAH (Fahrig, 2013).
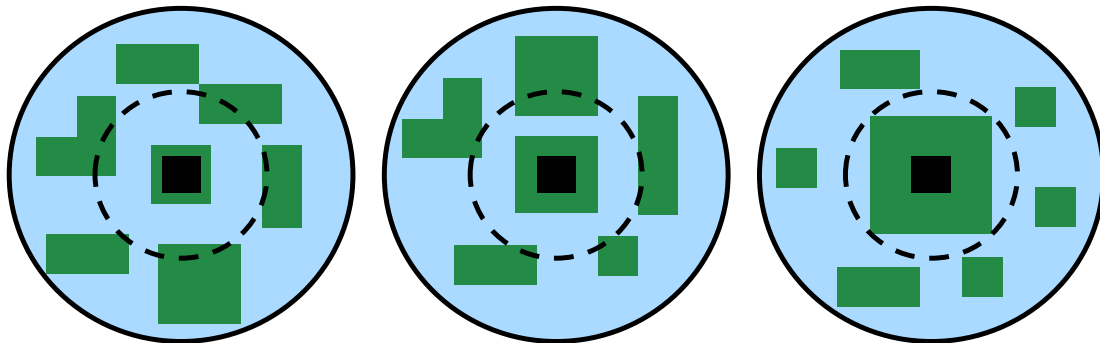
Figure 3.6: Sampling design established to control for the positive correlation between patch size and habitat amount. The solid black circle correspondents to the local landscape defined *a priori* (i.e. before the scale of effect is known). The dashed black circle corresponds to the local landscape derived from a multi-scale analysis (i.e. post data analysis). The difference between the two landscapes scales (solid and dashed black circles) implies that even a well-designed study may fail to control for the collinearity between predictors. Figure modified from Fahrig (2013).

First, if patch size *per se* does not affect the number of species, as predicted by the HAH, z-values derived from SARs are expected to be the same in landscapes with either low or high habitat amounts (Figure 3.2). We showed that the z-value for islands at landscapes isolated by low habitat amounts ($<55\%$) is statistically higher than that at landscapes connected by high habitat amounts ($>70\%$; Figure 3.3), which contradicts a prediction of the HAH. Such a difference could be attributed to the selected cutoff (55%) that distinguishes low from high habitat amounts. To test whether the difference in z-values is sensitive to this threshold, we reran test 2 (z-values) using different cutoffs for habitat amount (54%, 50%, 48.5%, 45%, 43.1%, 30%; see Supporting Information). The differences in z-values held true except when the SAR for islands surrounded by low habitat amounts was not significant. As all islands were smaller than 10 ha for the lower cutoffs ($\leqslant 45\%$), the lack of a relationship may be explained by the small island effect (i.e. for small islands, the variation in island size does not affect the

number of species; Lomolino & Weiser, 2001). In sum, forest islands surrounded by low habitat amounts had a steeper decline in species richness as a function of island size reduction (i.e. higher z-value) than those surrounded by high habitat amounts whenever the SAR models were significant.

The difference between z-values is attributed to an island effect, which is consistent with the IBT (Fahrig, 2013). Accordingly, larger and less isolated islands are more species-rich than smaller and more isolated islands because the former experience lower extinction (area effect) and higher colonisation rates (distance effect) (MacArthur & Wilson, 1967). Thus, if islands within their local landscapes were functionally connected as assumed by the HAH, the number of immigrants reaching focal islands would mainly depend on the amount of habitat surrounding those islands (sample area effect), thereby compensating species declines through rescue effects (Fahrig, 2013; see Seibold et al., 2017).

Second, the SACs from either small-to-large patches or from large-to-small patches should roughly coincide to support the HAH (Figure 3.2) as this is caused by a sample area effect, meaning that the long-celebrated dichotomy between a Single Large Or Several Small patches (SLOSS) should harbour a similar number of species. Alternatively, a faster accumulation in the number of species from large-to-small patches, compared to that from small-to-large patches, would be attributed to an island effect (IBT). The cumulative number of understorey insectivorous bird species at the Balbina forest archipelago rose faster from small-to-large patches than from large-to-small patches, which supports neither IBT nor HAH.

The fact that several small patches (islands) apparently harboured more species than a single large patch (Figure 3.4) is consistent with the literature (Fahrig, 2017). However, we did not find support for several possible explanations for this pattern at the Balbina forest archipelago. First, we strictly focused on forest species implying that the pattern was not confounded by the inclusion of disturbance-adapted species, which would increase the overall number of species across small patch sites (Lövei, Magura, Tóthmérész, & Ködöböcz, 2006). Second, habitat heterogeneity, regarding vegetation structure in Amazonian terra firme forests, is associated with elevation (Castilho et al., 2006), below-ground vertical distance to the water-table (Schietti et al., 2014) and horizontal distance to perennial streams (Drucker, Costa, & Magnusson, 2008). Thus, several small patches could harbour more species than a single large patch if they covered wider topographic and hydrologic gradients, resulting in higher levels of habitat heterogeneity (Báldi, 2008). However, our islands consist of upland habitat remnants resulting from hilltop terrains of the once continuous forest. As such, they span similar elevations and streams were missing from all but two very large islands

(Beco do Catitu and Mascote). As a result, surveyed islands shared relatively low levels of intra-patch habitat heterogeneity regarding closed-canopy forest structure. Third, the Balbina islands are isolated by a hostile water matrix which likely hampers the dynamic of colonisation and extinction (Palmeirim, Vieira, & Peres, 2017), particularly for species that are unable to either cross wide gaps or use dead tree snags as stepping stones. Indeed, the disappearance of understorey insectivorous birds from forest fragments has been largely attributed to dispersal limitation (Şekercioğlu et al., 2002), which along with a severely hostile water matrix explain patterns of bird species occupancy on forest islands (Moore, Robinson, Lovette, & Robinson, 2008). The relatively small local landscape threshold (200-m buffer) for understorey insectivorous birds at the Balbina forest archipelago provides additional evidence of such dispersal limitation (Jackson & Fahrig, 2012).

The most likely explanation for the observed SACs (Figure 3.4) relies on a bias associated with this method. In a SLOSS-type study, Gavish, Ziv, and Rosenzweig (2012) compared four methods to examine the effects of habitat loss and fragmentation on the species richness of spiders. They concluded that only SACs (Quinn & Harrison, 1988) should be avoided as this method was biased towards detecting more species in several small habitat patches than in a single large patch. This occurs because the method is sensitive to sampling intensity (i.e. proportion of patch area that is sampled), which could lead to an apparent higher number of species in small but more intensively sampled patches (Gavish et al., 2012). As the proportion of the island area we sampled in smaller islands was immensely higher than in larger islands, the result of the SACs is likely to be misleading. Moreover, the pattern of SACs was inconsistent with the HAH even in an experiment designed to decouple the independent effects of patch size and habitat amount on saproxylic beetles whose revealed strong support for HAH (Seibold et al., 2017).

Third, the extrapolation of the SAR model (Yaacobi et al., 2007) suggests that several small islands did not harbour more understorey insectivorous bird species than a single large island containing the same aggregate area of several small islands (Figure 3.5). This result is presumably consistent with the HAH (Fahrig, 2013; MacDonald, Anderson, Acorn, & Nielsen, 2018). However, had this method been suitable to test the HAH, oceanic archipelagos should harbour fewer species than that predicted by the extrapolation of SAR models derived from their constituent islands. Indeed, observed and extrapolated values of species richness for most oceanic archipelagos are statistically the same (75%–95% of 40 case studies; Santos et al., 2010). Collectively, this means that neither SACs nor an extrapolation of SAR models seem to be reliable methods to test the HAH.

Testing the HAH is by no means a trivial task for two main reasons. First,

as a general rule, habitat patch size and habitat amount tend to be positively correlated (Fahrig, 2003). However, these two predictors should be either largely orthogonal or negatively correlated to properly test predictions derived from the HAH (fig. 7 in Fahrig, 2013). Depending on the landscape, this constraint may however be overcome if the scale of effect (*sensu* Jackson & Fahrig, 2015) is known prior to the establishment of the experimental design. Second, species assemblages are comprised of species with varying degrees of dispersal ability, although within some groups, such as understorey insectivorous birds, such a trait is broadly similar across species (Laurance et al., 2004). Thus, the scale of effect for a given species assemblage will result from a combination of species with either lower or higher dispersal ability (Lees & Peres, 2009). Therefore, we believe the most robust way forward in testing the HAH would be to focus on individual species (Hanski, 2015) whose dispersal ability through the matrix (i.e. landscape vagility) is known *a priori* and derived from *in situ* studies (e.g. Awade & Metzger, 2008).

3.5.1 Moving beyond: a conceptual framework to assess the role of patch size and habitat amount in explaining species responses to habitat fragmentation

We can reasonably expect that local assemblage structure is primarily governed by patch-level characteristics in a hypothetical situation in which species seldom if ever exit the patch, due to low dispersal ability, low matrix permeability, or both (Moore et al., 2008). Conversely, landscape-level characteristics should matter most in a hypothetical situation in which species often move among patches within the local landscape, due to high dispersal ability, high matrix permeability, or both (Walter et al., 2017). Hence, the degree to which either a patch- or landscape-centric worldview is most pertinent in fragmentation ecology studies should be determined by the species vagility within the local landscape, which is largely a combination of matrix permeability (a landscape attribute) and dispersal ability (a species trait) (Figure 3.7). Accordingly, increasing support for IBT should be expected for a species assemblage with low dispersal ability in patches surrounded by an impermeable matrix (Figure 3.7c; this study; Palmeirim et al., 2017). Conversely, increasing support for HAH would be expected for a species assemblage in which high dispersal ability is prevalent and habitat patches are surrounded by a permeable matrix (Figure 3.7b; Melo et al., 2017). Under intermediate scenarios (Figure 3.7a,d), the most appropriate theoretical framework—IBT or HAH—would depend on the relative contributions of matrix permeability and species dispersal ability. For instance, support for

HAH would be expected if species successfully move among patches even if they are surrounded by an impermeable matrix (Figure 3.7a; Storck-Tonon & Peres, 2017), whereas support for IBT would be expected if species fail to move among patches even if they are surrounded by a relatively permeable matrix (Figure 3.7d; Munguía-Rosas & Montiel, 2014).
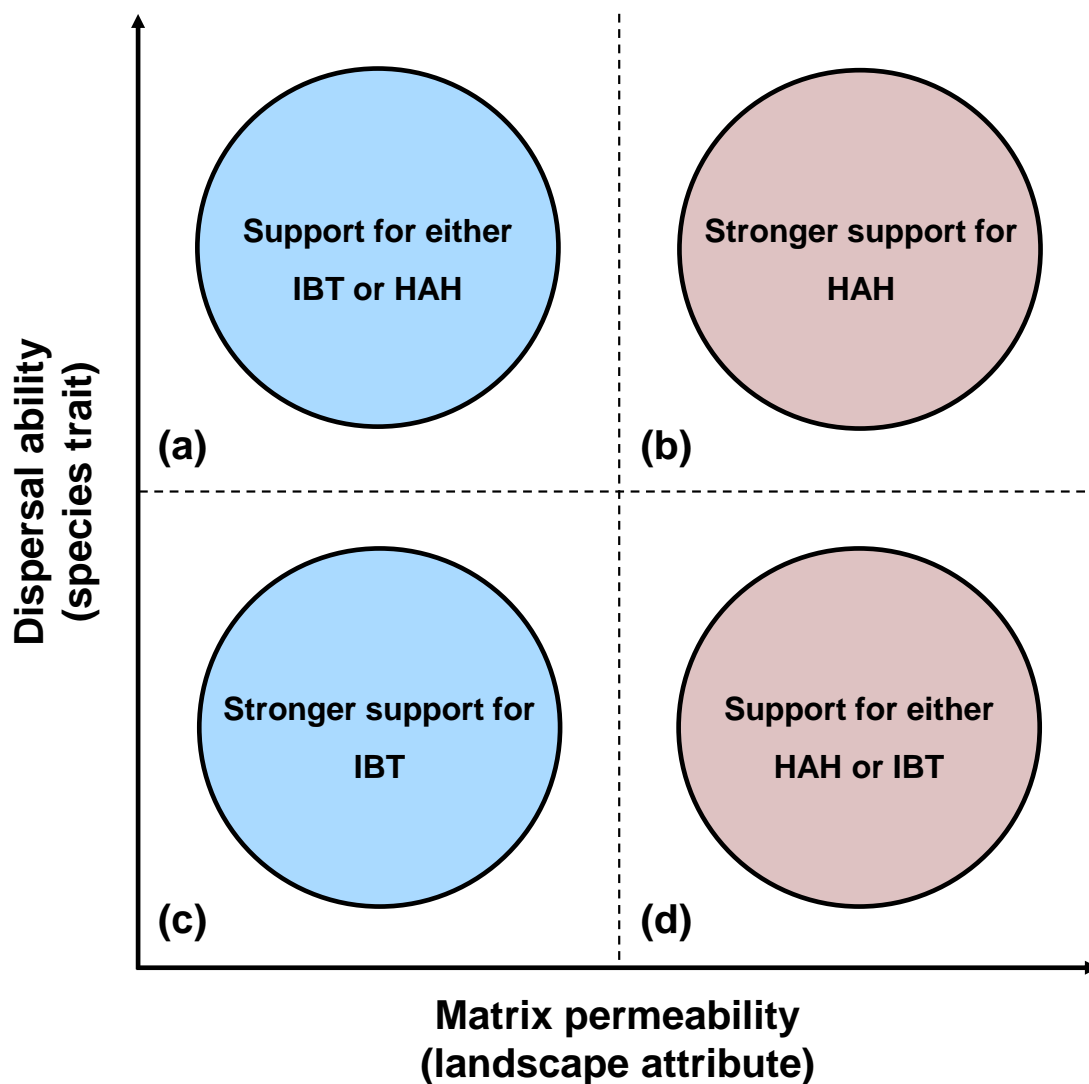
Figure 3.7: Conceptual framework based on the degree of matrix permeability and species dispersal ability in determining whether the island biogeography theory (IBT) or the habitat amount hypothesis (HAH) is the most appropriate guiding framework for biodiversity studies in fragmented landscapes.

A recent empirical study testing the HAH (MacDonald et al., 2018) provides further support for our conceptual framework. Accordingly, the inclusion of highly mobile species in the species pool led to stronger support for HAH in explaining the number of butterfly species on islands within a natural archipelagic landscape (Figure 3.7a), whereas excluding highly mobile species led to stronger

support for IBT (Figure 3.7c). Our conceptual framework also accounts for dynamic matrix habitats that change over time. As such, for a given forest landscape dominated by a regenerating vegetation matrix that accumulates aboveground phytomass, a patch-centric approach should be gradually replaced by a landscape-centric approach as the matrix becomes more permeable, ultimately enhancing species vagility of even the most sedentary species (Powell et al., 2013). The Biological Dynamics of Forest Fragments Project (BDFFP) in central Brazilian Amazonia is an iconic example of a dynamic tropical landscape, in which a cattle pasture matrix surrounding primary forest fragments has been fully replaced by an ageing secondary forest over the past *c.* 35 years (Stouffer, Johnson, Bierregaard, & Lovejoy, 2011). As the structural contrast between forest fragments and their adjacent matrix decreases, forest species can resume movements between forest fragments (Stouffer et al., 2011), exploit newly available matrix resources (Blake & Loiselle, 2001), and incorporate matrix habitats into their territories (Stouffer, Bierregaard, Strong, & Lovejoy, 2006). In such situation, a dichotomous classification of the landscape into either habitat or non-habitat is at best misleading (Stouffer et al., 2006), and a landscape-centric approach would be most appropriate.

At the Balbina forest archipelago, the structural contrast between habitat patches (forest islands) and the matrix (open-water) could not be greater, and is aggravated by the fact that matrix recovery, by definition, cannot occur within hydroelectric reservoirs with stable water levels. Such harsh landscape scenario restricts populations of species with low dispersal ability to fewer islands compared to species that can traverse the matrix. Indeed, the inherent swimming capacity—a measure of dispersal ability on open-water—of large vertebrate species at the Balbina forest archipelago is positively related to species island occupancy (Benchimol & Peres, 2015b). Based on both patch- and landscape-scale predictors, that study also found island size to be the single best predictor of island occupancy for most species. Likewise, island size was a powerful predictor of species richness of terrestrial and arboreal vertebrates (r^2 = 0.910, Benchimol & Peres, 2015c), birds (r^2 = 0.808, Aurélio-Silva et al., 2016), lizards (r^2 = 0.870, Palmeirim et al., 2017), and frogs (r^2 = 0.891, Lima et al., 2015) within Amazonian forest archipelagos. Given this bulk of evidence showing a strong island size effect on species richness, a patch-centric approach (IBT) is likely to be the most appropriate in true archipelagic landscapes. Nevertheless, species with high dispersal ability (e.g. orchid bees, Storck-Tonon & Peres, 2017; butterflies, MacDonald et al., 2018) may still be able to cross hostile expanses of water, which would justify a landscape-centric approach (HAH).

3.6   Conclusions

We tested the HAH under one extreme of the continuum of matrix permeability and species dispersal ability (Figure 3.7c) and found stronger support for IBT. Meanwhile, we hypothesize that stronger support for HAH is expected under the opposite extreme of this continuum (Figure 3.7b), and to either IBT or HAH under intermediate scenarios (Figure 3.7a,d). This notion implies that most species responses to habitat fragmentation lie somewhere along these extremes. Hence, IBT and HAH should not be seen as a mutually exclusive dichotomy, but instead a continuum in explaining patterns of species retention in habitat patches. The conceptual framework we propose (Figure 3.7) also considers fragmented landscapes with dynamic (e.g. vegetation re-growth following land abandonment) or managed matrices (e.g. restored habitats following human intervention). In such landscapes, patch-centric patterns of occupancy (IBT) should gradually transit into those dominated by entire landscapes (HAH) given the role of matrix type in mediating SARs (Freeman, Olivier, & van Aarde, 2018). Conversely, matrix habitat degradation would revert the emphasis back to prime habitat patches. Although ameliorating the harshness of water matrices is virtually impossible, other hostile matrix habitats, such as bauxite mining (Kennedy & Marra, 2010), can be managed to enhance functional connectivity among habitat patches (fig. 7 in Villard & Metzger, 2014). Finally, we conclude that the most appropriate worldview in fragmentation ecology (IBT or HAH) is not only context-dependent but also dynamic. Therefore, the best conservation strategy—focusing on either the spatial arrangement of remaining habitat patches or the overall habitat amount in the landscape—is neither static nor can be generalised to a wide spectrum of landscape scenarios and taxonomic groups.

3.7   Acknowledgments

3.8   References

1. Alvares, C. A., Stape, J. L., Sentelhas, P. C., Gonçalves, J. L. M., & Sparovek, G. (2013). Köppen's climate classification map for Brazil. Meteorologische Zeitschrift, 22, 711– 728. https://doi.org/10.1127/0941-2948/2013/0507

2. Aurélio-Silva, M., Anciães, M., Henriques, L. M. P., Benchimol, M., & Peres, C. A. (2016). Patterns of local extinction in an Amazonian archipelagic avifauna following 25 years of insularization. Biological Conservation, 199, 101– 109. https://doi.org/10.1016/j.biocon.2016.03.016

3. Awade, M., & Metzger, J. P. (2008). Using gap-crossing capacity to evaluate functional connectivity of two Atlantic rainforest birds and their response to fragmentation. Austral Ecology, 33, 863– 871. https://doi.org/10.1111/j.1442-9993.2008.01857.x

4. Báldi, A. (2008). Habitat heterogeneity overrides the species–area relationship. Journal of Biogeography, 35, 675– 681. https://doi.org/10.1111/j.1365-2699.2007.01825.x

5. Benchimol, M., & Peres, C. A. (2015a). Edge-mediated compositional and functional decay of tree assemblages in Amazonian forest islands after 26 years of isolation. Journal of Ecology, 103, 408– 420. https://doi.org/10.1111/1365-2745.12371

6. Benchimol, M., & Peres, C. A. (2015b). Predicting local extinctions of Amazonian vertebrates in forest islands created by a mega dam. Biological Conservation, 187, 61– 72. https://doi.org/10.1016/j.biocon.2015.04.005

7. Benchimol, M., & Peres, C. A. (2015c). Widespread forest vertebrate extinctions induced by a mega hydroelectric dam in lowland Amazonia. PLoS ONE, 10, e0129818. https://doi.org/10.1371/journal.pone.0129818

8. BirdLife International. (2018). IUCN Red List for birds. Retrieved from https://www.birdlife.org

9. Blake, J. G., & Loiselle, B. A. (2001). Bird assemblages in second-growth and old-growth forests, Costa Rica: Perspectives from mist nets and point counts. The Auk, 118, 304– 326. https://doi.org/10.1642/0004-8038(2001)118[0304:BAISGA]2.0.CO;2

10. Castilho, C. V., Magnusson, W. E., Araújo, R. N. O., Luizão, R. C. C., Luizão, F. J., Lima, A. P., & Higuchi, N. (2006). Variation in aboveground tree live biomass in a central Amazonian forest: Effects of soil and topography. Forest Ecology and Management, 234, 85– 96. https://doi.org/10.1016/j.foreco.2006.06.024

11. Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., … Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. Ecography, 36, 27– 46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

12. Drucker, D. P., Costa, F. R. C., & Magnusson, W. E. (2008). How wide is the riparian zone of small streams in tropical forests? A test with terrestrial herbs. Journal of Tropical Ecology, 24, 65– 74. https://doi.org/10.1017/S0266467407004701

13. Evju, M., & Sverdrup-Thygeson, A. (2016). Spatial configuration matters: A test of the habitat amount hypothesis for plants in calcareous grasslands. Landscape Ecology, 31, 1891– 1902. https://doi.org/10.1007/s10980-016-0405-7

14. Fahrig, L. (2003). Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, Evolution, and Systematics, 34, 487– 515. https://doi.org/10.1146/annurev.ecolsys.34.011802.132419

15. Fahrig, L. (2013). Rethinking patch size and isolation effects: The habitat amount hypothesis. Journal of Biogeography, 40, 1649– 1663. https://doi.org/10.1111/jbi.12130

16. Fahrig, L. (2015). Just a hypothesis: A reply to Hanski. Journal of Biogeography, 42, 993– 994. https://doi.org/10.1111/jbi.12504

17. Fahrig, L. (2017). Ecological responses to habitat fragmentation *per se*. Annual Review of Ecology, Evolution, and Systematics, 48, 1– 23. https://doi.org/10.1146/annurev-ecolsys-110316-022612

18. Fearnside, P. M. (2016). Environmental and social impacts of hydroelectric dams in Brazilian Amazonia: Implications for the aluminum industry. World Development, 77, 48– 65. https://doi.org/10.1016/j.worlddev.2015.08.015

19. Freeman, M. T., Olivier, P. I., & Aarde, R. J. (2018). Matrix transformation alters species-area relationships in fragmented coastal forests. Landscape Ecology, 33, 307– 322. https://doi.org/10.1007/s10980-017-0604-x

20. Gavish, Y., Ziv, Y., & Rosenzweig, M. L. (2012). Decoupling fragmentation from habitat loss for spiders in patchy agricultural landscapes. Conservation Biology, 26, 150– 159. https://doi.org/10.1111/j.1523-1739.2011.01799.x

21. Haddad, N. M., Gonzalez, A., Brudvig, L. A., Burt, M. A., Levey, D. J., & Damschen, E. I. (2016). Experimental evidence does not support the Habitat Amount Hypothesis. Ecography, 125, 336– 342. https://doi.org/10.1111/ecog.02535

22. Haila, Y. (2002). A conceptual genealogy of fragmentation research: From island biogeography to landscape ecology. Ecological Applications, 12, 321– 334. https://doi.org/10.1890/1051-0761(2002)012[0321:ACGOFR]2.0.CO;2

23. Hanski, I. (2015). Habitat fragmentation and species richness. Journal of Biogeography, 42, 989– 993. https://doi.org/10.1111/jbi.12478

24. Jackson, H. B., & Fahrig, L. (2012). What size is a biologically relevant landscape? Landscape Ecology, 27, 929– 941. https://doi.org/10.1007/s10980-012-9757-9

25. Jackson, H. B., & Fahrig, L. (2015). Are ecologists conducting research at the optimal scale? Global Ecology and Biogeography, 24, 52– 63. https://doi.org/10.1111/geb.12233

26. Jung, M. (2016). LecoS – A python plugin for automated landscape ecology analysis. Ecological Informatics, 31, 18– 21. https://doi.org/10.1016/j.ecoinf.2015.11.006

27. Karr, J. R. (1981). Surveying birds in the tropics. Studies in Avian Biology, 6, 548– 553. https://sora.unm.edu/node/139165

28. Kennedy, C. M., & Marra, P. P. (2010). Matrix mediates avian movements in tropical forested landscapes: Inference from experimental translocations. Biological Conservation, 143, 2136– 2145. https://doi.org/10.1016/j.biocon.2010.05.025

29. Laurance, W. F. (2008). Theory meets reality: How habitat fragmentation research has transcended island biogeographic theory. Biological Conservation, 141, 1731– 1744. https://doi.org/10.1016/j.biocon.2008.05.011

30. Laurance, S. G. W., Stouffer, P. C., & Laurance, W. F. (2004). Effects of road clearings on movement patterns of understory rainforest birds in central Amazonia. Conservation Biology, 18, 1099– 1109. https://doi.org/10.1111/j.1523-1739.2004.00268.x

31. Lees, A. C., & Peres, C. A. (2009). Gap-crossing movements predict species occupancy in Amazonian forest fragments. Oikos, 118, 280– 290. https://doi.org/10.1111/j.1600-0706.2008.16842.x

32. Legendre, P., & Legendre, L. (1998). Numerical ecology. Amsterdam, The Netherlands: Elsevier Science.

33. Lima, J. R., Galatti, U., Lima, C. J., Fáveri, S. B., Vasconcelos, H. L., & Neckel-Oliveira, S. (2015). Amphibians on Amazonian land-bridge islands are affected more by area than isolation. Biotropica, 47, 369– 376. https://doi.org/10.1111/btp.12205

34. Lomolino, V. M. & Weiser, M. D. (2001). Towards a more general species-area relationship: Diversity on all islands, great and small. Journal of Biogeography, 28, 431– 445. https://doi.org/10.1046/j.1365-2699.2001.00550.x

35. Lövei, G. L., Magura, T., Tóthmérész, B., & Ködöböcz, V. (2006). The influence of matrix and edges on species richness patterns of ground beetles (Coleoptera: Carabidae) in habitat islands. Global Ecology and Biogeography, 15, 283– 289. https://doi.org/10.1111/j.1466-8238.2005.00221.x

36. MacArthur, R. H., & Wilson, E. O. (1967). The theory of island biogeography. Princeton, NJ: Princeton University Press.

37. MacDonald, Z. G., Anderson, I. D., Acorn, J. H., & Nielsen, S. E. (2018). Decoupling habitat fragmentation from habitat loss: Butterfly species mobility obscures fragmentation effects in a naturally fragmented landscape of lake islands. Oecologia, 186, 11– 27. https://doi.org/10.1007/s00442-017-4005-2

38. Matthews, T. J., Guilhaumon, F., Triantis, K. A., Borregaard, M. K., & Whittaker, R. J. (2016). On the form of species-area relationships in habitat islands and true islands. Global Ecology and Biogeography, 25, 847– 858. https://doi.org/10.1111/geb.12269

39. Melo, G. L., Sponchiado, J., Cáceres, N. C., & Fahrig, L. (2017). Testing the habitat amount hypothesis for South American small mammals. Biological Conservation, 209, 304– 314. https://doi.org/10.1016/j.biocon.2017.02.031

40. Moore, R. P., Robinson, W. D., Lovette, I. J., & Robinson, T. R. (2008). Experimental evidence for extreme dispersal limitation in tropical forest birds. Ecology Letters, 11, 960– 968. https://doi.org/10.1111/j.1461-0248.2008.01196.x

41. 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. https://doi.org/10.1371/journal.pone.0128923

42. Munguía-Rosas, M. A., & Montiel, S. (2014). Patch size and isolation predict plant species density in a naturally fragmented forest. PLoS ONE, 9, e111742. https://doi.org/10.1371/journal.pone.0111742

43. Palmeirim, A. F., Vieira, M. V., & Peres, C. A. (2017). Non-random lizard extinctions in land-bridge Amazonian forest islands after 28 years of isolation. Biological Conservation, 214, 55– 65. https://doi.org/10.1016/j.biocon.2017.08.002

44. Powell, L. L., Cordeiro, N. J., & Stratford, J. A. (2015). Ecology and conservation of avian insectivores of the rainforest understory: A pantropical perspective. Biological Conservation, 188, 1– 10. https://doi.org/10.1016/j.biocon.2015.03.025

45. Powell, L. L., Stouffer, P. C., & Johnson, E. I. (2013). Recovery of understory bird movement across the interface of primary and secondary Amazon rainforest. The Auk, 130, 459– 468. https://doi.org/10.1525/auk.2013.12202

46. Prevedello, J. A., & Vieira, M. V. (2010). Does the type of matrix matter? A quantitative review of the evidence. Biodiversity and Conservation, 19, 1205– 1223. https://doi.org/10.1007/s10531-009-9750-z

47. Prugh, L. R., Hodges, K. E., Sinclair, A. R. E., & Brashares, J. S. (2008). Effect of habitat area and isolation on fragmented animal populations.

Proceedings of the National Academy of Sciences, 105, 20770– 20775. https://doi.org/10.1073/pnas.0806080105

48. QGIS Development Team. (2016). QGIS Geographic Information System. Version 2.14.22. Open Source Geospatial Foundation Project. Retrieved from https://www.qgis.org

49. Quinn, J. F., & Harrison, S. P. (1988). Effects of habitat fragmentation and isolation on species richness: Evidence from biogeographic patterns. Oecologia, 75, 132– 140. https://doi.org/10.1007/BF00378826

50. Rabelo, R. M., Bicca-Marques, J. C., Aragón, S., & Nelson, B. W. (2017). Are fluvial islands "real" islands for arboreal mammals? Uncovering the effect of patch size under the species-area relationship. Journal of Biogeography, 44, 1802– 1812. https://doi.org/10.1111/jbi.13034

51. Rosenzweig, M. L. (1995). Species diversity in space and time. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511623387

52. Santos, A. M. C., Whittaker, R. J., Triantis, K. A., Borges, P. A. V., Jones, O. R., Quicke, D. L. J., & Hortal, J. (2010). Are species-area relationships from entire archipelagos congruent with those of their constituent islands? Global Ecology and Biogeography, 19, 527– 540. https://doi.org/10.1111/j.1466-8238.2010.00536.x

53. Scheiner, S. M. (2003). Six types of species-area curves. Global Ecology and Biogeography, 12, 441– 447. https://doi.org/10.1046/j.1466-822X.2003.00061.x

54. Schietti, J., Emilio, T., Rennó, C. D., Drucker, D. P., Costa, F. R. C., Nogueira, A., … Magnusson, W. E. (2014). Vertical distance from drainage drives floristic composition changes in an Amazonian rainforest. Plant Ecology & Diversity, 7, 241– 253. https://doi.org/10.1080/17550874.2013.783642

55. Seibold, S., Bässler, C., Brandl, R., Fahrig, L., Förster, B., Heurich, M., … Müller, J. (2017). An experimental test of the habitat-amount hypothesis for saproxylic beetles in a forested region. Ecology, 98, 1613– 1622. https://doi.org/10.1002/ecy.1819

56. Şekercioğlu, Ç. H., Ehrlich, P. R., Daily, G. C., Aygen, D., Goehring, D., & Sandí, R. F. (2002). Disappearance of insectivorous birds from tropical forest fragments. Proceedings of the National Academy of Sciences, 99, 263– 267. https://doi.org/10.1073/pnas.012616199

57. Storck-Tonon, D., & Peres, C. A. (2017). Forest patch isolation drives local extinctions of Amazonian orchid bees in a 26 years old archipelago. Biological Conservation, 214, 270– 277. https://doi.org/10.1016/j.biocon.2017.07.018

58. Stouffer, P. C., Bierregaard, R. O., Strong, C., & Lovejoy, T. E. (2006). Long-term landscape change and bird abundance in Amazonian rainforest fragments. Conservation Biology, 20, 1212– 1223. https://doi.org/10.1111/j.1523-1739.2006.00427.x

59. Stouffer, P. C., Johnson, E. I., Bierregaard, R. O., & 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. https://doi.org/10.1371/journal.pone.0020543

60. Torrenta, R., & Villard, M.-A. (2017). A test of the habitat amount hypothesis as an explanation for the species richness of forest bird assemblages. Journal of Biogeography, 44, 1791– 1801. https://doi.org/10.1111/jbi.13022

61. Villard, M.-A., & Metzger, J. P. (2014). Beyond the fragmentation debate: A conceptual model to predict when habitat configuration really matters. Journal of Applied Ecology, 51, 309– 318. https://doi.org/10.1111/1365-2664.12190

62. Walter, S. T., Browne, L., Freile, J., Olivo, J., González, M., & Karubian, J. (2017). Landscape-level tree cover predicts species richness of large-bodied frugivorous birds in forest fragments. Biotropica, 49, 838– 847. https://doi.org/10.1111/btp.12469

63. Wiens, J. A. (2008). Habitat fragmentation: Island v landscape perspectives on bird conservation. Ibis, 137, S97– S104. https://doi.org/10.1111/j.1474-919X.1995.tb08464.x

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

65. Yaacobi, G., Ziv, Y., & Rosenzweig, M. L. (2007). Habitat fragmentation may not matter to species diversity. Proceedings of the Royal Society B: Biological Sciences, 274, 2409– 2412. https://doi.org/10.1098/rspb.2007.0674

3.9 Supporting Information

This supporting information contains:

- Table S3.1

- Table S3.2

- Figure S3.1

- Figure S3.2

- Figure S3.3

- R code

    Click here to download the R code in .Rmd file format

Table S3.1: Forest insectivorous bird species captured on the understorey of 33 forest islands. 'Islands' corresponds to the number of forest islands at which species were captured. 'Individuals' corresponds to the number of individuals captured. Taxonomy follows Piacentini et al. (2015).

Family	Species	Islands	Individuals
Cuculidae	*Coccycua minuta*	1	1
Trogonidae	*Trogon rufus*	2	2
Galbulidae	*Galbula albirostris*	7	19
Bucconidae	*Bucco tamatia*	1	1
	Monasa atra	7	11
Picidae	*Celeus elegans*	11	13
	Campephilus rubricollis	1	1
Thamnophilidae	*Epinecrophylla gutturalis*	1	1
	Myrmotherula axillaris	23	186
	Myrmotherula menetriesii	1	1
	Isleria guttata	13	44
	Thamnomanes ardesiacus	3	7
	Thamnomanes caesius	3	5
	Thamnophilus murinus	7	9
	Thamnophilus punctatus	1	2
	Myrmoderus ferrugineus	6	7
	Hylophylax naevius	1	1
	Percnostola rufifrons	16	54
	Cercomacroides tyrannina	3	10
	Hypocnemis cantator	14	72
	Pithys albifrons	1	2
	Willisornis poecilinotus	9	30
Conopophagidae	*Conopophaga aurita*	1	2
Formicariidae	*Formicarius colma*	4	4
	Formicarius analis	2	2
Scleruridae	*Sclerurus rufigularis*	3	3
Dendrocolaptidae	*Dendrocincla fuliginosa*	17	45
	Sittasomus griseicapillus	2	2
	Certhiasomus stictolaemus	3	8
	Glyphorynchus spirurus	16	84
	Xiphorhynchus pardalotus	21	73
	Xiphorhynchus obsoletus	5	8
	Nasica longirostris	1	1
	Dendrocolaptes certhia	19	39
	Hylexetastes perrotii	1	1
Xenopidae	*Xenops minutus*	4	5

Family	Species	Islands	Individuals
Furnariidae	*Automolus ochrolaemus*	1	1
	Automolus infuscatus	1	1
	Philydor erythrocercum	3	4
	Philydor pyrrhodes	2	3
	Synallaxis rutilans	1	3
Onychorhynchidae	*Onychorhynchus coronatus*	2	6
	Terenotriccus erythrurus	3	7
	Myiobius barbatus	4	8
Platyrinchidae	*Platyrinchus saturatus*	1	1
	Platyrinchus coronatus	2	5
Rhynchocyclidae	*Mionectes macconnelli*	6	12
	Rhynchocyclus olivaceus	1	1
	Tolmomyias poliocephalus	2	2
	Lophotriccus vitiosus	1	1
	Lophotriccus galeatus	3	3
Tyrannidae	*Ramphotrigon ruficauda*	2	2
	Myiarchus tuberculifer	1	1
	Rhytipterna simplex	3	3
Troglodytidae	*Pheugopedius coraya*	1	2
	Henicorhina leucosticta	2	8
Turdidae	*Turdus albicollis*	15	39
Thraupidae	*Saltator maximus*	1	2
	Saltator grossus	1	3

Reference

Piacentini, V. de Q., Aleixo, A., Agne, C.E., et al. (2015) Annotated checklist of the birds of Brazil by the Brazilian Ornithological Records Committee. Revista Brasileira de Ornitologia, 23, 91–298.

Table S3.2: Characteristics of the sample sites surveyed on 33 forest islands within the Balbina Hydroelectric Reservoir in central Brazilian Amazonia. 'Habitat amount' is defined as the percentage of forest cover within 200-m buffer around sample sites (i.e. mist-net lines).

Island name	Latitude (S)	Longitude (W)	Island size (ha)	Habitat amount (%)	Richness (n)	Abundance (n)
Joaninha	1°31'22"	59°49'44"	0.63	3.50	0	0
Xibé	1°28'08"	59°50'11"	0.90	14.69	1	1
Formiga	1°50'00"	59°25'16"	1.54	8.17	0	0
André	1°35'05"	59°52'19"	2.08	11.17	2	7
Cafundó	1°30'02"	59°36'37"	2.62	43.23	3	4
Panema	1°46'28"	59°41'33"	3.08	17.74	0	0
Torem	1°49'46"	59°37'57"	3.62	17.78	0	0
Pé Torto	1°45'59"	59°21'49"	4.98	22.44	4	17
Jiquitaia	1°50'08"	59°35'44"	6.79	29.30	2	4
Arrepiado	1°30'54"	59°44'21"	7.42	29.08	2	4
Garrafa	1°35'19"	59°50'08"	9.41	43.08	3	7
Piquiá	1°30'23"	59°47'20"	13.03	48.35	12	16
Abusado	1°45'44"	59°40'43"	13.30	54.04	3	26
Coata	1°29'18"	59°47'13"	16.19	48.72	10	18
Palhal	1°47'25"	59°26'52"	21.35	74.70	6	30
Neto	1°50'29"	59°21'08"	32.84	71.92	6	33
Adeus	1°52'26"	59°39'45"	49.40	53.44	10	21
Bacaba	1°30'15"	59°49'20"	52.38	74.10	4	9
Relógio	1°40'55"	59°39'09"	70.48	92.80	9	47
Sapupara	1°41'50"	59°36'45"	77.71	84.96	8	18
Moitá	1°33'36"	59°53'33"	91.19	78.74	5	12
Pontal	1°49'52"	59°41'12"	106.48	85.21	16	37
Furo de Santa Luzia	1°44'25"	59°26'31"	198.31	93.49	22	64
Cipoal	1°41'54"	59°47'05"	217.40	93.10	9	33
Jabuti	1°37'34"	59°45'24"	232.24	90.80	11	36
Tucumari	1°35'29"	59°25'47"	329.68	91.47	17	45
Martelo	1°39'51"	59°42'51"	460.04	84.76	7	19
Tristeza	1°46'03"	59°45'16"	497.86	95.99	19	61
Beco do Catitu	1°44'22"	59°42'18"	631.85	97.18	16	54
Mascote	1°38'56"	59°49'58"	663.06	100.00	18	46
Fuzaca	1°29'33"	59°51'36"	934.29	100.00	27	84
Porto Seguro	1°46'48"	59°31'02"	1323.13	100.00	21	74
Gavião-real	1°35'46"	59°37'58"	1678.96	100.00	17	47
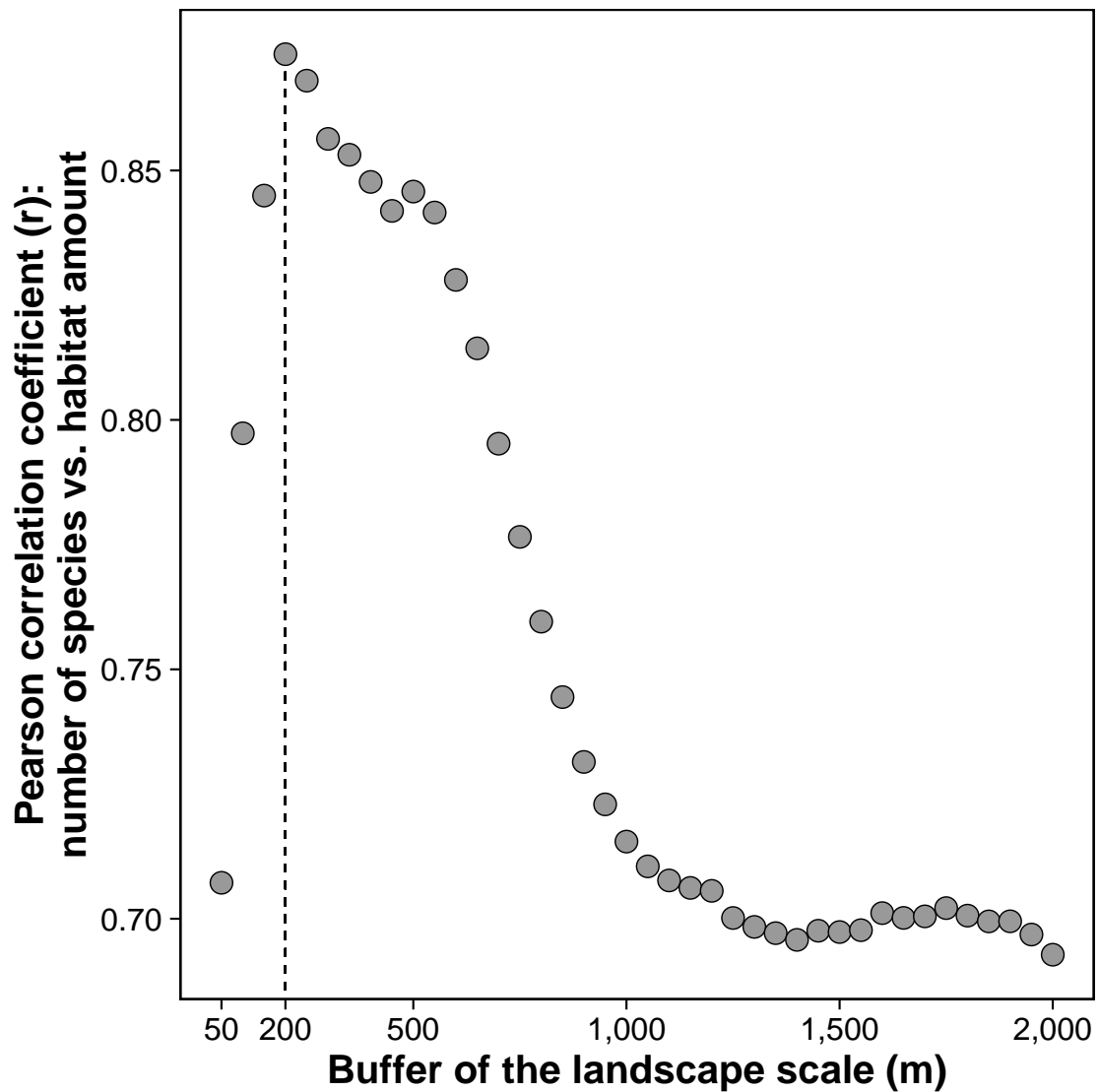
Figure S3.1: Correlation between number of forest insectivorous bird species and habitat amount (i.e. percentage of forest cover) across 40 landscape scales, ranging from 50 to 2,000-m buffer at 50-m intervals. The peak of correlation value (dashed line) corresponds to the size of the local landscape (i.e. the scale of effect).
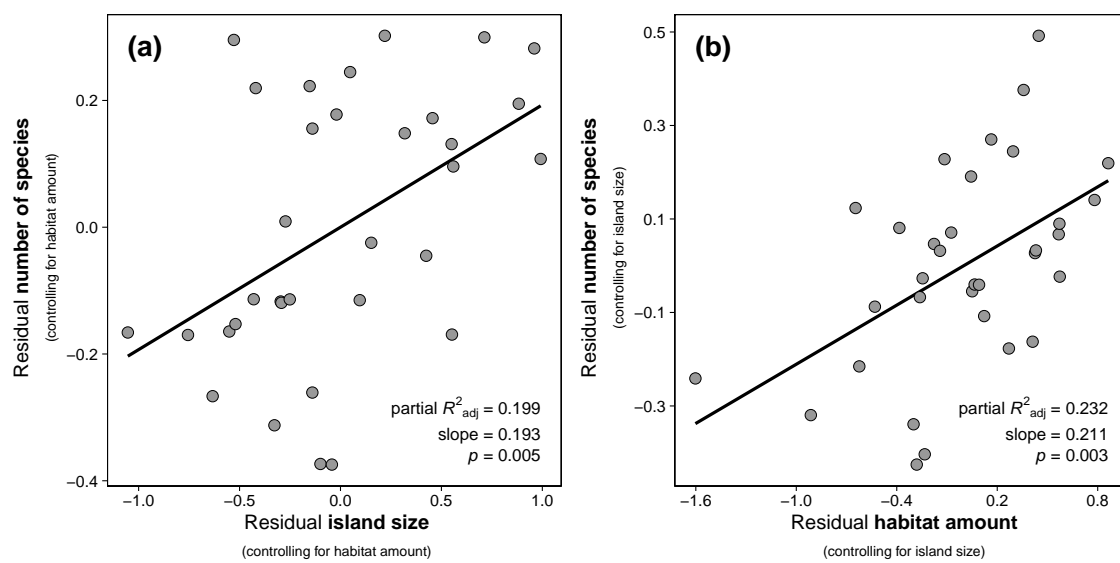
Figure S3.2: Partial regression plots illustrating the independent effects of (a) island size and (b) habitat amount on the number of forest insectivorous bird species occurring at 33 forest islands within the Balbina Hydroelectric Reservoir in central Brazilian Amazonia.
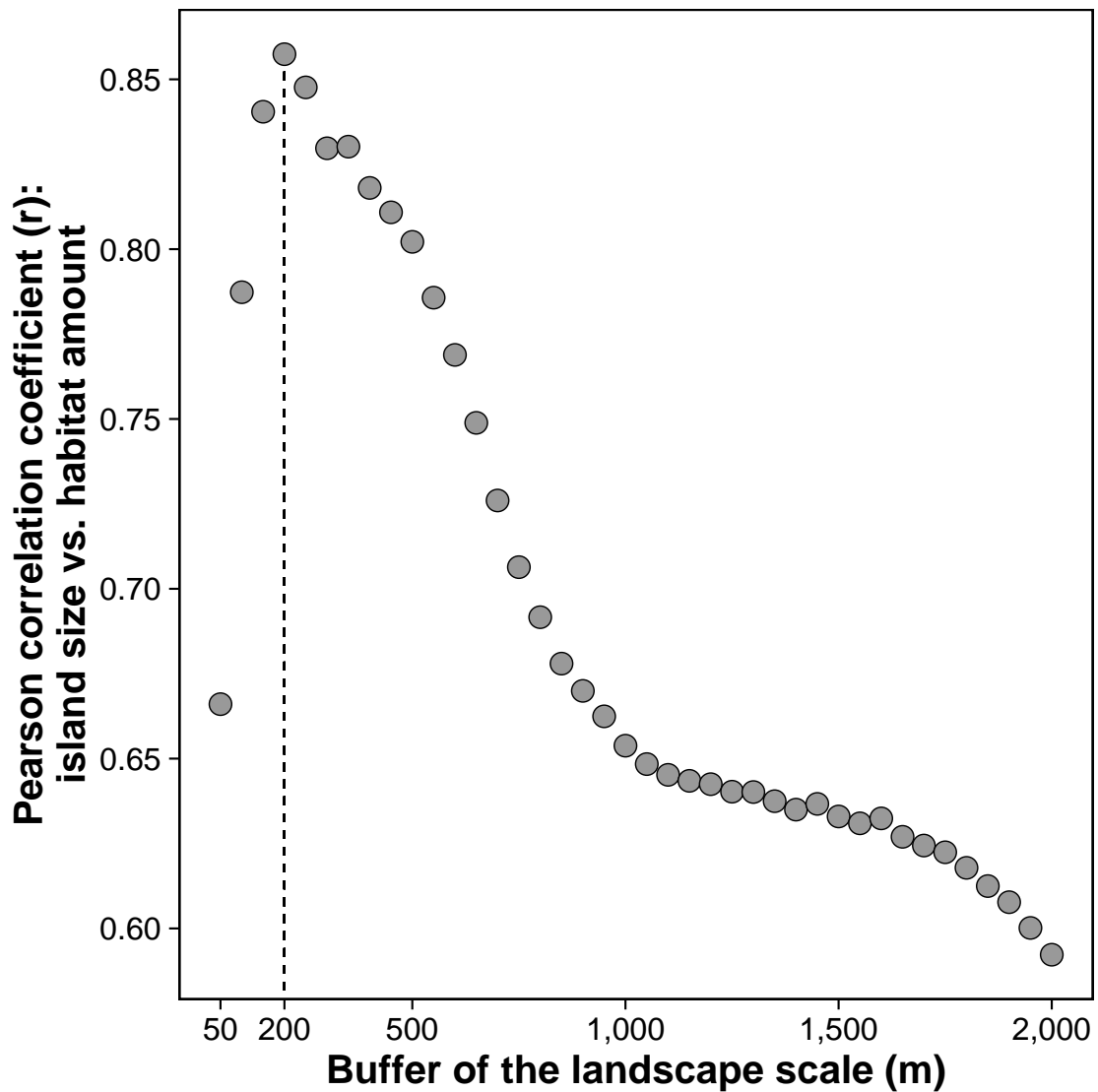
Figure S3.3: Correlation between island size and habitat amount (i.e. percentage of forest cover) across 40 landscape scales, ranging from 50 to 2,000-m buffer at 50-m intervals. Note that the maximum correlation value (dashed line) between island size and habitat amount occurs exactly at the scale of effect (200-m buffer; see Fig. S1).

Supporting Information

Patch-scale biodiversity retention in fragmented landscapes: Reconciling the habitat amount hypothesis with the island biogeography theory

Preamble

Any given analytical framework is at the heart of modern ecological studies, yet this is rarely presented in detail as supplementary guidelines in the interest of transparent reproducibility should anyone wish to replicate parts or the entirety of the analyses. This extended section was constructed with the explicit intent of walking interested students and future investigators through every sequential step of our analytical strategy to test the Habitat Amount Hypothesis – HAH (Fahrig, 2013). The section contains step-by-step R codes which may be used to generate all inferential statistics and supporting elements presented in this paper, herein disclosed as Supporting Information. We do this to both make every analytical procedure of this paper crystal-clear and to leave no ambiguity as to how our results were derived. We hope this is useful to our readership and can be put to use.

```r
# Clear workspace and disable scientific notation
remove(list = ls()); options(scipen = 999)
```

Packages

```r
# Load required packages
library(ggplot2)
library(vegan)
```

Dataset

We provide the raw data derived from original fieldwork conducted during the project **Ecological Impacts of River Damming on Forest Bird Assemblages in the Brazilian Amazon**. Find out more at the project webpage.

Data are available from the KNB repository and should be cited as:

Anderson Saldanha Bueno. 2018. Balbina Understory Bird Data from 2015 to 2016. Knowledge Network for Biocomplexity. urn:uuid:dbfd1504-2212-422c-8e04-610fb2327b7c.

Import data

- The file **balbina_understory_birds_captures.csv** contains information on birds captured in the fieldwork.
- The file **balbina_understory_birds_taxonomy_traits.csv** contains the taxonomy and traits of bird species captured.
- The file **balbina_environmental_variables.csv** contains environmental variables of sample sites.

```r
# Load the file "balbina_understorey_birds_captures.csv"
birds = read.csv("https://ndownloader.figshare.com/files/15158531")

# Load the file "balbina_understorey_birds_taxonomy_traits.csv"
traits = read.csv("https://ndownloader.figshare.com/files/15158534")

# Load the file "balbina_environmental_variables.csv"
env = read.csv("https://ndownloader.figshare.com/files/15158528", row.names = 1)
```

Data handling

The raw data are stored in the long table format and include all individuals captured. This means that each bird captured is represented by a row. However, the analysis performed in this study requires the short table format, where sample sites are in rows and species are in columns. Furthermore, here we focus only on **forest insectivorous birds** (referred to in the main text as **understorey insectivorous birds**).

To obtain the table we need, we first add to each row of the raw data (object *birds*) species traits (from the object *traits*) on **forest dependency**, **diet**, and **foraging time**, which will be used to select diurnal **forest insectivorous birds**.

```r
# Remove recaptures
birds = subset(birds, birds$new.individual != "no")

# Add a column with the forest dependency of each species
birds$forest.dependency = traits$forest.dependency[match(birds$species,
                                                    traits$cbro.2015)]

# Add a column with the diet of each species
birds$diet = traits$diet.5cat[match(birds$species, traits$cbro.2015)]

# Add a column with the foraging time of each species
birds$nocturnal = traits$nocturnal[match(birds$species, traits$cbro.2015)]

# Add a column indicating that each bird captured (row) corresponds to
# an individual record
birds$occurrence = 1
```

The next step is to produce the table in the short format containing both bird and environmental data.

```r
# Site-by-species abundance matrix
birds.matrix = tapply(birds$occurrence, list(birds$site, birds$species), sum)
# Species non detected in a given site are real zeros
birds.matrix[is.na(birds.matrix)] = 0

# Site-by-species abundance matrix with environmental data Add environmental
# data to the site-by-species abundance matrix
birds.env = cbind(birds.matrix, env)
```

We now make a species list of **forest insectivorous birds**.

```r
# Species list of forest insectivorous birds
ins = subset(traits,
            # Select only forest species
            c(traits$forest.dependency == "Medium" |
                traits$forest.dependency == "High") &
              # Select only insectivorous species
              traits$diet.5cat == "Invertebrate" &
              # Remove nocturnal species
              traits$nocturnal == 0)$cbro.2015
```

Finally, we produce a table containing data from all 33 surveyed islands, and omit data from continuous forest sites because they were not used in this study.

```r
# Island-by-species abundance matrix with environmental data Select only
# forest insectivorous birds captured on islands
islands.spp = birds.env[-c(8:12), ins][colSums(birds.env[-c(8:12), ins]) > 0]
# Select environmental data for islands
islands.env = birds.env[-c(8:12), c(131:ncol(birds.env))]
```

```
# Combine bird with environmental data
islands = cbind(islands.spp, islands.env)
```

Table S1

```
# Species frequency (number of islands) and abundance (number of individuals)
table.s1 = data.frame(Islands = colSums(ifelse(islands[1:59] > 0, 1, 0)),
                      Individuals = colSums(islands[1:59]))

# Species taxonomy
# Position of the species in the taxonomic sequence
table.s1$ID = traits$id[match(rownames(table.s1), traits$cbro.2015)]
# Species name
table.s1$Species = rownames(table.s1)
# Species family
table.s1$Family = traits$family[match(rownames(table.s1), traits$cbro.2015)]

# Rearrange the table
table.s1 = table.s1[order(table.s1$ID), c(5, 4, 1, 2)]
row.names(table.s1) = NULL

#table.s1
```

TABLE S1 Forest insectivorous bird species captured on the understorey of 33 forest islands within the Balbina Hydroelectric Reservoir in central Brazilian Amazonia. 'Islands' corresponds to the number of forest islands at which species were captured. 'Individuals' corresponds to the number of individuals captured. Taxonomy follows Piacentini et al. (2015).

Variables

Response variable

We used the **number of species** (also referred to as **species richness**) of **forest insectivorous birds** as the response variable.

```
# Calculate species richness of forest insectivorous birds
islands$richness = rowSums(ifelse(islands.spp > 0, 1, 0))
```

Scale of effect

We employed a multi-scale analysis to determine the 'scale of effect' - the landscape scale at which the relationship between the **number of species** and **habitat amount** peaks. We defined the scale of effect as the 'local landscape' for **forest insectivorous birds** at the Balbina forest archipelago. Our multi-scale analysis examined 40 different buffer sizes around sample sites (i.e. mist-net lines), ranging from 50 to 2,000 m at 50-m intervals.

```
# Pearson's product-moment correlation between number of species and habitat amount
# across 40 landscape scales
buffer = apply(islands[63:102], 2, function(z)
  cor.test(log10(z + 1), log10(islands$richness + 1)))

# Convert the results stored as a 'list' object into a 'matrix' object
buffer = do.call(rbind, lapply(buffer, function(z) z[4]))
```

```r
# Save the results in a 'data frame' object
buffer = data.frame(buffer = seq(50, 2000, 50), r = as.numeric(buffer))

# Scale of effect: 200-m buffer around mist-net lines
buffer[order(-buffer$r), ][1, ]
```

```
##    buffer        r
## 4     200 0.873322
```

Figure S1

```r
# Graph of the correlation between number of species and habitat amount
# across 40 landscape scales
fig.s1 =
  ggplot(data = buffer,
         aes(x = buffer, y = r)) +

  labs(x = "Buffer of the landscape scale (m)",
       y = "Pearson correlation coefficient (r):\nnumber of species vs. habitat amount") +

  scale_x_continuous(breaks = c(50, 200, 500, 1000, 1500, 2000),
                     labels = c("50", "200", "500", "1,000", "1,500", "2,000")) +

  geom_segment(x = 200, xend = 200, y = 0, yend = max(buffer$r), linetype = "dashed") +
  geom_point(shape = 21, colour = "black", fill = "#999999", size = 5) +

  theme_bw(base_size = 20) +
  theme(panel.grid = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.title = element_text(colour = "black", face = "bold"),
        axis.text = element_text(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5))

#fig.s1
```

FIGURE S1 Correlation between number of forest insectivorous bird species and habitat amount (i.e. percentage of forest cover) across 40 landscape scales, ranging from 50 to 2,000-m buffer at 50-m intervals. The peak of correlation value (dashed line) corresponds to the size of the local landscape (i.e. the scale of effect).

Predictor variables

We used **island size** (ha) and **habitat amount** (%) within 200-m buffer local landscapes as predictor variables. In addition, local landscapes were classified as surrounded by either 'low' (< 55%) or 'high' habitat amount (> 70%).

Relevant variables

For simplicity, we create a table with the relevant variables to perform the analyses. Geographic coordinates and the number of individuals (i.e. species abundance) are included for informative purpose.

```r
# Response and predictor variables used to perform the tests of the HAH
variables = data.frame(row.names = rownames(islands),
                       # Island size (ha)
```

```r
                        island.size = islands$area.ha,
                        # Habitat amount (%)
                        habitat.amount = islands$forest.cover.200,
                        # Habitat amount category
                        ha.cat.55 = ifelse(islands$forest.cover.200 < 55, "low", "high"),
                      # Number of species
                       richness = islands$richness)

# Order rows by island size
variables = variables[order(variables$island.size), ]
```

Table S2

```r
# Information to include in the table
table.s2 = data.frame("Island name" = rownames(islands),
                        # Y coordinate (datum WGS 84)
                        "Latitude" = islands$latitude.WGS84,
                        # X coordinate (datum WGS 84)
                        "Longitude" = islands$longitude.WGS84,
                        # Forest area in hectares
                        "Island size" = islands$area.ha,
                        # Forest cover within 200-m buffer (%)
                        "Habitat amount" = islands$forest.cover.200,
                        # Number of species captured (n)
                        "Richness" = islands$richness,
                        # Number of individuals captured (n)
                        "Abundance" = rowSums(islands[1:59]))

# Order rows by island size
table.s2 = table.s2[order(table.s2$Island.size), ]
row.names(table.s2) = NULL

#table.s2
```

TABLE S2 Characteristics of the sample sites surveyed at 33 forest islands within the Balbina Hydroelectric Reservoir in central Brazilian Amazonia.

IBT vs. HAH

We applied four tests to determine whether the **island biogeography theory (IBT)** or the **habitat amount hypothesis (HAH)** is the most appropriate theoretical framework to explain the number of forest insectivorous bird species on forest islands within the Balbina Hydroelectric Reservoir.

Test 1: Multiple linear regression

We used multiple linear regression analysis to examine the independent effects of **island size** and **habitat amount** within the local landscape on **species richness**. Prior to analysis, the response and predictor variables were log-transformed (log10 x + 1). The predictor variables were also standardised (mean = 0, SD = 1) to allow comparison of regression slopes (i.e. beta coefficients).

```
# To shorten the codes, we log-transform and standardise the variables
# accordingly and save them as 'vector' objects Species richness (sr),
# log-transformed
sr = log10(variables$richness + 1)
# Island size (is), log-transformed and standardised
is = scale(log10(variables$island.size + 1))
# Habitat amount (ha), log-transformed and standardised
ha = scale(log10(variables$habitat.amount + 1))
```

```
# Multiple linear regression model
mod = lm(sr ~ is + ha)
summary(mod)
```

```
##
## Call:
## lm(formula = sr ~ is + ha)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.36621 -0.08383 -0.02406  0.09675  0.39747
##
## Coefficients:
##              Estimate Std. Error t value            Pr(>|t|)
## (Intercept)   0.82547    0.03325  24.825 < 0.0000000000000002 ***
## is            0.19311    0.06561   2.943             0.00622 **
## ha            0.21082    0.06561   3.213             0.00313 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.191 on 30 degrees of freedom
## Multiple R-squared:  0.8159, Adjusted R-squared:  0.8036
## F-statistic: 66.46 on 2 and 30 DF,  p-value: 0.000000000009492
```

Partial regression plots

To visualise the results of the multiple linear regression model, we can produce partial regression plots to show the scatter of data points around the partial lines and their respective coefficients.

The next chunks of code describe how to produce partial regression plots, following Moya-Laraño & Corcobado (2008).

Partial regression between the number of species and island size

```
# 1) To calculate the residuals of the regression between number of species
# and habitat amount Number of species controlling for habitat amount
res.sr.ha = residuals(lm(sr ~ ha))

# 2) To calculate the residuals of the regression between island size and
# habitat amount Island size controlling for habitat amount
res.is.ha = residuals(lm(is ~ ha))

# 3) To regress the residuals from step 1 against the residuals from step 2
# to estimate the effect of island size on the number of species,
# controlling for the effect of habitat amount
```

```r
mod.sr.is = lm(res.sr.ha ~ res.is.ha)
summary(mod.sr.is)
```

```
##
## Call:
## lm(formula = res.sr.ha ~ res.is.ha)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.36621 -0.08383 -0.02406  0.09675  0.39747
##
## Coefficients:
##                             Estimate             Std. Error t value
## (Intercept) -0.000000000000000003732   0.032710467700918323397   0.000
## res.is.ha    0.193107547142374991189   0.064547803391814984053   2.992
##               Pr(>|t|)
## (Intercept)    1.0000
## res.is.ha      0.0054 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1879 on 31 degrees of freedom
## Multiple R-squared:  0.224,  Adjusted R-squared:  0.199
## F-statistic:  8.95 on 1 and 31 DF,  p-value: 0.005401
```

```r
# Coefficients of the partial regression between number of species and
# habitat amount (step 3) Adjusted R-squared = 0.199
round(summary(mod.sr.is)$adj.r.squared, 3)
```

```
## [1] 0.199
```

```r
# Slope = 0.193
round(summary(mod.sr.is)$coefficients[2, 1], 3)
```

```
## [1] 0.193
```

```r
# P-value = 0.005
round(summary(mod.sr.is)$coefficients[2, 4], 3)
```

```
## [1] 0.005
```

Partial regression between number of species and habitat amount

```r
# 4) To calculate the residuals of the regression between the number of
# species and island size Number of species controlling for island size
res.sr.is = residuals(lm(sr ~ is))

# 5) To calculate the residuals of the regression between habitat amount and
# island size Habitat amount controlling for island size
res.ha.is = residuals(lm(ha ~ is))

# 6) To regress the residuals from step 4 against the residuals from step 5
# to estimate the effect of habitat amount on the number of species,
# controlling for the effect of island size
mod.sr.ha = lm(res.sr.is ~ res.ha.is)
summary(mod.sr.ha)
```

```
##
## Call:
## lm(formula = res.sr.is ~ res.ha.is)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.36621 -0.08383 -0.02406  0.09675  0.39747
##
## Coefficients:
##                              Estimate             Std. Error t value
## (Intercept) -0.00000000000000002567  0.03271046770091832340   0.000
## res.ha.is    0.21082306275781120819  0.06454780339181498405   3.266
##                      Pr(>|t|)
## (Intercept)   1.00000
## res.ha.is     0.00266 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1879 on 31 degrees of freedom
## Multiple R-squared:  0.256,  Adjusted R-squared:  0.232
## F-statistic: 10.67 on 1 and 31 DF,  p-value: 0.002663
```

```
# Coefficients of the partial regression between number of species and
# habitat amount (step 6) Adjusted R-squared = 0.232
round(summary(mod.sr.ha)$adj.r.squared, 3)
```

```
## [1] 0.232
```

```
# Slope = 0.211
round(summary(mod.sr.ha)$coefficients[2, 1], 3)
```

```
## [1] 0.211
```

```
# P-value = 0.003
round(summary(mod.sr.ha)$coefficients[2, 4], 3)
```

```
## [1] 0.003
```

Figure S2

```
# Data to produce the plots
partials = data.frame(res.sr.ha, res.is.ha, res.sr.is, res.ha.is)

# Partial regression plot: number of species ~ island size
fig.s2a =
    ggplot(data = partials[order(partials$res.is.ha), ],
           aes(x = res.is.ha, y = res.sr.ha)) +

    labs(x = ~ atop(paste("Residual ", bold("island size")),
                    paste(scriptstyle("(controlling for habitat amount)"))),
         y = ~ atop(paste("Residual ", bold("number of species")),
                    paste(scriptstyle("(controlling for habitat amount)")))) +

    geom_smooth(method = "lm", se = FALSE, colour = "black", size = 1.5) +
    geom_point(shape = 21, colour = "black", fill = "#999999", size = 5) +
```

```r
    theme_bw(base_size = 20) +
    theme(panel.grid = element_blank(),
          panel.border = element_rect(colour = "black"),
          axis.title = element_text(colour = "black", face = "bold"),
          axis.text = element_text(colour = "black"),
          axis.ticks = element_line(colour = "black", size = 0.5)) +

  annotate("text", x = min(partials$res.is.ha), y = max(partials$res.sr.ha),
           hjust = 0, vjust = 1, fontface = "bold", size = 10, label = "(a)") +
  annotate("text", x = max(partials$res.is.ha), y = min(partials$res.sr.ha),
             hjust = 1, vjust = -2, size = 6,
             parse = T,
             label = as.character(expression(partial~italic(R)^{2}*""[adj]*" = 0.199"))) +
  annotate("text", x = max(partials$res.is.ha), y = min(partials$res.sr.ha),
           hjust = 1, vjust = -2, size = 6,  label = "slope = 0.193") +
  annotate("text", x = max(partials$res.is.ha), y = min(partials$res.sr.ha),
           hjust = 1, vjust = 0,  size = 6,
           parse = T, label = as.character(expression(italic(p)*" = 0.005")))

# Partial regression plot: number of species ~ habitat amount
fig.s2b =
    ggplot(data = partials[order(partials$res.ha.is), ],
           aes(x = res.ha.is, y = res.sr.is)) +

    labs(x = ~ atop(paste("Residual ", bold("habitat amount")),
                    paste(scriptstyle("(controlling for island size)"))),
         y = ~ atop(paste("Residual ", bold("number of species")),
                    paste(scriptstyle("(controlling for island size)")))) +

    scale_x_continuous(breaks = seq(-1.6, 0.8, length = 5)) +
    scale_y_continuous(breaks = seq(-0.3, 0.5, length = 5)) +

    geom_smooth(method = "lm", se = FALSE, colour = "black", size = 1.5) +
    geom_point(shape = 21, colour = "black", fill = "#999999", size = 5) +

    theme_bw(base_size = 20) +
    theme(panel.grid = element_blank(),
          panel.border = element_rect(colour = "black"),
          axis.title = element_text(colour = "black", face = "bold"),
          axis.text = element_text(colour = "black"),
          axis.ticks = element_line(colour = "black", size = 0.5)) +

  annotate("text", x = min(partials$res.ha.is), y = max(partials$res.sr.is),
             hjust = 0, vjust = 1, fontface = "bold", size = 10, label = "(b)") +
  annotate("text", x = max(partials$res.ha.is), y = min(partials$res.sr.is),
           hjust = 1, vjust = -2, size = 6,
           parse = T,
           label = as.character(expression(partial~italic(R)^{2}*""[adj]*" = 0.232"))) +
  annotate("text", x = max(partials$res.ha.is), y = min(partials$res.sr.is),
           hjust = 1, vjust = -2, size = 6, label = "slope = 0.211") +
  annotate("text", x = max(partials$res.ha.is), y = min(partials$res.sr.is),
           hjust = 1, vjust = 0, size = 6,
           parse = T, label = as.character(expression(italic(p)*" = 0.003")))
```

```
#fig.s2a; fig.s2b
```

FIGURE S2 Partial regression plots illustrating the independent effects of (a) island size and (b) habitat amount on the number of forest insectivorous bird species occurring at 33 forest islands within the Balbina Hydroelectric Reservoir in central Brazilian Amazonia.

Correlation between predictors

We examined the strength of correlation values between island size and habitat amount across the entire spectrum of 40 landscape scales (50-2,000 m) to assess the suitability of the multiple linear regression test.

```r
# Pearson's product-moment correlation between island size and habitat
# amount across 40 landscape scales
cor.predictors = apply(islands[63:102], 2, function(z) cor.test(log10(z + 1),
    log10(islands$area.ha + 1)))

# Convert the results stored as a 'list' object into a 'matrix' object
cor.predictors = do.call(rbind, lapply(cor.predictors, function(z) z[4]))

# Save the results in a 'data frame' object
cor.predictors = data.frame(buffer = seq(50, 2000, 50), r = as.numeric(cor.predictors))

# The highest correlation between island size and habitat amount occurred
# exactly at the scale of effect: 200-m buffer
cor.predictors[order(-cor.predictors$r), ][1, ]
```

```
##   buffer        r
## 4    200 0.8574179
```

Figure S3

```r
# Graph of the correlation between island size and habitat amount
# across 40 landscape scales
fig.s3 =
  ggplot(data = cor.predictors,
         aes(x = buffer, y = r)) +

  labs(x = "Buffer of the landscape scale (m)",
       y = "Pearson correlation coefficient (r):\nisland size vs. habitat amount") +

  scale_x_continuous(breaks = c(50, 200, 500, 1000, 1500, 2000),
                     labels = c("50", "200", "500", "1,000", "1,500", "2,000")) +
  scale_y_continuous(breaks = seq(0.60, 0.85, 0.05)) +

  geom_segment(x = 200, xend = 200, y = 0, yend = max(cor.predictors$r),
               linetype = "dashed") +
  geom_point(shape = 21, colour = "black", fill = "#999999", size = 5) +

  theme_bw(base_size = 20) +
  theme(panel.grid = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.title = element_text(colour = "black", face = "bold"),
```

```
        axis.text = element_text(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5))

#fig.s3
```

FIGURE S3 Correlation between island size and habitat amount (i.e. percentage of forest cover) across 40 landscape scales, ranging from 50 to 2,000-m buffer at 50-m intervals. Note that the maximum correlation value (dashed line) between island size and habitat amount occurs exactly at the scale of effect (200-m buffer; see Fig. S1).

Test 2: *Z*-values

We used the logarithmic form of the species-area relationship (SAR) to fit simple linear regression models for islands surrounded by either low habitat amounts (up to *c.* 50% of the landscape) or high habitat amounts. To assess whether the *z*-values (i.e. regression sloes) derived from either SARs were significantly different ($p < 0.05$), we performed an ANCOVA model with habitat amount (low or high) as an independent categorical variable. To support the IBT, the SAR for islands with low habitat amounts should have a higher *z*-value than those with high habitat amounts. Conversely, *z*-values should be statistically equivalent to support the HAH.

```
# SAR for forest islands surrounded by low habitat amount (SAR low)
sar.low.55 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.55 == "low"))
summary(sar.low.55)
```

```
##
## Call:
## lm(formula = log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
##     variables$ha.cat.55 == "low"))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.38960 -0.14095 -0.05168  0.21539  0.36381
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)             -0.1068     0.1437  -0.743 0.470462
## log10(island.size + 1)   0.7469     0.1600   4.667 0.000441 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2442 on 13 degrees of freedom
## Multiple R-squared:  0.6262, Adjusted R-squared:  0.5975
## F-statistic: 21.78 on 1 and 13 DF,  p-value: 0.0004406
```
```
# Z-value of the SAR low
summary(sar.low.55)$coefficients[2, 1]
```

```
## [1] 0.746865
```

```
# Is the SAR low significant?
summary(sar.low.55)$coefficients[2, 4] < 0.05
```

```
## [1] TRUE
```

```
# SAR for forest islands surrounded by high habitat amount (SAR high)
sar.high.55 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.55 == "high"))
summary(sar.high.55)
```

```
##
## Call:
## lm(formula = log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
##     variables$ha.cat.55 == "high"))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.297587 -0.080551 -0.001453  0.081922  0.275937
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.37142    0.16127   2.303 0.035030 *
## log10(island.size + 1)  0.31060    0.06698   4.637 0.000274 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1529 on 16 degrees of freedom
## Multiple R-squared:  0.5734, Adjusted R-squared:  0.5467
## F-statistic:  21.5 on 1 and 16 DF,  p-value: 0.0002739
```

```r
# Z-value of the SAR high
summary(sar.high.55)$coefficients[2, 1]
```

```
## [1] 0.3105972
```

```r
# Is the SAR high significant?
summary(sar.high.55)$coefficients[2, 4] < 0.05
```

```
## [1] TRUE
```

```r
# Compare the z-values
sar.ancova.55 = lm(log10(richness + 1) ~ log10(island.size + 1) * ha.cat.55,
    data = variables)
summary(sar.ancova.55)
```

```
##
## Call:
## lm(formula = log10(richness + 1) ~ log10(island.size + 1) * ha.cat.55,
##     data = variables)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.38960 -0.11950 -0.00623  0.15190  0.36381
##
## Coefficients:
##                                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)                         0.3714     0.2100   1.769  0.08741
## log10(island.size + 1)              0.3106     0.0872   3.562  0.00130
## ha.cat.55low                       -0.4782     0.2404  -1.989  0.05619
## log10(island.size + 1):ha.cat.55low  0.4363     0.1569   2.780  0.00944
##
## (Intercept)                      .
## log10(island.size + 1)           **
## ha.cat.55low                     .
## log10(island.size + 1):ha.cat.55low **
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1991 on 29 degrees of freedom
## Multiple R-squared:  0.8066, Adjusted R-squared:  0.7866
## F-statistic: 40.31 on 3 and 29 DF,  p-value: 0.0000000001799
# Are the z-values significantly different?
anova(sar.ancova.55)[3, 5] < 0.05
```

```
## [1] TRUE
```

Figure 3

```
# Graph showing the species-area curves for forest islands surrounded by either
# low and high habitat amounts.
fig.3 =
  ggplot(data = variables,
         aes(x = island.size + 1, y = richness + 1,
             size = habitat.amount, colour = ha.cat.55, fill = ha.cat.55)) +

  labs(x = "Island size + 1 (ha)",
       y = "Number of species + 1 (n)",
       colour = "Habitat amount") +

  scale_x_log10(breaks = c(3, 10, 30, 100, 300, 1000),
                labels = c("3", "10", "30", "100", "300", "1,000")) +
  scale_y_log10(limits = c(NA, 10^max(predict(sar.high.55, interval = "confidence"))),
                breaks = c(1, 3, 10, 30)) +
  annotation_logticks() +

  geom_smooth(method = "lm", linetype = 0) +
  geom_smooth(method = "lm", se = F, size = 1.5) +
  geom_point(shape = 21, colour = "black", show.legend = FALSE) +

  scale_colour_manual(values = c("blue", "red"),
                      labels = c("high (> 70%)", "low (< 55%)")) +
  scale_fill_manual(values = c("dodgerblue", "firebrick1"),
                    labels = c("high (> 70%)", "low (< 55%)")) +
  scale_size_continuous(range = c(3, 11)) +

  theme_bw(base_size = 20) +
  theme(panel.grid = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.title = element_text(colour = "black", face = "bold"),
        axis.text = element_text(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5)) +
  theme(legend.title = element_text(size = 16),
        legend.text = element_text(size = 14),
        legend.position = c(0.95, 0.05),
        legend.justification = c(0.95, 0.05),
        legend.background = element_rect(colour = "black", size = 0.5),
        legend.key = element_rect(fill = NA)) +
  guides(size = FALSE, fill = FALSE,
```

```
        colour = guide_legend(override.aes = list(fill = NA), reverse = TRUE)) +

  annotate("text", x = min(variables$island.size + 1),
           y = 10^max(predict(sar.high.55, interval = "confidence")),
           hjust = 0, vjust = 1, colour = "red", size = 6,
           parse = T, label = as.character(expression(italic(z)*"-value = 0.747"))) +
  annotate("text", x = min(variables$island.size + 1),
           y = 10^max(predict(sar.high.55, interval = "confidence")),
           hjust = 0, vjust = 3.5, colour = "blue", size = 6,
           parse = T, label = as.character(expression(italic(z)*"-value = 0.311"))))

#fig.3
```

FIGURE 3 Divergent species-area relationships for understorey insectivorous birds surveyed at 33 forest islands surrounded by either low (r^2_{adj} = 0.598, p < 0.001, n = 15) or high (r^2_{adj} = 0.547, p < 0.001, n = 18) habitat amounts. Circle sizes scale to the landscape-scale habitat amount, with smaller and larger circles representing islands surrounded by either low or high habitat amounts, respectively. Note the different z-values for these two landscape scenarios (p = 0.009) and the base 10 logarithmic scales along both axes.

Sensitivity analysis

The z-value at landscapes isolated by low habitat amounts (< 55%) is statistically higher than that at landscapes connected by high habitat amounts (> 70%). Such a difference could be attributed to the selected cut-off (55%) to distinguish low from high habitat amounts. To test whether the difference in z-values is sensitive to that, we reran the z-value test using different cut-offs for habitat amount.

Cut-offs: 54%, 50%, 48.5%, 45%, 43.1%, 30%

```
# Classify the habitat amount into low or high using different cut-offs
variables$ha.cat.54 = ifelse(variables$habitat.amount < 54, "low", "high")   # 54%
variables$ha.cat.50 = ifelse(variables$habitat.amount < 50, "low", "high")   # 50%
variables$ha.cat.48.5 = ifelse(variables$habitat.amount < 48.5, "low", "high")   # 48.5%
variables$ha.cat.45 = ifelse(variables$habitat.amount < 45, "low", "high")   # 45%
variables$ha.cat.43.1 = ifelse(variables$habitat.amount < 43.1, "low", "high")   # 43.1%
variables$ha.cat.30 = ifelse(variables$habitat.amount < 30, "low", "high")   # 30%
```

Cut-off of 54%

```
# SAR for forest islands surrounded by low habitat amount (SAR low)
sar.low.54 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.54 == "low"))
```

```
# Z-value of the SAR low
summary(sar.low.54)$coefficients[2, 1]
```

```
## [1] 0.7731002
```

```
# Is the SAR low significant?
summary(sar.low.54)$coefficients[2, 4] < 0.05
```

```
## [1] TRUE
```

```r
# SAR for forest islands surrounded by high habitat amount (SAR high)
sar.high.54 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.54 == "high"))
```

```r
# Z-value of the SAR high
summary(sar.high.54)$coefficients[2, 1]
```

```
## [1] 0.332675
```

```r
# Is the SAR high significant?
summary(sar.high.54)$coefficients[2, 4] < 0.05
```

```
## [1] TRUE
```

```r
# Compare the z-values
sar.ancova.54 = lm(log10(richness + 1) ~ log10(island.size + 1) * ha.cat.54,
    data = variables)
```

```r
# Are the z-values significantly different?
anova(sar.ancova.54)[3, 5] < 0.05
```

```
## [1] TRUE
```

Summary of the results for the cut-off of 54%

- SAR low is **significant**
- SAR high is **significant**
- *Z*-values are **different**

Cut-off of 50%

```r
# SAR for forest islands surrounded by low habitat amount (SAR low)
sar.low.50 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.50 == "low"))
```

```r
# Z-value of the SAR low
summary(sar.low.50)$coefficients[2, 1]
```

```
## [1] 0.9007251
```

```r
# Is the SAR low significant?
summary(sar.low.50)$coefficients[2, 4] < 0.05
```

```
## [1] TRUE
```

```r
# SAR for forest islands surrounded by high habitat amount (SAR high)
sar.high.50 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.50 == "high"))
```

```r
# Z-value of the SAR high
summary(sar.high.50)$coefficients[2, 1]
```

```
## [1] 0.3200293
```

```r
# Is the SAR high significant?
summary(sar.high.50)$coefficients[2, 4] < 0.05
```

```
## [1] TRUE
```

```r
# Compare the z-values
sar.ancova.50 = lm(log10(richness + 1) ~ log10(island.size + 1) * ha.cat.50,
    data = variables)
```

```r
# Are the z-values significantly different?
anova(sar.ancova.50)[3, 5] < 0.05
```

[1] TRUE

Summary of the results for the cut-off of 50%

- SAR low is **significant**
- SAR high is **significant**
- *Z*-values are **different**

Cut-off of 48.5%

```r
# SAR for forest islands surrounded by low habitat amount (SAR low)
sar.low.48.5 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.48.5 == "low"))
```

```r
# Z-value of the SAR low
summary(sar.low.48.5)$coefficients[2, 1]
```

[1] 0.8350526

```r
# Is the SAR low significant?
summary(sar.low.48.5)$coefficients[2, 4] < 0.05
```

[1] TRUE

```r
# SAR for forest islands surrounded by high habitat amount (SAR high)
sar.high.48.5 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.48.5 == "high"))
```

```r
# Z-value of the SAR high
summary(sar.high.48.5)$coefficients[2, 1]
```

[1] 0.2847645

```r
# Is the SAR high significant?
summary(sar.high.48.5)$coefficients[2, 4] < 0.05
```

[1] TRUE

```r
# Compare the z-values
sar.ancova.48.5 = lm(log10(richness + 1) ~ log10(island.size + 1) * ha.cat.48.5,
    data = variables)
```

```r
# Are the z-values significantly different?
anova(sar.ancova.48.5)[3, 5] < 0.05
```

[1] TRUE

Summary of the results for the cut-off of 48.5%

- SAR low is **significant**
- SAR high is **significant**
- *Z*-values are **different**

Cut-off of 45%

```r
# SAR for forest islands surrounded by low habitat amount (SAR low)
sar.low.45 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.45 == "low"))
```

```r
# Z-value of the SAR low
summary(sar.low.45)$coefficients[2, 1]
```

```
## [1] 0.598169
```

```r
# Is the SAR low significant?
summary(sar.low.45)$coefficients[2, 4] < 0.05
```

```
## [1] FALSE
```

```r
# SAR for forest islands surrounded by high habitat amount (SAR high)
sar.high.45 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.45 == "high"))
```

```r
# Z-value of the SAR high
summary(sar.high.45)$coefficients[2, 1]
```

```
## [1] 0.245553
```

```r
# Is the SAR high significant?
summary(sar.high.45)$coefficients[2, 4] < 0.05
```

```
## [1] TRUE
```

```r
# Compare the z-values
sar.ancova.45 = lm(log10(richness + 1) ~ log10(island.size + 1) * ha.cat.45,
    data = variables)
```

```r
# Are the z-values significantly different?
anova(sar.ancova.45)[3, 5] < 0.05
```

```
## [1] FALSE
```

Summary of the results for the cut-off of 45%

- SAR low is **not significant**
- SAR high is **significant**
- *Z*-values are **not different**

Cut-off of 43.1%

```r
# SAR for forest islands surrounded by low habitat amount (SAR low)
sar.low.43.1 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.43.1 == "low"))
```

```r
# Z-value of the SAR low
summary(sar.low.43.1)$coefficients[2, 1]
```

```
## [1] 0.6280358
```

```r
# Is the SAR low significant?
summary(sar.low.43.1)$coefficients[2, 4] < 0.05
```

```
## [1] FALSE
```

```r
# SAR for forest islands surrounded by high habitat amount (SAR high)
sar.high.43.1 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.43.1 == "high"))
```

```r
# Z-value of the SAR high
summary(sar.high.43.1)$coefficients[2, 1]
```

```
## [1] 0.2560844
```

```r
# Is the SAR high significant?
summary(sar.high.43.1)$coefficients[2, 4] < 0.05
```

```
## [1] TRUE
```

```r
# Compare the z-values
sar.ancova.43.1 = lm(log10(richness + 1) ~ log10(island.size + 1) * ha.cat.43.1,
    data = variables)
```

```r
# Are the z-values significantly different?
anova(sar.ancova.43.1)[3, 5] < 0.05
```

```
## [1] FALSE
```

Summary of the results for the cut-off of 43.1%

- SAR low is **not significant**
- SAR high is **significant**
- *Z*-values are **not different**

Cut-off of 30%

```r
# SAR for forest islands surrounded by low habitat amount (SAR low)
sar.low.30 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.30 == "low"))
```

```r
# Z-value of the SAR low
summary(sar.low.30)$coefficients[2, 1]
```

```
## [1] 0.584008
```

```r
# Is the SAR low significant?
summary(sar.low.30)$coefficients[2, 4] < 0.05
```

```
## [1] FALSE
```

```r
# SAR for forest islands surrounded by high habitat amount (SAR high)
sar.high.30 = lm(log10(richness + 1) ~ log10(island.size + 1), data = subset(variables,
    variables$ha.cat.30 == "high"))
```

```r
# Z-value of the SAR high
summary(sar.high.30)$coefficients[2, 1]
```

```
## [1] 0.2705981
```

```r
# Is the SAR high significant?
summary(sar.high.30)$coefficients[2, 4] < 0.05
```

```
## [1] TRUE
```

```
# Compare the z-values
sar.ancova.30 = lm(log10(richness + 1) ~ log10(island.size + 1) * ha.cat.30,
    data = variables)
```

```
# Are the z-values significantly different?
anova(sar.ancova.30)[3, 5] < 0.05
```

[1] FALSE

Summary of the results for the cut-off of 30%

- SAR low is **not significant**
- SAR high is **significant**
- *Z*-values are **not different**

Test 3: Species accumulation curves (SACs)

We compared the cumulative number of species across all 33 surveyed islands ordered according to either increasing (small-to-large) or decreasing (large-to-small) island sizes, which may lead to three possible outcomes. First, the small-to-large accumulation curve lies below the large-to-small, supporting IBT. Second, the curves overlap, supporting HAH. Third, the small-to-large accumulation curve lies above the large-to-small, supporting neither IBT nor HAH.

Small-to-large curve

```
# Cumulative number of species
small.large.richness = specaccum(islands[order(islands$area.ha), names(islands.spp)],
    method = "collector")$richness
```

```
# Cumulative island size
small.large.size = cumsum(islands[order(islands$area.ha), "area.ha"])
```

Large-to-small curve

```
# Cumulative number of species
large.small.richness = specaccum(islands[order(-islands$area.ha), names(islands.spp)],
    method = "collector")$richness
```

```
# Cumulative island size
large.small.size = cumsum(islands[order(-islands$area.ha), "area.ha"])
```

```
# Combine the results of small-to-large with large-to-small curves
curves = data.frame(cum.richness = c(small.large.richness, large.small.richness),
    cum.size = c(small.large.size, large.small.size), order = rep(c("small.large",
        "large.small"), each = 33))
```

Figure 4

```
# Graph comparing small-to-large with large-to-small species accumulation curves
fig.4 =
  ggplot(data = curves,
        aes(x = cum.size, y = cum.richness,
            fill = order, linetype = order)) +
```

```r
  labs(x = "Cumulative island size (ha)",
       y = "Cumulative number of species (n)",
       fill = "Island size order") +

  scale_x_continuous(breaks = c(0.63, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 7841),
                      labels = c("0.63", "1,000", "2,000", "3,000", "4,000", "5,000",
                                 "6,000", "7,000", "7,841")) +
  scale_y_continuous(breaks = c(0, 10, 20, 30, 40, 50, 59)) +

  geom_line(colour = "black", size = 1.5) +
  geom_point(shape = 21, colour = "black", size = 5) +

  scale_fill_manual(values = c("grey40", "grey80"),
                    labels = c("Large-to-small", "Small-to-large")) +
  scale_linetype_manual(values = c("solid", "dashed")) +

  theme_bw(base_size = 20) +
  theme(panel.grid = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.title = element_text(colour = "black", face = "bold"),
        axis.text = element_text(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5)) +
  theme(legend.title = element_text(size = 16),
        legend.text = element_text(size = 14),
        legend.position = c(0.95, 0.05),
        legend.justification = c(0.95, 0.05),
        legend.background = element_rect(colour = "black", size = 0.5),
        legend.key = element_rect(fill = NA)) +
  guides(fill = guide_legend(override.aes = list(size = 5), reverse = TRUE),
         linetype = FALSE)

#fig.4
```

FIGURE 4 Species accumulation curves of understorey insectivorous birds for islands ordered according to either increasing (light grey circles, dashed line) or decreasing (dark grey circles, solid line) island sizes.

Test 4: Extrapolation of SAR

We fit a SAR model (see Test 2: *Z*-values) to all 33 surveyed islands. We further extrapolated the number of species to a hypothetical island containing the area (+ 1) of all surveyed islands combined (7,874 ha), and compared the total number of species observed across surveyed islands to the extrapolated value. Compared to the extrapolated value, a lower observed number of species would support IBT; a statistically equivalent number, HAH; a higher number, neither IBT nor HAH.

```r
# SAR for all 33 surveyed islands
sar = lm(log10(richness + 1) ~ log10(island.size + 1), data = variables)
summary(sar)
```

```
##
## Call:
## lm(formula = log10(richness + 1) ~ log10(island.size + 1), data = variables)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.42540 -0.10776  0.02702  0.12345  0.49190
```

```
## 
## Coefficients:
##                          Estimate Std. Error t value         Pr(>|t|)    
## (Intercept)              0.15465    0.07882   1.962          0.0588 .  
## log10(island.size + 1)   0.40736    0.04196   9.708 0.0000000000647 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.2179 on 31 degrees of freedom
## Multiple R-squared:  0.7525, Adjusted R-squared:  0.7445 
## F-statistic: 94.25 on 1 and 31 DF,  p-value: 0.00000000006471
```

```r
# Calculate the extrapolated value and the respective confidence intervals
extrapolated = 10^predict(sar,
                          newdata = data.frame(island.size = sum(variables$island.size + 1)),
                          interval = "confidence")

extrapolated
```

```
##        fit      lwr      upr
## 1 55.18854 34.22766 88.98579
```

Figure 5

```r
# Graph showing the SAR for all 33 surveyed islands and the extrapolated value
fig.5 =
  ggplot() +

  labs(x = "Island size + 1 (ha)",
       y = "Number of species + 1 (n)",
       fill = "Number of species") +

  scale_x_log10(breaks = c(10, 100, 1000, 5000),
                labels = c("10", "100", "1,000", "5,000")) +
  scale_y_log10(breaks = c(1, 10, 100)) +
  annotation_logticks() +

  geom_line(size = 1.5,
            aes(x = variables$island.size + 1,
                y = 10^predict(sar))) +

  geom_line(size = 0.5, linetype = "dashed",
            aes(x = variables$island.size + 1,
                y = 10^predict(sar, interval = "confidence")[, 2])) +

  geom_line(size = 0.5, linetype = "dashed",
            aes(x = variables$island.size + 1,
                y = 10^predict(sar, interval = "confidence")[, 3])) +

  geom_line(size = 1.5, linetype = "dashed",
            aes(x = seq(700, sum(variables$island.size + 1), length = 10),
                y = 10^predict(sar,
                               newdata = data.frame(island.size =
```

```r
                                                    seq(700,
                                                        sum(variables$island.size + 1),
                                                        length = 10))))) +

geom_line(size = 0.5, linetype = "dashed",
          aes(x = seq(1700, sum(variables$island.size + 1), length = 10),
              y = 10^predict(sar,
                             newdata = data.frame(island.size =
                                                      seq(1700,
                                                          sum(variables$island.size + 1),
                                                          length = 10)),
                             interval = "confidence")[, 2])) +

geom_line(size = 0.5, linetype = "dashed",
          aes(x = seq(1700, sum(variables$island.size + 1), length = 10),
              y = 10^predict(sar,
                             newdata = data.frame(island.size =
                                                      seq(1700,
                                                          sum(variables$island.size + 1),
                                                          length = 10)),
                             interval = "confidence")[, 3])) +

geom_point(shape = 21, colour = "black", size = 5,
           data = data.frame(richness = c(variables$richness + 1,
                                          extrapolated[1],
                                          length(islands.spp) + 1),
                             island.size = c(variables$island.size + 1,
                                             rep(sum(variables$island.size + 1), 2)),
                             group = c(rep("Observed per island", 33),
                                       "Extrapolated",
                                       "Observed in total")),
           aes(x = island.size, y = richness, fill = group)) +

scale_fill_manual(values = c("white", "black", "#999999")) +

theme_bw(base_size = 20) +
theme(panel.grid = element_blank(),
      panel.border = element_rect(colour = "black"),
      axis.title = element_text(colour = "black", face = "bold"),
      axis.text = element_text(colour = "black"),
      axis.ticks = element_line(colour = "black", size = 0.5)) +
theme(legend.title = element_text(size = 16),
      legend.text = element_text(size = 14),
      legend.position = c(0.95, 0.05),
      legend.justification = c(0.95, 0.05),
      legend.background = element_rect(colour = "black", size = 0.5),
      legend.key = element_rect(fill = NA)) +
guides(fill = guide_legend(reverse = TRUE)) +

annotate("text", x = min(variables$island.size + 1), y = 100,
         hjust = 0, vjust = 1, size = 6,
         parse = T,
         label = as.character(expression(italic(r)^{2}*""[adj]*" = 0.744"))) +
```

```
  annotate("text", x = min(variables$island.size + 1), y = 100,
           hjust = 0, vjust = 3.5, size = 6,
           parse = T, label = as.character(expression(italic(z)*"-value = 0.407")))

#fig.5
```

FIGURE 5 Species-area relationship for understorey insectivorous birds surveyed at 33 forest islands. The white circle shows the extrapolated number of species (55.2) to a hypothetical island containing the area (+ 1) of all 33 surveyed islands combined (7,874 ha), whereas the black circle shows the total number of species observed in this study (59 + 1). Dashed lines show the 95% confidence intervals of the predicted line. Note the base 10 logarithmic scales along both axes.

Save the figures

```
library(gridExtra)

ggsave(fig.s1, file = "figs1.pdf", width = 20, height = 20, units = "cm")
ggsave(grid.arrange(fig.s2a, fig.s2b, ncol = 2),
       file = "figs2.pdf", width = 40, height = 20, units = "cm")
ggsave(fig.s3, file = "figs3.pdf", width = 20, height = 20, units = "cm")
ggsave(fig.3, file = "fig3.pdf", width = 20, height = 20, units = "cm")
ggsave(fig.4, file = "fig4.pdf", width = 20, height = 20, units = "cm")
ggsave(fig.5, file = "fig5.pdf", width = 20, height = 20, units = "cm")
```

Chapter 4

The role of baseline suitability in assessing the impacts of land-use change on biodiversity
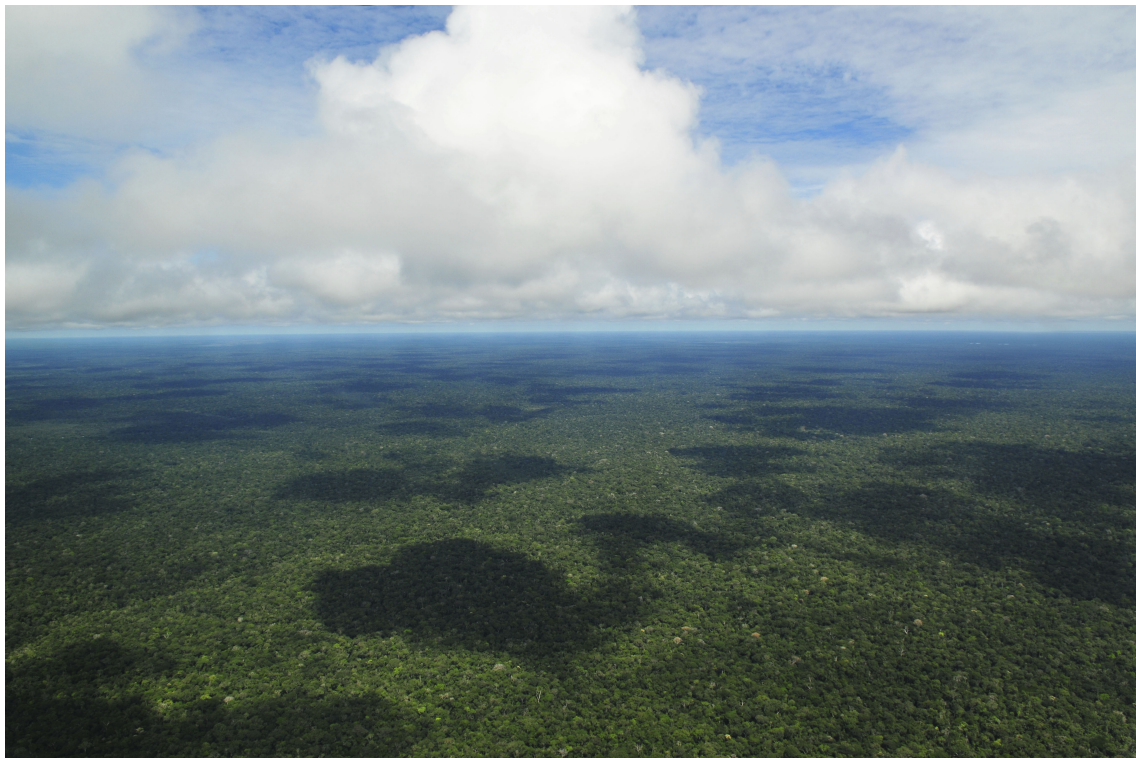
© Neil Palmer (CIAT)

4.1   Abstract

We examined changes in species richness from reference sites to impacted sites to illustrate the extent to which estimated impacts of land-use change on biodiversity can be affected by the degree of baseline suitability (intactness of references sites) and the species assemblage under consideration. We mist-netted birds at five continuous Amazonian forest sites and 33 forest islands (0.63-1,699 ha) within a large hydroelectric reservoir. We then produced a gradient of baseline suitability based on forest area of five sets of reference sites, namely *continuous forest*, *1,000 ha*, *500 ha*, *250 ha* and *100 ha*, and contrasted these with all smaller islands combined considering two types of species assemblages. The first comprised only species captured at reference sites, whereas the second comprised all species captured at all sites. We also used a biodiversity complementarity approach to define the minimum set of forest islands retaining the maximum combined number of species for the species assemblage occurring both at continuous forest sites and across all sites. A focus on the baseline species assemblage from the most suitable baseline (*continuous forest*) resulted in an estimated decrease of 67% in species richness (negative effect) at impacted sites, whereas a focus on the overall species assemblage and the use of the least suitable baseline (*100 ha*) as a reference condition reversed this trend resulting in an estimated increase of 43% (positive effect) at impacted sites. Furthermore, the solution for the minimum set problem targeting only species captured at continuous forest sites resulted in a decrease of 43% in the number of islands to be protected in relation to the solution targeting the overall species assemblage. We therefore underline the imperative of considering the intactness of reference sites to accurately assess the impacts of land-use change on biodiversity and define mitigation strategies.

Keywords: environmental impact assessment, habitat fragmentation, habitat loss, insularization, intact forest landscapes, shifting baseline

4.2   Introduction

Land-use change is the most pervasive threat to terrestrial biodiversity world-wide, particularly in the tropics where species-rich biotas are more sensitive to environmental changes (Gardner et al. 2009) that are projected to escalate (Sala et al. 2000). Therefore, the fate of native biodiversity in human-modified land-scapes depends on the ability of species to persist in fragmented, degraded, converted and/or regenerated habitats (i.e. modified habitats). Despite the potential value of structurally degraded forest habitats for biodiversity conservation, such as selectively logged (Edwards et al., 2012) and secondary forests (Chazdon et al., 2009), species losses will inevitably occur in modified habitats as primary old-growth forests are largely irreplaceable for sustaining continuous-forest species assemblages (Gibson et al., 2011). Thus, evaluating the costs and benefits of conservation strategies in human-modified landscapes is an urgent task in prioritising efforts in the most promising modified habitats available (Banks-Leite et al., 2014).

A central question in conservation biology is "What was the condition of ecosystems before significant human disruption, and how can this knowledge be used to improve current and future management?" (Sutherland et al., 2009). Ideally, paleoecological, archaeological, and historical data would be gathered to assess the impacts of anthropogenic disturbance on biodiversity, thereby supporting conservation planning (Jackson et al., 2001). Since these data are almost invariably non-existent, relatively intact contemporary baseline (i.e. control) sites within the same landscape can be employed in a space-for-time substitution (Ewers et al., 2009). However, shifting baselines (i.e. changing biological conditions induced by past disturbance; Papworth et al. 2009) may obscure the most severe impacts of land-use change on sensitive species as a result of extinction filters (Balmford, 1996) and local proliferation of disturbance-adapted species (Morante-Filho et al., 2015). Therefore, the credibility of environmental impact assessments largely depends on baseline suitability (Gardner et al., 2009; Ritter et al., 2017), which herein is defined as the reliability of reference sites in resembling some pre-disturbance condition. Accordingly, a continuous primary forest covering tens of thousands of hectares would be a suitable baseline for an adjacent fragmented landscape (Sigel et al. 2010).

If one focuses only on the species assemblage from reference sites, land-use change will almost inevitably result in species losses (i.e. decreases in species richness). Alternatively, a focus on the overall species assemblage may not only result in species losses, but also in compensatory dynamics whereby any loss of species exclusively found at reference sites is either compensated (Ewers et al.

2009) or even surpassed (Humphreys and Kitchener 1982) by any gain of species restricted to impacted sites (i.e. either no net loss or increase in species richness). Hence, environmental impact assessments relying on overall species assemblages are likely to yield optimistic but misleading outcomes (Lövei et al., 2006), which can misdirect conservation strategies. For instance, the minimum set problem is a commonly used tool to identify the most cost-effective set of sites (i.e. the maximum number of species retained in the fewest number of sites) to concentrate conservation efforts (Howard et al., 1998). Nevertheless, the possible solutions for the minimum set problem are target specific (Howard et al., 2000), implying that a focus on either the species assemblage typical of the baseline or the overall species assemblage may result in different solutions. Importantly, such issues would only be relevant if conservation and restoration programs target protecting either the "natural" state or those returning to pre-disturbance conditions (Wiens and Hobbs 2015), which will become increasingly daunting to achieve given the rapid development of 'novel ecosystems' induced by human activities (Hobbs et al. 2009).

Here, we use a gradient of baseline suitability and define two types of species assemblages derived from systematic avifaunal surveys carried out in continuous forest (control) sites and forest islands within an anthropogenic archipelago of central Amazonia. We then examine how the degree of baseline suitability and both types of species assemblages affect the estimated impact of forest fragmentation on species richness. Furthermore, we compare the solutions for the minimum set problem targeting either the continuous-forest species assemblage derived from only control sites or the overall species assemblage derived from both control sites and forest islands. In doing so, we provide empirical evidence that environmental impact assessments and conservation strategies can be severely biased by both the suitability of baseline sites and which set of focal species are considered.

4.3 Methods

4.3.1 Study area

This study was carried out within the vast Balbina Hydroelectric Reservoir (hereafter, BHR; 1°40′ S, 59°40′ W; Fig. 4.1) and adjacent areas of continuous intact forest, located in the State of Amazonas, central Brazilian Amazonia. The BHR spans *c.* 300,000 ha and was formed by the damming of the Uatumã River in 1987 (Fearnside, 2016), creating over 3,500 forest islands on higher elevation terrain of the once continuous intact forest (Benchimol and Peres, 2015). To offset the environmental impacts of the Balbina hydroelectric dam, an area of 940,358 ha was set-aside on the left bank of the former Uatumã River to create the strictly protected Uatumã Biological Reserve (IUCN category Ia), the largest of its category in Brazil. Moreover, the reservoir on the left bank, including all its islands, is also protected.
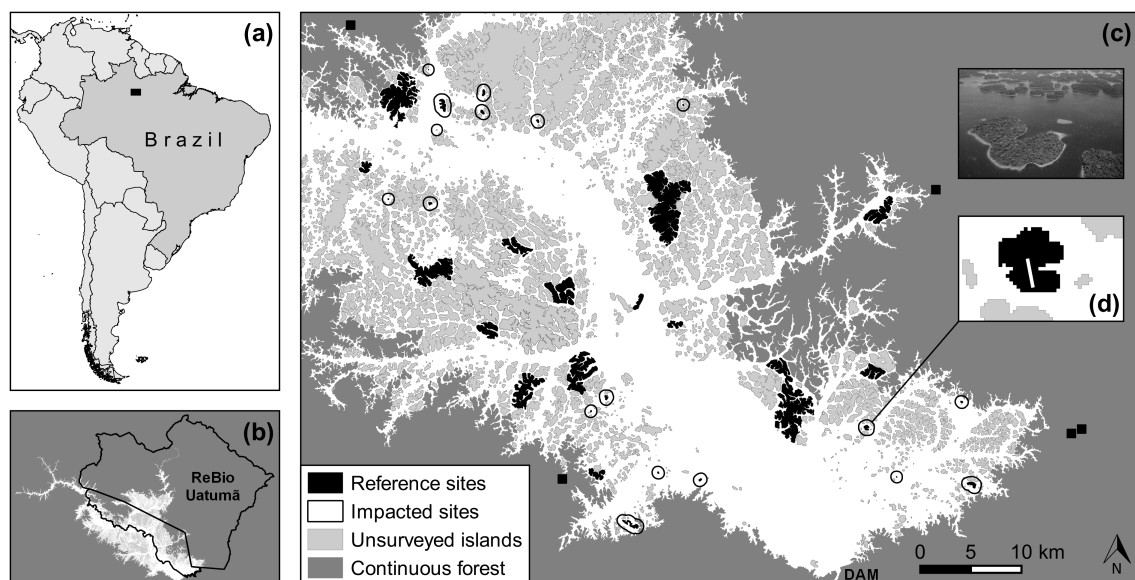
Figure 4.1: Location of the study area in central Brazilian Amazonia, indicated by a solid rectangle containing (b) the Balbina Hydroelectric Reservoir (BHR) landscape (water in white, forest in gray), showing the boundaries of the Uatumã Biological Reserve, a strictly-protected area safeguarding most of this landscape; (c) larger inset map showing the spatial distribution of the 38 sample sites within the BHR landscape, including all surveyed islands and continuous forest sites; and (d) detail of an understory mist-net line (white line) used to sample the avifauna. Reference sites correspond to continuous forest sites and forest islands larger than *c.* 100 ha, whereas impacted sites correspond to forest islands smaller than 55 ha (see Fig. 4.3 and Table S4.1). Photo credit: Eduardo M. Venticinque.

The vegetation is comprised primarily of submontane dense ombrophilous (*terra firme*) forest, although seasonally flooded *igapó* forest formerly occurred along the margins of the Uatumã River before damming. Forest islands

at the BHR range in size from 0.2 to 4,878 ha and are surrounded by an inhospitable open-water matrix punctuated by dead tree snags rising above the water level (Benchimol and Peres, 2015). Both island area and associated edge-mediated disturbance shape forest structure: smaller islands are species-poor and dominated by pioneer tree species, whereas larger islands are species-rich and contain a higher dominance of large-seeded canopy tree species (Benchimol and Peres, 2015). According to the Köppen classification, the climate is equatorial fully humid (Af), with mean annual precipitation and temperature of 2,464 mm and 26.5 °C, respectively (Alvares et al., 2013).

4.3.2 Sampling design

We selected 38 sample sites, five in continuous forest sites (hereafter, control sites) and 33 in forest islands (Table S4.1) distributed across an area of over 200,000 ha (Fig. 4.1). Sample sites were spaced apart by an average distance of 31.1 km (SD = 17.1 km, range = 1.1-82.6 km). Sixteen forest islands and four control sites were on the left bank, whereas 17 forest islands and one control site were on the right bank (Fig. 4.1). Forest area of surveyed islands was calculated in QGIS software (QGIS Development Team 2016) using a classified image (Collection 2, 2015, Amazon) derived from 30-m resolution LANDSAT imagery downloaded from the Brazilian Annual Land Use and Land Cover Mapping Project (available at http://mapbiomas.org). Forest cover was defined as 'dense forest' (pixel value 3), because other pixel values effectively represent either heavily degraded forests or non-forest land cover types. Accordingly, the forest area of surveyed islands ranged from 0.63 to 1,699 ha.

4.3.3 Avian surveys

Fieldwork was carried out over 12 months from July to December 2015 and 2016. At each sample site, we surveyed birds using 16 mist nets (12 × 2.5 m, *Ecotone* 1016/12) set in the understory along a continuous near-linear *c.* 200-m net-line whenever possible. In forest islands smaller than 4 ha, we used a cross-shaped net-line design, thereby ensuring the same sampling effort across all 38 sample sites. Herein, each mist-net line corresponds to one sample site. Mist nets were operated from 06:00 to 15:00 h for two days each year, amounting a sampling effort of 576 net-hours per sample site and 21,888 net-hours in total (16 mist nets × 9 hours × 2 days × 2 years × 38 sample sites). Since mist nets capture mainly understory birds and occasionally birds that walk on the ground or that fly above net level (Karr 1981), our inferences are drawn from a subset of the avifauna. To

avoid double-counting, we ringed birds with coded aluminum rings and subsequently excluded recaptures. Rings were provided by the Brazilian National Center for Bird Conservation and Research (CEMAVE) under research permits SISBIO 49068 and CEMAVE 3984.

4.3.4   Species assemblages

We defined two types of species assemblages based on the species occurrence across all sample sites ($n = 38$). The first comprises species captured only at reference sample sites (hereafter referred to as *baseline species assemblage*, although we also use the term *continuous-forest species assemblage* to refer to the species subset from control sites). The second comprises all species captured in both reference sample sites and forest islands (hereafter, *overall species assemblage*).

4.3.5   Estimated impact of forest fragmentation on species richness

We used five sets of reference sample sites to represent a gradient of baseline suitability, which was based on insular forest area (Table S4.2). We reasonably assumed that the avifauna in control sites was more intact than that in forest islands and that the avifauna on larger forest islands was more intact than on smaller ones (Aurélio-Silva et al. 2016). Accordingly, the avifauna of each set of reference sample sites comprises a different baseline, namely *continuous forest*, *1,000 ha*, *500 ha*, *250 ha* and *100 ha* (Table S4.2). The suitability of each baseline was inspected by comparing the number of species retained and gained in relation to the continuous forest baseline, so that the suitability of the continuous forest baseline was assigned the maximum biodiversity value (Fig. 4.2).

A previous study at the BHR landscape reported that bird species richness was remarkably reduced in forest islands smaller than 55 ha compared to larger forest islands of up to *c.* 1,700 ha (Aurélio-Silva et al., 2016). Therefore, surveyed islands smaller than 55 ha ($n = 18$; Table S4.1) were regarded as impacted sites due to forest fragmentation – the landscape process involving both forest loss and the breaking apart of forest (Fahrig, 2003).

Having defined the gradient of baseline suitability and the impacted sites, we compared the species richness of all impacted sites combined to that of each of the five baseline levels (Fig. 4.3) for both the *baseline* and the *overall* species assemblage. To accomplish this, we used the INEXT package (Hsieh et al., 2016) in R (R Core Team 2018) to calculate the rarefied number of species based on equal completeness (i.e. sample coverage) instead of equal size (i.e. number of
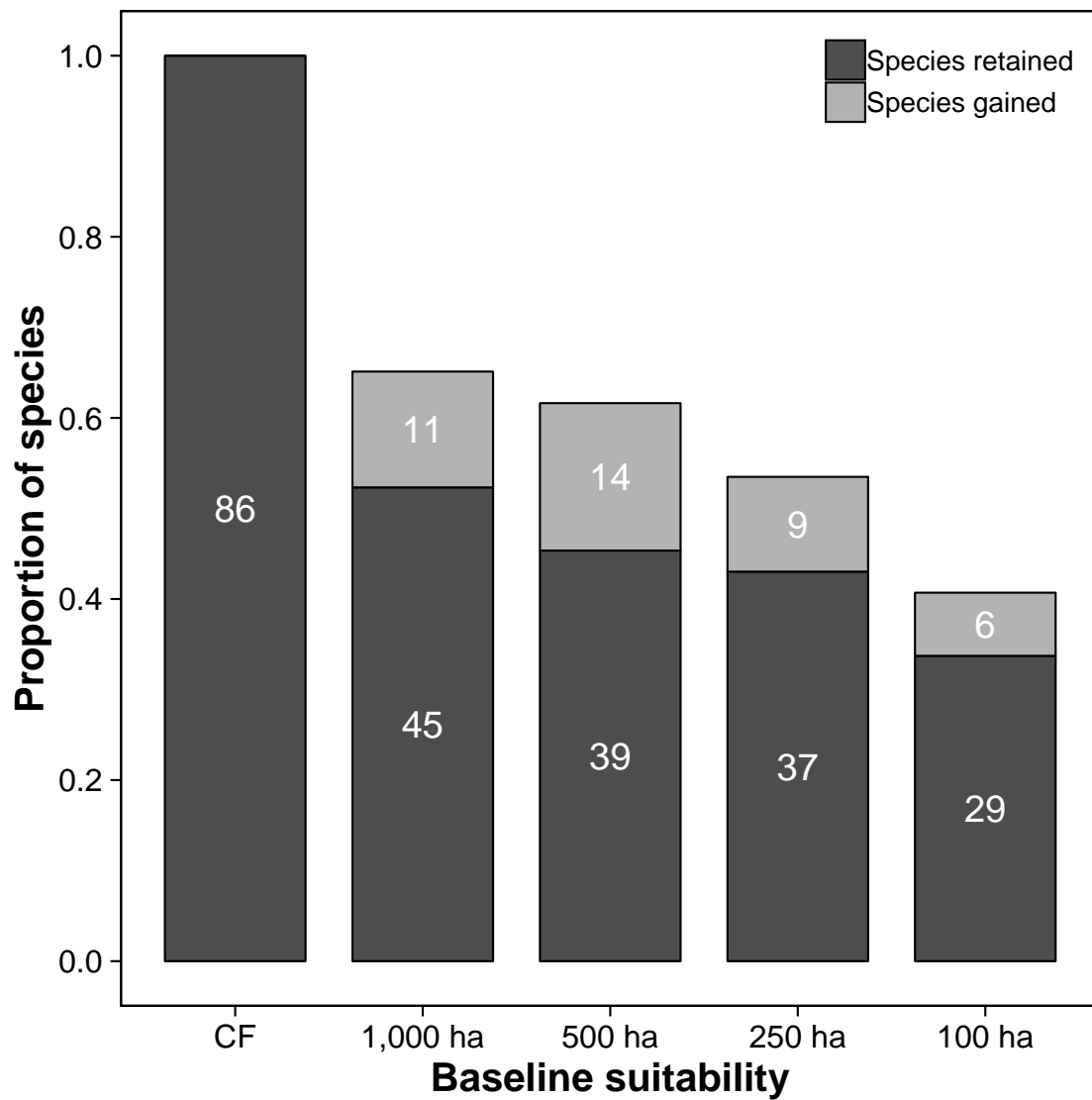
Figure 4.2: Gradient of baseline suitability showing the proportion of bird species retained and gained in relation to the understory bird assemblage captured in continuous forest sites (CF). Numbers within bars correspond to the total number of species in each of the five categories of baseline suitability.
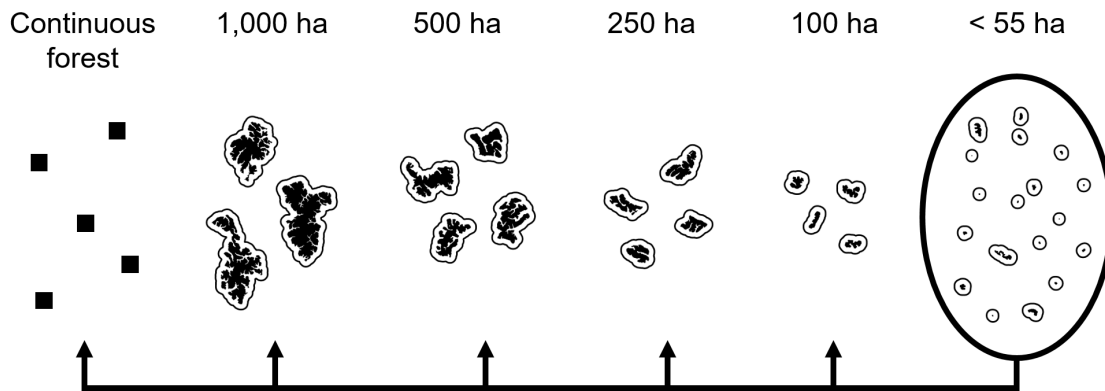
Figure 4.3: Schematic representation of the comparisons between the number of species at impacted sites (forest islands smaller than 55 ha) and at reference sites representing a gradient of baselines suitability, from the most (continuous forest) to the least suitable (100-ha forest islands).

individuals), because species-rich sites require a greater number of individuals to be fully characterized than species-poor sites (Chao and Jost 2012). Next, we calculated the proportional difference in species richness between the impacted sites combined and each of the five baseline levels to assess to what extent our estimated impacts of forest fragmentation are affected by the degree of baseline suitability and the composition of focal species assemblages. Estimated impact was measured as one minus the proportional difference between impacted and reference sites, so the higher the value, the more severe the estimated impact. Accordingly, if reference sites harbor 10 species and impacted sites 4 species, the estimated impact is 0.6 [1 − (4/10)] or a 60% reduction in species richness.

4.3.6   Minimum set problem

We used a biodiversity complementarity approach to determine the minimum number of forest islands (i.e. the 'solution' for the minimum set problem) that retained the maximum combined number of species (Howard et al., 1998) for both the *continuous-forest* and the *overall* species assemblage captured across the entire set of 33 surveyed islands. To do so, we used an algorithm that first selects the forest island containing the highest number of species and then sequentially selects the forest islands that add the highest number of previously unrepresented species (Howard et al., 1998). In the event of a tie, the more species-rich island was selected, and if that tie persisted, we selected the largest island. This procedure was repeated until all species captured in the entire set of 33 forest islands were represented by at least one individual. We then quantified the differences between the solutions for the minimum set problem targeting both the *continuous-forest* and the *overall* species assemblage.

4.4   Results

Considering all 38 sample sites, we captured a total of 2,115 birds representing 130 species, 103 genera, and 38 families. At the five continuous forest control sites, we captured 614 birds representing 86 species, 71 genera and 30 families; the number of individuals per control site ranged from 75 to 165 (mean [SD] = 122.8 [37.9]), and the number of species from 28 to 54 (42.8 [10.5]; Table S4.1). At the 33 forest islands, we captured 1,501 birds representing 109 species, 90 genera and 35 families; the number of individuals per forest island ranged from 5 to 121 (45.5 [28.3]), and the number of species from 3 to 43 (16.1 [9.72]; Table S4.1).

4.4.1   Estimated impact of forest fragmentation on species richness

The estimated impact of forest fragmentation on species richness, measured as the contrast between reference and impacted sites, was affected by both the degree of baseline suitability and the species assemblage under consideration (Fig. 4.4). Accordingly, the estimated impact was gradually reduced as a function of decreasing baseline suitability. Yet forest fragmentation significantly depressed the species richness of the *baseline species assemblage*, regardless of the degree of baseline suitability (Fig. 4.4; Table S4.3). In contrast, a focus on the *overall species assemblage* revealed a negative impact of forest fragmentation only when the most suitable baseline (*continuous forest*) was used as the reference condition (Fig. 4.4; Table S4.3). Furthermore, the *overall* species richness at impacted sites was significantly higher than that at the least suitable baseline (*100 ha*; Fig. 4.4; Table S4.3), which could be interpreted as a positive effect of forest fragmentation on species richness. Importantly, these results would not hold true if we had used only a one-year dataset from either 2015 or 2016 (see Supporting Information), which indicates that the reliability of environmental impact assessments also depends on the robustness of biodiversity inventories.
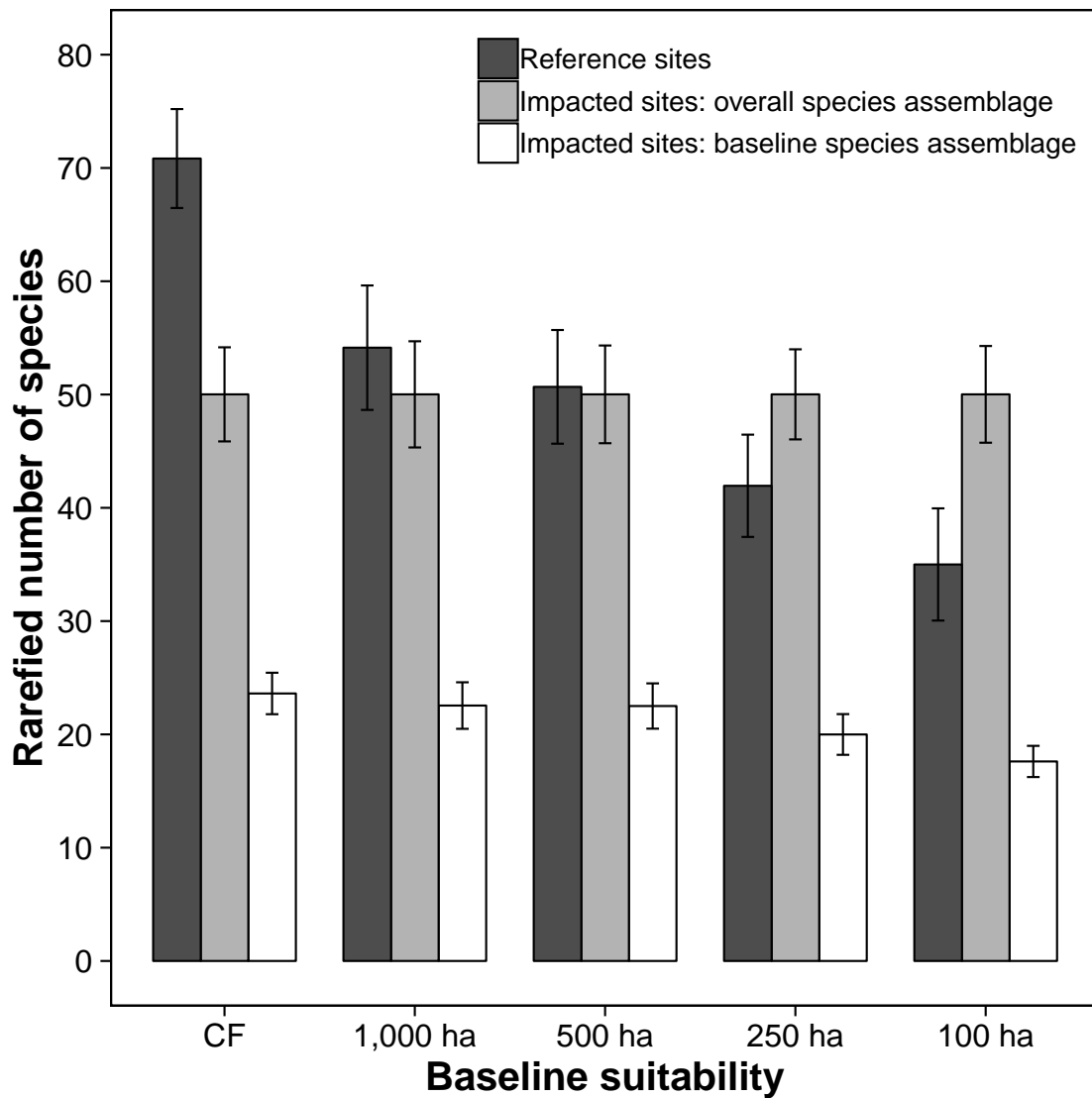
Figure 4.4: Comparisons between the rarefied number of understory bird species at reference and impacted sites across a gradient of baseline suitability for two types of species assemblages. The bar plot shows the contrast between reference and impacted sites from the most (CF – continuous forest) to the least suitable baselines (100-ha forest islands), thereby affecting the estimated impact of forest fragmentation on species richness of both types of species assemblages. Rarefied number of species were standardized by sample coverage (0.927). Error bars indicate the 95% confidence intervals.

4.4.2   Minimum set problem

Considering the *continuous-forest species assemblage*, the solution for the minimum set problem included 12 forest islands (36.4% of 33 islands; 71.9% of the aggregate area), which harbored 65 out of 86 species (75.6%; Fig. 4.5a). Considering the *overall species assemblage*, the solution for the minimum set problem included 21 forest islands (63.3% of 33 islands; 69.7% of the aggregate area), which included 109 out of 130 species (83.8%; Fig. 4.5b). Thus, a focus on the *continuous-forest species assemblage* would reduce conservation investments in terms of the number of forest islands to be protected from 21 to 12, but not in terms of the aggregate area. Nevertheless, even the entire set of 33 forest islands surveyed failed to include 21 species that were only represented at continuous forest sites.
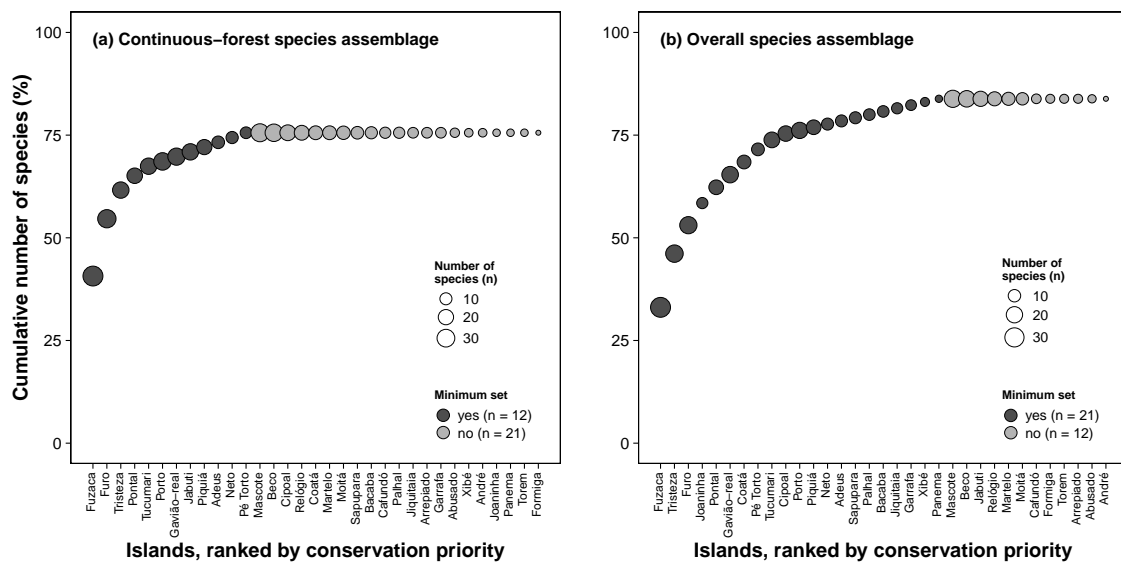
Figure 4.5: Cumulative number of understory bird species occurring on forest islands considering both the (a) continuous-forest and the (b) overall species assemblage. Forest islands are ranked from those adding the largest to the smallest number of unrepresented species in the entire metacommunity. Dark gray circles represent forest islands included within the solution of the minimum set problem: a set of forest islands whose avifauna complement each other and collectively capture the largest number of species within the fewest number of forest islands. Light gray circles represent forest islands that fail to add new species to the metacommunity.

4.5   Discussion

Up to 82% of the world's forests are degraded to some extent (Watson et al. 2018) and over 70% lie within 1 km of a forest edge, and are therefore exposed to edge-mediated disturbance (Haddad et al., 2015). Most remaining forest areas world-wide could be deemed as 'shifting baselines' compared to their once primeval ecological condition. Accordingly, both the richness and abundance of sensitive species are expected to decline, whereas those of disturbance-adapted species are expected to increase, thereby resulting in a shift in community structure (i.e. changes in species composition and population sizes) with some relict species populations retained. In particular, this has been shown for birds (Sigel et al. 2006), amphibians and reptiles (Whitfield et al., 2007), and dung beetles (Escobar et al., 2008) at La Selva Biological Station (*c.* 1,600 ha), an intensively studied tropical forest reserve in Costa Rica. Such shifts in community structure have been attributed to forest shrinkage and isolation induced by surrounding land-use change (Escobar et al., 2008; Sigel et al., 2006), and climate-driven reduction in microhabitat resources (Whitfield et al., 2007). Some authors have therefore cautioned against the use of La Selva as an "intact" tropical forest baseline for neighboring modified habitats, given its declining biodiversity conservation value over 35-40 years (Escobar et al., 2008; Sigel et al., 2006). Likewise, avian declines and extirpations have also been documented over 85 years in the Panamanian forest reserve of Barro Colorado Island (*c.* 1,600 ha), the longest and best studied tropical forest fragment (Robinson, 1999). These two examples of shifting baselines illustrate that isolated tropical forest reserves often fail to preserve "intact" species assemblages in the long run (Sigel et al., 2010), even if they are well protected.

Biodiversity comparisons between reference and impacted sites are likely to be severely biased if reference sites are significantly degraded or reduced in extent, rather than represented by a relatively suitable baseline of continuous primary habitat. As a result, shifting baselines reduce the contrast between reference and impacted sites, leading to unduly optimistic diagnostics of the magnitude of impacts as we have shown here (Fig. 4.4; Table S4.3). For instance, the amphibian species richness of a *c.* 250-ha shifting baseline forest reserve in a southern Costa Rican landscape was not significantly different from that of either countryside forest elements (e.g. small forest fragments, live fences, hedgerows, riparian strips) or crop fields and pastures (Mendenhall et al., 2016). In contrast, the amphibian species richness of primary forest sites larger than 1 Mha (i.e. suitable baseline) in northeastern Brazilian Amazonia was higher than that in either secondary forests or eucalypt plantations retaining a native understory (Barlow et al., 2007). Apart from the role of suitable baselines to properly assess the scale

of impacts, the authors of two meta-analyses have either considered sites larger than 100 ha as 'minimally altered forests' (Mendenhall et al., 2016) or have not defined their minimum size (Mendenhall et al., 2014). Nonetheless, forest intactness (i.e. baseline suitability) depends not only on levels of habitat disturbance (Barlow et al., 2016), but also on its overall spatial extent (Potapov et al., 2017), since protected areas larger than 1 Mha are required to support full complements of species and landscape-scale ecological processes in tropical forests (Peres, 2005).

We caution against the naïve use of the "best" locally available reference sites whenever those are not representative of a suitable baseline (Cardinale et al., 2018), and emphasize the need to describe in future studies the size and the level of structural and/or compositional habitat disturbance of reference sites. For example, Hannah et al. (2007) found no differences in bird species richness and composition between fragments (¡ 300 ha) and reference sites (2,500-40,000 ha) of eucalypt woodlands in Australia. However, these authors acknowledged that almost their entire study area had been subjected to about 150 years of habitat disturbance primarily associated with pastoralism, and that the species assemblage they sampled likely represented a relatively resilient relict subset of the once "intact" species assemblage. In other words, environmental impact assessments should be explicitly interpreted by considering the condition of the baseline.

Global meta-analyses reporting no net losses in local-scale species richness over time (Dornelas et al., 2014; Vellend et al., 2013) have been criticised because their generalities are spatially biased and lack appropriate baselines (Gonzalez et al., 2016). For instance, most of the data collated by Vellend et al. (2013) were derived from studies carried out in the United States and Europe (Gonzalez et al., 2016). In fact, land-use change is not the main threat to biodiversity in northern temperate ecosystems where most primary habitat has already been lost (Sala et al. 2000). In the case of Europe, the lack of intact temperate forests is so dramatic that only the *c.* 150,000-ha Białowieża Forest in the Polish-Belarussian border remains in near-primeval conditions (Wesołowski, 2007). Collectively, this means that worldwide claims of no net loss in local species richness are heavily influenced by studies lacking an appropriate baseline (Dornelas et al., 2014; Vellend et al., 2013).

Given the pace of habitat loss and degradation across the globe, most remaining suitable baseline sites will likely become shifting baselines (Watson et al., 2018), ultimately preventing realistic assessments of the impacts of land-use change on biodiversity. As previously pointed out, "what we need are datasets that have clear baselines that tell us what expected values of biodiversity are" (Cardinale et al., 2018). We therefore urge prioritizing biodiversity surveys in

the last remaining undisturbed areas of any biome to establish solid baselines prior to anthropogenic disturbance (Bobrowiec and Tavares, 2017; Ritter et al., 2017), especially in the tropics where an unknown but vast number of undescribed species exist, and large forest tracts still remain intact (Watson et al., 2018). However, wherever suitable baselines are unavailable or located far away from impacted sites, baseline species assemblages can be coarsely determined using data from natural history collections (Lister 2011); both formal and gray literature, and species databases (e.g. Bogoni et al. 2018); species distribution range maps, interviews with local people, and expert information (e.g. Canale et al. 2012). Alternatively, we recommend environmental impacts to be estimated from species groups of high conservation concern (e.g. habitat specialists, sensitive to disturbance, threatened with extinction).

Despite the steady or increasing species richness following anthropogenic disturbance in some regions (e.g. northern temperate ecosystems), such trends are often driven by a replacement of sensitive species by disturbance-adapted species (Tabarelli et al., 2012). Non-random species responses to land-use change (e.g. patch- and landscape-scale forest loss) has led some researchers to either disentangle the responses of habitat specialists from non-habitat specialists (Lövei et al., 2006) or apply metrics of compositional changes (Banks-Leite et al., 2012), since the overall species richness may obscure the effects of habitat conversion on biodiversity.

Here, we used the number of species at any given baseline that was retained within impacted sites (*number of relict species*) as a response variable. This avoids the non-trivial task of assigning species specificity to different habitat types and the nuisance of including disturbance-adapted species that often proliferates at impacted sites. The *number of relict species* is still a measure of species richness, with several advantages: it is simple to collect, intuitive to interpret, and easy to compare across studies (Banks-Leite et al., 2012). In particular, we showed a significant difference between reference and impacted sites if the *number of relict species* is used as a metric of biodiversity, regardless of the baseline condition (e.g. continuous forest and larger islands). In contrast, by focusing on the *overall species assemblage*, the negative effect of forest fragmentation on species richness was only apparent if continuous forest sites were defined as the baseline (Fig. 4.4). A combination of shifting baselines and the inclusion of disturbance-adapted species from impacted sites can lead to a higher perceived conservation value for impacted sites than for reference sites, if species richness is used as a biodiversity metric (Fig. 4.4). For example, Blake and Loiselle (2001) found an overall higher number of understory bird species (including disturbance-adapted species) in young second-growth at La Selva, northern Costa Rica, compared to

neighboring old-growth forest. We reanalysed their data using the *number of relict species* as the response variable to compare young second-growth, old second-growth and old-growth forests at La Selva. In doing so, we found that old-growth forest retained the highest number of species compared to either age classes of second-growth forest patches (Fig. S4.1). Essentially, the *number of relict species* is a measure of species richness lost from the reference baseline site, which by definition holds the maximum biodiversity value. Therefore, we believe the *number of relict species* is a reliable and straightforward biodiversity metric to quantify the residual conservation value of human-modified habitat patches in anthropogenic landscapes (fig. 6 in Gardner et al. 2009), even when baseline sites that could be deemed as intact are unavailable. Notwithstanding, other components of biological diversity (e.g. functional and phylogenetic diversity) should also be incorporated into environmental impact assessments to inform decision-making and fine-tune conservation strategies (Moreno et al., 2017).

A cost-effective virtue of focusing on the *continuous-forest species assemblage* is that it reduces the conservation effort that could be allocated to highly complementary priority sites – here represented by the minimum number of forest islands that could be protected to maximize the number of species retained across the entire landscape – compared to the number of target sites prioritized based on the *overall species assemblage*. This can be simply explained by the fact that the former (*continuous-forest*) species assemblage is a subset of the latter (*overall*). Furthermore, solutions for the minimum set problem depend on the target species group (Howard et al., 2000). For example, our smallest surveyed island (Joaninha, 0.63 ha) was ranked as the fourth most important in the minimum set selection based on the *overall species assemblage*, but it was not included in the minimum set based on the *continuous-forest species assemblage* (Fig. 4.5), essentially because this island failed to retain any species assigned to 'high' forest dependency (*sensu* BirdLife International 2018). In other words, the number, identity and conservation priority of the forest islands included in solution of the minimum set problem were all affected by which species assemblage was targeted. However, we underline that even a forest fragment larger than 7,500 ha (the total area of all 33 surveyed islands) failed to sustain the entire avifauna recorded at our continuous forest sites, reinforcing the notion that there is no substitute for large areas of unbroken primary forest to safeguard primary forest biodiversity (Gibson et al., 2011; Meyer et al., 2015). Notwithstanding, these outcomes support our claim that a focus on species assemblages derived from reference sites provides the best possible impact assessment approach in examining the effects of land-use change on biodiversity, while also establishing conservation strategies to compensate for those impacts.

The magnitude of the impacts of land-use change on biodiversity is measured as the extent to which impacted sites deviates from an assumed reference condition (i.e. baseline), with greater deviations leading to higher impact diagnostics. The suitability of the baseline is therefore likely to affect the estimated magnitude of the impacts, thereby resulting in realistic assessments whenever the baseline is suitable, but unduly optimistic assessments whenever the baseline was historically degraded. However, genuine optimistic assessments based on suitable baselines may arise if impacted sites are embedded in biodiversity-friendly landscapes (e.g. primary forest fragments surrounded by old-growth secondary forest; Stouffer et al. 2011) and/or the focal species assemblage is not sensitive to the impact being assessed (e.g. temperate birds are less likely to be negatively affected by forest fragmentation than tropical birds; Lindell et al. 2007).

Although the issue on shifting baselines has already been raised (Gardner et al., 2009), here we provide quantitative evidence on the pivotal role of suitable baselines in deriving reliable assessments of the impacts of land-use change on biodiversity. Accordingly, defining even well-preserved forest islands as large as 1,000 ha as a reference condition significantly reduced impact estimates compared to baselines consisting of undisturbed continuous forest, which reinforces the notion that the most pronounced biodiversity losses occur at the initial phases of deforestation in relatively intact landscapes (Betts et al. 2017). Since suitable baselines are regrettably no longer available in most regions worldwide, we emphasize the need to consider the size and level of habitat disturbance of comparable reference sites to interpret the outcomes of environmental impact assessments (Hannah et al., 2007; Ritter et al., 2017). We also showed that a focus on species assemblages derived from reference sites alone, rather than the overall species assemblage, has two consequences. First, there was a negative impact of forest fragmentation on species richness even when we defined the most degraded reference sites (100-ha forest islands) as the comparative baseline. Second, allocation of conservation investments could be considerably reduced if one targets only those species that presumably occupied any given site prior to habitat disturbance, which are likely those of higher conservation concern. We therefore conclude that environmental impact assessments should focus on species assemblages derived from suitable baselines if they are to be reliable.

4.6 Acknowledgments

4.7 References

1. Alvares, C.A., Stape, J.L., Sentelhas, P.C., Gonçalves, J.L. de M., Sparovek, G., 2013. Köppen's climate classification map for Brazil. Meteorol. Zeitschrift 22, 711–728. https://doi.org/10.1127/0941-2948/2013/0507

2. Aurélio-Silva, M., Anciães, M., Henriques, L.M.P., Benchimol, M., Peres, C.A., 2016. Patterns of local extinction in an Amazonian archipelagic avifauna following 25 years of insularization. Biol. Conserv. 199, 101–109. https://doi.org/http://dx.doi.org/10.1016/j.biocon.2016.03.016

3. Balmford, A., 1996. Extinction filters and current resilience: The significance of past selection pressures for conservation biology. Trends Ecol. Evol. 11, 193–196. https://doi.org/10.1016/0169-5347(96)10026-4

4. Banks-Leite, C., Ewers, R.M., Metzger, J.P., 2012. Unraveling the drivers of community dissimilarity and species extinction in fragmented landscapes. Ecology 93, 2560–2569. https://doi.org/10.1890/11-2054.1

5. Banks-Leite, C., Pardini, R., Tambosi, L.R., Pearse, W.D., Bueno, A.A., Bruscagin, R.T., Condez, T.H., Dixo, M., Igari, A.T., Martensen, A.C., Metzger, J.P., 2014. Using ecological thresholds to evaluate the costs and benefits of set-asides in a biodiversity hotspot. Science (80-.). 345, 1041–1045. https://doi.org/10.1126/science.1255768

6. Barlow, J., Gardner, T.A., Araujo, I.S., Avila-Pires, T.C., Bonaldo, A.B., Costa, J.E., Esposito, M.C., Ferreira, L. V, Hawes, J., Hernandez, M.I.M., Hoogmoed, M.S., Leite, R.N., Lo-Man-Hung, N.F., Malcolm, J.R., Martins, M.B., Mestre, L. a M., Miranda-Santos, R., Nunes-Gutjahr, a L., Overal, W.L.,

Parry, L., Peters, S.L., Ribeiro-Junior, M.A., da Silva, M.N.F., da Silva Motta, C., Peres, C.A., 2007. Quantifying the biodiversity value of tropical primary, secondary, and plantation forests. Proc. Natl. Acad. Sci. 104, 18555–18560. https://doi.org/10.1073/pnas.0703333104

7. Barlow, J., Lennox, G.D., Ferreira, J., Berenguer, E., Lees, A.C., Nally, R. Mac, Thomson, J.R., Ferraz, S.F.D.B., Louzada, J., Oliveira, V.H.F., Parry, L., Ribeiro de Castro Solar, R., Vieira, I.C.G., Aragão, L.E.O.C., Begotti, R.A., Braga, R.F., Cardoso, T.M., Jr, R.C.D.O., Souza Jr, C.M., Moura, N.G., Nunes, S.S., Siqueira, J.V., Pardini, R., Silveira, J.M., Vaz-de-Mello, F.Z., Veiga, R.C.S., Venturieri, A., Gardner, T.A., 2016. Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. Nature 535, 144–147. https://doi.org/10.1038/nature18326

8. Benchimol, M., Peres, C.A., 2015. Edge-mediated compositional and functional decay of tree assemblages in Amazonian forest islands after 26 years of isolation. J. Ecol. 103, 408–420. https://doi.org/10.1111/1365-2745.12371

9. BirdLife International, 2018. IUCN Red List for birds. Downloaded from http//www.birdlife.org April 2018.

10. Blake, J.G., Loiselle, B.A., 2001. Bird assemblages in second-growth and old-growth forests, Costa Rica: Perspectives from mist nets and point counts. Auk 118, 304–326. https://doi.org/10.1642/0004-8038(2001)118[0304:BAISGA]2.0.CO;2

11. Bobrowiec, P.E.D., Tavares, V.D.C., 2017. Establishing baseline biodiversity data prior to hydroelectric dam construction to monitoring impacts to bats in the Brazilian Amazon. PLoS One 12, e0183036. https://doi.org/10.1371/journal.pone.0183036

12. Boyle, W.A., Sigel, B.J., 2015. Ongoing changes in the avifauna of La Selva Biological Station, Costa Rica: Twenty-three years of Christmas Bird Counts. Biol. Conserv. 188, 11–21. https://doi.org/10.1016/j.biocon.2015.01.004

13. Cardinale, B.J., Gonzalez, A., Allington, G.R.H., Loreau, M., 2018. Is local biodiversity declining or not? A summary of the debate over analysis of species richness time trends. Biol. Conserv. 219, 175–183. https://doi.org/10.1016/j.biocon.2017.12.021

14. Chazdon, R.L., Peres, C.A., Dent, D., Sheil, D., Lugo, A.E., Lamb, D., Stork, N.E., Miller, S.E., 2009. The potential for species conser-

vation in tropical secondary forests. Conserv. Biol. 23, 1406–1417. https://doi.org/10.1111/j.1523-1739.2009.01338.x

15. Dornelas, M., Gotelli, N.J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C., Magurran, A.E., 2014. Assemblage time series reveal biodiversity change but not systematic loss. Science (80-.). 344, 296–299. https://doi.org/10.1126/science.1248484

16. Edwards, D.P., Woodcock, P., Edwards, F.A., Larsen, T.H., Hsu, W.W., Benedick, S., Wilcove, D.S., 2012. Reduced-impact logging and biodiversity conservation: A case study from Borneo. Ecol. Appl. 22, 561–571. https://doi.org/10.1890/11-1362.1

17. Escobar, F., Halffter, G., Solís, Á., Halffter, V., Navarrete, D., 2008. Temporal shifts in dung beetle community structure within a protected area of tropical wet forest: A 35-year study and its implications for long-term conservation. J. Appl. Ecol. 45, 1584–1592. https://doi.org/10.1111/j.1365-2664.2008.01551.x

18. 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. https://doi.org/10.1016/j.biocon.2009.06.022

19. Fahrig, L., 2003. Effects of habitat fragmentation on biodiversity. Annu. Rev. Ecol. Evol. Syst. 34, 487–515. https://doi.org/10.1146/annurev.ecolsys.34.011802.132419

20. Fearnside, P.M., 2016. Environmental and social impacts of hydroelectric dams in Brazilian Amazonia: Implications for the aluminum industry. World Dev. 77, 48–65. https://doi.org/10.1016/j.worlddev.2015.08.015

21. Gardner, T.A., Barlow, J., Chazdon, R., Ewers, R.M., Harvey, C.A., Peres, C.A., Sodhi, N.S., 2009. Prospects for tropical forest biodiversity in a human-modified world. Ecol. Lett. 12, 561–582. https://doi.org/10.1111/j.1461-0248.2009.01294.x

22. Gibson, L., Lee, T.M., Koh, L.P., Brook, B.W., Gardner, T.A., Barlow, J., Peres, C.A., Bradshaw, C.J.A., Laurance, W.F., Lovejoy, T.E., Sodhi, N.S., 2011. Primary forests are irreplaceable for sustaining tropical biodiversity. Nature 478, 378–381. https://doi.org/10.1038/nature10425

23. Gonzalez, A., Cardinale, B.J., Allington, G.R.H., Byrnes, J., Arthur Endsley, K., Brown, D.G., Hooper, D.U., Isbell, F., O'Connor, M.I., Loreau, M., 2016.

Estimating local biodiversity change: a critique of papers claiming no net loss of local diversity. Ecology 97, 1949–1960. https://doi.org/10.1890/15-1759.1

24. Haddad, N.M., Brudvig, L.A., Clobert, J., Davies, K.F., Gonzalez, A., Holt, R.D., Lovejoy, T.E., Sexton, J.O., Austin, M.P., Collins, C.D., Cook, W.M., Damschen, E.I., Ewers, R.M., Foster, B.L., Jenkins, C.N., King, A.J., Laurance, W.F., Levey, D.J., Margules, C.R., Melbourne, B.A., Nicholls, A.O., Orrock, J.L., Song, D.-X., Townshend, J.R., 2015. Habitat fragmentation and its lasting impact on Earth's ecosystems. Sci. Adv. 1, e1500052–e1500052. https://doi.org/10.1126/sciadv.1500052

25. Hannah, D., Woinarski, J.C.Z., Catterall, C.P., Mccosker, J.C., Thurgate, N.Y., Fensham, R.J., 2007. Impacts of clearing, fragmentation and disturbance on the bird fauna of Eucalypt savanna woodlands in central Queensland, Australia. Austral Ecol. 32, 261–276. https://doi.org/10.1111/j.1442-9993.2007.01683.x

26. Howard, P.C., Davenport, T.R.B., Kigenyi, F.W., Viskanic, P., Baltzer, M.C., Dickinson, C.J., Lwanga, J., Matthews, R.A., Mupada, E., 2000. Protected area planning in the tropics: Uganda's National System of Forest Nature Reserves. Conserv. Biol. 14, 858–875. https://doi.org/10.1046/j.1523-1739.2000.99180.x

27. Howard, P.C., Viskanic, P., Davenport, T.R.B., Kigenyi, F.W., Baltzer, M., Dickinson, C.J., Lwanga, J.S., Matthews, R.A., Balmford, A., 1998. Complementarity and the use of indicator groups for reserve selection in Uganda. Nature 394, 472–475. https://doi.org/10.1038/28843

28. Hsieh, T.C., Ma, K.H., Chao, A., 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods Ecol. Evol. 7, 1451–1456. https://doi.org/10.1111/2041-210X.12613

29. Jackson, J.B.C., Kirby, M.X., Berger, W.H., Bjornda, K.A., Botsford, L.W., Bourque, B.J., Bradbury, R.H., Cooke, R., Erlandson, J., Estes, J.A., Hughes, T.P., Kidwell, S., Lange, C.B., Lenihan, H.S., Pandolfi, J.M., Peterson, C.H., Steneck, R.S., Tegner, M.J., Warner, R.R., 2001. Historical overfishing and the recent collapse of coastal ecosystems. Science (80-.). 293, 629–637. https://doi.org/10.1126/science.1059199

30. Lövei, G.L., Magura, T., Tóthmérész, B., Ködöböcz, V., 2006. The influence of matrix and edges on species richness patterns of ground beetles

(Coleoptera: Carabidae) in habitat islands. Glob. Ecol. Biogeogr. 15, 283–289. https://doi.org/10.1111/j.1466-822X.2005.00221.x

31. Mendenhall, C.D., Karp, D.S., Meyer, C.F.J., Hadly, E.A., Daily, G.C., 2014. Predicting biodiversity change and averting collapse in agricultural landscapes. Nature 509, 213–217. https://doi.org/10.1038/nature13139

32. Mendenhall, C.D., Shields-Estrada, A., Krishnaswami, A.J., Daily, G.C., 2016. Quantifying and sustaining biodiversity in tropical agricultural landscapes. Proc. Natl. Acad. Sci. 113, 14544–14551. https://doi.org/10.1073/pnas.1604981113

33. Meyer, N.F.V., Esser, H.J., Moreno, R., van Langevelde, F., Liefting, Y., Ros Oller, D., Vogels, C.B.F., Carver, A.D., Nielsen, C.K., Jansen, P.A., 2015. An assessment of the terrestrial mammal communities in forests of Central Panama, using camera-trap surveys. J. Nat. Conserv. 26, 28–35. https://doi.org/10.1016/j.jnc.2015.04.003

34. 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. https://doi.org/10.1371/journal.pone.0128923

35. Moreno, C.E., Calderón-Patrón, J.M., Arroyo-Rodríguez, V., Barragán, F., Escobar, F., Gómez-Ortiz, Y., Martín-Regalado, N., Martínez-Falcón, A.P., Martínez-Morales, M.Á., Mendoza, E., Ortega-Martínez, I.J., Pérez-Hernández, C.X., Pineda, E., Pineda-López, R., Rios-Díaz, C.L., Rodríguez, P., Rosas, F., Schondube, J.E., Zuria, I., 2017. Measuring biodiversity in the Anthropocene: a simple guide to helpful methods. Biodivers. Conserv. 1–6. https://doi.org/10.1007/s10531-017-1401-1

36. Papworth, S.K., Rist, J., Coad, L., Milner-Gulland, E.J., 2009. Evidence for shifting baseline syndrome in conservation. Conserv. Lett. 2, 93–100. https://doi.org/10.1111/j.1755-263X.2009.00049.x

37. Peres, C.A., 2005. Why we need megareserves in Amazonia. Conserv. Biol. 19, 728–733. https://doi.org/10.1111/j.1523-1739.2005.00691.x

38. Potapov, P., Hansen, M.C., Laestadius, L., Turubanova, S., Yaroshenko, A., Thies, C., Smith, W., Zhuravleva, I., Komarova, A., Minnemeyer, S., Esipova, E., 2017. The last frontiers of wilderness: Tracking loss of intact forest landscapes from 2000 to 2013. Sci. Adv. 3, e1600821. https://doi.org/10.1126/sciadv.1600821

39. R Core Team, 2018. R: A language and environment for statistical computing. Version 3.5.1. R Foundation for Statistical Computing, Vienna, Austria.

40. Ritter, C.D., McCrate, G., Nilsson, R.H., Fearnside, P.M., Palme, U., Antonelli, A., 2017. Environmental impact assessment in Brazilian Amazonia: Challenges and prospects to assess biodiversity. Biol. Conserv. 206, 161–168. https://doi.org/10.1016/j.biocon.2016.12.031

41. Robinson, W.D., 1999. Long-term changes in the avifauna of Barro Colorado Island, Panama, a tropical forest isolate. Conserv. Biol. 13, 85–97. https://doi.org/10.1046/j.1523-1739.1999.97492.x

42. Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., Leemans, R., Lodge, D.M., Mooney, H.A., Oesterheld, M., Poff, N.L., Sykes, M.T., Walker, B.H., Walker, M., Wall, D.H., 2000. Global biodiversity scenarios for the year 2100. Science (80-.). 287, 1770–1774. https://doi.org/10.1126/science.287.5459.1770

43. Sigel, B.J., Douglas Robinson, W., Sherry, T.W., 2010. Comparing bird community responses to forest fragmentation in two lowland Central American reserves. Biol. Conserv. 143, 340–350. https://doi.org/10.1016/j.biocon.2009.10.020

44. Sigel, B.J., Sherry, T.W., Young, B.E., 2006. Avian community response to lowland tropical rainforest isolation: 40 years of change at La Selva Biological Station, Costa Rica. Conserv. Biol. 20, 111–121. https://doi.org/10.1111/j.1523-1739.2006.00293.x

45. Sutherland, W.J., Adams, W.M., Aronson, R.B., Aveling, R., Blackburn, T.M., Broad, S., Ceballos, G., Côté, I.M., Cowling, R.M., Da fonseca, G.A.B., Dinerstein, E., Ferraro, P.J., Fleishman, E., Gascon, C., Hunter Jr., M., Hutton, J., Kareiva, P., Kuria, A., Macdonald, D.W., Mackinnon, K., Madgwick, F.J., Mascia, M.B., Mcneely, J., Milner-Gulland, E.J., Moon, S., Morley, C.G., Nelson, S., Osborn, D., Pai, M., Parsons, E.C.M., Peck, L.S., Possingham, H., Prior, S. V., Pullin, A.S., Rands, M.R.W., Ranganathan, J., Redford, K.H., Rodriguez, J.P., Seymour, F., Sobel, J., Sodhi, N.S., Stott, A., Vance-Borland, K., Watkinson, A.R., 2009. One hundred questions of importance to the conservation of global biological diversity. Conserv. Biol. 23, 557–567. https://doi.org/10.1111/j.1523-1739.2009.01212.x

46. Tabarelli, M., Peres, C.A., Melo, F.P.L., 2012. The 'few winners and many losers' paradigm revisited: Emerging prospects

for tropical forest biodiversity. Biol. Conserv. 155, 136–140. https://doi.org/10.1016/j.biocon.2012.06.020

47. Vellend, M., Baeten, L., Myers-Smith, I.H., Elmendorf, S.C., Beausejour, R., Brown, C.D., De Frenne, P., Verheyen, K., Wipf, S., 2013. Global meta-analysis reveals no net change in local-scale plant biodiversity over time. Proc. Natl. Acad. Sci. 110, 19456–19459. https://doi.org/10.1073/pnas.1312779110

48. Watson, J.E.M., Evans, T., Venter, O., Williams, B., Tulloch, A., Stewart, C., Thompson, I., Ray, J.C., Murray, K., Salazar, A., McAlpine, C., Potapov, P., Walston, J., Robinson, J.G., Painter, M., Wilkie, D., Filardi, C., Laurance, W.F., Houghton, R.A., Maxwell, S., Grantham, H., Samper, C., Wang, S., Laestadius, L., Runting, R.K., Silva-Chávez, G.A., Ervin, J., Lindenmayer, D., 2018. The exceptional value of intact forest ecosystems. Nat. Ecol. Evol. 2, 599–610. https://doi.org/10.1038/s41559-018-0490-x

49. Wesołowski, T., 2007. Primeval conditions - what can we learn from them? Ibis (Lond. 1859). 149, 64–77. https://doi.org/10.1111/j.1474-919X.2007.00721.x

50. Whitfield, S.M., Bell, K.E., Philippi, T., Sasa, M., Bolanos, F., Chaves, G., Savage, J.M., Donnelly, M.A., 2007. Amphibian and reptile declines over 35 years at La Selva, Costa Rica. Proc. Natl. Acad. Sci. 104, 8352–8356. https://doi.org/10.1073/pnas.0611256104

4.8 Supporting Information

This supporting information contains:

- Table S4.1

- Table S4.2

- Table S4.3

- Figure S4.1

- R code

  Click here to download the R code in .Rmd file format

Table S4.1: Description of the 38 sample sites surveyed at the Balbina Hydroelectric Reservoir landscape, including 33 forest islands and five continuous forest sites, and the species richness and abundance of understory forest birds captured therein.

Site name	Latitude (S)	Longitude (W)	Forest area (ha)	Richness (n)	Abundance (n)
Joaninha	1°31'22"	59°49'44"	0.63	10	54
Xibé	1°28'08"	59°50'11"	0.91	6	9
Formiga	1°50'00"	59°25'16"	1.54	6	59
André	1°35'05"	59°52'19"	2.08	3	8
Cafundó	1°30'02"	59°36'37"	2.63	7	10
Panema	1°46'28"	59°41'33"	3.08	4	5
Torem	1°49'46"	59°37'57"	3.62	6	20
Pé Torto	1°45'59"	59°21'49"	4.98	14	38
Jiquitaia	1°50'08"	59°35'44"	6.79	10	30
Arrepiado	1°30'54"	59°44'21"	7.43	6	10
Garrafa	1°35'19"	59°50'08"	9.42	9	15
Piquiá	1°30'23"	59°47'20"	13.04	20	30
Abusado	1°45'44"	59°40'43"	13.31	5	29
Coatá	1°29'18"	59°47'13"	16.94	17	33
Palhal	1°47'25"	59°26'52"	21.37	11	44
Neto	1°50'29"	59°21'08"	32.87	12	56
Adeus	1°52'26"	59°39'45"	50.08	12	26
Bacaba	1°30'15"	59°49'20"	53.35	11	22
Relógio	1°40'55"	59°39'09"	70.55	18	70
Sapupara	1°41'50"	59°36'45"	77.8	12	32
Moitá	1°33'36"	59°53'33"	91.3	13	27
Pontal	1°49'52"	59°41'12"	108.76	20	49
Furo de Santa Luzia	1°44'25"	59°26'31"	198.52	30	79
Cipoal	1°41'54"	59°47'05"	217.63	23	71
Jabuti	1°37'34"	59°45'24"	232.49	22	58
Tucumari	1°35'29"	59°25'47"	336.02	24	63
Martelo	1°39'51"	59°42'51"	466.6	15	39
Tristeza	1°46'03"	59°45'16"	499.91	31	90
Beco do Catitu	1°44'22"	59°42'18"	638.66	27	74
Mascote	1°38'56"	59°49'58"	668.03	30	77
Fuzaca	1°29'33"	59°51'36"	941.71	43	121
Porto Seguro	1°46'48"	59°31'02"	1350.56	26	85
Gavião-real	1°35'46"	59°37'58"	1698.84	27	68
CF Waba	1°25'44"	59°54'19"	∞	28	75
CF Loreno	1°50'06"	59°43'05"	∞	40	105
CF Tucumari	1°34'36"	59°23'09"	∞	40	111
CF Grid-NS2	1°47'41"	59°15'57"	∞	54	158
CF Grid-NS3	1°47'27"	59°15'24"	∞	52	165

Table S4.2: Description and number of reference sites used as a 'baseline condition' to estimate the impacts of forest fragmentation on the number of bird species.

Reference sites	Sample sites (n)	Forest area (ha)	Richness (n)
Continuous forest	5	∞	86
1,000 ha	3	942–1,699	56
500 ha	4	467–668	53
250 ha	4	199–336	46
100 ha	4	71–109	35

Table S4.3: Estimated impact of forest fragmentation on species richness measured for five reference sites (Fig. 4.3), representing a gradient of baseline suitability from the most (continuous forest) to the least suitable (100-ha forest islands). Asterisks (*) denote a significant difference in species richness between reference and impacted sites. Higher values indicate more severe impacts of forest fragmentation on species richness (Fig. 4.4).

Reference sites	Species assemblage	
	Baseline	Overall
Continuous forest	0.67*	0.29*
1,000 ha	0.58*	0.08
500 ha	0.56*	0.01
250 ha	0.52*	−0.19
100 ha	0.50*	−0.43*
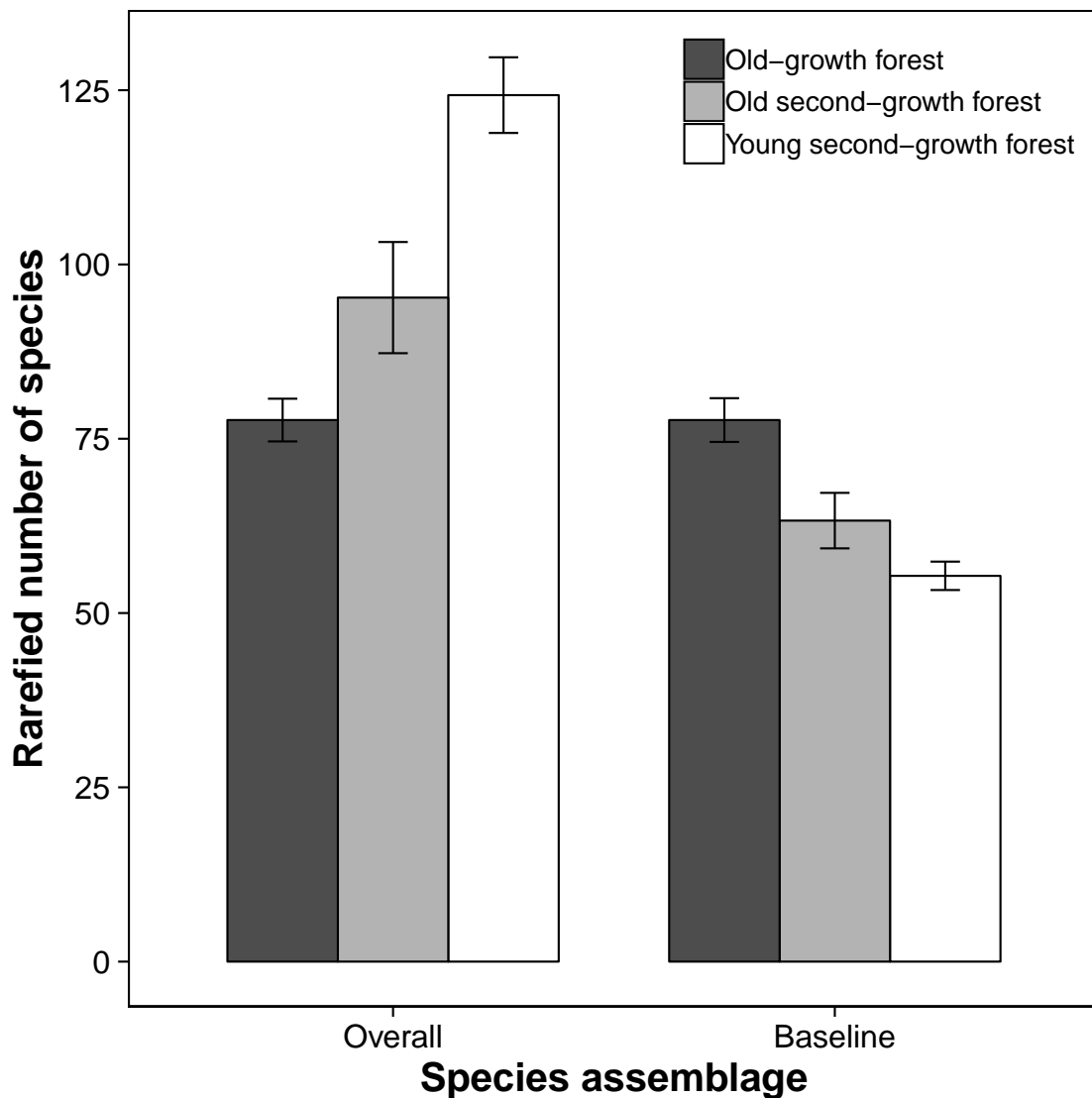
Figure S4.1: Comparisons between the rarefied number of understory bird species in northern Costa Rica across three successional forest types considering both the overall and the baseline species assemblage. The latter is derived from the neighboring La Selva Biological Station, whose old-growth forest was used as the reference site as in the original study (data from Blake and Loiselle 2001). The graph shows that a focus on the overall species assemblage results in a positive effect of land-use change (from old-growth forest to young second-growth forest) on the number of species (as reported in Blake and Loiselle 2001). In contrast, a focus on the baseline species assemblage results in a negative effect with either old second-growth or young second-growth forest harboring a more depauperate set species captured at the reference site. Rarefied number of species were standardized by sample coverage (0.981). Error bars indicate the 95% confidence intervals.

Supporting Information

The role of baseline suitability in assessing the impacts of land-use change on biodiversity

```
# Clear workspace and disable scientific notation
remove(list = ls()); options(scipen = 999)
```

Packages

```
# Load required packages
library(vegan)
library(iNEXT)
library(ggplot2)
library(gridExtra)
```

Dataset

We provide the raw data derived from original fieldwork conducted during the project **Ecological Impacts of River Damming on Forest Bird Assemblages in the Brazilian Amazon**. Find out more at the project webpage.

Data are available from the KNB repository and should be cited as:

Anderson Saldanha Bueno. 2018. Balbina Understory Bird Data from 2015 to 2016. Knowledge Network for Biocomplexity. urn:uuid:dbfd1504-2212-422c-8e04-610fb2327b7c.

Import data

- The file **balbina_understory_birds_captures.csv** contains information on birds captured in the fieldwork.
- The file **balbina_understory_birds_taxonomy_traits.csv** contains the taxonomy and traits of bird species captured.
- The file **balbina_environmental_variables.csv** contains environmental variables of sample sites.

```
# Load the file "balbina_understorey_birds_captures.csv"
birds = read.csv("https://ndownloader.figshare.com/files/15158552")

# Load the file "balbina_understorey_birds_taxonomy_traits.csv"
traits = read.csv("https://ndownloader.figshare.com/files/15158543")

# Load the file "balbina_environmental_variables.csv"
env = read.csv("https://ndownloader.figshare.com/files/15158549", row.names = 1)
```

Data handling

The raw data are stored in the long table format and include all individuals captured. This means that each bird captured is represented by a row. However, the analysis performed in this study requires the short table format, where sample sites are in rows and species are in columns.

```
# Add a column indicating that each bird captured (row) corresponds to an individual record
birds$occurrence = 1
```

```
# To facilitate sorting, we assigned area values of one order of magnitude
# greater than the largest surveyed island for continuous forest sites
env[is.na(env)] = 1698.84 * 10
```

The next step is to produce the table in the short format containing both bird and environmental data.

Three datasets will be used:

1. Birds captured in 2015 and 2016 (also referred to as "full").
2. Birds captured only in 2015 (also referred to as "2015").
3. Birds captured only in 2016 (also referred to as "2016").

Birds captured in 2015 and 2016

```
# Remove recaptures
birds.temp = subset(birds, birds$new.individual != "no")

# Site-by-species abundance matrix
birds.temp.matrix = tapply(birds.temp$occurrence,
                           list(birds.temp$site, birds.temp$species), sum)
# Species non detected in a given site are real zeros
birds.temp.matrix[is.na(birds.temp.matrix)] = 0

# Order rows by site name
#birds.temp.matrix = birds.temp.matrix[c(1:6, 12, 7:11, 13:38), ]

# Site-by-species abundance matrix with environmental data
# Add environmental data to the site-by-species abundance matrix
# 'balbina.full' stands for the full dataset (i.e. 2015 and 2016 combined)
balbina.full = cbind(birds.temp.matrix, env)
# Order rows by island size (i.e. forest area in hectares)
balbina.full = balbina.full[order(balbina.full$area.ha), ]
# Select the site-by-species abundance matrix and island size
balbina.full = balbina.full[c(1:ncol(birds.temp.matrix),
                              match("area.ha", names(balbina.full)))]
```

Species list

```
# Species frequency (number of sites) and abundance (number of individuals)
spp.list = data.frame(sites = colSums(ifelse(balbina.full[1:(ncol(balbina.full) -
    1)] > 0, 1, 0)), individuals = colSums(balbina.full[1:(ncol(balbina.full) -
    1)]))

# species taxonomy position of the species in the taxonomic sequence
spp.list$id = traits$id[match(rownames(spp.list), traits$cbro.2015)]
# species name
spp.list$species = rownames(spp.list)
# species family
spp.list$family = traits$family[match(rownames(spp.list), traits$cbro.2015)]

# rearrange the table
```

```
spp.list = spp.list[order(spp.list$id), c(5, 4, 1, 2)]
row.names(spp.list) = NULL

spp.list
```

```
##              family                 species sites individuals
## 1          Ardeidae      Zebrilus_undulatus     1           1
## 2       Accipitridae   Leucopternis_melanops     2           2
## 3        Columbidae     Leptotila_rufaxilla    20          26
## 4        Columbidae      Geotrygon_montana     5           6
## 5         Cuculidae        Coccycua_minuta     1           1
## 6         Strigidae      Glaucidium_hardyi     3           3
## 7       Trochilidae        Glaucis_hirsutus     7          11
## 8       Trochilidae       Phaethornis_ruber    15          27
## 9       Trochilidae   Phaethornis_bourcieri     8          23
## 10      Trochilidae Phaethornis_superciliosus    15          40
## 11      Trochilidae Campylopterus_largipennis     3           3
## 12      Trochilidae       Florisuga_mellivora     3           3
## 13      Trochilidae       Thalurania_furcata    23          54
## 14      Trochilidae       Amazilia_versicolor     4           5
## 15      Trochilidae       Heliothryx_auritus     4           4
## 16      Trogonidae          Trogon_viridis     1           1
## 17      Trogonidae           Trogon_rufus     3           4
## 18      Alcedinidae      Chloroceryle_aenea     3           4
## 19      Alcedinidae  Chloroceryle_americana     1           1
## 20       Momotidae          Momotus_momota    12          16
## 21      Galbulidae      Galbula_albirostris    11          35
## 22      Galbulidae       Jacamerops_aureus     1           1
## 23      Bucconidae          Bucco_tamatia     1           1
## 24      Bucconidae          Bucco_capensis     1           1
## 25      Bucconidae       Malacoptila_fusca     3           4
## 26      Bucconidae           Monasa_atra     8          14
## 27     Ramphastidae   Ramphastos_vitellinus     2           2
## 28     Ramphastidae    Pteroglossus_viridis     1           1
## 29         Picidae      Veniliornis_cassini     1           2
## 30         Picidae          Celeus_elegans    13          15
## 31         Picidae      Dryocopus_lineatus     1           1
## 32         Picidae  Campephilus_rubricollis     3           3
## 33      Falconidae     Micrastur_ruficollis     1           1
## 34      Falconidae    Micrastur_gilvicollis     3           4
## 35     Psittacidae          Pyrrhura_picta     1           3
## 36  Thamnophilidae      Myrmornis_torquata     2           2
## 37  Thamnophilidae Epinecrophylla_gutturalis     5          11
## 38  Thamnophilidae    Myrmotherula_axillaris    28         203
## 39  Thamnophilidae  Myrmotherula_longipennis     1           7
## 40  Thamnophilidae  Myrmotherula_menetriesii     4           5
## 41  Thamnophilidae         Isleria_guttata    18          66
## 42  Thamnophilidae    Thamnomanes_ardesiacus     7          24
## 43  Thamnophilidae      Thamnomanes_caesius     6          25
## 44  Thamnophilidae     Thamnophilus_murinus    10          18
## 45  Thamnophilidae   Thamnophilus_punctatus     2           3
## 46  Thamnophilidae Thamnophilus_melanothorax     1           1
## 47  Thamnophilidae     Cymbilaimus_lineatus     1           1
## 48  Thamnophilidae    Myrmoderus_ferrugineus     7           8
```

```
## 49     Thamnophilidae       Hylophylax_naevius        3          5
## 50     Thamnophilidae     Percnostola_rufifrons       20         69
## 51     Thamnophilidae   Cercomacroides_tyrannina       3         10
## 52     Thamnophilidae       Hypocnemis_cantator       18         86
## 53     Thamnophilidae         Pithys_albifrons         6         33
## 54     Thamnophilidae    Willisornis_poecilinotus      12         42
## 55     Thamnophilidae     Gymnopithys_rufigula         4         13
## 56     Conopophagidae        Conopophaga_aurita         2          6
## 57       Grallariidae     Myrmothera_campanisona        1          1
## 58       Formicariidae       Formicarius_colma         6          6
## 59       Formicariidae       Formicarius_analis        4          4
## 60        Scleruridae      Sclerurus_macconnelli        2          5
## 61        Scleruridae      Sclerurus_rufigularis        6          6
## 62        Scleruridae      Sclerurus_caudacutus        1          1
## 63    Dendrocolaptidae   Dendrocincla_fuliginosa       21         58
## 64    Dendrocolaptidae  Sittasomus_griseicapillus       2          2
## 65    Dendrocolaptidae  Certhiasomus_stictolaemus       7         17
## 66    Dendrocolaptidae    Glyphorynchus_spirurus       21        135
## 67    Dendrocolaptidae   Xiphorhynchus_pardalotus      26         91
## 68    Dendrocolaptidae    Xiphorhynchus_obsoletus       5          8
## 69    Dendrocolaptidae       Nasica_longirostris        1          1
## 70    Dendrocolaptidae     Dendrocolaptes_certhia      23         47
## 71    Dendrocolaptidae      Hylexetastes_perrotii       2          2
## 72        Xenopidae           Xenops_minutus          8         13
## 73        Furnariidae      Automolus_ochrolaemus        3          3
## 74        Furnariidae       Automolus_infuscatus        4          9
## 75        Furnariidae      Philydor_erythrocercum       3          4
## 76        Furnariidae       Philydor_pyrrhodes         5         12
## 77        Furnariidae       Synallaxis_rutilans        1          3
## 78         Pipridae      Neopelma_chrysocephalum       1          1
## 79         Pipridae       Tyranneutes_virescens        1          2
## 80         Pipridae    Ceratopipra_erythrocephala     10         16
## 81         Pipridae         Lepidothrix_serena        4         12
## 82         Pipridae          Manacus_manacus          3          3
## 83         Pipridae           Dixiphia_pipra         30        193
## 84     Onychorhynchidae   Onychorhynchus_coronatus      3          7
## 85     Onychorhynchidae   Terenotriccus_erythrurus      5          9
## 86     Onychorhynchidae       Myiobius_barbatus        8         19
## 87        Tityridae        Schiffornis_turdina        6         11
## 88        Cotingidae       Lipaugus_vociferans        4          4
## 89       Platyrinchidae    Platyrinchus_saturatus       3          5
## 90       Platyrinchidae    Platyrinchus_coronatus       6         22
## 91       Platyrinchidae  Platyrinchus_platyrhynchos     1          1
## 92     Rhynchocyclidae     Mionectes_oleagineus       12         23
## 93     Rhynchocyclidae     Mionectes_macconnelli      10         27
## 94     Rhynchocyclidae      Corythopis_torquatus        3          6
## 95     Rhynchocyclidae     Rhynchocyclus_olivaceus       3          3
## 96     Rhynchocyclidae      Tolmomyias_assimilis        1          1
## 97     Rhynchocyclidae    Tolmomyias_poliocephalus      2          2
## 98     Rhynchocyclidae      Hemitriccus_zosterops       1          1
## 99     Rhynchocyclidae      Lophotriccus_vitiosus       1          1
## 100    Rhynchocyclidae      Lophotriccus_galeatus       3          3
## 101       Tyrannidae         Attila_spadiceus        13         16
## 102       Tyrannidae      Ramphotrigon_ruficauda       2          2
```

```
## 103        Tyrannidae       Myiarchus_tuberculifer      2          2
## 104        Tyrannidae              Myiarchus_ferox      2          3
## 105        Tyrannidae           Rhytipterna_simplex      3          3
## 106        Tyrannidae           Pitangus_sulphuratus     1          1
## 107         Vireonidae    Tunchiornis_ochraceiceps       4          9
## 108      Troglodytidae          Troglodytes_musculus    11         89
## 109      Troglodytidae            Pheugopedius_coraya     3          4
## 110      Troglodytidae     Henicorhina_leucosticta       4         10
## 111      Troglodytidae             Cyphorhinus_arada      2          3
## 112      Polioptilidae     Ramphocaenus_melanurus        1          1
## 113          Turdidae           Catharus_fuscescens      2          2
## 114          Turdidae             Turdus_albicollis     19         62
## 115      Passerellidae         Ammodramus_humeralis      1          1
## 116      Passerellidae         Ammodramus_aurifrons      2          7
## 117      Passerellidae          Arremon_taciturnus       2          2
## 118          Icteridae                  Cacicus_cela      1          1
## 119        Thraupidae              Paroaria_gularis      1          2
## 120        Thraupidae              Tangara_palmarum      3          8
## 121        Thraupidae              Lanio_surinamus     18         45
## 122        Thraupidae                 Lanio_fulvus      1          1
## 123        Thraupidae            Ramphocelus_carbo      7         46
## 124        Thraupidae            Cyanerpes_caeruleus     1          2
## 125        Thraupidae              Coereba_flaveola      3          3
## 126        Thraupidae Sporophila_castaneiventris        1          4
## 127        Thraupidae          Sporophila_angolensis     8         12
## 128        Thraupidae             Saltator_maximus      1          2
## 129        Thraupidae             Saltator_grossus      2          4
## 130       Cardinalidae     Cyanoloxia_rothschildii      5          7
```

Species list. Bird species captured on the understorey of 33 forest islands and 5 continuous forest sites (n = 38 sites). 'Sites' corresponds to the number of sites at which species were captured. 'Individuals' corresponds to the number of individuals captured. Taxonomy follows Piacentini et al. (2015).

Table S1

```
# Information to include in the table
table.s1 = data.frame("Site name" = rownames(env[order(env$area.ha), ]),
                      # Y coordinate (datum WGS 84)
                      "Latitude" = env[order(env$area.ha), ]$latitude.WGS84,
                      # X coordinate (datum WGS 84)
                      "Longitude" = env[order(env$area.ha), ]$longitude.WGS84,
                      # Forest area in hectares
                      "Forest area" = env[order(env$area.ha), ]$area.ha,
                      # Observed number of species
                      "Richness" = specnumber(balbina.full[1:(ncol(balbina.full)-1)]),
                      # Total number of individuals captured
                      "Abundance" = rowSums(balbina.full[1:(ncol(balbina.full)-1)]))

# Order rows by island size
table.s1 = table.s1[order(table.s1$Forest.area, table.s1$Richness), ]
row.names(table.s1) = NULL
```

```
table.s1
```

```
##              Site.name Latitude Longitude Forest.area Richness Abundance
## 1            Joaninha -1.52272 -59.82891        0.63       10        54
## 2                Xibe -1.46889 -59.83630        0.91        6         9
## 3             Formiga -1.83338 -59.42115        1.54        6        59
## 4               Andre -1.58461 -59.87198        2.08        3         8
## 5             Cafundo -1.50042 -59.61020        2.63        7        10
## 6              Panema -1.77448 -59.69241        3.08        4         5
## 7               Torem -1.82947 -59.63244        3.62        6        20
## 8            Pe_Torto -1.76646 -59.36351        4.98       14        38
## 9           Jiquitaia -1.83569 -59.59568        6.79       10        30
## 10          Arrepiado -1.51512 -59.73919        7.43        6        10
## 11            Garrafa -1.58859 -59.83555        9.42        9        15
## 12             Piquia -1.50647 -59.78897       13.04       20        30
## 13            Abusado -1.76221 -59.67875       13.31        5        29
## 14               Coata -1.48824 -59.78702      16.94       17        33
## 15              Palhal -1.79036 -59.44790      21.37       11        44
## 16                Neto -1.84149 -59.35228      32.87       12        56
## 17               Adeus -1.87378 -59.66240      50.08       12        26
## 18              Bacaba -1.50413 -59.82236      53.35       11        22
## 19             Relogio -1.68189 -59.65259      70.55       18        70
## 20            Sapupara -1.69713 -59.61263      77.80       12        32
## 21               Moita -1.55987 -59.89247      91.30       13        27
## 22              Pontal -1.83105 -59.68658     108.76       20        49
## 23 Furo_de_Santa_Luzia -1.74040 -59.44206     198.52       30        79
## 24               Cipoal -1.69830 -59.78477     217.63       23        71
## 25              Jabuti -1.62625 -59.75667     232.49       22        58
## 26            Tucumari -1.59151 -59.42986     336.02       24        63
## 27             Martelo -1.66407 -59.71429     466.60       15        39
## 28             Tristeza -1.76752 -59.75447     499.91       31        90
## 29       Beco_do_Catitu -1.73954 -59.70500     638.66       27        74
## 30             Mascote -1.64884 -59.83285     668.03       30        77
## 31              Fuzaca -1.49238 -59.85987     941.71       43       121
## 32        Porto_Seguro -1.78009 -59.51723    1350.56       26        85
## 33         Gaviao_real -1.59599 -59.63279    1698.84       27        68
## 34             CF_Waba -1.42884 -59.90534   16988.40       28        75
## 35           CF_Loreno -1.83490 -59.71799   16988.40       40       105
## 36          CF_Tucumari -1.57660 -59.38594   16988.40       40       111
## 37          CF_Grid_NS3 -1.79088 -59.25671   16988.40       52       165
## 38          CF_Grid_NS2 -1.79480 -59.26587   16988.40       54       158
```

Table S1. Description of the 38 sample sites surveyed at the Balbina Hydroelectric Reservoir landscape, including 33 forest islands and five continuous forest sites, and the species richness and abundance of understory forest birds captured therein.

Species assemblages: 2015 and 2016

We defined two types of species assemblages based on the species occurrence across all sample sites (n = 38). The first comprises species captured only at reference sample sites (hereafter referred to as **baseline species assemblage,** although we also use the term **continuous-forest species assemblage** to refer to the species subset from control sites). The second comprises all species captured in both reference sample sites and forest islands (hereafter, **overall species assemblage**).

```r
spp.full.all = names(balbina.full[1:(ncol(balbina.full) - 1)])  # all sample sites (n = 38)
# Continuous forest sites (n = 5)
spp.full.control = names(balbina.full[34:38, 1:(ncol(balbina.full) - 1)][colSums(balbina.full[34:38,
    1:(ncol(balbina.full) - 1)]) > 0])
# 1000-ha forest islands (n = 3)
spp.full.1000 = names(balbina.full[31:33, 1:(ncol(balbina.full) - 1)][colSums(balbina.full[31:33,
    1:(ncol(balbina.full) - 1)]) > 0])
# 500-ha forest islands (n = 4)
spp.full.500 = names(balbina.full[27:30, 1:(ncol(balbina.full) - 1)][colSums(balbina.full[27:30,
    1:(ncol(balbina.full) - 1)]) > 0])
# 250-ha forest islands (n = 4)
spp.full.250 = names(balbina.full[23:26, 1:(ncol(balbina.full) - 1)][colSums(balbina.full[23:26,
    1:(ncol(balbina.full) - 1)]) > 0])
# 100-ha forest islands (n = 4)
spp.full.100 = names(balbina.full[19:22, 1:(ncol(balbina.full) - 1)][colSums(balbina.full[19:22,
    1:(ncol(balbina.full) - 1)]) > 0])
# Forest islands smaller than 55 ha (n = 18)
spp.full.impact = names(balbina.full[1:18, 1:(ncol(balbina.full) - 1)][colSums(balbina.full[1:18,
    1:(ncol(balbina.full) - 1)]) > 0])
```

Baseline suitability

The suitability of each baseline was inspected by comparing the number of species retained and gained in relation to the *continuous forest* baseline, so that the suitability of the *continuous forest* baseline was assigned the maximum biodiversity value.

1,000-ha forest islands

```r
# Number of species from continuous forest sites ('continuous-forest species
# assemblage') captured on 1,000-ha islands (i.e. species retained)
specnumber(colSums(balbina.full[31:33, spp.full.control]))

# Total number of species ('overall species assemblage') captured on
# 1,000-ha islands
specnumber(colSums(balbina.full[31:33, spp.full.all]))

# Number of species gained (i.e. absent in continuous forest sites)
56 - 45
```

```
## [1] 11
```

500-ha forest islands

```r
specnumber(colSums(balbina.full[27:30, spp.full.control]))
specnumber(colSums(balbina.full[27:30, spp.full.all]))

53 - 39
```

```
## [1] 14
```

250-ha forest islands

```
specnumber(colSums(balbina.full[23:26, spp.full.control]))
specnumber(colSums(balbina.full[23:26, spp.full.all]))
```

```
46 - 37
```

```
## [1] 9
```

100-ha forest islands

```
specnumber(colSums(balbina.full[19:22, spp.full.control]))
specnumber(colSums(balbina.full[19:22, spp.full.all]))
```

```
35 - 29
```

```
## [1] 6
```

```
# Data preparation to produce a graph to represent the gradient of baseline suitability
retained = c(86, 45, 39, 37, 29)
gained = c(0, 11, 14, 9, 6)
total = retained + gained

bas.suit = data.frame(spp = c(retained, gained),
                      baseline = rep(1:5, 2),
                      group = c("r","r","r","r","r","a","a","a","a","a"))
```

Figure 2

```
# Graph of baseline suitability
fig2 =
  ggplot(data = bas.suit, aes(x = baseline, y = spp/86, fill = group)) +

  geom_bar(stat = "identity", width = 0.75, colour = "black") +

  geom_text(aes(label = c(86, 45, 39, 37, 29, "", 11, 14, 9, 6)),
            colour = "white", size = 6.8,
            position = position_stack(vjust = 0.5)) +

  labs(x = "Baseline suitability",
       y = "Proportion of species") +

  scale_x_continuous(breaks = 1:5,
                     labels = c("CF","1,000 ha","500 ha","250 ha","100 ha")) +

  scale_y_continuous(breaks = c(0.0, 0.2, 0.4, 0.6, 0.8, 1.0),
                     labels = c("0.0","0.2","0.4","0.6","0.8","1.0")) +

  scale_fill_manual(values = c("grey70", "grey30"),
                    labels = c("Species gained", "Species retained")) +

  theme_bw(base_size = 20) +
```

```r
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5),
        axis.title = element_text(face = "bold"),
        axis.text = element_text(colour = "black")) +
  theme(legend.title = element_blank(),
        legend.text = element_text(size = 14),
        legend.justification = c(0.99, 0.99),
        legend.position = c(0.99, 0.99),
        legend.key.size = unit(0.75, "cm")) +
  guides(fill = guide_legend(reverse = TRUE))

ggsave(fig2, file = "fig2.pdf", width = 20, height = 20, units = "cm")


#fig2
```

Figure 2. Gradient of baseline suitability showing the number of bird species retained and gained in relation to the understory bird assemblage captured in continuous forest sites (CF).

Surveyed islands smaller than 55 ha (n = 18; Table S1) were regarded as impacted sites due to forest fragmentation.

Rarefied number of species

Having defined the gradient of baseline suitability and the impacted sites, we compared the species richness of all impacted sites combined to that of each of the five baseline levels (Fig. 3) for both the baseline and the overall species assemblage. To accomplish this, we used the INEXT package (Hsieh et al. 2016) in R (R Core Team 2018) to calculate the rarefied number of species based on equal completeness (i.e. sample coverage) instead of equal size (i.e. number of individuals), because species-rich sites require a greater number of individuals to be fully characterized than species-poor sites (Chao and Jost 2012).

Data preparation

```r
# Continuous forest sites
# "b" stands for "baseline species assemblage"
comm.full.cf.b = colSums(balbina.full[34:38, 1:(ncol(balbina.full)-1)])
# "io" stands for "impacted sites" and "overall species assemblage"
comm.full.cf.io = colSums(balbina.full[1:18, 1:(ncol(balbina.full)-1)])

# "ib" stands for "impacted sites" and "baseline species assemblage"
comm.full.cf.ib = colSums(balbina.full[1:18, 1:(ncol(balbina.full)-1)])
comm.full.cf.ib[-c(which(match(names(comm.full.cf.ib), spp.full.control) != "NA"))] = 0

# 1000-ha forest islands
comm.full.1000.b = colSums(balbina.full[31:33, 1:(ncol(balbina.full)-1)])
comm.full.1000.io = colSums(balbina.full[1:18, 1:(ncol(balbina.full)-1)])

comm.full.1000.ib = colSums(balbina.full[1:18, 1:(ncol(balbina.full)-1)])
comm.full.1000.ib[-c(which(match(names(comm.full.1000.ib), spp.full.1000) != "NA"))] = 0

# 500-ha forest islands
comm.full.500.b = colSums(balbina.full[27:30, 1:(ncol(balbina.full)-1)])
```

```
comm.full.500.io = colSums(balbina.full[1:18, 1:(ncol(balbina.full)-1)])

comm.full.500.ib = colSums(balbina.full[1:18, 1:(ncol(balbina.full)-1)])
comm.full.500.ib[-c(which(match(names(comm.full.500.ib), spp.full.500) != "NA"))] = 0

# 250-ha forest islands
comm.full.250.b = colSums(balbina.full[23:26, 1:(ncol(balbina.full)-1)])
comm.full.250.io = colSums(balbina.full[1:18, 1:(ncol(balbina.full)-1)])

comm.full.250.ib = colSums(balbina.full[1:18, 1:(ncol(balbina.full)-1)])
comm.full.250.ib[-c(which(match(names(comm.full.250.ib), spp.full.250) != "NA"))] = 0

# 100-ha forest islands
comm.full.100.b = colSums(balbina.full[19:22, 1:(ncol(balbina.full)-1)])
comm.full.100.io = colSums(balbina.full[1:18, 1:(ncol(balbina.full)-1)])

comm.full.100.ib = colSums(balbina.full[1:18, 1:(ncol(balbina.full)-1)])
comm.full.100.ib[-c(which(match(names(comm.full.100.ib), spp.full.100) != "NA"))] = 0

# Combine data in a data frame
comm.full = data.frame(comm.full.cf.b, comm.full.cf.io, comm.full.cf.ib,
                       comm.full.1000.b, comm.full.1000.io, comm.full.1000.ib,
                       comm.full.500.b, comm.full.500.io, comm.full.500.ib,
                       comm.full.250.b, comm.full.250.io, comm.full.250.ib,
                       comm.full.100.b, comm.full.100.io, comm.full.100.ib)
```

Calculation of the rarefied number of species

```
# Rarefied number of species standardized by sample coverage (0.927)
# "SC" stands for "sample coverage"
diversity.SC.full = estimateD(comm.full, datatype = "abundance", base = "coverage",
                              level = NULL, conf = 0.95) # includes q = 0, 1, 2
# includes only q = 0
richness.SC.full = subset(diversity.SC.full, diversity.SC.full$order == "0")

richness.SC.full$baseline = rep(c("a.control", "b.1000", "c.500", "d.250", "e.100"), each = 3)
richness.SC.full$assemblage = rep(c("a.baseline", "b.impact.overall", "c.impact.baseline"), 5)

colnames(richness.SC.full) = c("site", "individuals", "method", "q", "coverage",
                               "richness", "lowerCI", "upperCI", "baseline", "assemblage")
rownames(richness.SC.full) = 1:nrow(richness.SC.full)
```

Figure 4

```
fig4 =
  ggplot(data = richness.SC.full, aes(x = baseline, y = richness, fill = assemblage)) +

  labs(x = "Baseline suitability", y = "Rarefied number of species") +

  scale_x_discrete(breaks = c("a.control", "b.1000", "c.500", "d.250", "e.100"),
                   labels = c("CF","1,000 ha","500 ha","250 ha","100 ha")) +
```

```r
  scale_y_continuous(limits = c(0, 80), breaks = seq(0, 80, 10)) +

  scale_fill_manual(values = c("grey30", "grey70", "white"),
                    labels = c("Reference sites",
                               "Impacted sites: overall species assemblage",
                               "Impacted sites: baseline species assemblage")) +

  geom_bar(colour = "black", stat = "identity",
           width = 0.75, position = position_dodge()) +
  geom_errorbar(aes(ymin = lowerCI, ymax = upperCI),
                width = 0.2, position = position_dodge(0.75)) +

  theme_bw(base_size = 20) +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5),
        axis.title = element_text(face = "bold"),
        axis.text = element_text(colour = "black")) +
  theme(legend.title = element_blank(),
        legend.text = element_text(size = 14),
        legend.justification = c(0.99, 0.99),
        legend.position = c(0.99, 0.99),
        legend.key.size = unit(0.75, "cm"))

ggsave(fig4, file = "fig4.pdf", width = 20, height = 20, units = "cm")

#fig4
```

Figure 4. Comparisons between the rarefied number of understory bird species at reference and impacted sites across a gradient of baseline suitability for two types of species assemblages. The bar plot shows the contrast between reference and impacted sites from the most (CF - continuous forest) to the least suitable baselines (100-ha forest islands), thereby affecting the estimated impact of forest fragmentation on species richness of both types of species assemblages. Rarefied number of species were standardized by sample coverage (0.927). Error bars indicate the 95% confidence intervals.

Estimated impact of forest fragmentation on species richness

We calculated the proportional difference in species richness between the impacted sites combined and each of the five baseline levels to assess to what extent our estimated impacts of forest fragmentation are affected by the degree of baseline suitability and the composition of focal species assemblages. Estimated impact was measured as one minus the proportional difference between impacted and reference sites, so the higher the value, the more severe the estimated impact. Accordingly, if reference sites harbor 10 species and impacted sites 4 species, the estimated impact is 0.6 [1 - (4/10)] or a 60% reduction in species richness.

```r
# Continuous forest sites
# "io" stands for "impacted sites" and "overall species assemblage"
impact.cf.io = 1 - (richness.SC.full[2, "richness"] / richness.SC.full[1, "richness"])
# "ib" stands for "impacted sites" and "baseline species assemblage"
impact.cf.ib = 1 - (richness.SC.full[3, "richness"] / richness.SC.full[1, "richness"])

# 1000-ha forest islands
impact.1000.io = 1 - (richness.SC.full[5, "richness"] / richness.SC.full[4, "richness"])
impact.1000.ib = 1 - (richness.SC.full[6, "richness"] / richness.SC.full[4, "richness"])
```

```
# 500-ha forest islands
impact.500.io = 1 - (richness.SC.full[8, "richness"] / richness.SC.full[7, "richness"])
impact.500.ib = 1 - (richness.SC.full[9, "richness"] / richness.SC.full[7, "richness"])

# 250-ha forest islands
impact.250.io = 1 - (richness.SC.full[11, "richness"] / richness.SC.full[10, "richness"])
impact.250.ib = 1 - (richness.SC.full[12, "richness"] / richness.SC.full[10, "richness"])

# 100-ha forest islands
impact.100.io = 1 - (richness.SC.full[14, "richness"] / richness.SC.full[13, "richness"])
impact.100.ib = 1 - (richness.SC.full[15, "richness"] / richness.SC.full[13, "richness"])
```

Table S3

```
table.s3 = as.data.frame(cbind(
  round(rbind(impact.cf.ib, impact.1000.ib, impact.500.ib,
              impact.250.ib, impact.100.ib), 2),
  round(rbind(impact.cf.io, impact.1000.io, impact.500.io,
              impact.250.io, impact.100.io), 2)))

table.s3$V1 = c("0.67*", "0.58*", "0.56*", "0.52*", "0.50*")
table.s3$V2 = c("0.29*", "0.08", "0.01", "-0.19", "-0.43*")

table.s3$Reference_sites = c("Continuous forest", "1,000 ha", "500 ha", "250 ha", "100 ha")
colnames(table.s3) = c("Baseline species assemblage",
                       "Overall species assemblage",
                       "Reference sites")
rownames(table.s3) = NULL
table.s3 = table.s3[c(3, 1, 2)]

table.s3
```

```
##       Reference sites Baseline species assemblage Overall species assemblage
## 1 Continuous forest                         0.67*                      0.29*
## 2          1,000 ha                         0.58*                       0.08
## 3            500 ha                         0.56*                       0.01
## 4            250 ha                         0.52*                      -0.19
## 5            100 ha                         0.50*                      -0.43*
```

Table S3. Estimated impact of forest fragmentation on species richness measured for five reference sites (Fig. 3), representing a gradient of baseline suitability from the most (continuous forest) to the least suitable (100-ha forest islands).

Asterisks (*) denote a significant difference in species richness between reference and impacted sites. Higher values indicate more severe impacts of forest fragmentation on species richness (Fig. 4).

Minimum set problem

We used a biodiversity complementarity approach to determine the minimum number of forest islands (i.e. the 'solution' for the minimum set problem) that retained the maximum combined number of species (Howard et al. 1998) for both the **continuous-forest** and the **overall** species assemblage captured across the entire set of 33 surveyed islands. To do so, we used an algorithm (function `minimum.set`) that first selects the forest island containing the highest number of species and then sequentially selects

the forest islands that add the highest number of previously unrepresented species (Howard et al. 1998). In the event of a tie, the more species-rich island was selected, and if that tie persisted, we selected the largest island. This procedure was repeated until all species captured in the entire set of 33 forest islands were represented by at least one individual. We then quantified the differences between the solutions for the minimum set problem targeting both the **continuous-forest** and the **overall** species assemblage.

The function `minimum.set` was kindly developed by Cristian Dambros for the purpose of this study.

```r
# Load the function "minimum.set"
source("https://raw.githubusercontent.com/csdambros/R-functions/master/minimum.set")
```

Solutions for the minimum set problem

```r
# Data frame including only islands
islands = balbina.full[1:33, ]
```

```r
# continuous-forest species assemblage
# Species subset from control sites (i.e. continuous forest)

ms.t = data.frame(minimum.set(islands[spp.full.control]))
names(ms.t) = sub("\\.", " ", names(ms.t))

ms.t$richness = rowSums(ms.t[spp.full.control])
ms.t$add = c(max(ms.t$richness), diff(as.vector(
  specaccum(ms.t[spp.full.control], method = "collector")$richness)))
ms.t$psr = rowSums(ms.t[spp.full.control])/length(spp.full.control)*100
ms.t$psr.cum = specaccum(ms.t[spp.full.control],
                    method = "collector")$richness/length(spp.full.control)*100
ms.t$minimum.set = "yes"
ms.t[as.numeric(which(grepl(max(ms.t$psr.cum),
                      ms.t$psr.cum))[1]+1):nrow(ms.t),]$minimum.set = "no"
ms.t$rank = 1:nrow(ms.t)
```

```r
# Overall species assemblage
# All species

ms.f = data.frame(minimum.set(islands[spp.full.all]))
names(ms.f) = sub("\\.", " ", names(ms.f))

ms.f$richness = rowSums(ms.f[spp.full.all])
ms.f$add = c(max(ms.f$richness), diff(as.vector(
  specaccum(ms.f[spp.full.all], method = "collector")$richness)))
ms.f$psr = rowSums(ms.f[spp.full.all])/length(spp.full.all)*100
ms.f$psr.cum = specaccum(ms.f[spp.full.all],
                    method = "collector")$richness/length(spp.full.all)*100
ms.f$minimum.set = "yes"
ms.f[as.numeric(which(grepl(max(ms.f$psr.cum),
                      ms.f$psr.cum))[1]+1):nrow(ms.f),]$minimum.set = "no"
ms.f$rank = 1:nrow(ms.f)
```

Figure 5

```r
# continuous-forest species assemblage
ms.t.graph =
  ggplot(data = ms.t,
         aes(x = rank, y = psr.cum, fill = minimum.set, size = psr)) +

  labs(x = "Islands, ranked by conservation priority",
       y = "Cumulative number of species (%)",
       fill = "Minimum set",
       size = "Number of\nspecies (n)") +

    scale_x_continuous(breaks = 1:length(ms.t$rank),
                       labels = c("Fuzaca", "Furo", "Tristeza", "Pontal", "Tucumari",
                                  "Porto", "Gavião-real", "Jabuti", "Piquiá", "Adeus",
                                  "Neto", "Pé Torto", "Mascote", "Beco", "Cipoal",
                                  "Relógio", "Coatá", "Martelo", "Moitá", "Sapupara",
                                  "Bacaba", "Cafundó", "Palhal", "Jiquitaia", "Arrepiado",
                                  "Garrafa", "Abusado", "Xibé", "André", "Joaninha",
                                  "Panema", "Torem", "Formiga")) +
  scale_y_continuous(limits = c(0, 100),
                     breaks = c(0, 25, 50, 75, 100)) +
  scale_fill_manual(values = c("grey70", "grey30"),
                    labels = c("no (n = 21)","yes (n = 12)")) +
  scale_size_continuous(range = c(2, 9),
                        breaks = c(10, 20, 30)) +

  geom_point(shape = 21, colour = "black") +

  theme_bw(base_size = 20) +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5),
        axis.title = element_text(face = "bold"),
        axis.text = element_text(colour = "black"),
        axis.text.x = element_text(angle = 90, vjust = 0.275, hjust = 1, size = 12)) +
  theme(legend.title = element_text(size = 12, face = "bold"),
        legend.text = element_text(size = 14),
        legend.justification = c(0.95, 0.05), legend.position = c(0.95, 0.05)) +
  guides(fill = guide_legend(reverse = TRUE,
                             order = 2, override.aes = list(size = c(5.5,5.5))),
         size = guide_legend(order = 1)) +

  annotate("text", x = 1, y = 100, label = "(a) Continuous-forest species assemblage",
           hjust = "left", vjust = "top", fontface = "bold", size = 6)


# Overall species assemblage
ms.f.graph =
  ggplot(data = ms.f,
         aes(x = rank, y = psr.cum, fill = minimum.set, size = psr)) +
```

```r
labs(x = "Islands, ranked by conservation priority",
    y = "",
    fill = "Minimum set",
    size = "Number of\nspecies (n)") +

  scale_x_continuous(breaks = 1:length(ms.f$rank),
                     labels = c("Fuzaca", "Tristeza", "Furo", "Joaninha", "Pontal",
                                "Gavião-real", "Coatá", "Pé Torto", "Tucumari",
                                "Cipoal", "Porto", "Piquiá", "Neto", "Adeus",
                                "Sapupara", "Palhal", "Bacaba", "Jiquitaia",
                                "Garrafa", "Xibé", "Panema", "Mascote", "Beco",
                                "Jabuti", "Relógio", "Martelo", "Moitá", "Cafundó",
                                "Formiga", "Torem", "Arrepiado", "Abusado", "André")) +
  scale_y_continuous(limits = c(0, 100),
                     breaks = c(0, 25, 50, 75, 100)) +
  scale_fill_manual(values = c("grey70", "grey30"),
                    labels = c("no (n = 12)","yes (n = 21)")) +
  scale_size_continuous(range = c(2, 9),
                        breaks = c(10, 20, 30)) +

  geom_point(shape = 21, colour = "black") +

theme_bw(base_size = 20) +
theme(panel.grid.major = element_blank(),
      panel.grid.minor = element_blank(),
      panel.border = element_rect(colour = "black"),
      axis.ticks = element_line(colour = "black", size = 0.5),
      axis.title = element_text(face = "bold"),
      axis.text = element_text(colour = "black"),
      axis.text.x = element_text(angle = 90, vjust = 0.275, hjust = 1, size = 12)) +
theme(legend.title = element_text(size = 12, face = "bold"),
      legend.text = element_text(size = 14),
      legend.justification = c(0.95, 0.05), legend.position = c(0.95, 0.05)) +
guides(fill = guide_legend(reverse = TRUE,
                           order = 2, override.aes = list(size = c(5.5,5.5))),
       size = guide_legend(order = 1)) +

annotate("text", x = 1, y = 100, label = "(b) Overall species assemblage",
         hjust = "left", vjust = "top", fontface = "bold", size = 6)

# Combine and save the graphs
ggsave(grid.arrange(ms.t.graph, ms.f.graph, ncol = 2),
       file = "fig5.pdf", h = 20, w = 40, units = "cm")
```

Figure S1

Blake and Loiselle (2001) found an overall higher number of understory bird species (including disturbance-adapted species) in young second-growth at La Selva, northern Costa Rica, compared to neighboring old-growth forest. We reanalysed their data using the **number of relict species** (see main text for definition) as the response variable to compare young second-growth, old second-growth and old-growth forests at La Selva. In doing so, we found that old-growth forest retained the highest number of species compared to either age classes of second-growth forest patches.

Import data

We used data from Appendix 1 in Blake and Loiselle (2001).

```
blake = read.csv("https://ndownloader.figshare.com/files/15158546", row.names = 1)
```

Data preparation

In Appendix 1, Blake and Loiselle (2001) reported the percentage of totals for species captured in mist nets or recorded during point counts in young second-growth (YSG), old second-growth (OSG), and old-growth forest (OGF) at La Selva Biological Station.

We extracted only the relevant information (mist-net data) and converted percentage of totals to number of captures. To do so, we used the total number of individuals captured per successional forest type as reported in Table 1 Blake and Loiselle (2001): 4526 in YSG, 1463 in OSG and 4030 in OGF.

```
# Extract only the relevant information
blake = blake[2:4]

# Rename the columns
colnames(blake) = c("YSG", "OSG", "OGF")

# Exclude species not captured in mist nets
blake = blake[rowSums(blake) > 0, ]

# Convert percentage of totals to number of captures
blake$YSG = round(blake$YSG * 4526 / 100)
blake$OSG = round(blake$OSG * 1463 / 100)
blake$OGF = round(blake$OGF * 4030 / 100)

# Create a table only with species captured in OGF (i.e. the baseline)
# "bas" stands for "baseline"
blake.bas = blake[blake$OGF > 0, ]
```

Calculation of the rarefied number of species

```
# Calculate the rarefied number of species standardized by sample coverage (0.981)
# (overall species assemblage)
diversity.SC.blake = estimateD(blake, datatype = "abundance", base = "coverage",
                               level = 0.981, conf = 0.95) # includes q = 0, 1, 2
# includes only q = 0
richness.SC.blake = subset(diversity.SC.blake, diversity.SC.blake$order == "0")

colnames(richness.SC.blake) = c("site", "individuals", "method", "q", "coverage",
                                "richness", "lowerCI", "upperCI")
rownames(richness.SC.blake) = 1:nrow(richness.SC.blake)

# Calculate the rarefied number of species standardized by sample coverage (0.981)
# (baseline species assemblage)
diversity.SC.blake.bas = estimateD(blake.bas, datatype = "abundance", base = "coverage",
                                   level = 0.981, conf = 0.95) # includes q = 0, 1, 2
# includes only q = 0
richness.SC.blake.bas = subset(diversity.SC.blake.bas, diversity.SC.blake.bas$order == "0")
```

```r
colnames(richness.SC.blake.bas) = c("site", "individuals", "method", "q", "coverage",
                                    "richness", "lowerCI", "upperCI")
rownames(richness.SC.blake.bas) = 1:nrow(richness.SC.blake.bas)

# Combine results in a data frame
blake.values = rbind(richness.SC.blake[6:8],
                     richness.SC.blake.bas[6:8])

blake.values$forest = rep(c("c.YSG", "b.OSG", "a.OGF"), 2)
blake.values$assemblage = rep(c("a.overall", "b.baseline"), each = 3)

colnames(blake.values) = c("richness", "lowerCI", "upperCI", "forest", "assemblage")
rownames(blake.values) = 1:nrow(blake.values)

blake.values
```

```
##    richness lowerCI upperCI forest assemblage
## 1  124.292 118.867 129.718  c.YSG  a.overall
## 2   95.248  87.272 103.223  b.OSG  a.overall
## 3   77.683  74.619  80.747  a.OGF  a.overall
## 4   55.339  53.302  57.376  c.YSG b.baseline
## 5   63.259  59.278  67.239  b.OSG b.baseline
## 6   77.683  74.545  80.820  a.OGF b.baseline
```

Figure S1

```r
figs1 =
  ggplot(data = blake.values, aes(x = assemblage, y = richness, fill = forest)) +

  labs(x = "Species assemblage",
       y = "Rarefied number of species") +

  scale_x_discrete(breaks = c("a.overall", "b.baseline"),
                   labels = c("Overall", "Baseline")) +
  scale_y_continuous(limits = c(0, 130),
                     breaks = c(0, 25, 50, 75, 100, 125)) +

  scale_fill_manual(values = c("grey30", "grey70", "white"),
                    labels = c("Old-growth forest",
                               "Old second-growth forest",
                               "Young second-growth forest")) +

  geom_bar(colour = "black", stat = "identity",
           width = 0.75, position = position_dodge()) +
  geom_errorbar(aes(ymin = lowerCI, ymax = upperCI),
                width = 0.2, position = position_dodge(0.75)) +

  theme_bw(base_size = 20) +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5),
        axis.title = element_text(face = "bold"),
```

```
        axis.text = element_text(colour = "black")) +
  theme(legend.title = element_blank(),
        legend.text = element_text(size = 14),
        legend.justification = c(0.99, 0.99),
        legend.position = c(0.99, 0.99),
        legend.key.size = unit(0.75, "cm"))

ggsave(figs1, file = "figs1.pdf", width = 20, height = 20, units = "cm")

#figs1
```

Figure S1. Comparisons between the rarefied number of understory bird species in northern Costa Rica across three successional forest types considering both the overall and the baseline species assemblage. The latter is derived from the neighboring La Selva Biological Station, whose old-growth forest was used as the reference site as in the original study (data from Blake and Loiselle 2001). The graph shows that a focus on the overall species assemblage results in a positive effect of land-use change (from old-growth forest to young second-growth forest) on the number of species (as reported in Blake and Loiselle 2001). In contrast, a focus on the baseline species assemblage results in a negative effect with either old second-growth or young second-growth forest harboring a more depauperate set species captured at the reference site. Rarefied number of species were standardized by sample coverage (0.981). Error bars indicate the 95% confidence intervals.

Estimated impact using one-year data

We examined whether the estimated impacts of forest fragmentation on species richness using two-year data (2015 and 2016) differ from those using one-year data from either 2015 or 2016.

Birds captured only in 2015

```
# Extract only data from 2015
birds.temp = birds[grep("2015", birds$date), ]

# Remove recaptures
birds.temp = subset(birds.temp, birds.temp$new.individual != "Recaptures")
birds.temp = subset(birds.temp, birds.temp$id != 376 &
                      birds.temp$id != 409 &
                      birds.temp$id != 600 &
                      birds.temp$id != 773 &
                      birds.temp$id != 822 &
                      birds.temp$id != 829)

# Site-by-species abundance matrix
birds.temp.matrix = tapply(birds.temp$occurrence,
                          list(birds.temp$site, birds.temp$species), sum)
# Species non detected in a given site are real zeros
birds.temp.matrix[is.na(birds.temp.matrix)] = 0

# Remove species that were not captured in 2015
birds.temp.matrix = birds.temp.matrix[, colSums(birds.temp.matrix) > 0]

# Order rows by site name
#birds.temp.matrix = birds.temp.matrix[c(1:6, 12, 7:11, 13:38), ]
```

```r
# Site-by-species abundance matrix with environmental data
# Add environmental data to the site-by-species abundance matrix
# 'balbina.2015' stands for the data collected only in 2015
balbina.2015 = cbind(birds.temp.matrix, env)
# Order rows by island size (i.e. forest area in hectares)
balbina.2015 = balbina.2015[order(balbina.2015$area.ha), ]
# Select the site-by-species abundance matrix and island size
balbina.2015 = balbina.2015[c(1:ncol(birds.temp.matrix),
                              match("area.ha", names(balbina.2015)))]
```

Species assemblages: 2015

```r
spp.2015.all = names(balbina.2015[1:(ncol(balbina.2015) - 1)])  # all sample sites (n = 38)
# Continuous forest sites (n = 5)
spp.2015.control = names(balbina.2015[34:38, 1:(ncol(balbina.2015) - 1)][colSums(balbina.2015[34:38,
    1:(ncol(balbina.2015) - 1)]) > 0])
# 1000-ha forest islands (n = 3)
spp.2015.1000 = names(balbina.2015[31:33, 1:(ncol(balbina.2015) - 1)][colSums(balbina.2015[31:33,
    1:(ncol(balbina.2015) - 1)]) > 0])
# 500-ha forest islands (n = 4)
spp.2015.500 = names(balbina.2015[27:30, 1:(ncol(balbina.2015) - 1)][colSums(balbina.2015[27:30,
    1:(ncol(balbina.2015) - 1)]) > 0])
# 250-ha forest islands (n = 4)
spp.2015.250 = names(balbina.2015[23:26, 1:(ncol(balbina.2015) - 1)][colSums(balbina.2015[23:26,
    1:(ncol(balbina.2015) - 1)]) > 0])
# 100-ha forest islands (n = 4)
spp.2015.100 = names(balbina.2015[19:22, 1:(ncol(balbina.2015) - 1)][colSums(balbina.2015[19:22,
    1:(ncol(balbina.2015) - 1)]) > 0])
# Forest islands smaller than 55 ha (n = 18)
spp.2015.impact = names(balbina.2015[1:18, 1:(ncol(balbina.2015) - 1)][colSums(balbina.2015[1:18,
    1:(ncol(balbina.2015) - 1)]) > 0])
```

Data preparation: 2015

```r
# Continuous forest sites
# "b" stands for "baseline species assemblage"
comm.2015.cf.b = colSums(balbina.2015[34:38, 1:(ncol(balbina.2015)-1)])
# "io" stands for "impacted sites" and "overall species assemblage"
comm.2015.cf.io = colSums(balbina.2015[1:18, 1:(ncol(balbina.2015)-1)])

# "ib" stands for "impacted sites" and "baseline species assemblage"
comm.2015.cf.ib = colSums(balbina.2015[1:18, 1:(ncol(balbina.2015)-1)])
comm.2015.cf.ib[-c(which(match(names(comm.2015.cf.ib), spp.2015.control) != "NA"))] = 0

# 1000-ha forest islands
comm.2015.1000.b = colSums(balbina.2015[31:33, 1:(ncol(balbina.2015)-1)])
comm.2015.1000.io = colSums(balbina.2015[1:18, 1:(ncol(balbina.2015)-1)])

comm.2015.1000.ib = colSums(balbina.2015[1:18, 1:(ncol(balbina.2015)-1)])
comm.2015.1000.ib[-c(which(match(names(comm.2015.1000.ib), spp.2015.1000) != "NA"))] = 0
```

```r
# 500-ha forest islands
comm.2015.500.b = colSums(balbina.2015[27:30, 1:(ncol(balbina.2015)-1)])
comm.2015.500.io = colSums(balbina.2015[1:18, 1:(ncol(balbina.2015)-1)])

comm.2015.500.ib = colSums(balbina.2015[1:18, 1:(ncol(balbina.2015)-1)])
comm.2015.500.ib[-c(which(match(names(comm.2015.500.ib), spp.2015.500) != "NA"))] = 0

# 250-ha forest islands
comm.2015.250.b = colSums(balbina.2015[23:26, 1:(ncol(balbina.2015)-1)])
comm.2015.250.io = colSums(balbina.2015[1:18, 1:(ncol(balbina.2015)-1)])

comm.2015.250.ib = colSums(balbina.2015[1:18, 1:(ncol(balbina.2015)-1)])
comm.2015.250.ib[-c(which(match(names(comm.2015.250.ib), spp.2015.250) != "NA"))] = 0

# 100-ha forest islands
comm.2015.100.b = colSums(balbina.2015[19:22, 1:(ncol(balbina.2015)-1)])
comm.2015.100.io = colSums(balbina.2015[1:18, 1:(ncol(balbina.2015)-1)])

comm.2015.100.ib = colSums(balbina.2015[1:18, 1:(ncol(balbina.2015)-1)])
comm.2015.100.ib[-c(which(match(names(comm.2015.100.ib), spp.2015.100) != "NA"))] = 0

# Combine data in a data frame
comm.2015 = data.frame(comm.2015.cf.b, comm.2015.cf.io, comm.2015.cf.ib,
                       comm.2015.1000.b, comm.2015.1000.io, comm.2015.1000.ib,
                       comm.2015.500.b, comm.2015.500.io, comm.2015.500.ib,
                       comm.2015.250.b, comm.2015.250.io, comm.2015.250.ib,
                       comm.2015.100.b, comm.2015.100.io, comm.2015.100.ib)
```

Calculation of the rarefied number of species: 2015

```r
# Rarefied number of species standardized by sample coverage (0.898)
# "SC" stands for "sample coverage"
diversity.SC.2015 = estimateD(comm.2015, datatype = "abundance", base = "coverage",
                              level = NULL, conf = 0.95) # includes q = 0, 1, 2
# includes only q = 0
richness.SC.2015 = subset(diversity.SC.2015, diversity.SC.2015$order == "0")

richness.SC.2015$baseline = rep(c("a.control", "b.1000", "c.500", "d.250", "e.100"), each = 3)
richness.SC.2015$assemblage = rep(c("a.baseline", "b.impact.overall", "c.impact.baseline"), 5)

colnames(richness.SC.2015) = c("site", "individuals", "method", "q", "coverage",
                               "richness", "lowerCI", "upperCI", "baseline", "assemblage")
rownames(richness.SC.2015) = 1:nrow(richness.SC.2015)

richness.obs.2015 = as.data.frame(colSums(ifelse(comm.2015 >= 1, 1, 0)))

richness.obs.2015$baseline = rep(c("a.control", "b.1000", "c.500", "d.250", "e.100"), each = 3)
richness.obs.2015$assemblage = rep(c("a.baseline", "b.impact.overall", "c.impact.baseline"), 5)

colnames(richness.obs.2015) = c("richness", "baseline", "assemblage")
rownames(richness.obs.2015) = 1:nrow(richness.obs.2015)
```

Figure: 2015

```
fig.2015 =
  ggplot(data = richness.SC.2015, aes(x = baseline, y = richness, fill = assemblage)) +

  labs(x = "Baseline suitability", y = "Rarefied number of species") +

  scale_x_discrete(breaks = c("a.control", "b.1000", "c.500", "d.250", "e.100"),
                   labels = c("CF","1,000 ha","500 ha","250 ha","100 ha")) +

  scale_y_continuous(limits = c(0, 80), breaks = seq(0, 80, 10)) +

  scale_fill_manual(values = c("grey30", "grey70", "white"),
                    labels = c("Reference sites",
                               "Impacted sites: overall species assemblage",
                               "Impacted sites: baseline species assemblage")) +

  geom_bar(colour = "black", stat = "identity",
           width = 0.75, position = position_dodge()) +
  geom_errorbar(aes(ymin = lowerCI, ymax = upperCI),
                width = 0.2, position = position_dodge(0.75)) +

  theme_bw(base_size = 20) +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5),
        axis.title = element_text(face = "bold"),
        axis.text = element_text(colour = "black")) +
  theme(legend.title = element_blank(),
        legend.text = element_text(size = 14),
        legend.justification = c(0.99, 0.99),
        legend.position = c(0.99, 0.99),
        legend.key.size = unit(0.75, "cm"))

ggsave(fig.2015, file = "fig.2015.pdf", width = 20, height = 20, units = "cm")

#fig.2015
```

Birds captured only in 2016

```
# Extract only data from 2016
birds.temp = birds[grep("2016", birds$date), ]

# Remove recaptures
birds.temp = subset(birds.temp, birds.temp$new.individual != "Recaptures")
birds.temp = subset(birds.temp, birds.temp$id != 1647 &
                    birds.temp$id != 1776 &
                    birds.temp$id != 2007 &
                    birds.temp$id != 2052 &
                    birds.temp$id != 2132 &
                    birds.temp$id != 2439)
```

```r
# Site-by-species abundance matrix
birds.temp.matrix = tapply(birds.temp$occurrence,
                           list(birds.temp$site, birds.temp$species), sum)
# Species non detected in a given site are real zeros
birds.temp.matrix[is.na(birds.temp.matrix)] = 0

# Remove species that were not captured in 2016
birds.temp.matrix = birds.temp.matrix[, colSums(birds.temp.matrix) > 0]

# Order rows by site name
#birds.temp.matrix = birds.temp.matrix[c(1:6, 12, 7:11, 13:38), ]

# Site-by-species abundance matrix with environmental data
# Add environmental data to the site-by-species abundance matrix
# 'balbina.2016' stands for the data collected only in 2016
balbina.2016 = cbind(birds.temp.matrix, env)
# Order rows by island size (i.e. forest area in hectares)
balbina.2016 = balbina.2016[order(balbina.2016$area.ha), ]
# Select the site-by-species abundance matrix and island size
balbina.2016 = balbina.2016[c(1:ncol(birds.temp.matrix),
                              match("area.ha", names(balbina.2016)))]
```

Species assemblages: 2016

```r
spp.2016.all = names(balbina.2016[1:(ncol(balbina.2016) - 1)])  # all sample sites (n = 38)
# Continuous forest sites (n = 5)
spp.2016.control = names(balbina.2016[34:38, 1:(ncol(balbina.2016) - 1)][colSums(balbina.2016[34:38,
    1:(ncol(balbina.2016) - 1)]) > 0])
# 1000-ha forest islands (n = 3)
spp.2016.1000 = names(balbina.2016[31:33, 1:(ncol(balbina.2016) - 1)][colSums(balbina.2016[31:33,
    1:(ncol(balbina.2016) - 1)]) > 0])
# 500-ha forest islands (n = 4)
spp.2016.500 = names(balbina.2016[27:30, 1:(ncol(balbina.2016) - 1)][colSums(balbina.2016[27:30,
    1:(ncol(balbina.2016) - 1)]) > 0])
# 250-ha forest islands (n = 4)
spp.2016.250 = names(balbina.2016[23:26, 1:(ncol(balbina.2016) - 1)][colSums(balbina.2016[23:26,
    1:(ncol(balbina.2016) - 1)]) > 0])
# 100-ha forest islands (n = 4)
spp.2016.100 = names(balbina.2016[19:22, 1:(ncol(balbina.2016) - 1)][colSums(balbina.2016[19:22,
    1:(ncol(balbina.2016) - 1)]) > 0])
# Forest islands smaller than 55 ha (n = 18)
spp.2016.impact = names(balbina.2016[1:18, 1:(ncol(balbina.2016) - 1)][colSums(balbina.2016[1:18,
    1:(ncol(balbina.2016) - 1)]) > 0])
```

Data preparation: 2016

```r
# Continuous forest sites
# "b" stands for "baseline species assemblage"
comm.2016.cf.b = colSums(balbina.2016[34:38, 1:(ncol(balbina.2016)-1)])
# "io" stands for "impacted sites" and "overall species assemblage"
```

```r
comm.2016.cf.io = colSums(balbina.2016[1:18, 1:(ncol(balbina.2016)-1)])

# "ib" stands for "impacted sites" and "baseline species assemblage"
comm.2016.cf.ib = colSums(balbina.2016[1:18, 1:(ncol(balbina.2016)-1)])
comm.2016.cf.ib[-c(which(match(names(comm.2016.cf.ib), spp.2016.control) != "NA"))] = 0

# 1000-ha forest islands
comm.2016.1000.b = colSums(balbina.2016[31:33, 1:(ncol(balbina.2016)-1)])
comm.2016.1000.io = colSums(balbina.2016[1:18, 1:(ncol(balbina.2016)-1)])

comm.2016.1000.ib = colSums(balbina.2016[1:18, 1:(ncol(balbina.2016)-1)])
comm.2016.1000.ib[-c(which(match(names(comm.2016.1000.ib), spp.2016.1000) != "NA"))] = 0

# 500-ha forest islands
comm.2016.500.b = colSums(balbina.2016[27:30, 1:(ncol(balbina.2016)-1)])
comm.2016.500.io = colSums(balbina.2016[1:18, 1:(ncol(balbina.2016)-1)])

comm.2016.500.ib = colSums(balbina.2016[1:18, 1:(ncol(balbina.2016)-1)])
comm.2016.500.ib[-c(which(match(names(comm.2016.500.ib), spp.2016.500) != "NA"))] = 0

# 250-ha forest islands
comm.2016.250.b = colSums(balbina.2016[23:26, 1:(ncol(balbina.2016)-1)])
comm.2016.250.io = colSums(balbina.2016[1:18, 1:(ncol(balbina.2016)-1)])

comm.2016.250.ib = colSums(balbina.2016[1:18, 1:(ncol(balbina.2016)-1)])
comm.2016.250.ib[-c(which(match(names(comm.2016.250.ib), spp.2016.250) != "NA"))] = 0

# 100-ha forest islands
comm.2016.100.b = colSums(balbina.2016[19:22, 1:(ncol(balbina.2016)-1)])
comm.2016.100.io = colSums(balbina.2016[1:18, 1:(ncol(balbina.2016)-1)])

comm.2016.100.ib = colSums(balbina.2016[1:18, 1:(ncol(balbina.2016)-1)])
comm.2016.100.ib[-c(which(match(names(comm.2016.100.ib), spp.2016.100) != "NA"))] = 0

# Combine data in a data frame
comm.2016 = data.frame(comm.2016.cf.b, comm.2016.cf.io, comm.2016.cf.ib,
                       comm.2016.1000.b, comm.2016.1000.io, comm.2016.1000.ib,
                       comm.2016.500.b, comm.2016.500.io, comm.2016.500.ib,
                       comm.2016.250.b, comm.2016.250.io, comm.2016.250.ib,
                       comm.2016.100.b, comm.2016.100.io, comm.2016.100.ib)
```

Calculation of the rarefied number of species: 2016

```r
# Rarefied number of species standardized by sample coverage (0.868)
# "SC" stands for "sample coverage"
diversity.SC.2016 = estimateD(comm.2016, datatype = "abundance", base = "coverage",
                              level = NULL, conf = 0.95) # includes q = 0, 1, 2
# includes only q = 0
richness.SC.2016 = subset(diversity.SC.2016, diversity.SC.2016$order == "0")

richness.SC.2016$baseline = rep(c("a.control", "b.1000", "c.500", "d.250", "e.100"), each = 3)
richness.SC.2016$assemblage = rep(c("a.baseline", "b.impact.overall", "c.impact.baseline"), 5)
```

```
colnames(richness.SC.2016) = c("site", "individuals", "method", "q", "coverage",
                               "richness", "lowerCI", "upperCI", "baseline", "assemblage")
rownames(richness.SC.2016) = 1:nrow(richness.SC.2016)

richness.obs.2016 = as.data.frame(colSums(ifelse(comm.2016 >= 1, 1, 0)))

richness.obs.2016$baseline = rep(c("a.control", "b.1000", "c.500", "d.250", "e.100"), each = 3)
richness.obs.2016$assemblage = rep(c("a.baseline", "b.impact.overall", "c.impact.baseline"), 5)

colnames(richness.obs.2016) = c("richness", "baseline", "assemblage")
rownames(richness.obs.2016) = 1:nrow(richness.obs.2016)
```

Figure: 2016

```
fig.2016 =
  ggplot(data = richness.SC.2016, aes(x = baseline, y = richness, fill = assemblage)) +

  labs(x = "Baseline suitability", y = "Rarefied number of species") +

  scale_x_discrete(breaks = c("a.control", "b.1000", "c.500", "d.250", "e.100"),
                   labels = c("CF","1,000 ha","500 ha","250 ha","100 ha")) +

  scale_y_continuous(limits = c(0, 80), breaks = seq(0, 80, 10)) +

  scale_fill_manual(values = c("grey30", "grey70", "white"),
                    labels = c("Reference sites",
                               "Impacted sites: overall species assemblage",
                               "Impacted sites: baseline species assemblage")) +

  geom_bar(colour = "black", stat = "identity",
           width = 0.75, position = position_dodge()) +
  geom_errorbar(aes(ymin = lowerCI, ymax = upperCI),
                width = 0.2, position = position_dodge(0.75)) +

  theme_bw(base_size = 20) +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5),
        axis.title = element_text(face = "bold"),
        axis.text = element_text(colour = "black")) +
  theme(legend.title = element_blank(),
        legend.text = element_text(size = 14),
        legend.justification = c(0.99, 0.99),
        legend.position = c(0.99, 0.99),
        legend.key.size = unit(0.75, "cm"))

ggsave(fig.2016, file = "fig.2016.pdf", width = 20, height = 20, units = "cm")

#fig.2016
```

Interpretation of the results using one-year data

The results from the two-year dataset were different to those from either 2015 or 2016. Such contrasting results likely arose due to poor characterization of sample sites within a one-year survey given the smaller number of captures. The number of captures was 1,264 in 2015; 1,179 in 2016; and 2,115 in both years combined. (Recall that we have excluded recaptures, so individuals captured in 2015 and 2016 were not double-counted to tally the total of both years.) Overall, this indicates that the reliability of environmental impact assessments also depends on the robustness of biodiversity inventories.

Chapter 5

Sampling design may obscure species-area relationships in landscape-scale field studies
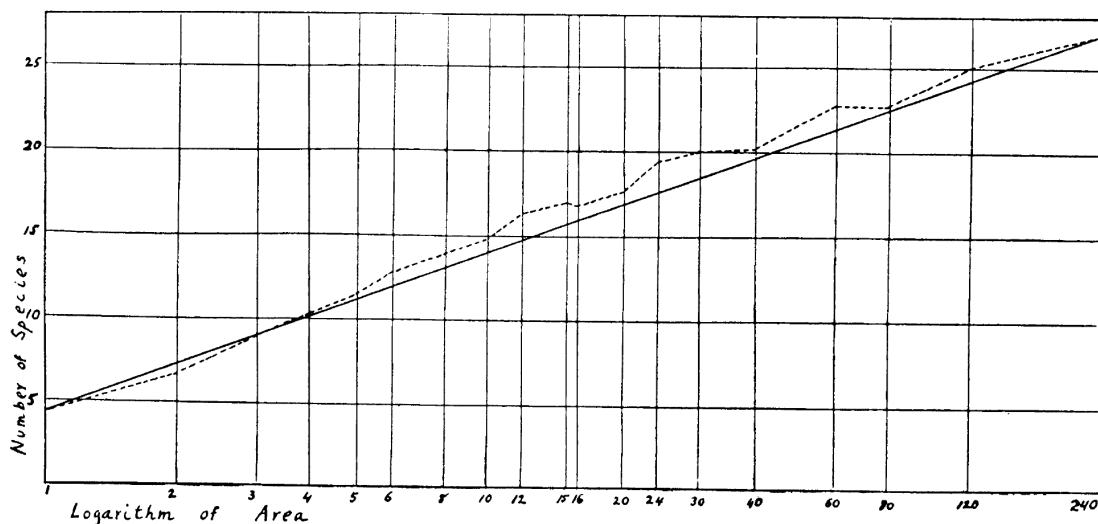
FIG. 1. Relation of species to area in the aspen association of northern Michigan. The horizontals indicate the area in square meters spaced according to the logarithm of the area. The verticals indicate the number of species. The dotted line indicates the number of species observed; the solid line indicates the number of species expected according to the logarithmic exponential equation.

For almost 100 years, the species-area relationship has proved to be a timeless topic in ecology. Plot from Gleason (1922).

5.1 Abstract

We investigated (1) the role of area *per se* in explaining anuran species richness on reservoir forest islands, after controlling for several confounding factors. We also assessed (2) how sampling design affects the inferential power of island species-area relationships (ISARs) aiming to (3) provide guidelines to yield reliable estimates of area-induced species losses in patchy systems. We surveyed anurans with autonomous recording units at 151 plots located on 74 islands and four continuous forest sites at the Balbina Hydroelectric Reservoir landscape, central Brazilian Amazonia. We applied semi-log ISAR models to assess the effect of sampling design on the fit and slope of species-area curves. To do so, we subsampled our surveyed islands following both a (i) stratified and (ii) non-stratified random selection of 5, 10, 15, 20 and 25 islands covering (i) the full range in island size (0.45-1,699 ha) and (ii) only islands smaller than 100 ha, respectively. We also compiled 25 datasets from the literature to assess the generality of our findings. Island size explained *c.* half of the variation in species richness. The fit and slope of species-area curves were affected mainly by the range in island size considered, and to a very small extent by the number of islands surveyed. In our literature review, all datasets covering a range of patch sizes larger than 300 ha yielded a positive ISAR, whereas the number of patches alone did not affect the detection of ISARs. We conclude that (1) area *per se* plays a major role in explaining anuran species richness on forest islands within an Amazonian anthropogenic archipelago; (2) the inferential power of island species-area relationships is severely degraded by sub-optimal sampling designs; (3) at least 10 habitat patches spanning three orders of magnitude in size should be surveyed to yield reliable species-area estimates in patchy systems.

Keywords: Amazonia, amphibians, environmental gradients, frogs, habitat fragmentation, hydroelectric dam, insularization, island biogeography theory, species-area relationship, tropical forest

5.2   Introduction

The species-area relationship (SAR) is the earliest and best-documented pattern in spatial ecology (Rosenzweig 1995, Tjørve et al. 2018). The "obvious fact that the larger the area taken the greater the number of species" (Arrhenius 1921) has led to a number of refinements to understand such a pattern. Many mathematical models (Tjørve 2003, Triantis et al. 2012) in tandem with several types of sampling schemes (Scheiner 2003) and concurrent underlying mechanisms have been invoked to explain SARs (Hill et al. 1994).

SAR models were firstly developed in the 1920s for contiguous areas (i.e. mainland SAR; Arrhenius 1921, Gleason 1922). Accordingly, larger areas are more species-rich because they have more individuals and contain a wider spectrum of habitats (Rosenzweig 1995). Thus, given a random abundance distribution, the larger number of individuals recorded over larger areas should imply more species (i.e. sampling effect; Hill et al. 1994). Meanwhile, the greater variety of habitats encompassed by larger areas supports species restricted to specific habitats and those requiring a combination thereof (i.e. habitat diversity effect; Connor and McCoy 2001).

Subsequently, Wilson (1961) showed that the rate of species loss as a function of area reduction is higher for archipelagos (i.e. island SAR) than for contiguous areas. This occurs because the number of species on islands is also mediated by the dynamic of extinction and colonisation as postulated in the Island Biogeography Theory (MacArthur and Wilson 1963, 1967), a paradigm also attributed to E. G. Munroe in 1948 (Tjørve et al. 2018). Accordingly, larger islands have larger population sizes resulting in lower extinction rates (i.e. area effect), and islands closer to a mainland source of species experience higher immigration rates (i.e. distance effect). Less isolated larger islands are therefore more species-rich than more isolated smaller islands (fig. 5 in MacArthur and Wilson 1963).

Since any reduction in island area depresses species richness more than a similar reduction in contiguous areas, mainland and island SAR can be seen as extremes of a continuum that is determined by the "islandness" (i.e. functional connectivity) of surveyed areas (Rosenzweig 1995). Such a property is arguably mediated by the dispersal ability of any given species group and the hostility of the intervening matrix in patchy systems (Bueno and Peres 2019). For example, in a forest archipelago induced by a hydroelectric dam in French Guiana, raptors were more prone to move between islands by traversing the water matrix than small mammals (Cosson et al. 1999), so the same archipelago is more functionally connected for raptors than for small mammals. Meanwhile, forest fragments surrounded by cattle pastures (i.e. less hostile matrix; Lees and Peres 2008) expe-

rience lower rates of bird species loss as a function of area reduction than forest islands within a hydroelectric reservoir (i.e. more hostile matrix; Bueno et al. 2018). Collectively, this means that islands and mainland areas will converge in their SARs at lower levels of landscape "islandness". Moreover, matrix type (Kennedy et al. 2010), history of disturbance (e.g. clear-cut or burned forests; Stouffer and Bierregaard, 1995), time since habitat patch isolation (Jones et al. 2016), and direct human disturbance (e.g. hunting pressure; Canale et al. 2012) all mediate the number of species in habitat remnants embedded in human-modified landscapes.

Even though positive SARs appear to be ubiquitous (Connor and McCoy 2001), some studies have found a non-significant or even negative relationship (Lövei et al. 2006), with smaller patches harbouring more species than larger ones. Such unexpected results can emerge for several reasons. First, surveyed patches often span a modest size range (Watling and Donnelly 2008, Lion et al. 2014) and are, therefore, exposed to the 'small island effect' (Wang et al. 2018), where a modest variation in patch size does not affect species richness (Lomolino and Weiser 2001). Second, few patches are surveyed, thereby reducing SAR model fits (Triantis et al. 2012). Finally, the species assemblage under consideration includes both habitat (e.g. forest dwellers) and non-habitat affiliated species (e.g. matrix dwellers), resulting in compensatory dynamics whereby any loss of the former is either compensated for (Russildi et al. 2016) or exceeded (Lövei et al. 2006) by any gain of the latter.

Here, we investigated (1) the role of area *per se* in explaining anuran species richness on Amazonian forest islands induced by river damming, after controlling for several confounding factors (Table 5.1). We also assessed (2) how sampling design affects the inferential power of island species-area relationships (type IV curve *sensu* Scheiner 2003) aiming to (3) provide guidelines to yield reliable estimates of area-induced species losses in patchy systems. We took advantage of passive acoustic monitoring (Deichmann et al. 2018) to survey anurans at a large number of forest islands (n = 74) covering a wide size range (0.45-1,699 ha) within the Balbina Hydroelectric Reservoir in central Brazilian Amazonia. We used anurans as a model group because each species has a distinct, simple, and relatively stereotyped vocalisation, thereby permitting reliable species identification even in megadiverse regions (Marques et al. 2013, Ribeiro et al. 2017). Moreover, anurans generally show pronounced site fidelity and limited dispersal ability (Smith and Green 2005), allowing us to largely control for distance effects (Palmeirim et al. 2017). As a vast 'real-world' experimental landscape, the Balbina forest archipelago is a unique setting to examine habitat area *per se* effects on species assemblages because (1) it provides over 3,500 replicated forest islands varying widely in size (Benchimol and Peres 2015a); (2) all forest islands were

created simultaneously 28 years ago (Fearnside 2016), having therefore been subjected to an uniform and relatively long relaxation time; (3) the open-water matrix is equally hostile; (4) forest islands span similar elevations, are restricted to upland forest and lack perennial streams, ultimately reducing habitat diversity; (5) adjacent control sites in undisturbed continuous primary forest are widely available; and the (6) *de facto* protection from any human disturbance covering most of the archipelago.

Table 5.1: Confounding factors affecting area *per se* effects on species richness and how they were controlled for in this study. Note that, as an observational study exploiting a natural field experiment, confounding factors could not be entirely removed but were minimised to a large extent.

Factor	How the factor was controlled for
Sampling effect	We used the rarefied number of species as our response variable, rather than the observed number of species.
Habitat diversity effect	Forest islands resulted from the rise of the reservoir floodwaters. Because lowlands are invariably flooded, only upland areas lacking streams persist in any one isolate.
Distance effect	We focused on anurans because they show high site fidelity and limited dispersal ability. Also, the hostility of the open-water matrix further hampers anurans from moving across forest islands.
Matrix, history, time since isolation, direct human disturbance	All forest island islands are surrounded by a lentic-water matrix and were created at the same time (1987). The study region has a low human population density and negligible direct anthropogenic impact, and most islands are within a large strictly-protected area.
Species assemblage	Only forest species were recorded across surveyed sites – i.e. species with "Forest - Subtropical/Tropical Moist Lowland" listed as a "Suitable" habitat according to IUCN (2018).

5.3   Methods

5.3.1   Study area

This study was carried out within the vast Balbina Hydroelectric Reservoir (BHR) and adjacent areas of continuous forest, located in central Brazilian Amazonia (1°40′ S, 59°40′ W; Fig. 5.1). The BHR was formed in 1987 by the damming of the Uatumã River, a tributary of the Amazon River, and covers an area of *c.* 300,000 ha (Fearnside 2016). The aftermath of dam construction created over 3,500 islands (Benchimol and Peres 2015a) derived from former hilltops of the once primary continuous forest. To offset the environmental impacts of the dam, 938,720 ha were set aside as the Uatumã Biological Reserve (IUCN category Ia; Fig. 5.1b), the largest biological reserve in Brazil. Moreover, the left bank of the former Uatumã River, including all islands, has also been effectively protected (Fig. 5.1b).
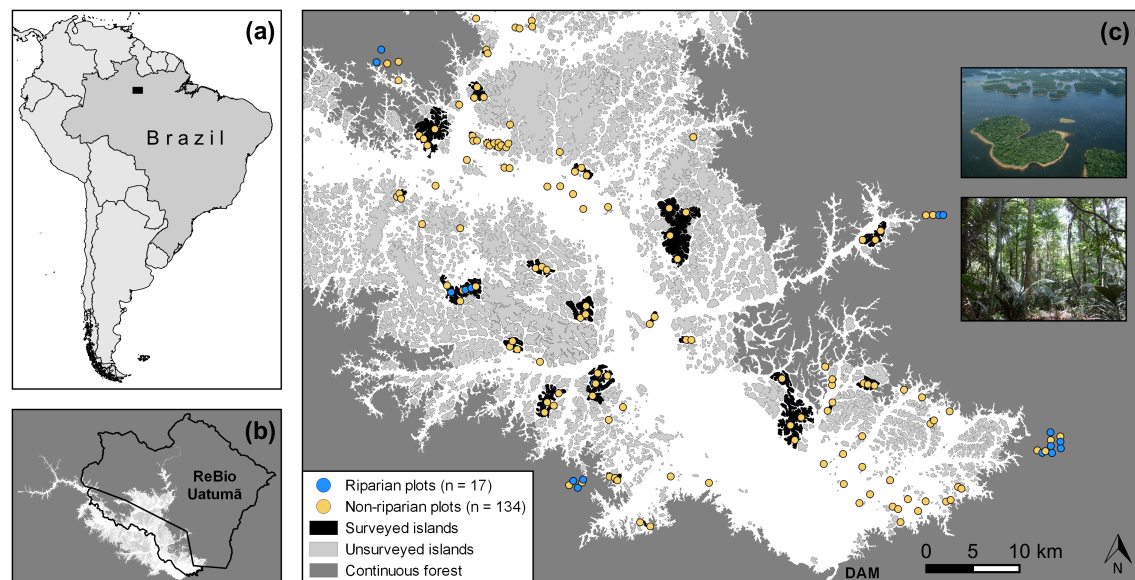
Figure 5.1: (a) Location of the study area in central Brazilian Amazonia, indicated by a solid rectangle containing (b) the Balbina Hydroelectric Reservoir (BHR) landscape, showing the boundaries of the Uatumã Biological Reserve, a strictly-protected area safeguarding most of this landscape; (c) larger inset map showing the spatial distribution of the 151 survey plots on 74 forest islands and in four continuous forest sites. Photographs represent the BHR landscape (credit: Eduardo M. Venticinque) and the forest interior of a surveyed island (credit: ASB).

The vegetation is characterised by submontane dense ombrophilous (*terra firme*) forest, although seasonally flooded *igapó* forest formerly occurred along the margins of the Uatumã River before damming. Islands span a wide range in size (0.2-4,878 ha; Benchimol and Peres 2015), virtually all of which lack perennial streams because lowland areas were submerged following the rise of floodwaters. Forest structure in larger islands resembles that of the continuous forest with

a higher dominance of large-seeded and canopy tree species, whereas smaller islands are dominated by pioneer tree species due to unavoidable edge-mediated forest disturbance (Benchimol and Peres 2015a). According to the Köppen classification, the climate is equatorial fully humid (Af), with mean annual precipitation and temperature of 2,464 mm and 26.5 °C, respectively (Alvares et al. 2013).

5.3.2   Sampling design

We surveyed 151 plots located on 78 sites, including 74 islands and four continuous forest sites (Fig. 5.1c). We attempted to survey a similar number of plots in riparian (i.e. along streams) and non-riparian habitats (i.e. away from streams) within continuous forest, and all available riparian habitats on islands, but only two islands had streams. Accordingly, we surveyed 13 riparian and 10 non-riparian plots in continuous forest, and 4 riparian and 124 non-riparian plots on islands (Fig. 5.1c). The number of plots per survey site was defined according to island size and presence of streams and varied from 4 to 10 in continuous forest sites and from 1 to 7 on islands (Fig. 5.1c; Table S5.1). The overall study meta-landscape encompassed 253,951 ha in which plots were spaced apart by an average distance of 32.63 km (SD = 18.83 km; range = 0.19-84.60 km; Fig. 5.1c).

5.3.3   Frog surveys

We surveyed anurans from July to December 2015 using autonomous recording units (ARUs) developed by the Automated Remote Biodiversity Monitoring Network (ARBIMON, <https://www.sieve-analytics.com>). Each ARU consists of an LG smartphone enclosed within a waterproof case with an external connector linked to an omnidirectional microphone. At each of the 151 plots, we deployed one ARU attached to a tree trunk 1.5 m above ground with the microphone pointing downward. ARUs were left unattended at each plot for five consecutive days and programmed to record 1 min in every 5-min interval using the ARBIMON Touch application. All recordings are archived at the ARBIMON II web platform and are freely available at <https://arbimon.sieve-analytics.com/project/balbina>.

   We selected a subset of 62 1-min recordings per plot (n = 151) to identify all anuran species occurring therein, totalling 9,362 1-min recordings. These recordings were derived from the following schedule: the first 1-min recording segment every 10 min over a 5-hour period (from 17:00 to 22:01) during sample days 2 and 4. Anuran species were identified by GSM who inspected all the recordings both aurally and visually using the ARBIMON II Visualizer tool. Species identifications were validated thereafter by ILK as a procedure to ensure accuracy. Dur-

ing this validation procedure, species records had to be either readily identified by hearing or inspecting the sonograms to be included in the analysis. Species records were discarded if they could not readily identified and/or if clearly audible sonogram acquisitions were inadequate (e.g. faint vocalisations too far away from the microphone).

5.3.4   Response variable

Before accounting for differences in sampling effort (i.e. number of 1-min recordings) across survey sites, we inspected the degree to which the observed number of species was correlated with sampling effort. We then calculated the rarefied number of anuran species using sampling-unit-based incidence data with 1,000 bootstrap replicates using the iNEXT package (Hsieh et al. 2016) within the R software (R Core Team 2018). To accomplish this, we created a species-by-sampling-unit matrix per survey site, in which each species (row) was assigned as present (1) or absent (0) and each sampling unit (column) corresponded to a 1-min recording. We standardised the sampling effort to the statistical mode, the most frequent number of 1-min recordings across survey sites ($n = 62$). We did so because iNEXT calculates both the interpolated and extrapolated number of species. Accordingly, we used the interpolated, observed and extrapolated number of species for sites allocated a sampling effort above ($n = 33$), equal to ($n = 43$) and below ($n = 2$) the statistical mode, respectively (all of which hereafter referred to as the rarefied number of species). We also used iNEXT to calculate the sample coverage to assess whether survey sites were sufficiently inventoried on the basis of 62 1-min recordings.

5.3.5   Predictor variable

Island size corresponds to the total insular forest cover and was calculated in QGIS software (QGIS Development Team 2016) using a classified image (Collection 2, 2015, Amazon) derived from 30-m resolution *Landsat* imagery downloaded from the Brazilian Annual Land Use and Land Cover Mapping Project (available at <http://mapbiomas.org>). Forest cover was defined as 'dense forest' (pixel value 3), because other pixel values effectively represent either heavily degraded forests or non-forest land cover types. Accordingly, the size of our 74 surveyed islands ranged from 0.45 to 1,699 ha (Table S5.1).

5.3.6 Island species-area relationships

To depict island species-area relationships (ISARs), we used the exponential equation (semi-log model; Gleason, 1922) to fit simple linear regression models as follows:

$$S = z \times log_{10}(A) + c,$$

where S = rarefied number of anuran species, z = regression slope, A = island size (ha), c = regression intercept. In this equation, z indicates the rate of species loss as a function of island size reduction, whereas c indicates the carrying capacity per unit area (Triantis et al. 2012). Despite the fact that ISARs can be fitted with dozens of alternative models (Triantis et al. 2012), the semi-log model was chosen because it is widely used (Tjørve 2003), easy to interpret, and allows the inclusion of sites at which no species was recorded ($S = 0$; Table S5.1).

To assess the degree to which shortening the range in island size changes the fit (r^2) and the slope (z) of the ISAR for anurans at the BHR landscape, we first classified the survey islands into five size categories: very small (< 4 ha, $n = 23$), small (4-20 ha, $n = 20$), medium sized (20-100 ha, $n = 17$), large (100-400 ha, $n = 7$), very large (> 400 ha, $n = 7$). We then fitted semi-log models to islands classified as (1) very small + small + medium sized + large + very large ($n = 74$); (2) very small + small + medium sized + large ($n = 67$); (3) very small + small + medium sized ($n = 60$); (4) very small + small ($n = 43$); and (5) very small ($n = 23$).

5.3.7 Tradeoff between replication power and extent of the gradient

Ideally, biodiversity surveys should include many sites covering a wide variation along any given gradient. However, logistical, financial or landscape (e.g. few and small habitat patches remaining) constraints, or combinations thereof, prevent attempts to adopt an ideal sampling design. With this in mind, we investigated the role of island-scale replication and the range in island size in detecting a positive ISAR for anurans at the BHR landscape. To do so, we subsampled our surveyed islands ($n = 74$) following both a stratified and non-stratified random selection of 5, 10, 15, 20 and 25 islands. In the stratified random selection, an equal number of islands belonging to each size category (see above) was selected to cover the *full range in island size* (0.45-1,699 ha). Accordingly, for 5, 10, 15, 20, 25 islands selected, each island size category was represented by 1, 2, 3, 4, and 5 islands, respectively. In the non-stratified random selection, subsets of 5, 10, 15, 20 and 25 islands smaller than 100 ha ($n = 60$) were selected, thereby covering a

short range in island size.

We also carried out a literature review (see the R code in the Supporting Information; Appendix S5.1) focused on both tropical and temperate anuran studies worldwide to assess (1) how prevalent positive ISARs are at a global scale, and (2) the role of the number of patches and range in patch size (largest minus the smallest) in detecting ISARs. Herein, the term 'patch' is used to encompass both 'fragment' and 'island'. Since results may be affected by the analytical approach employed (Bueno et al. 2018), we reanalysed data from each study based on our literature review using the semi-log model as described above. Note that for the global analysis, the response variable is the observed number of anuran species.

5.3.8   Data deposition

Data will be available from the KNB Repository, and are currently accessible through the links provided in the Supporting Information.

5.4   Results

Considering all 151 plots at 78 survey sites, we recorded 37 anuran species representing 18 genera and nine families (Table S5.2). The most ubiquitous species was *Ameerega trivittata* (n = 54 sites), whereas five species were only recorded at one site (Table S5.2). At the four continuous forest sites (n = 23 plots), we recorded 27 species from 15 genera and eight families; the number of species per continuous forest site ranged from 13 to 20 (mean $\pm$ SD = 15.75 $\pm$ 3.10; Table S5.1). On the 74 islands (n = 128 plots), we recorded 34 species from 18 genera and nine families, and the number of species per island ranged from 0 to 21 (6.12 $\pm$ 4.46; Table S5.1).

5.4.1   Species richness and sampling effort

The observed number of species was strongly and positively correlated with sampling effort (i.e. number of 1-min recordings; r = 0.82; Fig. 5.2a). However, the observed and rarefied number of species standardised to 62 1-min recordings were also strongly and positively correlated (r = 0.98; Fig. 5.2b). Therefore, high levels of local species packing was not artificially inflated by higher sampling effort, and sample coverage was adequate, exceeding 90% at 75 survey sites (Fig. 5.2c), indicating that a sampling effort per site of 62 1-min recordings was overall sufficient.
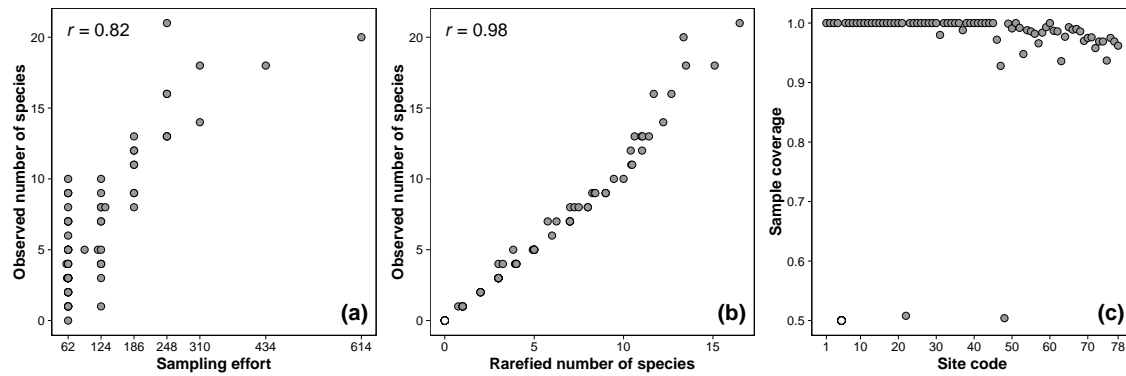
Figure 5.2: Correlation between (a) observed numbers of species and sampling effort, measured as the number of 1-min recordings per survey site; and (b) observed and rarefied number of species. (c) Sample coverage for 62 1-min recordings per survey site (see Table S5.1 for site codes). The open circle in (b) and (c) represents a small forest island where no anuran species was detected, so neither the rarefied number of species nor the sample coverage could be calculated.

5.4.2   Anuran species-area relationships at Balbina

Island size explained *c.* half of the variation ($r^2 = 0.49$) in the rarefied number of species considering all 74 islands (Fig. 5.3). Regression slopes were flattened, and model fits dramatically reduced as the range in island size was narrowed down, leading to a non-significant species-area relationship for islands smaller than 4 ha ($p = 0.90$; Fig. 5.3). Importantly, the similar rarefied number of species calculated for very large islands (> 400 ha) and continuous forest sites (Fig. 5.3) suggests that a further increase in island size would not imply more species, as long as the sampling effort is standardised.
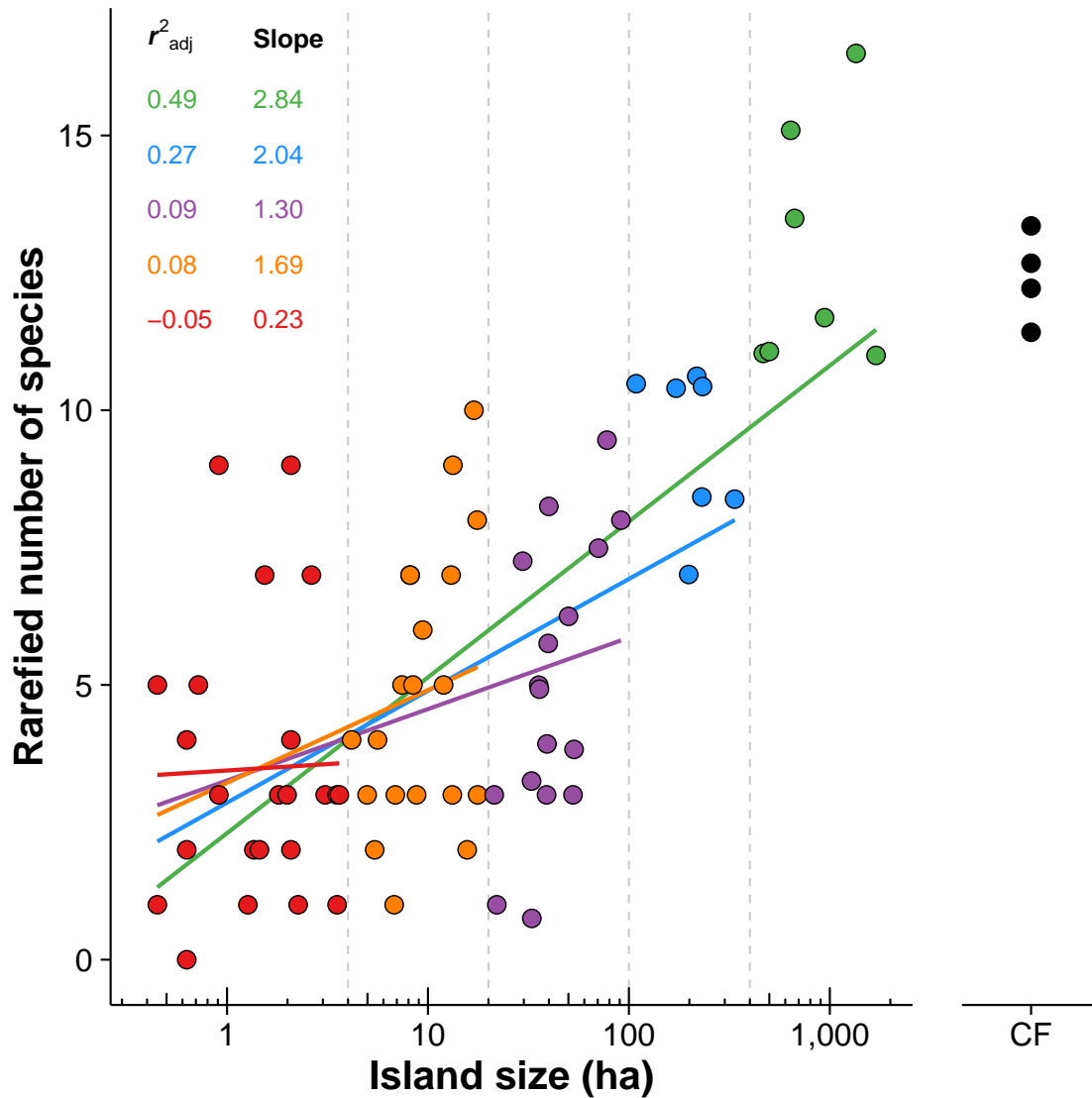
Figure 5.3: Anuran species-area relationships on forest islands surveyed at the Balbina Hydroelectric Reservoir landscape across five sets of islands (indicated by grey dashed lines and colour circles). The regression line for islands smaller than 4 ha ($n = 23$) is represented in red; up to 20 ha ($n = 43$) in orange; up to 100 ha ($n = 60$) in purple; up to 400 ha ($n = 67$) in blue; and up to 1,699 ha in green ($n = 74$). Continuous forest sites (CF, black circles) were not included in the regression fits. Note that model fits (r^2) and regression slopes tend to be reduced as the range in island size is narrowed down, thereby decreasing the estimated impact and the explanatory power of forest shrinkage on the rarefied number of species. Islands larger than 400 ha yielded similar values of species richness as continuous forest sites, indicating that, by controlling for sampling effort, further increases in the range of island size would not necessarily increase the number of species detected on islands.

For the stratified random selection of islands covering the *full range in island size* (0.45-1,699 ha), positive ISARs always held true for 15, 20 or 25 islands selected, usually for 10 islands, but only sometimes for 5 islands (Fig. 5.4). Slope deviances, measured as the degree to which the angle of a regression line deviates from that derived from all 74 islands, approximated one (i.e. no deviance) on average (Fig. 5.4a), while model fits (r^2) were increased to about 62% on average (Fig. 5.4c), regardless of the number of islands (5 to 25) selected. In none of the cases, ISARs were significantly negative.

For the non-stratified random selection of islands covering the *short range in island size* (< 100 ha), positive ISARs failed to hold true in the vast majority of cases, regardless of replication power; i.e. the number of islands selected (5 to 25; Fig. 5.4). Slope deviances were reduced in 35% to 65% on average (Fig. 5.4b), while model fits (r^2) were reduced to about 8% on average (Fig. 5.4d), regardless of the number of islands (5 to 25) selected. In none of the cases, ISARs were significantly negative.
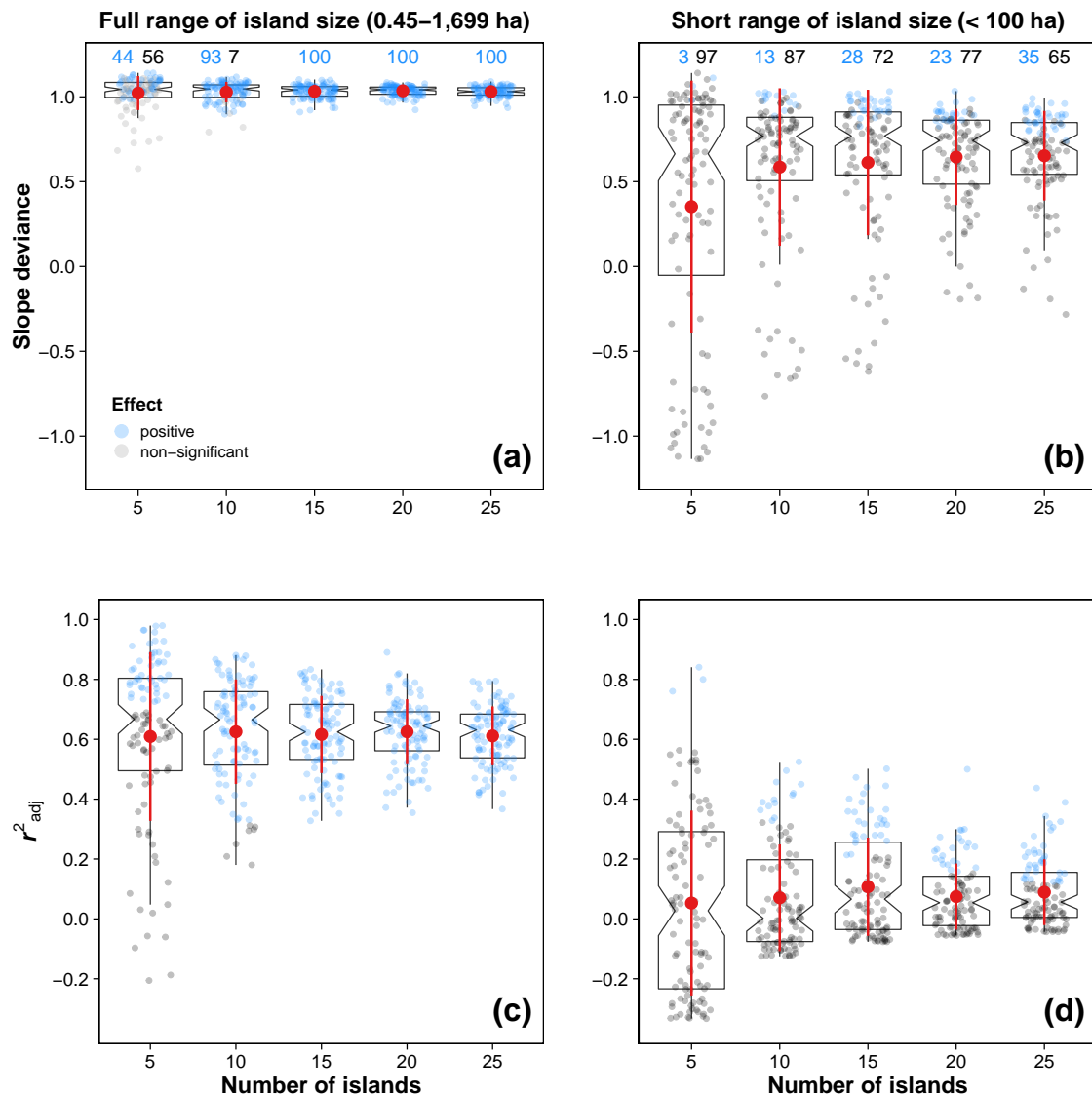
Figure 5.4: Results of anuran species-area relationships (SARs) derived from random sampling of 74 forest islands (see Fig. 5.3) surveyed at the Balbina Hydroelectric Reservoir landscape (see *Methods: Tradeoff between replication power and extent of the gradient* for further clarification). Each dot corresponds to a single SAR. Slope deviance was measured as the degree to which the angle of a regression line deviates from that derived from all 74 islands (green line in Fig. 5.3): 1.0 indicates no deviance (i.e. same slope), and values smaller and larger than 1.0 indicate lower and higher slopes, respectively. Red circles and red lines show means and standard deviations. Box-and-whisker plots show median (at notch), lower and upper quartiles and $1.5 \times$ interquartile ranges. Note that plots on the same row (a and b; c and d) are on the same scale to allow direct comparisons.

5.4.3 Prevalence of island species-area relationships for anurans worldwide

We compiled 25 datasets from 23 anuran studies in fragmented landscapes representing 12 countries worldwide (Fig. 5.5). Our reanalysis of the original data using the semi-log model revealed a positive ISAR for 18 datasets (mean r^2_{adj} $\pm$ SD = 0.45 $\pm$ 0.32) and a non-significant ISAR for seven datasets (Fig. 5.6). In none of the cases, ISARs were significantly negative. Remarkably, all 17 datasets that spanned a range in patch size (largest minus the smallest) larger than 300 ha yielded a positive ISAR, but the number of patches alone, which was widely variable (5 to 24), did not affect the detection of ISARs (Fig. 5.6).
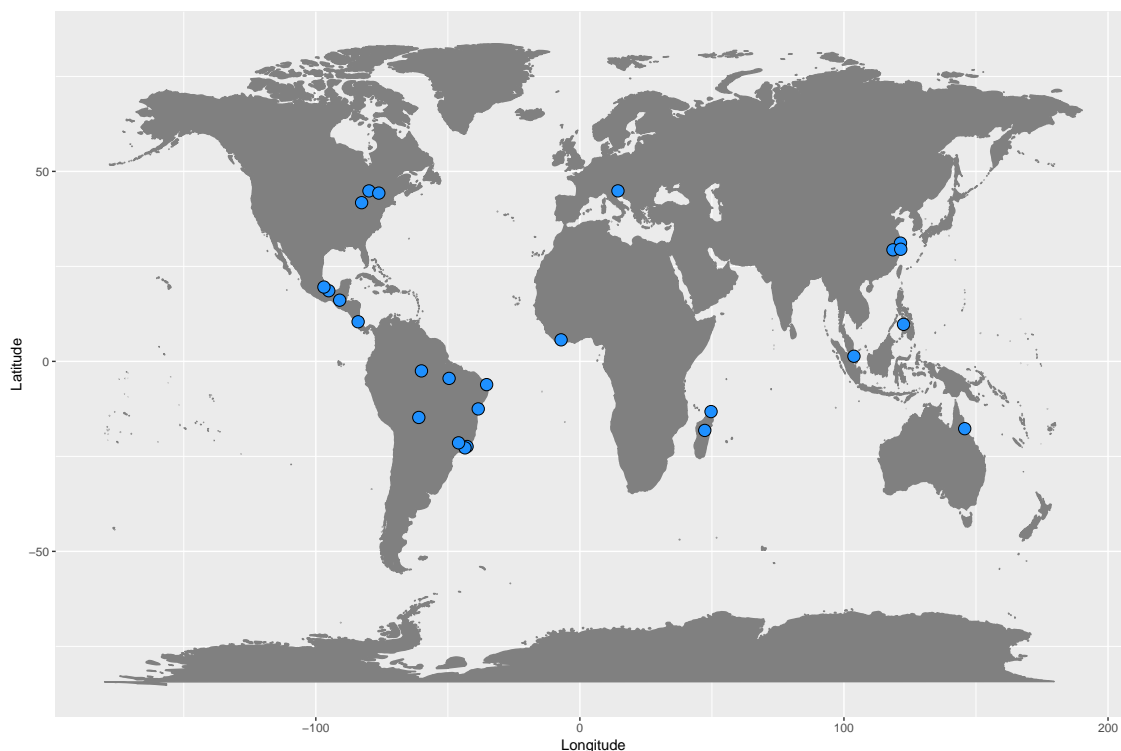
Figure 5.5: Location of the 25 datasets (blue circles) included in our analytical review of the prevalence of island species-area relationships for anurans worldwide.
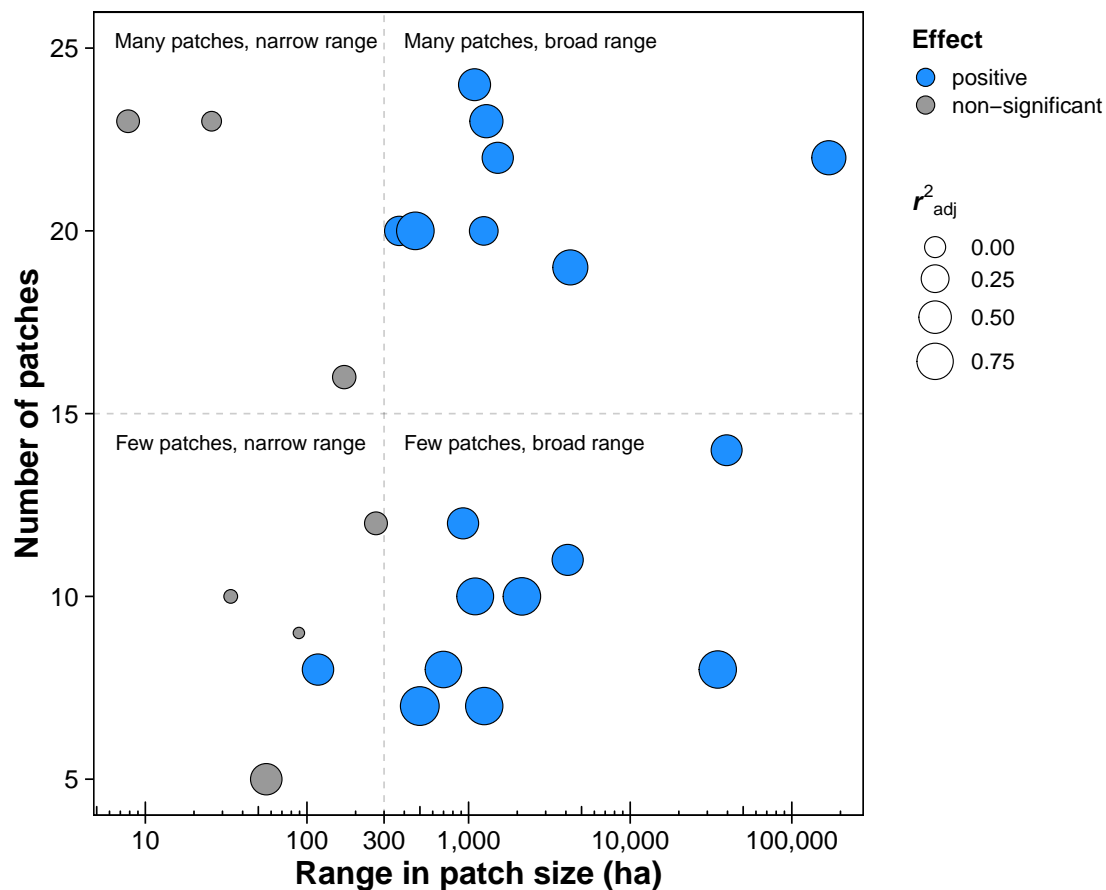
Figure 5.6: Prevalence of island species-area relationships (ISARs) derived from 25 anuran datasets worldwide (see Fig. 5.5), showing how the number of habitat patches surveyed and the range in patch size (largest minus the smallest) affect the significance and fit of ISAR models. Blue and grey circles indicate positive and non-significant relationships, respectively. Circle sizes are proportional to the magnitude of r^2-values. All studies covering a sufficiently wide range in patch size (¿ 300 ha; dashed vertical grey line) resulted in positive ISAR semi-log models.

5.5   Discussion

Our results indicate that habitat area *per se* plays a major role in explaining anuran species richness on forest islands induced by a large hydroelectric dam in lowland Amazonia. However, the fit and the slope of the island species-area relationship (ISAR) derived from the semi-log model [$S \sim log_{10}(A)$] were affected mainly by the range in island size considered, and to a very small extent by the number of islands surveyed. Hence, reported failures in detecting positive ISARs (e.g. Watling and Donnelly 2008, Lion et al. 2014) could be attributed to a sub-optimal sampling design, rather than an inherent absence of area effects on species richness.

5.5.1 Effect of area *per se* on species richness

Area represents a supra-variable

Habitat patch size or forest remnant area can explain virtually the entire variation ($r = 0.99$) of anuran species richness across Amazonian *terra firme* forest fragments, but this does not necessarily imply that area *per se* is the only underlying mechanism driving ISARs (Zimmerman and Bierregaard 1986). Because larger areas usually accommodate more individuals and more habitats, more species are expected to be recorded therein due to both the (1) sampling effect and (2) habitat diversity effect (Rosenzweig 1995). Accordingly, once these two effects are controlled for, ISARs tend to become weaker (lower fit and slope), relegating area *per se* to a lesser role in explaining overall species richness. Since the sampling effect is purely governed by the laws of probability, it should be refuted before ecological processes can be examined (Hill et al. 1994). These authors reported a strong fit ($r^2 = 0.80$ in a second-order polynomial regression) of the ISAR for birds in forest fragments in Ghana, which was largely attributed ($r^2 = 0.16$-0.37) to a sampling effect. Likewise, the shallow slope of the ISAR (log-log model) for birds in an anthropogenic forest archipelago in China was attributed to the low habitat diversity among islands, which were dominated by Masson pine (Yu et al. 2012).

Controlling for the sampling effect

Since poorly standardised sampling effort in biodiversity inventories is likely to produce misleading results, one should compare species richness among sites using either *individual-based* (abundance data) or *sample-based* (presence-absence data in a given sample) rarefaction curves (Gotelli and Colwell 2001). Here, we used a *sample-based* rarefaction procedure because of the nature of passive acoustic monitoring data, in which individuals recorded in consecutive samples (1-min recordings) are not independent. Accordingly, a rarefaction procedure (Hsieh et al. 2016) based on species incidence in 1-min recordings (samples) was used to divorce the sampling effect from the area effect. Meanwhile, we assessed the degree to which a standardised sampling effort of 62 1-min recordings yielded sufficient sampling effort to quantify the number of anuran species across survey sites. The fact that sampling effort was standardised to calculate species richness and that sample coverage was over 90% in virtually all surveyed sites (Fig. 5.2) provide robust support for an area effect that is independent of sampling effect.

Controlling for habitat diversity

In Amazonian *terra firme* forests, habitat diversity in terms of vegetation structure is associated with hydrological features of the terrain including elevation (Castilho et al. 2006), below-ground vertical distance to the water table (Schietti et al. 2014) and horizontal distance to perennial streams (Drucker et al. 2008). Both elevation and distance from streams – two variables that are typically correlated - have been shown to shape anuran assemblages in continuous forest adjacent to our study landscape (Condrati 2009). Accordingly, low-elevation sites near streams are more species-rich and harbour a distinct anuran species composition compared to high-elevation sites far away from streams (Condrati 2009). However, our forest islands are, by definition, upland remnants induced by the flooding of lowland areas in the once continuous forest, thereby lacking perennial streams. Therefore, within-island habitat diversity associated with area is greatly reduced, therefore providing evidence of an area effect that is independent of habitat diversity.

The mechanisms underlying ecological patterns can only be properly inferred from field experiments, despite their limitations in isolating co-varying mechanisms (McGarigal and Cushman 2002). As a mensurative experiment (*sensu* McGarigal and Cushman 2002), the relationship between island size and habitat diversity could not be entirely removed. For example, one key feature of how habitat structure affects anuran assemblages is the presence of breeding sites (Hillers et al. 2008, Bickford et al. 2010), so a higher species richness at larger habitat patches (i.e. fragments or islands) often results from higher diversity of breeding sites (Almeida-Gomes and Rocha 2015, Almeida-Gomes et al. 2016). In Amazonian *terra firme* forests, peccary wallows, which are used by some species as a breeding site (Zimmerman and Bierregaard 1986), are unlikely to be found on islands smaller than 100 ha where the occupancy probability of peccaries is much lower (Benchimol and Peres 2015b), thereby decreasing the number of breeding sites. Moreover, the number of anuran reproductive modes (*sensu* Haddad and Prado 2005) – a proxy for breeding sites – was positively related to log-transformed island size (r^2 = 0.45, Fig. S5.1), but not to the same extent as reported for in Atlantic Forest fragments of southeastern Brazil (r^2 = 0.87; Almeida-Gomes and Rocha 2015). Therefore, the effect of area *per se* on species richness was significant but probably weaker than that estimated in general (r^2 = 0.49 considering all 74 islands; Fig. 5.3).

Stranded on islands

The dynamics of local extinction and colonisation mediates the number of species on islands at any given time. According to the Island Biogeography Theory (MacArthur and Wilson 1963, 1967), extinction rates are determined by island size (area effect) and immigration rates by island isolation (distance effect). Concurrently, extinction rates may also be determined by island isolation (rescue effect; Brown and Kodric-Brown 1977) and immigration rates by island size (target effect; Lomolino 1990). Since three of these four effects result in variable colonisation rates, only the area effect would remain operational if species colonisation events are hampered. Given the pronounced site fidelity and limited dispersal ability of anuran species (Smith and Green 2005), coupled with the hostility of an often vast water matrix, we believe that island colonisation rates at Balbina are at best minimal (Jones et al. 2016, Palmeirim et al. 2017). Not surprisingly, the number of anuran species on forest islands within a hydroelectric reservoir in eastern Brazilian Amazonia was positively related to island size, regardless of island isolation (Lima et al. 2015). In other words, anurans are likely stranded on islands and local extinctions, if any, are rarely if ever rescued by new immigrants. At present, however, such assertion lacks empirical evidence, which could be derived from long-term studies aided by passive acoustic monitoring (Deichmann et al. 2018).

In fragmented landscapes, species composition in habitat patches is comprised of relict species (that were present before fragmentation), matrix-derived species and inter-patch dispersers (Watson 2002). The relative contribution of these groups depends on the type of matrix surrounding habitat patches. On the one hand, species richness in forest fragments embedded in terrestrial matrices is the result of (1) loss of relict species (Stouffer and Bierregaard 1995), (2) influx of matrix-derived species (Lövei et al. 2006) and (3) colonisation of inter-patch disperses following matrix regeneration (Stouffer et al. 2009). On the other hand, on land-bridge forest islands induced by hydroelectric dams, there are no matrix-derived species and the colonisation of inter-patch disperses is largely prohibitive in the case of anurans. Therefore, we can only surmise that anuran assemblages inhabiting our forest islands largely consist of a subset of relict species whose fate is determined by the area effect, with larger islands generally experiencing both lower and slower extinction rates than smaller islands (Jones et al. 2016).

5.5.2 The form of island species-area relationships

Regardless of the underlying mechanisms, larger areas tend to harbour more species. Combining data from two extensive syntheses on island species-area relationships (Triantis et al. 2012, Matthews et al. 2016), there were 584 out of 808 cases of significant ISARs (log-log model) and highly variable fits (r^2) and slopes (z) among those relationships that were significant. Many factors could account for such a variation in significance and model parameters (fit and slope). For example, the patch:matrix contrast (Kennedy et al. 2010), disturbance severity (Stouffer and Bierregaard 1995), relaxation time (Robinson 1999) and accessibility for hunters (Canale et al. 2012) can all affect the number of species in habitat patches, thereby modulating ISARs. Furthermore, the influx of matrix-derived species into habitat patches may either attenuate (Matthews et al. 2014) or even reverse (Lövei et al. 2006) the estimated impact of patch size on species richness. However, at the Balbina forest archipelago, all of these confounding factors were controlled for (Table 5.1), allowing us to depict unbiased patterns of ISARs for anurans in an Amazonian fragmented landscape.

The form of ISARs is also affected by issues of sampling design. At Balbina, both the model fit and the regression slope of the ISAR were reduced as the largest islands were progressively removed from the analysis (Fig. 5.3), thereby jeopardising inferences on area-driven anuran species losses. Thus, the truncated size range of habitat patches surveyed in any give study is likely the main reason for the lack of a significant ISAR for anurans in Neotropical forests in Brazil (n = 23 fragments, range = 1.71-27.41 ha; Lion et al. 2014) and Bolivia (n = 24 fragments, range = 0.6-8.5 ha; Watling and Donnelly 2008). Accordingly, species richness below a certain threshold can vary independently of area because smaller patches are more susceptible to environmental stochasticity (i.e. the small island effect; Lomolino and Weiser 2001). Although not necessarily ubiquitous (Wang et al. 2016), the small island effect was detected for anurans in an anthropogenic forest archipelago in China under a threshold of c. 40 ha (n = 23 islands, range = 0.59-1,289 ha; (Wang et al. 2018). Likewise, the ISAR for anurans at Balbina was either non-significant or yielded a very weak inferential power ($r^2 \leqslant 0.09$) considering only islands smaller than 100 ha.

Two studies in the Brazilian Atlantic Forest illustrate the role of sampling design in detecting ISARs for anurans. In the first (Almeida-Gomes and Rocha 2014), the authors surveyed 12 patches ranging from 4.7 to 272 ha and failed to detect a significant ISAR. In the second (Almeida-Gomes et al. 2016), they surveyed 21 patches ranging from 1.9 to 619 ha and subsequently detected a significant ISAR. These authors then concluded that failures to detect ISARs should be

attributed to an insufficient number of patches *and* range in patch size. However, our results suggest that the number of patches is not as critical as the range in patch size (Fig. 5.4). At Balbina, 10 islands covering the *full range in island size* (0.45-1,699 ha) yielded similar ISARs (in terms of model fit and regression slope) compared to 15, 20 or 25 islands. Conversely, the vast majority of ISARs were not significant regardless of the number of islands (5 to 25) if only a *short range in island size* (< 100 ha) had been sampled. Such a pattern was corroborated in our global review, which revealed that all datasets yielding a significant ISAR spanned a meaningful range in patch size larger than 300 ha, whereas all but one dataset covering a shorter range (< 300 ha) yielded non-significant ISARs (Fig. 5.6). Given the realities of field studies, the sampling tradeoff between number of sample replicates and extent of the gradient covered should therefore favour the latter to derive more reliable inferential relationships (Eigenbrod et al. 2011, Kreyling et al. 2018). This is in fact good news for field investigators who often face severely limited human, time and/or financial resources and can only survey a small number of sites.

5.6   Conclusions

We conclude that (1) habitat area *per se* plays a major role in explaining anuran species richness on Amazonian forest islands within one of the largest anthropogenic archipelagos on Earth; (2) the inferential power of island species-area relationships is clearly weakened by sub-optimal sampling designs; and (3) at least 10 habitat patches spanning three orders of magnitude in size should be surveyed to yield reliable estimates of area-driven species losses in patchy systems.

5.7 References

1. Almeida-Gomes, M. and Rocha, C. F. D. 2014. Landscape connectivity may explain anuran species distribution in an Atlantic forest fragmented area. - Landsc. Ecol. 29: 29–40.

2. Almeida-Gomes, M. and Rocha, C. F. D. 2015. Habitat loss reduces the diversity of frog reproductive modes in an Atlantic forest fragmented landscape. - Biotropica 47: 113–118.

3. Almeida-Gomes, M. et al. 2016. Patch size matters for amphibians in tropical fragmented landscapes. - Biol. Conserv. 195: 89–96.

4. Alvares, C. A. et al. 2013. Köppen's climate classification map for Brazil. - Meteorol. Zeitschrift 22: 711–728.

5. Arrhenius, O. 1921. Species and area. - J. Ecol. 9: 95–99.

6. Benchimol, M. and Peres, C. A. 2015a. Edge-mediated compositional and functional decay of tree assemblages in Amazonian forest islands after 26 years of isolation. - J. Ecol. 103: 408–420.

7. Benchimol, M. and Peres, C. A. 2015b. Predicting local extinctions of Amazonian vertebrates in forest islands created by a mega dam. - Biol. Conserv. 187: 61–72.

8. Bickford, D. et al. 2010. Forest fragment and breeding habitat characteristics explain frog diversity and abundance in Singapore. - Biotropica 42: 119–125.

9. Brown, J. H. and Kodric-Brown, A. 1977. Turnover rates in insular biogeography: Effect of immigration on extinction. - Ecology 58: 445–449.

10. Bueno, A. S. and Peres, C. A. 2019. Patch-scale biodiversity retention in fragmented landscapes: Reconciling the habitat amount hypothesis with the island biogeography theory. - J. Biogeogr. 46: 621–632.

11. Bueno, A. S. et al. 2018. Ecological traits modulate bird species responses to forest fragmentation in an Amazonian anthropogenic archipelago. - Divers. Distrib. 24: 387–402.

12. Canale, G. R. et al. 2012. Pervasive defaunation of forest remnants in a tropical biodiversity hotspot. - PLoS One 7: e41671.

13. Castilho, C. V. de et al. 2006. Variation in aboveground tree live biomass in a central Amazonian forest: Effects of soil and topography. - For. Ecol. Manage. 234: 85–96.

14. Condrati, L. H. 2009. Padrões de distribuição e abundância de anuros em áreas ripárias e não ripárias de floresta de terra firme na Reserva Biológica do Uatumã - Amazônia Central. MSc Thesis. Instituto Nacional de Pesquisas da Amazônia, Brazil.

15. Connor, E. F. and McCoy, E. D. 2001. Species-area relationships. - In: Levin, S. (ed), Encyclopedia of Biodiversity, Volume 5. Academic Press, San Diego, USA, pp. 397–411.

16. Cosson, J. F. et al. 1999. Ecological changes in recent land-bridge islands in French Guiana, with emphasis on vertebrate communities. - Biol. Conserv. 91: 213–222.

17. Deichmann, J. L. et al. 2018. It's time to listen: there is much to be learned from the sounds of tropical ecosystems. - Biotropica 0: 1–6.

18. Drucker, D. P. et al. 2008. How wide is the riparian zone of small streams in tropical forests? A test with terrestrial herbs. - J. Trop. Ecol. 24: 65–74.

19. Eigenbrod, F. et al. 2011. Sub-optimal study design has major impacts on landscape-scale inference. - Biol. Conserv. 144: 298–305.

20. Fearnside, P. M. 2016. Environmental and social impacts of hydroelectric dams in Brazilian Amazonia: Implications for the aluminum industry. - World Dev. 77: 48–65.

21. Gleason, H. A. 1922. On the relation between species and area. - Ecology 3: 158–162.

22. Gotelli, N. J. and Colwell, R. K. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. - Ecol. Lett. 4: 379–391.

23. Haddad, C. F. B. and Prado, C. P. D. A. 2005. Reproductive modes in frogs and their unexpected diversity in the Atlantic Forest of Brazil. - Bioscience 55: 207.

24. Hill, J. L. et al. 1994. The effect of sampling on the species-area curve. - Glob. Ecol. Biogeogr. Lett. 4: 97–106.

25. Hillers, A. et al. 2008. Effects of forest fragmentation and habitat degradation on west African leaf-litter frogs. - Conserv. Biol. 22: 762–772.

26. Hsieh, T. C. et al. 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). - Methods Ecol. Evol. 7: 1451–1456.

27. Jones, I. L. et al. 2016. Extinction debt on reservoir land-bridge islands. - Biol. Conserv. 199: 75–83.

28. Kennedy, C. M. et al. 2010. Landscape matrix and species traits mediate responses of Neotropical resident birds to forest fragmentation in Jamaica. - Ecol. Monogr. 80: 651–669.

29. Kreyling, J. et al. 2018. To replicate, or not to replicate - that is the question: how to tackle nonlinear responses in ecological experiments. - Ecol. Lett. in press.

30. Lees, A. C. and Peres, C. A. 2008. Avian life-history determinants of local extinction risk in a hyper-fragmented neotropical forest landscape. - Anim. Conserv. 11: 128–137.

31. Lima, J. R. et al. 2015. Amphibians on Amazonian land-bridge islands are affected more by area than isolation. - Biotropica 47: 369–376.

32. Lion, M. B. et al. 2014. Split distance: a key landscape metric shaping amphibian populations and communities in forest fragments. - Divers. Distrib. 20: 1245–1257.

33. Lomolino, M. V. 1990. The target area hypothesis: The influence of island area on immigration rates of non-volant mammals. - Oikos 57: 297–300.

34. Lomolino and Weiser 2001. Towards a more general species-area relationship: diversity on all islands, great and small. - J. Biogeogr. 28: 431–445.

35. Lövei, G. L. et al. 2006. The influence of matrix and edges on species richness patterns of ground beetles (Coleoptera: Carabidae) in habitat islands. - Glob. Ecol. Biogeogr. 15: 283–289.

36. MacArthur, R. H. and Wilson, E. O. 1963. An equilibrium theory of insular zoogeography. - Evolution (N. Y). 17: 373–387.

37. MacArthur, R. H. and Wilson, E. O. 1967. The Theory of Island Biogeography. - Princeton University Press.

38. Marques, T. A. et al. 2013. Estimating animal population density using passive acoustics. - Biol. Rev. 88: 287–309.

39. Matthews, T. J. et al. 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.

40. Matthews, T. J. et al. 2016. On the form of species-area relationships in habitat islands and true islands. - Glob. Ecol. Biogeogr. 25: 847–858.

41. McGarigal, K. and Cushman, S. A. 2002. Comparative evaluation of experimental approaches to the study of habitat fragmentation effects. - Ecol. Appl. 12: 335–345.

42. Palmeirim, A. F. et al. 2017. Non-random lizard extinctions in land-bridge Amazonian forest islands after 28 years of isolation. - Biol. Conserv. 214: 55–65.

43. QGIS Development Team 2016. QGIS Geographic Information System. Open Source Geospatial Foundation Project. Version 2.14.22. URL ¡http://qgis.osgeo.org¿

44. R Core Team 2018. R: A language and environment for statistical computing. Version 3.5.1. R Foundation for Statistical Computing, Vienna, Austria. URL ¡https://www.R-project.org¿

45. Ribeiro, J. W. et al. 2017. Passive acoustic monitoring as a complementary strategy to assess biodiversity in the Brazilian Amazonia. - Biodivers. Conserv. 26: 2999–3002.

46. Robinson, W. D. 1999. Long-term changes in the avifauna of Barro Colorado Island, Panama, a tropical forest isolate. - Conserv. Biol. 13: 85–97.

47. Rosenzweig, M. L. 1995. Species Diversity in Space and Time. - Cambridge University Press.

48. Russildi, G. et al. 2016. Species- and community-level responses to habitat spatial changes in fragmented rainforests: assessing compensatory dynamics in amphibians and reptiles. - Biodivers. Conserv. 25: 375–392.

49. Scheiner, S. M. 2003. Six types of species-area curves. - Glob. Ecol. Biogeogr. 12: 441–447.

50. Schietti, J. et al. 2014. Vertical distance from drainage drives floristic composition changes in an Amazonian rainforest. - Plant Ecol. Divers. 7: 241–253.

51. Smith, M. A. and Green, D. M. 2005. Dispersal and the metapopulation paradigm in amphibian ecology and conservation: are all amphibian populations metapopulations? - Ecography. 28: 110–128.

52. Stouffer, P. C. and Bierregaard, R. O. 1995. Use of Amazonian forest fragments by understory insectivorous birds. - Ecology 76: 2429–2445.

53. Stouffer, P. C. et al. 2009. Twenty years of understorey bird extinctions from Amazonian rain forest fragments: consistent trends and landscape-mediated dynamics. - Divers. Distrib. 15: 88–97.

54. Tjørve, E. 2003. Shapes and functions of species-area curves: a review of possible models. - J. Biogeogr. 30: 827–835.

55. Tjørve, E. et al. 2018. Great theories of species diversity in space and why they were forgotten: The beginnings of a spatial ecology and the Nordic early 20th-century botanists. - J. Biogeogr. 45: 530–540.

56. Triantis, K. A. et al. 2012. The island species-area relationship: biology and statistics. - J. Biogeogr. 39: 215–231.

57. Wang, Y. et al. 2016. On empty islands and the small-island effect. - Glob. Ecol. Biogeogr. 25: 1333–1345.

58. Wang, Y. et al. 2018. The small-island effect in amphibian assemblages on subtropical land-bridge islands of an inundated lake. - Curr. Zool. 64: 303–309.

59. Watling, J. I. and Donnelly, M. A. 2008. Species richness and composition of amphibians and reptiles in a fragmented forest landscape in northeastern Bolivia. - Basic Appl. Ecol. 9: 523–532.

60. Watson, D. M. 2002. A conceptual framework for studying species composition in fragments, islands and other patchy ecosystems. - J. Biogeogr. 29: 823–834.

61. Wilson, E. O. 1961. The nature of the taxon cycle in the Melanesian ant fauna. - Am. Nat. 95: 169–193.

62. Yu, M. et al. 2012. Richness and composition of plants and birds on land-bridge islands: effects of island attributes and differential responses of species groups. - J. Biogeogr. 39: 1124–1133.

63. Zimmerman, B. L. and Bierregaard, R. O. 1986. Relevance of the equilibrium theory of island biogeography and species-area relations to conservation with a case from Amazonia. - J. Biogeogr. 13: 133.

5.8   Supporting Information

This supporting information contains:

- Table S5.1

- Table S5.2

- Figure S5.1

- Appendix S5.1

- R code

    Click here to download the R code in .Rmd file format

Table S5.1: Description of the 78 sites surveyed at the Balbina Hydroelectric Reservoir landscape, including 74 forest islands and four continuous forest (CF) sites. 'Sampling effort' corresponds to the number of 1-min recordings.

Site code	Island size (ha)	Number of recorders	Sampling effort	Observed richness	Rarefied richness	Sampling completeness
1	0.45	1	62	5	5.000	1.000
2	0.45	1	62	1	1.000	1.000
3	0.63	1	62	2	2.000	1.000
4	0.63	1	59	4	4.000	1.000
5	0.63	1	62	0	NA	NA
6	0.72	1	62	5	5.000	1.000
7	0.91	1	62	3	3.000	1.000
8	0.91	1	62	9	9.000	1.000
9	1.27	1	62	1	1.000	1.000
10	1.36	1	62	2	2.000	1.000
11	1.45	1	62	2	2.000	1.000
12	1.54	1	62	7	7.000	1.000
13	1.81	1	62	3	3.000	1.000
14	1.99	1	62	3	3.000	1.000
15	2.08	1	62	2	2.000	1.000
16	2.08	1	62	4	4.000	1.000
17	2.08	1	62	9	9.000	1.000
18	2.26	1	62	1	1.000	1.000
19	2.63	1	62	7	7.000	1.000
20	3.08	1	62	3	3.000	1.000
21	3.53	1	62	1	1.000	1.000
22	3.53	1	62	3	3.000	0.508
23	3.62	1	62	3	3.000	1.000
24	4.17	1	62	4	4.000	1.000
25	4.98	1	62	3	3.000	1.000
26	5.43	1	62	2	2.000	1.000
27	5.61	1	62	4	4.000	1.000
28	6.79	1	62	1	1.000	1.000
29	6.88	1	62	3	3.000	1.000
30	7.43	1	62	5	5.000	1.000
31	8.15	1	62	7	7.000	0.980
32	8.15	1	62	7	7.000	1.000
33	8.42	1	62	5	5.000	1.000
34	8.78	1	62	3	3.000	1.000
35	9.42	1	62	6	6.000	1.000
36	11.96	1	62	5	5.000	1.000
37	13.04	1	62	7	7.000	0.988
38	13.22	1	60	3	3.000	1.000
39	13.31	1	62	9	9.000	1.000
40	15.67	1	62	2	2.000	1.000

Site code	Island size (ha)	Number of recorders	Sampling effort	Observed richness	Rarefied richness	Sampling completeness
41	16.94	1	62	10	10.000	1.000
42	17.57	1	62	8	8.000	1.000
43	17.66	1	62	3	3.000	1.000
44	21.37	1	62	3	3.000	1.000
45	22.01	1	62	1	1.000	1.000
46	29.62	2	124	8	7.253	0.972
47	32.78	2	124	4	3.251	0.928
48	32.87	2	124	1	0.752	0.504
49	35.60	2	93	5	4.996	0.999
50	35.87	2	124	5	4.924	0.991
51	38.94	2	124	3	3.000	1.000
52	39.12	2	124	4	3.927	0.992
53	39.67	2	124	7	5.756	0.948
54	39.94	2	124	9	8.250	0.988
55	50.08	2	124	7	6.249	0.986
56	52.71	2	124	4	3.000	0.982
57	53.35	2	118	5	3.827	0.966
58	70.55	2	124	8	7.490	0.984
59	77.80	2	124	10	9.456	0.993
60	91.30	3	186	8	7.999	1.000
61	108.76	3	186	11	10.483	0.987
62	171.73	3	186	12	10.401	0.986
63	198.52	3	132	8	7.011	0.936
64	217.63	3	186	13	10.618	0.977
65	230.70	3	186	9	8.421	0.993
66	232.49	3	186	11	10.432	0.989
67	336.02	3	186	9	8.382	0.990
68	466.60	3	186	12	11.034	0.986
69	499.91	4	248	13	11.069	0.970
70	638.66	5	310	18	15.095	0.975
71	668.03	7	434	18	13.491	0.976
72	941.71	4	248	16	11.685	0.958
73	1350.56	4	248	21	16.494	0.969
74	1698.84	4	248	13	10.998	0.969
75	CF	10	614	20	13.355	0.937
76	CF	5	310	14	12.219	0.975
77	CF	4	248	13	11.419	0.969
78	CF	4	248	16	12.676	0.962

Table S5.2: Anuran species recorded across 78 surveyed sites at the Balbina Hydroelectric Reservoir landscape, including 74 forest islands and four continuous forest sites. 'Sites' corresponds to the number of sites at which species were captured. Species reproductive modes were assigned according to Haddad and Prado (2005). Taxonomy follows Frost et al. (2019).

Family	Species	Sites	Reproductive mode
Craugastoridae	*Pristimantis fenestratus*	10	23
	Pristimantis ockendeni	18	23
	Pristimantis zimmermanae	46	23
Bufonidae	*Atelopus hoogmoedi*	15	1
	Rhaebo guttatus	1	1
	Rhinella marina	2	1
	Rhinella merianae	2	1
Ceratophryidae	*Ceratophrys cornuta*	6	1
Aromobatidae	*Anomaloglossus stepheni*	28	21
Dendrobatidae	*Ameerega hahneli*	12	20
	Ameerega trivittata	54	20
Hylidae	*Boana boans*	34	4
	Boana calcarata	19	1
	Boana cinerascens	24	1
	Boana lanciformis	5	1
	Boana fasciata	3	1
	Dendropsophus brevifrons	39	24
	Dendropsophus minusculus	45	1
	Dendropsophus parviceps	8	1
	Osteocephalus buckleyi	1	2
	Osteocephalus oophagus	12	6
	Osteocephalus taurinus	6	1
	Trachycephalus coriaceus	4	1
	Trachycephalus resinifictrix	2	26
	Scinax garbei	1	1
	Scinax ruber	4	1
Leptodactylidae	*Adenomera andreae*	35	32
	Adenomera hylaedactyla	12	32
	Leptodactylus knudseni	7	13
	Leptodactylus longirostris	1	13
	Leptodactylus pentadactylus	14	13
	Leptodactylus stenodema	5	13
Microhylidae	*Chiasmocleis shudikarensis*	9	1
	Elachistocleis bicolor	21	1
	Synapturanus mirandaribeiroi	8	23
Phyllomedusidae	*Phyllomedusa tarsius*	1	24
	Phyllomedusa vaillantii	2	24

References

Frost, D. R. 2019. Amphibian Species of the World: An Online Reference. Version 6.0 (27 February 2019). - Am. Museum Nat. Hist. New York, USA. URL <http//research.amnh.org/herpetology/amphibia/index.html>

Haddad, C. F. B. and Prado, C. P. D. A. 2005. Reproductive modes in frogs and their unexpected diversity in the Atlantic Forest of Brazil. - Bioscience 55: 207.
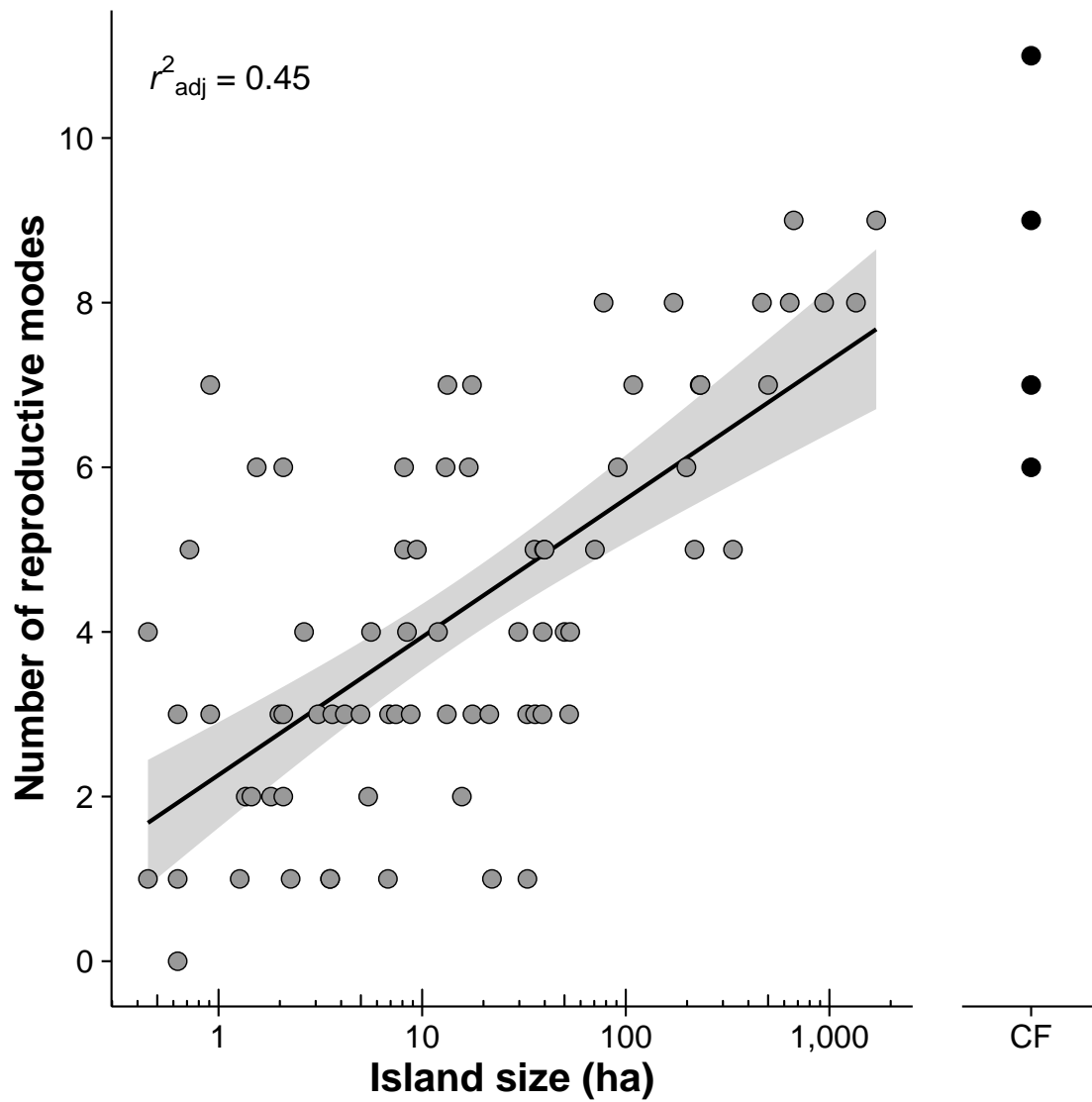
Figure S5.1: Relationship between the number of reproductive modes represented by at least one anuran species and island size across 74 forest islands (grey circles) surveyed at the Balbina Hydroelectric Reservoir landscape. Continuous forest sites (CF, black circles) were not included in the linear fit.

Appendix S5.1: Data sources

1. Alcala, E. L. et al. 2004. Amphibians and reptiles in tropical rainforest fragments on Negros Island, the Philippines. - Environ. Conserv. 31: 254–261.

2. Almeida-Gomes, M. and Rocha, C. F. D. 2014. Landscape connectivity may explain anuran species distribution in an Atlantic forest fragmented area. - Landsc. Ecol. 29: 29–40.

3. Almeida-Gomes, M. et al. 2016. Patch size matters for amphibians in tropical fragmented landscapes. - Biol. Conserv. 195: 89–96.

4. Bell, K. E. 2005. Conservation ecology of amphibians and reptiles in Sarapiqui, Costa Rica: forest fragmentation and long-term population change. MSc Thesis. Florida International University, USA.

5. Bickford, D. et al. 2010. Forest fragment and breeding habitat characteristics explain frog diversity and abundance in Singapore. - Biotropica 42: 119–125.

6. Bittencourt-Silva, G. B. and Silva, H. R. 2014. Effects of fragmentation and sea-level changes upon frog communities of land-bridge islands off the southeastern coast of Brazil. - PLoS One 9: e103522.

7. Cabrera-Guzmán, E. and Reynoso, V. H. 2012. Amphibian and reptile communities of rainforest fragments: minimum patch size to support high richness and abundance. - Biodivers. Conserv. 21: 3243–3265.

8. Faria, L. F. de 2017. A síndrome das florestas vazias e a importância dos pequenos fragmentos para a conservação dos anfíbios. MSc Thesis. Instituto Nacional de Pesquisas da Amazônia, Brazil.

9. Hager, H. A. 1998. Area-sensitivity of reptiles and amphibians: Are there indicator species for habitat fragmentation? - Écoscience 5: 139–147.

10. Herrera, J. B. 2011. Efeitos da heterogeneidade do ambiente, área e variáveis ambientais sobre anfíbios anuros em paisagem fragmentada de Floresta Atlântica. MSc Thesis. Universidade Federal da Bahia, Brazil.

11. Hillers, A. et al. 2008. Effects of forest fragmentation and habitat degradation on west African leaf-litter frogs. - Conserv. Biol. 22: 762–772.

12. Kryštufek, B. and Kletečki, E. 2007. Biogeography of small terrestrial vertebrates on the Adriatic landbridge islands. - Folia Zool. 56: 225–234.

13. Li, B. et al. 2018. Influence of breeding habitat characteristics and landscape heterogeneity on anuran species richness and abundance in urban parks of Shanghai, China. - Urban For. Urban Green. 32: 56–63.

14. Lima, J. R. et al. 2015. Amphibians on Amazonian land-bridge islands are affected more by area than isolation. - Biotropica 47: 369–376.

15. Lion, M. B. et al. 2014. Split distance: a key landscape metric shaping amphibian populations and communities in forest fragments. - Divers. Distrib. 20: 1245–1257.

16. Palmeirim, A. F. et al. 2017. Herpetofaunal responses to anthropogenic forest habitat modification across the neotropics: insights from partitioning β-diversity. - Biodivers. Conserv. 26: 2877–2891.

17. Pineda, E. and Halffter, G. 2004. Species diversity and habitat fragmentation: frogs in a tropical montane landscape in Mexico. - Biol. Conserv. 117: 499–508.

18. Rakotondravony, H. A. 2007. Conséquences de la variation des superficies forestières sur les communautés de reptiles et d'amphibiens dans la région Loky-Manambato, extrême nord-est de Madagascar. - Rev. d'Écologie (La Terre La Vie) 62: 209–227.

19. Russildi, G. et al. 2016. Species- and community-level responses to habitat spatial changes in fragmented rainforests: assessing compensatory dynamics in amphibians and reptiles. - Biodivers. Conserv. 25: 375–392.

20. Vallan, D. 2000. Influence of forest fragmentation on amphibian diversity in the nature reserve of Ambohitantely, highland Madagascar. - Biol. Conserv. 96: 31–43.

21. Wang, Y. et al. 2018. The small-island effect in amphibian assemblages on subtropical land-bridge islands of an inundated lake. - Curr. Zool. 64: 303–309.

22. Watling, J. I. et al. 2009. Nested species subsets of amphibians and reptiles on Neotropical forest islands. - Anim. Conserv. 12: 467–476.

23. Williams, S. E. and Pearson, R. G. 1997. Historical rainforest contractions, localized extinctions and patterns of vertebrate endemism in the rainforests of Australia's wet tropics. - Proc. R. Soc. B Biol. Sci. 264: 709–716.

24. Yiming, L. et al. 1998. Nested distribution of amphibians in the Zhoushan archipelago, China: can selective extinction cause nested subsets of species? - Oecologia 113: 557–564.

25. Zimmerman, B. L. and Bierregaard, R. O. 1986. Relevance of the equilibrium Theory of Island Biogeography and species-area relations to conservation with a case from Amazonia. - J. Biogeogr. 13: 133.

Supporting Information

Sampling design may obscure species-area relationships in landscape-scale field studies

```r
# Clear workspace and disable scientific notation
remove(list = ls()); options(scipen = 999)
```

Packages

```r
# Load required packages
library(vegan)
library(dplyr)
library(ggplot2)
library(gridExtra)
library(iNEXT)
```

Dataset

Data are available from the KNB repository and should be cited as:

Anderson Saldanha Bueno. 2019. Balbina Frog Data, 2015. Knowledge Network for Biocomplexity. urn:uuid:fb6c7193-eca5-41ba-89dd-146c31c9dbe0.

Import data

```r
# Fieldwork data derived from autonomous recordings units
rawdata = read.csv("https://ndownloader.figshare.com/files/15158558")
# Remove "Malfunctioning" recordings
rawdata = subset(rawdata, rawdata$species != "Malfunctioning")
# Remove the level "Malfunctioning" from "rawdata$species"
rawdata$species = droplevels(rawdata$species)
# Add a column indicating that each record corresponds to one detection
rawdata$occurrence = 1

# Site area (hectares)
area = read.csv("https://ndownloader.figshare.com/files/15158564")
rownames(area) = area$site
area = area[-1]

# Reproductive modes
traits = read.csv("https://ndownloader.figshare.com/files/15158561")
# Add a column with the reproductive of each species
rawdata$rep.mod = traits$reproductive_mode[match(rawdata$species, traits$species)]

# Site-by-reproductive mode matrix
rep.mod = tapply(rawdata$occurrence, list(rawdata$site, rawdata$rep.mod), sum)
# Species non detected in a given site are real zeros
rep.mod[is.na(rep.mod)] = 0
```

```r
# Coordinates of the surveyed plots (n = 151)
coordinates = read.csv("https://ndownloader.figshare.com/files/15158555", row.names = 1)
```

Figure S1

```r
# Data to draw the graph
rm.area = data.frame(rm = specnumber(rep.mod), area = area$area)

# Relationship between reproductive modes and island size
# Include only forest islands
summary(lm(rm ~ log10(area), data = rm.area[-c(47:50), ]))
```

```
##
## Call:
## lm(formula = rm ~ log10(area), data = rm.area[-c(47:50), ])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.8047 -1.1534 -0.0975  1.0647  4.8068
##
## Coefficients:
##             Estimate Std. Error t value        Pr(>|t|)
## (Intercept)   2.2618     0.3216   7.032 0.0000000009647 ***
## log10(area)   1.6765     0.2163   7.751 0.0000000000447 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.699 on 72 degrees of freedom
## Multiple R-squared:  0.4549, Adjusted R-squared:  0.4473
## F-statistic: 60.08 on 1 and 72 DF,  p-value: 0.0000000000447
```

```r
# Draw the graph of the relationship between reproductive modes and island size
graph.rm.isl =
  ggplot(aes(x = area, y = rm),
         # Include only forest islands
         data = rm.area[order(rm.area$area, rm.area$rm), ][-c(75:78), ]) +

  labs(x = "Island size (ha)",
       y = "Number of reproductive modes") +

  scale_x_log10(breaks = c(1, 10, 100, 1000),
                labels = c("1", "10", "100", "1,000")) +
  scale_y_continuous(limits = c(0, 11),
                     breaks = seq(0, 10, 2)) +
  annotation_logticks(base = 10, sides = "b") +

  geom_smooth(method = "lm", colour = "black") +
  geom_point(shape = 21, size = 4, colour = "black", fill = "#999999") +

  theme_classic(base_size = 20) +
  theme(axis.title = element_text(colour = "black", face = "bold"),
        axis.text = element_text(colour = "black"),
        axis.ticks = element_line(size = 0.5, colour = "black"),
```

```r
        axis.line = element_line(size = 0.5)) +

  annotate("text", x = min(rm.area$area), y = max(rm.area$rm),
           hjust = 0, vjust = 1, fontface = "bold", size = 6,
           parse = T, label = as.character(expression(italic(r)^{2}*""[adj]*" = 0.45")))

# Draw the graph for continuous forest sites
graph.rm.cf =
  ggplot() +

  scale_x_discrete(labels = c("16988.4" = "CF")) +
  scale_y_continuous(limits = c(min(rm.area$rm), max(rm.area$rm))) +

  geom_point(shape = 21, size = 4, colour = "black", fill = "black",
             aes(x = as.factor(area), y = rm),
             # Include only continuous forest sites
             data = subset(rm.area, rm.area$area > 2000)) +

    theme_classic(base_size = 20) +
    theme(axis.title.x = element_text(colour = "white", face = "bold"),
          axis.text.x = element_text(colour = "black"),
          axis.ticks.x = element_line(size = 0.5, colour = "black"),
          axis.line.x = element_line(size = 0.5),
          axis.title.y = element_blank(),
          axis.text.y = element_blank(),
          axis.ticks.y = element_blank(),
          axis.line.y = element_blank())

# Combine and save the graphs
ggsave(grid.arrange(graph.rm.isl, graph.rm.cf, ncol = 2, widths = c(5, 1)),
       file = "figs1.pdf", width = 20, height = 20, units = "cm")

# Display the graphs
#grid.arrange(graph.rm.isl, graph.rm.cf, ncol = 2, widths = c(5, 1))
```

Frog data

```r
# Create site-by-species matrix for 78 sites
frogs = as.data.frame(tapply(rawdata$occurrence,
                             list(rawdata$site, rawdata$species), sum))
# Species not detected in a site are real zeros
frogs[is.na(frogs)] = 0
# Remove the column "None"
frogs = frogs[-which(colnames(frogs) == "None")]
```

Number of species

The number of species recorded depends on the sampling effort – here measured as the number of 1-min recordings. Thus, the number of species increases with sampling effort until the asymptote of the species accumulation curve is reached. Since the sampling effort among sites was different (the larger the site area, the higher the number of recording stations and by extension of

1-min recordings), we calculated the **rarefied number of species** (to standardise sampling effort) as well as **sample coverage** to determine how close the observed number of species was to the "true"/estimated number of species.

We standardised the sampling effort to the most frequent number of 1-min recordings across surveyed sites (i.e. statistical mode; n = 62). We did so because the iNEXT package (Hsieh, Ma, & Chao, 2016) calculates both the **interpolated** and **extrapolated** number of species. Accordingly, we used the **interpolated**, **observed** and **extrapolated** number of species for sites with a sampling effort above (n = 33), equal to (n = 43) and below (n = 2) the statistical mode, respectively (hereafter, we refer to them all as the **rarefied number of species**). Still in the iNEXT package, we calculated sample coverage to assess whether surveyed sites were satisfactorily inventoried with 62 1-min recordings.

```
# Calculate the sampling effort (i.e. number of 1-min recordings) of each site (n = 78)
effort = as.vector(colSums(table(unique(rawdata[c("recording", "site", "occurrence")]))))
sort(effort) # sort sampling effort from the lowest to the highest
```

```
##  [1]  59  60  62  62  62  62  62  62  62  62  62  62  62  62  62  62  62
## [18]  62  62  62  62  62  62  62  62  62  62  62  62  62  62  62  62  62
## [35]  62  62  62  62  62  62  62  62  62  93 118 124 124 124 124
## [52] 124 124 124 124 124 124 124 124 132 186 186 186 186 186 186 186 186
## [69] 248 248 248 248 248 248 310 310 434 614
```

```
# Standardised sampling effort (i.e. statistical mode)
effort.mode = as.numeric(names(which(table(effort) == max(table(effort)))))
effort.mode
```

```
## [1] 62
```

Rarefied number of species

Data preparation

```
# Create recording-by-species matrices
frogs.recording = as.data.frame(tapply(rawdata$occurrence,
                                        list(rawdata$recording, rawdata$species), sum))
# Species not detected in a recording are real zeros
frogs.recording[is.na(frogs.recording)] = 0
# Remove the column "None"
frogs.recording = frogs.recording[-which(colnames(frogs.recording) == "None")]
```

```
# Create a recording-by-species matrix per site (n = 78)
# This format (matrices stored in a list object) is required
# to calculted the rarefied number of species per site using incidence
# (i.e. presence-absence) data
inext.site = split(frogs.recording, rawdata$site[match(rownames(frogs.recording),
                                                       rawdata$recording)])
```

```
# Transpose the data frames stored in the list objec
# This step creates species-by-sampling-unit matrices (i.e. species-by-recording matrices)
inext.site = lapply(inext.site, function(z) {t(z)})
```

```
# Remove sites where no species was recorded
inext.site = inext.site[-which(lapply(inext.site, sum) == 0)] # 1 site removed (#74)
```

Calculation of the rarefied number of species

```r
# Rarefied number of species per site standardised to 62 recordings
richness.site = matrix(nrow = 77, ncol = 5)
for (i in 1:77){
  # Remove site "Toquinho" (#74) where no species was recorded
  richness.site[i, 1] = rownames(frogs)[-74][i]
  richness.site[i, 2:5] = as.numeric(iNEXT(inext.site[[i]], q = 0,
                                      datatype = "incidence_raw", # iNEXT::iNEXT()
                                      endpoint = effort.mode, knot = effort.mode,
                                      nboot = 1000)$iNextEst[effort.mode, 4:7])
}

# Add site "Toquinho" (#74)
richness.site = rbind(richness.site, cbind("Toquinho", "0", "0", "0", "NA"))

# Order table by site names
richness.site = rbind(richness.site[1:73, ],
                      cbind("Toquinho", "0", "0", "0", "NA"),
                      richness.site[74:77, ])

# Check if the site "Toquinho" was placed back in the right row (#74)
richness.site[74, 1] == "Toquinho"
```

```r
# Finish the data frame
# Name rows according to site names
rownames(richness.site) = richness.site[, 1]
# Remove the column "site"
richness.site = richness.site[, -1]
# Rename columns
colnames(richness.site) = c("richness", "lowerCI", "upperCI", "coverage")
# Convert the matrix "richness.site" into a data frame
richness.site = as.data.frame(richness.site)

# Convert factor data into numeric data
richness.site$richness = as.numeric(as.character(richness.site$richness))
# Convert factor data into numeric data
richness.site$lowerCI = as.numeric(as.character(richness.site$lowerCI))
# Convert factor data into numeric data
richness.site$upperCI = as.numeric(as.character(richness.site$upperCI))
# Convert factor data into numeric data
richness.site$coverage = as.numeric(as.character(richness.site$coverage))
```

Correlation between observed number of species and sampling effort

```r
# Pearson's product-moment correlation between
# observed number of species and sampling effort
cor.test(specnumber(frogs), effort, method = "pearson")
```

```
##
##  Pearson's product-moment correlation
##
```

```
## data:  specnumber(frogs) and effort
## t = 12.452, df = 76, p-value < 0.00000000000000022
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##   0.7296882 0.8811067
## sample estimates:
##      cor
## 0.819199
```

Figure 2a

```
# Data to draw the graph
obs.effort = data.frame(obs = specnumber(frogs), effort = effort)

# Draw the graph of the correlation between
# observed number of species and sampling effort
graph.obs.effort =
  ggplot(aes(x = effort, y = obs),
         data = obs.effort[order(obs.effort$effort, obs.effort$obs), ]) +

  labs(x = "Sampling effort",
       y = "Observed number of species") +

  scale_x_continuous(breaks = c(62, 124, 186, 248, 310, 434, 614)) +

  geom_point(shape = 21, size = 4, colour = "black", fill = "#999999") +

  theme_bw(base_size = 20) +
  theme(panel.grid = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.title = element_text(colour = "black", face = "bold"),
        axis.text = element_text(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5)) +

  annotate("text", x = min(obs.effort$effort), y = max(obs.effort$obs),
           hjust = 0, vjust = 1, fontface = "bold", size = 8,
           parse = T, label = as.character(expression(italic(r)*" = 0.82"))) +
  annotate("text", x = Inf, y = -Inf,
           hjust = 1.3, vjust = -1, fontface = "bold", size = 10, label = "(a)")

#graph.obs.effort
```

Correlation between observed and rarefied number of species

```
# Pearson's product-moment correlation between observed and rarefied number of species
cor.test(specnumber(frogs), richness.site$richness, method = "pearson")
```

```
##
##  Pearson's product-moment correlation
##
## data:  specnumber(frogs) and richness.site$richness
```

```
## t = 47.348, df = 76, p-value < 0.00000000000000022
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.9741280 0.9894553
## sample estimates:
##       cor
## 0.9834688
```

Figure 2b

```r
# Data to draw the graph
# The rarefied number of speices for the site "Toquinho" was assigned to 0 (zero)
obs.raref = data.frame(obs = specnumber(frogs), raref = richness.site$richness)

# Draw the graph of the correlation between observed and rarefied number of species
graph.obs.raref =
  ggplot(aes(x = raref, y = obs),
         data = obs.raref[order(obs.raref$raref, obs.raref$obs), ]) +

  labs(x = "Rarefied number of species",
       y = "Observed number of species") +

  geom_point(shape = 21, size = 4, colour = "black", fill = "#999999") +
  geom_point(shape = 21, size = 4, colour = "black", fill = "white",
             aes(x = 0, y = 0)) +

  theme_bw(base_size = 20) +
  theme(panel.grid = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.title = element_text(colour = "black", face = "bold"),
        axis.text = element_text(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5)) +

  annotate("text", x = min(obs.raref$raref), y = max(obs.raref$obs),
           hjust = 0, vjust = 1, fontface = "bold", size = 8,
           parse = T, label = as.character(expression(italic(r)*" = 0.98"))) +
  annotate("text", x = Inf, y = -Inf,
           hjust = 1.3, vjust = -1, fontface = "bold", size = 10, label = "(b)")

#graph.obs.raref
```

Sample coverage

Figure 2c

```r
# Data to draw the graph
coverage = data.frame(site = NA, coverage = richness.site[4])
# Order the data frame according to site size
coverage = coverage[order(area$area), ]
# Add site number
coverage$site = 1:nrow(coverage)
```

```r
# Draw the graph showing sample coverage per site
graph.coverage =
  ggplot(aes(x = site, y = coverage),
         data = coverage) +

  labs(x = "Site code",
       y = "Sample coverage") + # sample coverage in 62 1-min recordings

  scale_x_continuous(breaks = c(1, 10, 20, 30, 40, 50, 60, 70, 78)) +
  scale_y_continuous(limits = c(0.5, 1)) +

  geom_point(shape = 21, size = 4, colour = "black", fill = "#999999") +
  geom_point(shape = 21, size = 4, colour = "black", fill = "white",
             aes(x = 5, y = 0.5)) +

  theme_bw(base_size = 20) +
  theme(panel.grid = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.title = element_text(colour = "black", face = "bold"),
        axis.text = element_text(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5)) +

  annotate("text", x = Inf, y = -Inf,
           hjust = 1.3, vjust = -1, fontface = "bold", size = 10, label = "(c)")

#graph.coverage
```

Combine graphs

Figure 2

```r
# Combine and save the graphs
ggsave(grid.arrange(graph.obs.effort, graph.obs.raref, graph.coverage, ncol = 3),
       file = "fig2.pdf", width = 17*3, height = 17, units = "cm")
```

Sample coverage was above 90% for 75 out of 78 sites, indicating that our sampling effort was overall satisfactory.

Table S1

Site attributes

```r
tables1 = data.frame(Island.size = area$area, # forest area in hectares
                     Recorders = area$recorders, # number of recording stations
                     Effort = effort, # number of 1-min recordings
                     S.observed = specnumber(frogs),
                     S.rarefied = richness.site$richness,
                     # lower bound of the 95% confidence intervals
                     S.rarefied.lowerCI = richness.site$lowerCI,
                     # upper bound of the 95% confidence intervals
                     S.rarefied.upperCI = richness.site$upperCI,
                     Coverage = richness.site$coverage)
```

```r
# Order table by site area
tables1 = tables1[order(tables1$Island.size), ]
# Number surveyed sites (the numbers match those of Fig. 2c)
tables1$Site.code = 1:nrow(tables1)
# Move "site.code" to the first column
tables1 = tables1[, c(9, 1:8)]
rownames(tables1) = NULL

tables1$Island.size[tables1$Island.size == 16988.40] = "Continuous forest"

#tables1
```

Table S2

List of frog species recorded across all 151 plots in 78 surveyed sites and the number of sites occupied per species. Taxonomy follows Frost (2018).

- Because of taxonomic revision, *Boana fasciata* is no longer considered present in Brazil according to Frost (2018). However, the species present in Brazil (i.e. *Boana aff. fasciata*) has not yet been assign to a new taxon. Therefore, we held the name *Boana fasciata*.

- *Leptodactylus knudseni* and *L. pentadactylus* are hardly distinguished through vocalisation by both hearing and inspecting sonograms. In the fieldwork data (available online at KNB repository), we assigned *Leptodactylus knudseni pentadactylus* whenever our best guess was *Leptodactylus knudseni*, and *Leptodactylus pentadactylus knudseni* whenever our best guess was *L. pentadactylus*. Thus, we acknowledge that one species maybe the other in some instances.

```r
tables2 = names(frogs)

# Rename two species
tables2[tables2 == "Leptodactylus knudseni pentadactylus"] = "Leptodactylus knudseni"
tables2[tables2 == "Leptodactylus pentadactylus knudseni"] = "Leptodactylus pentadactylus"

# Split genus and specific epithet into separete columns
tables2 = data.frame(do.call(rbind, strsplit(as.character(tables2), " ", fixed = TRUE)))

# Species family
families = c("Leptodactylidae", "Leptodactylidae", "Dendrobatidae", "Dendrobatidae",
             "Aromobatidae", "Bufonidae", "Hylidae", "Hylidae", "Hylidae",
             "Hylidae", "Hylidae", "Ceratophryidae", "Microhylidae", "Hylidae",
             "Hylidae", "Hylidae", "Microhylidae", "Leptodactylidae", "Leptodactylidae",
             "Leptodactylidae", "Leptodactylidae", "Hylidae", "Hylidae", "Hylidae",
             "Phyllomedusidae", "Phyllomedusidae", "Craugastoridae", "Craugastoridae",
             "Craugastoridae", "Bufonidae", "Bufonidae", "Bufonidae", "Hylidae",
             "Hylidae", "Microhylidae", "Hylidae", "Hylidae")

# Taxonomic sequence
id = c(860, 867, 333, 337, 326, 153, 401, 407, 409, 436, 428, 251, 971,
       498, 526, 534, 980, 895, 901, 911, 923, 572, 578, 581, 1077, 1079,
       89, 99, 114, 193, 221, 223, 695, 717, 999, 601, 610)

# Species reproductive modes
spp.rep.mod = c(32, 32, 20, 20, 21, 1, 4, 1, 1, 1, 1, 1, 1, 24, 1, 1, 1,
    13, 13, 13, 13, 2, 6, 1, 24, 24, 23, 23, 23, 1, 1, 1, 1, 1, 23, 1, 26)
```

```
# Species list and number of sites where each species was recorded
tables2 = data.frame(id = id,
                     Family = families,
                     Species = as.character(paste(tables2$X1, tables2$X2, sep = " ")),
                     Sites = cbind(colSums(decostand(frogs, method = "pa"))),
                     Reproductive.mode = spp.rep.mod)

tables2 = tables2[order(tables2$id), ] # order table according to the taxonomic sequence
tables2 = tables2[-1] # remove column "id"
rownames(tables2) = 1:nrow(tables2) # number the rows sequentially

#tables2
```

Species-area relationships (ISAR)

```
# Data to create ISAR models and draw the graph
sar = data.frame(richness = richness.site$richness, area)
# Order table by site area followed by species richness
sar = sar[order(sar$area, sar$richness), ]

# Create a column with site size category
sar$class[sar$area < 4] = "very_small" # up to 4.00 ha = very small
sar$class[sar$area > 4 & sar$area <= 20 ] = "small" # 4.01 - 20.00 ha = small
sar$class[sar$area > 20 & sar$area <= 100 ] = "medium" # 20.01 - 100.00 ha = medium
sar$class[sar$area > 100 & sar$area <= 400 ] = "large" # 100.01 - 400.00 = large
sar$class[sar$area > 400 & sar$area < 2000 ] = "very_large" # > 400.01 = very large
sar$class[sar$area > 2000 ] = "continuous" # continuous forest sites

# Create a column with an unique number per site
sar$id = 1:nrow(sar)
```

ISAR – semi-log models

```
# ISAR models
sar.74 = lm(richness ~ log10(area), data = subset(sar, sar$area < 2000))
sar.67 = lm(richness ~ log10(area), data = subset(sar, sar$area <= 400))
sar.60 = lm(richness ~ log10(area), data = subset(sar, sar$area <= 100))
sar.43 = lm(richness ~ log10(area), data = subset(sar, sar$area <= 20))
sar.23 = lm(richness ~ log10(area), data = subset(sar, sar$area < 4))

# ISAR results
sar.results = rbind(
  c(sar.74$coefficients[2], confint(sar.74)[2],
    confint(sar.74)[4], summary(sar.74)$adj.r.squared),
  c(sar.67$coefficients[2], confint(sar.67)[2],
    confint(sar.67)[4], summary(sar.67)$adj.r.squared),
  c(sar.60$coefficients[2], confint(sar.60)[2],
    confint(sar.60)[4], summary(sar.60)$adj.r.squared),
  c(sar.43$coefficients[2], confint(sar.43)[2],
    confint(sar.43)[4], summary(sar.43)$adj.r.squared),
```

```
  c(sar.23$coefficients[2], confint(sar.23)[2],
    confint(sar.23)[4], summary(sar.23)$adj.r.squared))

# Convert the matrix "sar.results" in to a data frame
sar.results = as.data.frame(sar.results)
# Name the columns
colnames(sar.results) = c("slope", "lowerCI", "upperCI", "r2")

sar.results
```

```
##        slope     lowerCI  upperCI           r2
## 1 2.8364991  2.17011698 3.502881   0.49305716
## 2 2.0359273  1.23228702 2.839568   0.27151241
## 3 1.3001936  0.31134217 2.289045   0.09129034
## 4 1.6885078  0.09017261 3.286843   0.07797120
## 5 0.2327528 -3.67518220 4.140688  -0.04685428
```

```
# Draw the graph for forest islands
graph.sar =
  ggplot() +

  labs(x = "Island size (ha)",
       y = "Rarefied number of species") +

  scale_x_log10(limits = c(0.4, NA),
                breaks = c(1, 10, 100, 1000),
                labels = c("1", "10", "100", "1,000")) +
  scale_y_continuous(limits = c(min(sar$richness), max(sar$richness))) +
  annotation_logticks(base = 10, sides = "b") +

  geom_vline(xintercept = 4, alpha = 0.2, linetype = "dashed") +
  geom_vline(xintercept = 20, alpha = 0.2, linetype = "dashed") +
  geom_vline(xintercept = 100, alpha = 0.2, linetype = "dashed") +
  geom_vline(xintercept = 400, alpha = 0.2, linetype = "dashed") +

  geom_smooth(size = 1, method = "lm", colour = "#4daf4a", se = FALSE,
              aes(x = area, y = richness),
              data = subset(sar, sar$area < 2000)) +
  geom_smooth(size = 1, method = "lm", colour = "dodgerblue", se = FALSE,
              aes(x = area, y = richness),
              data = subset(sar, sar$area <= 400)) +
  geom_smooth(size = 1, method = "lm", colour = "#984ea3", se = FALSE,
              aes(x = area, y = richness),
              data = subset(sar, sar$area <= 100)) +
  geom_smooth(size = 1, method = "lm", colour = "#ff7f00", se = FALSE,
              aes(x = area, y = richness),
              data = subset(sar, sar$area <= 20)) +
  geom_smooth(size = 1, method = "lm", colour = "#e41a1c", se = FALSE,
              aes(x = area, y = richness),
              data = subset(sar, sar$area < 4)) +

  geom_point(shape = 21, size = 4, colour = "black", fill = "#e41a1c",
             aes(x = area, y = richness),
             data = subset(sar, sar$class == "very_small")) +
```

```
geom_point(shape = 21, size = 4, colour = "black", fill = "#ff7f00",
           aes(x = area, y = richness),
           data = subset(sar, sar$class == "small")) +
geom_point(shape = 21, size = 4, colour = "black", fill = "#984ea3",
           aes(x = area, y = richness),
           data = subset(sar, sar$class == "medium")) +
geom_point(shape = 21, size = 4, colour = "black", fill = "dodgerblue",
           aes(x = area, y = richness),
           data = subset(sar, sar$class == "large")) +
geom_point(shape = 21, size = 4, colour = "black", fill = "#4daf4a",
           aes(x = area, y = richness),
           data = subset(sar, sar$class == "very_large")) +

theme_classic(base_size = 20) +
theme(axis.title = element_text(colour = "black", face = "bold"),
      axis.text = element_text(colour = "black"),
      axis.ticks = element_line(size = 0.5, colour = "black"),
      axis.line = element_line(size = 0.5)) +

annotate("text", x = 0.4, y = max(sar$richness),
         hjust = 0, vjust = 0, fontface = "bold", size = 4.5,
         parse = T, label = as.character(expression(bolditalic(r)^{2}*""[adj]))) +
annotate("text", x = 0.4, y = max(sar$richness) - 1,
         hjust = 0, vjust = 0, size = 4.5, colour = "#4daf4a",
         label = round(sar.results$r2[1], 2)) +
annotate("text", x = 0.4, y = max(sar$richness) - 2,
         hjust = 0, vjust = 0, size = 4.5, colour = "dodgerblue",
         label = round(sar.results$r2[2], 2)) +
annotate("text", x = 0.4, y = max(sar$richness) - 3,
         hjust = 0, vjust = 0, size = 4.5, colour = "#984ea3",
         label = round(sar.results$r2[3], 2)) +
annotate("text", x = 0.4, y = max(sar$richness) - 4,
         hjust = 0, vjust = 0, size = 4.5, colour = "#ff7f00",
         label = round(sar.results$r2[4], 2)) +
annotate("text", x = 0.4, y = max(sar$richness) - 5,
         hjust = 0, vjust = 0, size = 4.5, colour = "#e41a1c",
         label = round(sar.results$r2[5], 2)) +

annotate("text", x = min(sar$area) + 0.9, y = max(sar$richness),
         hjust = 0, vjust = 0, fontface = "bold", size = 4.5,
         parse = T, label = as.character(expression(bold(Slope)^{}*""[]))) +
annotate("text", x = min(sar$area) + 0.9, y = max(sar$richness) - 1,
         hjust = 0, vjust = 0, size = 4.5, colour = "#4daf4a",
         label = round(sar.results$slope[1], 2)) +
annotate("text", x = min(sar$area) + 0.9, y = max(sar$richness) - 2,
         hjust = 0, vjust = 0, size = 4.5, colour = "dodgerblue",
         label = round(sar.results$slope[2], 2)) +
annotate("text", x = min(sar$area) + 0.9, y = max(sar$richness) - 3,
         hjust = 0, vjust = 0, size = 4.5, colour = "#984ea3",
         label = paste(round(sar.results$slope[3], 2), "0", sep = "")) +
annotate("text", x = min(sar$area) + 0.9, y = max(sar$richness) - 4,
         hjust = 0, vjust = 0, size = 4.5, colour = "#ff7f00",
         label = round(sar.results$slope[4], 2)) +
```

```r
  annotate("text", x = min(sar$area) + 0.9, y = max(sar$richness) - 5,
          hjust = 0, vjust = 0, size = 4.5, colour = "#e41a1c",
          label = round(sar.results$slope[5], 2))

# Draw the graph for continuous forest sites
graph.cf =
  ggplot() +

  scale_x_discrete(labels = c("16988.4" = "CF")) +
  scale_y_continuous(limits = c(min(sar$richness), max(sar$richness))) +

  geom_point(shape = 21, size = 4, colour = "black", fill = "black",
            aes(x = as.factor(area), y = richness),
            data = subset(sar, sar$class == "continuous")) +

    theme_classic(base_size = 20) +
    theme(axis.title.x = element_text(colour = "white", face = "bold"),
          axis.text.x = element_text(colour = "black"),
          axis.ticks.x = element_line(size = 0.5, colour = "black"),
          axis.line.x = element_line(size = 0.5),
          axis.title.y = element_blank(),
          axis.text.y = element_blank(),
          axis.ticks.y = element_blank(),
          axis.line.y = element_blank())
```

Figure 3

```r
# Combine and save the graphs
ggsave(grid.arrange(graph.sar, graph.cf, ncol = 2, widths = c(5, 1)),
       file = "fig3.pdf", width = 20, height = 20, units = "cm")

# Display the graphs
#grid.arrange(graph.sar, graph.cf, ncol = 2, widths = c(5, 1))
```

Stratified random selection (full range)

```r
# Exclude continuous forest sites as they will not be used to produce species-area curves
sar.isl = sar[-c(75:78), ]

#############################
############# 25 #############
#############################

# Generate and save 100 random numbers
# sample(1:.Machine$integer.max, 100) # This command generated the following values
seed.full.25 = c(983130207, 383464058, 2017550925, 324031272, 1355246268,
    1466882829, 135869051, 2003688700, 670295328, 917206259, 1510315544,
    464385672, 1143533469, 1377184171, 272370526, 68508786, 865409067,
    572626257, 216861211, 1525481127, 1817804462, 1005149655, 844395643,
    115244292, 1577066482, 1362947697, 142092602, 299929719, 1392054026,
    1763778950, 1370019851, 1185576929, 606262959, 639263150, 1622779910,
```

```
    98441430, 1390362571, 670120600, 949563928, 186954742, 1574383261,
    986984798, 667612850, 1258974800, 512876766, 1427057290, 1651618970,
    219655950, 820545953, 387449755, 478465250, 853418124, 850681705, 273795464,
    364350513, 490331164, 1124787011, 21106045, 721221771, 2100270184,
    1759723902, 918280273, 886231669, 397522777, 696864331, 1907841509,
    696786153, 2097466496, 833093004, 1191889090, 2048925380, 642337978,
    2067077169, 971378045, 872419266, 424610369, 2001795318, 214703074,
    77460136, 1199542182, 25340488, 709315005, 508980850, 364845407, 674333206,
    748936353, 633356058, 120463230, 1498777378, 220758228, 883591132,
    796722650, 1119550301, 1328061235, 475824336, 687856599, 1846199925,
    233731714, 145666729, 470558232)


# Select 25 islands 100 times
sample.full.25 = matrix(nrow = 25, ncol = 100)

for (i in 1:100) {
  set.seed(seed.full.25[i])
  sample.full.25[, i] = data.frame(sar.isl %>% group_by(class) %>% sample_n(size = 5))$id
}

sample.full.25 = as.data.frame(sample.full.25) # convert the matrix into a data frame

# ISAR for each subset of 25 sites (n = 100)
slope.full.25 = r2.full.25 = p.full.25 = c()

for (i in 1:100) {
  slope.full.25[i] = lm(richness ~ log10(area),
                        # Regression slope
                        data = sar.isl[sample.full.25[, i], ])$coefficients[2]
  r2.full.25[i] = summary(lm(richness ~ log10(area),
                             # Model fit (r2)
                             data = sar.isl[sample.full.25[, i], ]))$adj.r.squared
  p.full.25[i] = summary(lm(richness ~ log10(area),
                            # p-value
                            data = sar.isl[sample.full.25[, i], ]))$coefficients[8]
}

# Store results in a data frame
sar.full.25 = data.frame(slope = slope.full.25,
                         r2 = r2.full.25, p = p.full.25, n.isl = 25)


###############################
############# 20 #############
###############################

# Generate and save 100 random numbers
# sample(1:.Machine$integer.max, 100) # This command generated the following values
seed.full.20 = c(1528803611, 1565203937, 1378459504, 1974162804, 1485726889,
    1722217419, 1083945832, 309442241, 2135586820, 904398258, 630425221,
    1571984846, 1881240840, 93071931, 1709954502, 1003892181, 242471852,
    374885596, 1471200150, 694386905, 1523256789, 1959936107, 102932741,
```

```
      1517589256, 2065922824, 1584981789, 815772258, 1290188023, 1479248051,
      1003253307, 1738192827, 1249221035, 592567972, 4847476, 241637366,
      856766312, 1250916602, 2097921953, 371046614, 1778211159, 732954012,
      2025770722, 1314949296, 357269182, 1399243432, 1255779923, 4390814,
      1154419746, 291235419, 1054835586, 980654265, 1066732189, 994545268,
      1009814846, 1894401743, 244382523, 766067747, 253772578, 1454223960,
      1007622885, 999021627, 1380324380, 1774930409, 214573555, 1492109997,
      1015109461, 747951269, 467889625, 785792844, 1613356238, 1152406342,
      780090909, 1582859756, 243340419, 955561798, 942675456, 1967585248,
      2116004579, 710299037, 1773143935, 360606364, 622966034, 1925227232,
      1285169228, 176215505, 572024367, 1009649767, 1424840105, 2053775408,
      690646673, 528845198, 398899374, 1986016719, 1228361023, 83457039,
      984747662, 578408909, 700617890, 1037532108, 776063694)

# Select 20 islands 100 times
sample.full.20 = matrix(nrow = 20, ncol = 100)

for (i in 1:100) {
  set.seed(seed.full.20[i])
  sample.full.20[, i] = data.frame(sar.isl %>% group_by(class) %>% sample_n(size = 4))$id
}

sample.full.20 = as.data.frame(sample.full.20) # convert the matrix into a data frame

# ISAR for each subset of 20 sites (n = 100)
slope.full.20 = r2.full.20 = p.full.20 = c()

for (i in 1:100) {
  slope.full.20[i] = lm(richness ~ log10(area),
                        # Regression slope
                        data = sar.isl[sample.full.20[, i], ])$coefficients[2]
  r2.full.20[i] = summary(lm(richness ~ log10(area),
                             # Model fit (r2)
                             data = sar.isl[sample.full.20[, i], ]))$adj.r.squared
  p.full.20[i] = summary(lm(richness ~ log10(area),
                            #  p-value
                            data = sar.isl[sample.full.20[, i], ]))$coefficients[8]
}

# Store results in a data frame
sar.full.20 = data.frame(slope = slope.full.20,
                         r2 = r2.full.20, p = p.full.20, n.isl = 20)


###############################
############# 15 #############
###############################

# Generate and save 100 random numbers
# sample(1:.Machine$integer.max, 100) # This command generated the following values
seed.full.15 = c(1032401521, 673699241, 1139750126, 142050191, 368795250,
    704181501, 663119545, 1012593324, 1282130941, 510145834, 1521414786,
    542345656, 99629308, 760242161, 156335171, 1443509149, 602490744, 1365239361,
```

```
      279730440, 182891111, 1948787162, 1446143351, 582764462, 211891494,
      264602481, 720226263, 1152228529, 535122942, 501508315, 1423455275,
      1938960957, 915067793, 1363188482, 516054131, 1718048524, 1558759182,
      1074540444, 1961880847, 197450330, 1958961253, 1987503820, 1470673232,
      872996307, 998258619, 24518139, 2035293581, 341314754, 659447199, 1505443998,
      286454269, 804362813, 945106698, 2003549731, 1365988459, 1213480577,
      1326706029, 550594977, 699060407, 1998830047, 56226447, 1142154175,
      668901331, 2063475684, 643829539, 945402271, 2124625018, 1690827344,
      948726833, 456447170, 866335614, 1032176747, 1794769238, 1958945420,
      2079412266, 1827338298, 645632015, 905251202, 1532502124, 1411979196,
      1184782471, 725942691, 1212558348, 1930511492, 921147012, 1364134716,
      1878989525, 721655347, 10415448, 125429553, 1711864830, 1294311706,
      484985294, 1710229680, 1927940813, 156218122, 543966632, 1557837957,
      245772028, 1153814762, 739406449)

# Select 15 islands 100 times
sample.full.15 = matrix(nrow = 15, ncol = 100)

for (i in 1:100) {
  set.seed(seed.full.15[i])
  sample.full.15[, i] = data.frame(sar.isl %>% group_by(class) %>% sample_n(size = 3))$id
}

sample.full.15 = as.data.frame(sample.full.15) # convert the matrix into a data frame

# ISAR for each subset of 15 sites (n = 100)
slope.full.15 = r2.full.15 = p.full.15 = c()

for (i in 1:100) {
  slope.full.15[i] = lm(richness ~ log10(area),
                        # Regression slope
                        data = sar.isl[sample.full.15[, i], ])$coefficients[2]
  r2.full.15[i] = summary(lm(richness ~ log10(area),
                             # Model fit (r2)
                             data = sar.isl[sample.full.15[, i], ]))$adj.r.squared
  p.full.15[i] = summary(lm(richness ~ log10(area),
                            # p-value
                            data = sar.isl[sample.full.15[, i], ]))$coefficients[8]
}

# Store results in a data frame
sar.full.15 = data.frame(slope = slope.full.15,
                         r2 = r2.full.15, p = p.full.15, n.isl = 15)


##############################
############# 10 #############
##############################

# Generate and save 100 random numbers
# sample(1:.Machine$integer.max, 100) # This command generated the following values
seed.full.10 = c(1836448795, 1602227588, 1258488269, 566114368, 1776020126,
    894391246, 939975163, 83309568, 507790578, 1992867076, 71644900, 1168011998,
```

```
    2207129, 617292461, 1404656791, 1491500252, 283127481, 1486520641,
    4151586, 350243671, 1600104000, 636115113, 2087365855, 944554374, 613956427,
    195147222, 1653167479, 203780141, 1135993341, 429403506, 1752383166,
    1965071688, 1334116186, 1253729813, 773317255, 2101729052, 2093757280,
    1932699623, 128678624, 1034608016, 2073892305, 121390581, 1391591900,
    747430867, 906679344, 435651795, 1774937385, 2029771941, 59277239,
    2128512526, 2008467108, 192065477, 853393523, 960774878, 1438757988,
    1804104337, 1527600430, 108492762, 1115622710, 289274517, 179332826,
    1910766374, 626250152, 1642845096, 502806666, 1562003476, 1781175387,
    809230528, 490314009, 1402021001, 368756102, 732395523, 84496965, 352270870,
    375811602, 1971730604, 146631313, 2060064172, 1198951728, 1563389336,
    1726472535, 1674206838, 475363694, 1517318119, 1865912176, 303053240,
    1361078265, 267293077, 171185485, 44395409, 143330223, 921602179, 2060765949,
    92596341, 2131167233, 683355650, 728572135, 987009280, 389601379, 1214888602)

# Select 10 islands 100 times
sample.full.10 = matrix(nrow = 10, ncol = 100)

for (i in 1:100) {
  set.seed(seed.full.10[i])
  sample.full.10[, i] = data.frame(sar.isl %>% group_by(class) %>% sample_n(size = 2))$id
}

sample.full.10 = as.data.frame(sample.full.10) # convert the matrix into a data frame

# ISAR for each subset of 10 sites (n = 100)
slope.full.10 = r2.full.10 = p.full.10 = c()

for (i in 1:100) {
  slope.full.10[i] = lm(richness ~ log10(area),
                        # Regression slope
                        data = sar.isl[sample.full.10[, i], ])$coefficients[2]
  r2.full.10[i] = summary(lm(richness ~ log10(area),
                            # Model fit (r2)
                            data = sar.isl[sample.full.10[, i], ]))$adj.r.squared
  p.full.10[i] = summary(lm(richness ~ log10(area),
                           # p-value
                           data = sar.isl[sample.full.10[, i], ]))$coefficients[8]
}

# Store results in a data frame
sar.full.10 = data.frame(slope = slope.full.10,
                         r2 = r2.full.10, p = p.full.10, n.isl = 10)


###############################
############# 5 ##############
###############################

# Generate and save 100 random numbers
# sample(1:.Machine$integer.max, 100) # This command generated the following values
seed.full.5 = c(156379320, 1690806060, 1629423097, 1414352733, 1676991359,
    976730569, 1863137538, 1725172610, 1245716941, 2144377928, 1892644476,
```

```
     902300213, 744495280, 1713546279, 1774310168, 1851413986, 553780345,
     1405284858, 108015246, 1673947428, 2097155812, 1056236678, 1721697030,
     2049545505, 603117198, 947333476, 2122239900, 1440513547, 1927913765,
     454943762, 290176211, 1153586770, 1180513862, 1453353980, 1040440908,
     1536727835, 941783144, 11760277, 1505622231, 787314459, 1802471460,
     792349783, 608428158, 1074061213, 1162948024, 424745217, 1828967737,
     1108684111, 603935622, 1659622245, 68119010, 125836783, 756516015,
     265540766, 1278239777, 336513225, 1804560728, 2138761109, 478843279,
     951316917, 1179275132, 710946839, 1149044041, 887556158, 462340394,
     406182211, 2039798191, 1659142329, 178380253, 1751210418, 781822530,
     991432089, 1517513217, 252335960, 977309456, 264568238, 1563478621,
     1169537062, 1679809152, 7449048, 708562664, 1081425022, 842892613,
     1489221701, 965928975, 964844400, 503102680, 1179685632, 846189621,
     570467263, 813254137, 209997858, 1093310310, 723541863, 192976009,
     252496991, 1843360057, 493987252, 1851015650, 1690781569)

# Select 5 islands 100 times
sample.full.5 = matrix(nrow = 5, ncol = 100)

for (i in 1:100) {
  set.seed(seed.full.5[i])
  sample.full.5[, i] = data.frame(sar.isl %>% group_by(class) %>% sample_n(size = 1))$id
}

sample.full.5 = as.data.frame(sample.full.5) # convert the matrix into a data frame

# ISAR for each subset of 5 sites (n = 100)
slope.full.5 = r2.full.5 = p.full.5 = c()

for (i in 1:100) {
  slope.full.5[i] = lm(richness ~ log10(area),
                       # Regression slope
                       data = sar.isl[sample.full.5[, i], ])$coefficients[2]
  r2.full.5[i] = summary(lm(richness ~ log10(area),
                            # Model fit (r2)
                            data = sar.isl[sample.full.5[, i], ]))$adj.r.squared
  p.full.5[i] = summary(lm(richness ~ log10(area),
                           # p-value
                           data = sar.isl[sample.full.5[, i], ]))$coefficients[8]
}

# Store results in a data frame
sar.full.5 = data.frame(slope = slope.full.5,
                        r2 = r2.full.5, p = p.full.5, n.isl = 5)


#############################
########### All #############
#############################

# Combine the results
sar.full = rbind(sar.full.25, sar.full.20, sar.full.15, sar.full.10, sar.full.5)
```

```r
# Indicate the direction of the relationship
sar.full$dir = NA
sar.full$dir[sar.full$p > 0.05] = "non-significant"
sar.full$dir[sar.full$p < 0.05 & sar.full$slope > 0] = "positive"
sar.full$dir[sar.full$p < 0.05 & sar.full$slope < 0] = "negative"

# Colour code
sar.full$col = NA
sar.full$col[sar.full$p > 0.05] = "black"
sar.full$col[sar.full$p < 0.05 & sar.full$slope > 0] = "dodgerblue"
sar.full$col[sar.full$p < 0.05 & sar.full$slope < 0] = "#e41a1c"
```

```r
# Function to produce summary statistics (mean and +/- sd)
data_summary = function(z) {
  m <- mean(z)
  ymin <- m-sd(z)
  ymax <- m+sd(z)
  return(c(y = m, ymin = ymin, ymax = ymax))
}
```

Figure 4a

```r
box.full.slope =
  ggplot(aes(x = as.factor(n.isl),
             y = (atan(sar.full$slope) * 180 / pi) /
               (atan(sar.74$coefficients[2]) * 180 / pi)),
         data = sar.full) +

  ggtitle("Full range of island size (0.45-1,699 ha)") +

  labs(x = "",
       y = "Slope deviance",
       colour = expression(bold(Effect))) +

  scale_y_continuous(limits = c(-1.2, 1.2)) +

  scale_colour_manual(values = c("#999999", "dodgerblue")) +

  geom_boxplot(fatten = NULL, outlier.shape = NA, notch = TRUE) +
  geom_jitter(shape = 19, size = 2, aes(colour = sar.full$dir),
              alpha = 0.25, width = 0.25, height = 0) +

  stat_summary(fun.data = data_summary, colour = "#e41a1c", size = 1) +

  theme_bw(base_size = 20) +
    theme(panel.grid = element_blank(),
          panel.border = element_rect(colour = "black"),
          axis.title = element_text(colour = "black", face = "bold"),
          axis.text = element_text(colour = "black"),
          axis.ticks = element_line(colour = "black", size = 0.5)) +
  theme(plot.title = element_text(size = 20, face = "bold", hjust = 0.5)) +
  theme(legend.title = element_text(size = 16),
```

```
        legend.text = element_text(size = 14),
        legend.position = c(0.05, 0.05),
        legend.justification = c(0.05, 0.05),
        legend.background = element_rect(colour = NULL),
        legend.key = element_rect(fill = NA)) +
  guides(colour = guide_legend(override.aes = list(size = 5),
                               reverse = TRUE, order = 1)) +

  annotate("text", x = "5", y = 1.2,
           hjust = 1.2, vjust = 0, size = 6, colour = "dodgerblue",
           label = table(subset(sar.full, sar.full$n.isl == 5)$dir)[2]) +
  annotate("text", x = "5", y = 1.2,
           hjust = -0.2, vjust = 0, size = 6, colour = "black",
           label = table(subset(sar.full, sar.full$n.isl == 5)$dir)[1]) +

  annotate("text", x = "10", y = 1.2,
           hjust = 1.2, vjust = 0, size = 6, colour = "dodgerblue",
           label = table(subset(sar.full, sar.full$n.isl == 10)$dir)[2]) +
  annotate("text", x = "10", y = 1.2,
           hjust = -0.2, vjust = 0, size = 6, colour = "black",
           label = table(subset(sar.full, sar.full$n.isl == 10)$dir)[1]) +

  annotate("text", x = "15", y = 1.2,
           hjust = 0.5, vjust = 0, size = 6, colour = "dodgerblue",
           label = table(subset(sar.full, sar.full$n.isl == 15)$dir)[1]) +

  annotate("text", x = "20", y = 1.2,
           hjust = 0.5, vjust = 0, size = 6, colour = "dodgerblue",
           label = table(subset(sar.full, sar.full$n.isl == 20)$dir)[1]) +

  annotate("text", x = "25", y = 1.2,
           hjust = 0.5, vjust = 0, size = 6, colour = "dodgerblue",
           label = table(subset(sar.full, sar.full$n.isl == 25)$dir)[1]) +
  annotate("text", x = Inf, y = -Inf,
           hjust = 1.3, vjust = -1, fontface = "bold", size = 10, label = "(a)")

#box.full.slope
```

Figure 4c

```
box.full.r2 =
  ggplot(aes(x = as.factor(n.isl),
             y = r2),
         data = sar.full) +

  ggtitle("") +

  labs(x = "Number of islands",
       y = (expression(bolditalic(r)^{2}*""[adj]*""))) +
  scale_y_continuous(limits = c(-0.35, 1),
                     breaks = seq(-0.2, 1, 0.2)) +
```

```r
  geom_boxplot(fatten = NULL, outlier.shape = NA, notch = TRUE) +
  geom_jitter(shape = 19, size = 2, colour = sar.full$col,
              alpha = 0.25, width = 0.25, height = 0) +

  stat_summary(fun.data = data_summary, colour = "#e41a1c", size = 1) +

  theme_bw(base_size = 20) +
    theme(panel.grid = element_blank(),
          panel.border = element_rect(colour = "black"),
          axis.title = element_text(colour = "black", face = "bold"),
          axis.text = element_text(colour = "black"),
          axis.ticks = element_line(colour = "black", size = 0.5)) +
  theme(plot.title = element_text(hjust = 0.5)) +

  annotate("text", x = Inf, y = -Inf,
           hjust = 1.3, vjust = -1, fontface = "bold", size = 10, label = "(c)")

#box.full.r2
```

Non-stratified random selection (short range)

```r
################################
############# 25 #############
################################

# Generate and save 100 random numbers
# sample(1:.Machine$integer.max, 100) # This command generated the following values
seed.short.25 = c(546470929, 1854569192, 472457862, 1519942303, 895556555,
    1919452227, 1930033853, 2063048861, 895797531, 671426628, 689697640,
    293680556, 1933282389, 1821079978, 1498918864, 1394247354, 153968628,
    271206201, 92592477, 2081273213, 188826510, 698750039, 947001660, 1773674421,
    1410373373, 2104025889, 1143083811, 1033814086, 1769894669, 872329229,
    1163073271, 1148528459, 560434502, 1224909623, 1031145262, 582656922,
    308392546, 67389807, 263459132, 391259468, 1576323070, 1212875871,
    2070690239, 633301053, 1672178552, 2073513731, 56937991, 1549026011,
    1198479649, 525813008, 151116584, 197262112, 1051865805, 1207593570,
    452395609, 728516306, 366933408, 1457426627, 1353714277, 2100197408,
    1818641172, 1132032800, 2046009994, 1110071781, 623413129, 985610257,
    344982842, 1741372437, 213322852, 132985715, 319609380, 1389669427,
    1940295069, 1497755854, 758207049, 1601760415, 2143143284, 1661429508,
    1548933504, 1475819423, 955483790, 3214408, 622301579, 235955175, 741148058,
    637739887, 1904814664, 330630631, 687650001, 1881465793, 1563203045,
    730846447, 916417468, 88085882, 1532756237, 1813116892, 1486690846,
    611013103, 21024891, 1152909322)

# Select 25 islands 100 times
sample.short.25 = matrix(nrow = 25, ncol = 100)

for (i in 1:100) {
  set.seed(seed.short.25[i])
  sample.short.25[, i] = sample(subset(sar.isl, sar.isl$area < 100)$id, 25)
}
```

```r
sample.short.25 = as.data.frame(sample.short.25) # convert the matrix into a data frame

# ISAR for each subset of 25 sites (n = 100)
slope.short.25 = r2.short.25 = p.short.25 = c()

for (i in 1:100) {
  slope.short.25[i] = lm(richness ~ log10(area),
                         # Regression slope
                         data = sar.isl[sample.short.25[, i], ])$coefficients[2]
  r2.short.25[i] = summary(lm(richness ~ log10(area),
                          # Model fit (r2)
                          data = sar.isl[sample.short.25[, i], ]))$adj.r.squared
  p.short.25[i] = summary(lm(richness ~ log10(area),
                         # p-value
                         data = sar.isl[sample.short.25[, i], ]))$coefficients[8]
}

# Store results in a data frame
sar.short.25 = data.frame(slope = slope.short.25,
                          r2 = r2.short.25, p = p.short.25, n.isl = 25)



###############################
############# 20 #############
###############################

# Generate and save 100 random numbers
# sample(1:.Machine$integer.max, 100) # This command generated the following values
seed.short.20 = c(666307173, 769548065, 537021976, 279055760, 62172685,
    1367337426, 957009295, 1989292791, 1853230961, 1911443969, 1261353721,
    1027415282, 241034004, 161718608, 1275703625, 2111439781, 1700598712,
    1592916179, 43374530, 232151158, 209340666, 1085935855, 285882049,
    2091206380, 1682878078, 859603848, 1697635446, 1635861806, 1022456555,
    2024697679, 2101397481, 1344500639, 1822426816, 1311649218, 543324895,
    2087581645, 526123100, 1680050699, 566775562, 345167024, 1602664518,
    64968048, 1639855157, 1241354938, 1485552218, 1861526072, 612335669,
    1633123308, 1459100985, 1158121107, 580497679, 306604827, 1332928838,
    286243345, 2141373777, 2037212311, 1506619446, 521308150, 1420686872,
    338641856, 260635652, 1252933919, 15446332, 2052072612, 241882135,
    807303067, 393868183, 1032293094, 1926166158, 909554276, 1742475855,
    663146752, 1736829215, 423988712, 180014558, 844018892, 1480232952,
    1405965368, 313990093, 1332559669, 347747447, 436081704, 1530297586,
    447891509, 1992335844, 881663723, 533522389, 1968307086, 719892721,
    87469310, 1778822274, 1736369187, 2082516166, 979182080, 1413199227,
    489656294, 49916018, 1372516049, 1458701386, 302606507)

# Select 20 islands 100 times
sample.short.20 = matrix(nrow = 20, ncol = 100)

for (i in 1:100) {
  set.seed(seed.short.20[i])
  sample.short.20[, i] = sample(subset(sar.isl, sar.isl$area < 100)$id, 20)
}
```

```r
sample.short.20 = as.data.frame(sample.short.20) # convert the matrix into a data frame

# ISAR for each subset of 20 sites (n = 100)
slope.short.20 = r2.short.20 = p.short.20 = c()

for (i in 1:100) {
  slope.short.20[i] = lm(richness ~ log10(area),
                         # Regression slope
                         data = sar.isl[sample.short.20[, i], ])$coefficients[2]
  r2.short.20[i] = summary(lm(richness ~ log10(area),
                              # Model fit (r2)
                              data = sar.isl[sample.short.20[, i], ]))$adj.r.squared
  p.short.20[i] = summary(lm(richness ~ log10(area),
                             # p-value
                             data = sar.isl[sample.short.20[, i], ]))$coefficients[8]
}

# Store results in a data frame
sar.short.20 = data.frame(slope = slope.short.20,
                          r2 = r2.short.20, p = p.short.20, n.isl = 20)



##############################
############# 15 #############
##############################

# Generate and save 100 random numbers
# sample(1:.Machine$integer.max, 100) # This command generated the following values
seed.short.15 = c(1797233886, 1477817566, 1812472833, 282658037, 159644218,
    1058820179, 30092270, 1430848344, 1802764155, 1546478896, 498495623,
    1573434690, 450846291, 1393940053, 653030952, 261493363, 1890241833,
    982501422, 1746246577, 1405205347, 1995988109, 120882561, 1747207129,
    327387968, 2020831209, 1331284056, 1745849496, 1292633655, 203322744,
    41782417, 37050654, 2067894056, 635559297, 479796905, 145114616, 291741248,
    184365873, 71118960, 2124533111, 534668116, 1357261360, 1113033225,
    1044605132, 813411959, 1743315622, 1255166571, 341128038, 781834300,
    233647969, 1726759231, 706151582, 1130525245, 1875316554, 1415337554,
    137326230, 217638676, 1906033574, 2146460365, 466812727, 113915874,
    24889785, 1401542266, 660202450, 314097483, 779164117, 1082807616,
    1390074298, 1416012482, 1012072300, 2025426536, 536217283, 2101306655,
    18189273, 1811934711, 1312938722, 854283991, 1818428581, 505361758,
    240179728, 935751634, 1969290289, 840928766, 781350302, 698273657,
    278682066, 1923288556, 200496280, 8906851, 1251617670, 419489047, 707522070,
    2117554166, 94965522, 1283543567, 1167658624, 160287416, 487054397,
    1478979793, 28832598, 704526487)

# Select 15 islands 100 times
sample.short.15 = matrix(nrow = 15, ncol = 100)

for (i in 1:100) {
  set.seed(seed.short.15[i])
  sample.short.15[, i] = sample(subset(sar.isl, sar.isl$area < 100)$id, 15)
}
```

```r
sample.short.15 = as.data.frame(sample.short.15) # convert the matrix into a data frame

# ISAR for each subset of 15 sites (n = 100)
slope.short.15 = r2.short.15 = p.short.15 = c()

for (i in 1:100) {
  slope.short.15[i] = lm(richness ~ log10(area),
                         # Regression slope
                         data = sar.isl[sample.short.15[, i], ])$coefficients[2]
  r2.short.15[i] = summary(lm(richness ~ log10(area),
                         # Model fit (r2)
                         data = sar.isl[sample.short.15[, i], ]))$adj.r.squared
  p.short.15[i] = summary(lm(richness ~ log10(area),
                         # p-value
                         data = sar.isl[sample.short.15[, i], ]))$coefficients[8]
}

# Store results in a data frame
sar.short.15 = data.frame(slope = slope.short.15,
                          r2 = r2.short.15, p = p.short.15, n.isl = 15)



##############################
############# 10 #############
##############################

# Generate and save 100 random numbers
# sample(1:.Machine$integer.max, 100) # This command generated the following values
seed.short.10 = c(1490864286, 2047178273, 1728145608, 281818095, 2087499922,
    564366804, 224610476, 1411050063, 281673670, 49572642, 1132043984,
    555201282, 716813516, 1280508556, 1589730837, 1534601435, 146351227,
    514405340, 1698506167, 1554576283, 318444025, 2093255895, 783318015,
    2013258027, 168402407, 576605944, 657233474, 1584910838, 1318819952,
    1820689937, 1249616156, 904786729, 1332431854, 812038017, 1885297493,
    1187631930, 928731614, 1051752178, 1138261024, 108586588, 1343434796,
    839610588, 2063078918, 1104421074, 759849601, 1713214309, 1858689334,
    450407697, 313256126, 1457307133, 1009155087, 505753525, 335191415,
    848433713, 1526577160, 2093789644, 386663266, 684043014, 2010058334,
    767706419, 1800026661, 735429860, 552159751, 1774679058, 2071490816,
    114035331, 1362076937, 452982224, 88437574, 824344601, 1771460610,
    85404261, 1045361323, 133918837, 1955817682, 1861748225, 1024056941,
    1068930095, 855334022, 618819368, 282124553, 131763589, 1642882943,
    1585028290, 1376992579, 133003909, 1692076196, 258150875, 791640343,
    392187150, 828995879, 761964529, 1665508994, 720215031, 1482990698,
    774879948, 1187796754, 288976470, 925947713, 1968126902)

# Select 10 islands 100 times
sample.short.10 = matrix(nrow = 10, ncol = 100)

for (i in 1:100) {
  set.seed(seed.short.10[i])
  sample.short.10[, i] = sample(subset(sar.isl, sar.isl$area < 100)$id, 10)
}
```

```r
sample.short.10 = as.data.frame(sample.short.10) # convert the matrix into a data frame

# ISAR for each subset of 10 sites (n = 100)
slope.short.10 = r2.short.10 = p.short.10 = c()

for (i in 1:100) {
  slope.short.10[i] = lm(richness ~ log10(area),
                         # Regression slope
                         data = sar.isl[sample.short.10[, i], ])$coefficients[2]
  r2.short.10[i] = summary(lm(richness ~ log10(area),
                          # Model fit (r2)
                          data = sar.isl[sample.short.10[, i], ]))$adj.r.squared
  p.short.10[i] = summary(lm(richness ~ log10(area),
                         # p-value
                         data = sar.isl[sample.short.10[, i], ]))$coefficients[8]
}

# Store results in a data frame
sar.short.10 = data.frame(slope = slope.short.10,
                          r2 = r2.short.10, p = p.short.10, n.isl = 10)



##############################
############# 5 ##############
##############################

# Generate and save 100 random numbers
# sample(1:.Machine$integer.max, 100) # This command generated the following values
seed.short.5 = c(1818899806, 572588167, 1792177407, 1659406357, 961103864,
    1530683178, 1134007730, 937830870, 1066204421, 1804204256, 503063391,
    1748500366, 474255682, 1214858577, 1348666140, 852326043, 1085653185,
    1412958084, 337382050, 1819034604, 347289743, 1832647990, 185514679,
    1223109694, 1973347031, 1800567392, 552920802, 830448278, 373898856,
    424348329, 235826459, 1805765409, 2046082354, 746323328, 1748767641,
    1189739601, 923159733, 1850158210, 1356612017, 1053329651, 1474626586,
    1300879879, 2122733924, 2083408007, 536239387, 706011796, 1637424859,
    657349614, 11260282, 17681210, 342174539, 1576634915, 1796853954, 1786492628,
    495993392, 896512922, 676915201, 532818466, 826704236, 662021449, 954280518,
    262499102, 1067107507, 318192681, 674273262, 1473103292, 481573782,
    2077582954, 1927740882, 801457600, 1996286074, 1030628649, 408608466,
    706757404, 927495176, 296551639, 875969451, 725602398, 265102155, 514037893,
    948860582, 758135234, 1709858334, 1720550012, 847904281, 842872595,
    620195336, 1792612878, 286096883, 490370245, 2094825162, 2078260721,
    1757203739, 1944745230, 1532469947, 1698773672, 753595436, 437582555,
    1745664625, 867798478)

# Select 5 islands 100 times
sample.short.5 = matrix(nrow = 5, ncol = 100)

for (i in 1:100) {
  set.seed(seed.short.5[i])
  sample.short.5[, i] = sample(subset(sar.isl, sar.isl$area < 100)$id, 5)
}
```

```r
sample.short.5 = as.data.frame(sample.short.5) # convert the matrix into a data frame

# ISAR for each subset of 5 sites (n = 100)
slope.short.5 = r2.short.5 = p.short.5 = c()

for (i in 1:100) {
  slope.short.5[i] = lm(richness ~ log10(area),
                        # Regression slope
                        data = sar.isl[sample.short.5[, i], ])$coefficients[2]
  r2.short.5[i] = summary(lm(richness ~ log10(area),
                             # Model fit (r2)
                             data = sar.isl[sample.short.5[, i], ]))$adj.r.squared
  p.short.5[i] = summary(lm(richness ~ log10(area),
                            # p-value
                            data = sar.isl[sample.short.5[, i], ]))$coefficients[8]
}

# Store results in a data frame
sar.short.5 = data.frame(slope = slope.short.5,
                         r2 = r2.short.5, p = p.short.5, n.isl = 5)



###############################
############ All #############
###############################

# Combine the results
sar.short = rbind(sar.short.25, sar.short.20, sar.short.15, sar.short.10, sar.short.5)

# Indicate the direction of the relationship
sar.short$dir = NA
sar.short$dir[sar.short$p > 0.05] = "non-significant"
sar.short$dir[sar.short$p < 0.05 & sar.short$slope > 0] = "positive"
sar.short$dir[sar.short$p < 0.05 & sar.short$slope < 0] = "negative"

# Colour code
sar.short$col = NA
sar.short$col[sar.short$p > 0.05] = "black"
sar.short$col[sar.short$p < 0.05 & sar.short$slope > 0] = "dodgerblue"
sar.short$col[sar.short$p < 0.05 & sar.short$slope < 0] = "#e41a1c"
```

Figure 4b

```r
box.short.slope =
  ggplot(aes(x = as.factor(n.isl),
             y = (atan(sar.short$slope) * 180 / pi) /
               (atan(sar.74$coefficients[2]) * 180 / pi)),
         data = sar.short) +

  ggtitle("Short range of island size (< 100 ha)") +

  labs(x = "",
```

```r
    y = "") +

scale_y_continuous(limits = c(-1.2, 1.2)) +

geom_boxplot(fatten = NULL, outlier.shape = NA, notch = TRUE) +
geom_jitter(shape = 19, size = 2, colour = sar.short$col,
            alpha = 0.25, width = 0.25, height = 0) +

stat_summary(fun.data = data_summary, colour = "#e41a1c", size = 1) +

theme_bw(base_size = 20) +
  theme(panel.grid = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.title = element_text(colour = "black", face = "bold"),
        axis.text = element_text(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5)) +
theme(plot.title = element_text(size = 20, face = "bold", hjust = 0.5)) +

annotate("text", x = "5", y = 1.2,
         hjust = 1.2, vjust = -0, size = 6, colour = "dodgerblue",
         label = table(subset(sar.short, sar.short$n.isl == 5)$dir)[2]) +
annotate("text", x = "5", y = 1.2,
         hjust = -0.2, vjust = 0, size = 6, colour = "black",
         label = table(subset(sar.short, sar.short$n.isl == 5)$dir)[1]) +

annotate("text", x = "10", y = 1.2,
         hjust = 1.2, vjust = 0, size = 6, colour = "dodgerblue",
         label = table(subset(sar.short, sar.short$n.isl == 10)$dir)[2]) +
annotate("text", x = "10", y = 1.2,
         hjust = -0.2, vjust = 0, size = 6, colour = "black",
         label = table(subset(sar.short, sar.short$n.isl == 10)$dir)[1]) +

annotate("text", x = "15", y = 1.2,
         hjust = 1.2, vjust = 0, size = 6, colour = "dodgerblue",
         label = table(subset(sar.short, sar.short$n.isl == 15)$dir)[2]) +
annotate("text", x = "15", y = 1.2,
         hjust = -0.2, vjust = 0, size = 6, colour = "black",
         label = table(subset(sar.short, sar.short$n.isl == 15)$dir)[1]) +

annotate("text", x = "20", y = 1.2,
         hjust = 1.2, vjust = 0, size = 6, colour = "dodgerblue",
         label = table(subset(sar.short, sar.short$n.isl == 20)$dir)[2]) +
annotate("text", x = "20", y = 1.2,
         hjust = -0.2, vjust = 0, size = 6, colour = "black",
         label = table(subset(sar.short, sar.short$n.isl == 20)$dir)[1]) +

annotate("text", x = "25", y = 1.2,
         hjust = 1.2, vjust = 0, size = 6, colour = "dodgerblue",
         label = table(subset(sar.short, sar.short$n.isl == 25)$dir)[2]) +
annotate("text", x = "25", y = 1.2,
         hjust = -0.2, vjust = 0, size = 6, colour = "black",
         label = table(subset(sar.short, sar.short$n.isl == 25)$dir)[1]) +
annotate("text", x = Inf, y = -Inf,
```

```
           hjust = 1.3, vjust = -1, fontface = "bold", size = 10, label = "(b)")

#box.short.slope
```

Figure 4d

```
box.short.r2 =
  ggplot(aes(x = as.factor(n.isl),
             y = r2),
         data = sar.short) +

  ggtitle("") +

  labs(x = "Number of islands",
       y = "") +

  scale_y_continuous(limits = c(-0.35, 1),
                     breaks = seq(-0.2, 1, 0.2)) +

  geom_boxplot(fatten = NULL, outlier.shape = NA, notch = TRUE) +
  geom_jitter(shape = 19, size = 2, colour = sar.short$col,
              alpha = 0.25, width = 0.25, height = 0) +

  stat_summary(fun.data = data_summary, colour = "#e41a1c", size = 1) +

  theme_bw(base_size = 20) +
    theme(panel.grid = element_blank(),
          panel.border = element_rect(colour = "black"),
          axis.title = element_text(colour = "black", face = "bold"),
          axis.text = element_text(colour = "black"),
          axis.ticks = element_line(colour = "black", size = 0.5)) +
  theme(plot.title = element_text(hjust = 0.5)) +

  annotate("text", x = Inf, y = -Inf,
           hjust = 1.3, vjust = -1, fontface = "bold", size = 10, label = "(d)")

#box.short.r2
```

Combine graphs

Figure 4

```
# Combine and save the graphs
ggsave(grid.arrange(box.full.slope, box.short.slope, box.full.r2, box.short.r2,
                    nrow = 2, ncol = 2),
       filename = "fig4.pdf", width = 17*2, height = 17*2, units = "cm")
```

Literaure review

We carried out a literature review focused on tropical and temperate frog studies worldwide to assess (1) how prevalent positive ISARs are at a global scale, and (2) the role of the number of patches and range in patch size in detecting ISARs.

Our literature review involved four steps as follows:

1. We searched for data (patch size and observed number of frog species) in all studies listed in Table 1 in Almeida-Gomes, Vieira, Rocha, Metzger, & De Coster (2016) and in Table S1 in Palmeirim, Vieira, & Peres (2017).

2. We updated the search carried out by Almeida-Gomes, Vieira, Rocha, Metzger, & De Coster (2016) in *Web of Science* to include additional studies published since their compilation cut-off in 2015. We searched for the same keywords they used – (fragment size* AND amphibia*) OR (fragment size* AND anura*) OR (fragment size* AND frog*) OR (patch size* AND amphibia*) OR (patch size* AND anura*) OR (patch size* AND frog*). These search terms (in *Topic* on 29 August 2018) resulted in 101 hits.

3. We searched for the keywords (("species-area relation*" OR "species-area curve*") AND (amphibia* OR anura* OR frog*) OR (nestedness AND amphibia*) OR (nestedness AND anura*) OR (nestedness AND frog*)) OR (("nested subset*" AND amphibia*) OR ("nested subset*" AND anura*) OR ("nested subset*" AND frog*) OR ("nested species" AND amphibia*) OR ("nested species" AND anura*) OR ("nested species" AND frog*)) in *Topic* using the *Web of Science* database on 29 August 2018 (*Timespan: All years*). This search resulted in 109 hits.

4. We also screened the titles in the reference list of two recent global syntheses on species-area relationships (Matthews, Guilhaumon, Triantis, Borregaard, & Whittaker, 2016) and vertebrate species responses to habitat fragmentation (Keinath et al., 2017).

Island species-area relationships across studies

Despite of being an expected pattern, the positive island species-area relationship (ISAR) has not held true in some studies (for amphibians, see Almeida-Gomes, Vieira, Rocha, Metzger, & De Coster (2016); for reptiles, see Lion, Garda, Santana, & Fonseca (2016)). Different than these authors, who presented a summary of the results (negative, non-significant, positive) reported in the previous studies, we reanalysed data from the literaure using the logarithmic exponential equation (semi-log model) proposed by Gleason (1922).

$$S = z \times log_{10}(A) + c$$

where S = number of species, z = regression slope, A = site area (ha), c = regression intercept.

We used the semi-log model to depict the species-area relationships because it allows the inclusion of sites with $S = 0$.

Import data compiled from the literature

```r
# Study attributes
studies = read.csv("https://ndownloader.figshare.com/files/15158570")
```

```r
# Import data of each study (site area and species richness)
sar.lit = read.csv("https://ndownloader.figshare.com/files/15158567")
```

```r
# Create a list object to store each study separately
sar.list = list()

for (i in 1:length(unique(sar.lit$studyID))) { # number of studies
  sar.list[[i]] = subset(sar.lit, sar.lit$studyID == unique(sar.lit$studyID)[[i]])
}
```

Location of the studies

Figure 5

```r
# Draw the map
map =
  ggplot() +

  labs(x = "Longitude", y = "Latitude") +

  borders("world", colour = "#808080", fill = "#808080") +

  geom_point(data = studies,
             aes(x = longitude, y = latitude),
             shape = 21, colour = "black", fill = "dodgerblue", size = 4)

# Save the map
ggsave(map, file = "fig5.pdf", width = 30, height = 20, units = "cm")

# Display the map
#map
```

ISAR – semi-log models

```r
# Fit the semi-log model to each dataset and save the results
semilog.out = matrix(nrow = length(sar.list), ncol = 9)

for (i in 1:length(sar.list)) {
  semilog.out[i, 1] = unique(sar.list[[i]]$studyID)
  semilog.out[i, 2] = as.character(unique(sar.list[[i]]$reference))
  semilog.out[i, 3] = lm(richness ~ log10(area),
                         data = sar.list[[i]])$coefficients[1] # regression intercept
  semilog.out[i, 4] = lm(richness ~ log10(area),
                         data = sar.list[[i]])$coefficients[2] # regression slope
  semilog.out[i, 5] = summary(lm(richness ~ log10(area),
                                 data = sar.list[[i]]))$adj.r.squared # model fit (r2)
  semilog.out[i, 6] = summary(lm(richness ~ log10(area),
                                 data = sar.list[[i]]))$coefficients[8] # p-value
  semilog.out[i, 7] = nrow(sar.list[[i]])
  semilog.out[i, 8] = max(sar.list[[i]]$area) - min(sar.list[[i]]$area)
  semilog.out[i, 9] = max(sar.list[[i]]$richness)
}

# Finish the data frame
# Name columns
colnames(semilog.out) = c("studyID", "reference", "intercept", "slope", "r2", "p",
                          "n", "range", "smax")
# Convert the matrix "semilog.out" into a data frame
semilog.out = as.data.frame(semilog.out)

# Convert factor data into numeric data
semilog.out$intercept = as.numeric(as.character(semilog.out$intercept))
```

```r
# Convert factor data into numeric data
semilog.out$slope = as.numeric(as.character(semilog.out$slope))
# Convert factor data into numeric data
semilog.out$r2 = as.numeric(as.character(semilog.out$r2))
# Convert factor data into numeric data
semilog.out$p = as.numeric(as.character(semilog.out$p))
# Convert factor data into integer data
semilog.out$n = as.integer(as.character(semilog.out$n))
# Convert factor data into numeric data
semilog.out$range = as.numeric(as.character(semilog.out$range))
# Convert factor data into integer data
semilog.out$smax = as.integer(as.character(semilog.out$smax))

# Indicate the direction of the relationship
semilog.out$dir = NA
semilog.out$dir[semilog.out$p > 0.05] = "non-significant"
semilog.out$dir[semilog.out$p < 0.05 & semilog.out$slope > 0] = "positive"
semilog.out$dir[semilog.out$p < 0.05 & semilog.out$slope < 0] = "negative"
```

Figure 6

```r
# Draw the graph summarising the results across all 25 datasets
# (number of patches vs. range in size)
graph.semilog =
  ggplot(aes(x = range, y = n,
             fill = as.factor(dir), size = r2),
         data = semilog.out) +

  labs(x = "Range in patch size (ha)",
       y = "Number of patches",
       fill = expression(bold(Effect)),
       size = expression(bolditalic(r)^{2}*""[adj])) +

  scale_x_log10(breaks = c(1, 10, 100, 300, 1000, 10000, 100000),
                labels = c("1", "10", "100", "300", "1,000", "10,000", "100,000")) +
  scale_y_continuous(limits = c(5, 25)) +
  annotation_logticks(base = 10, sides = "b") +

  scale_fill_manual(values = c("#999999", "dodgerblue")) +
  scale_size_continuous(range = c(3, 11)) +

  geom_vline(xintercept = 300, size = 0.5, alpha = 0.2, linetype = "dashed") +
  geom_hline(yintercept = 15, size = 0.5, alpha = 0.2, linetype = "dashed") +
  geom_point(shape = 21, colour = "black") +

  theme_bw(base_size = 20) +
  theme(panel.grid = element_blank(),
        panel.border = element_rect(colour = "black"),
        axis.title = element_text(colour = "black", face = "bold"),
        axis.text = element_text(colour = "black"),
        axis.ticks = element_line(colour = "black", size = 0.5)) +
  theme(legend.title = element_text(size = 16),
```

```
        legend.text = element_text(size = 14),
        legend.justification = "top",
        legend.background = element_rect(colour = NULL),
        legend.key = element_rect(fill = NA)) +
  guides(fill = guide_legend(override.aes = list(size = 5),
                             reverse = TRUE, order = 1)) +

  annotate("text", x = 7.82, y = 25,
           hjust = 0.05, vjust = 0, size = 4.5, label = "Many patches, narrow range") +
  annotate("text", x = 7.82, y = 14,
           hjust = 0.05, vjust = 0, size = 4.5, label = "Few patches, narrow range") +
  annotate("text", x = 400, y = 25,
           hjust = 0, vjust = 0, size = 4.5, label = "Many patches, broad range") +
  annotate("text", x = 400, y = 14,
           hjust = 0, vjust = 0, size = 4.5, label = "Few patches, broad range")

# Save the graph
ggsave(graph.semilog, file = "fig6.pdf", width = 25, height = 20, units = "cm")

# Display the graph
#graph.semilog
```

Chapter 6

General discussion

Habitat fragmentation is a long tangled tale (Fahrig 2019). A topic-based search for all articles and reviews in Web of Science (1945-2018) using the terms "habitat fragmentation" OR "forest fragmentation" resulted in 11,373 hits. Despite this impressive number of publications and a growing interest in the topic over the years (Fig. 6.1), a consensus on the impacts of habitat fragmentation on biodiversity is yet to be reached (Fahrig 2017; Fletcher et al. 2018; Fahrig et al. 2019; Miller-Rushing et al. 2019). Below, I show how the findings of the four data chapters that form the body of this thesis shed light on habitat fragmentation research, particularly in the tropics.

A primary reason for such a lack of consensus in the direction (positive, neutral or negative) of habitat fragmentation effects on biodiversity rests on the fact that habitat fragmentation leads to two distinct but interrelated processes: habitat loss (any reduction in available habitat) and habitat fragmentation *per se* (the subdivision of habitat or the creation of a number of smaller and more isolated habitat patches). While habitat loss is both a patch- and landscape-scale process – and undoubted a major threat to biodiversity –, habitat fragmentation *per se* is, by definition, a landscape-scale process (Fahrig 2003). Therefore, when the number of species in variable-sized habitat patches (i.e. the species-area relationship) is invoked to assess how habitat fragmentation affects biodiversity (Matthews et al. 2014), only the effects of patch-scale habitat loss are actually assessed. Moreover, habitat fragmentation has been used as a broad umbrella term for many patterns and processes that accompany landscape change, such as habitat patch shrinkage and increasing isolation (Fletcher et al. 2018), bringing about widespread usage of vague and/or ambiguous terms (Lindenmayer & Fischer 2007).
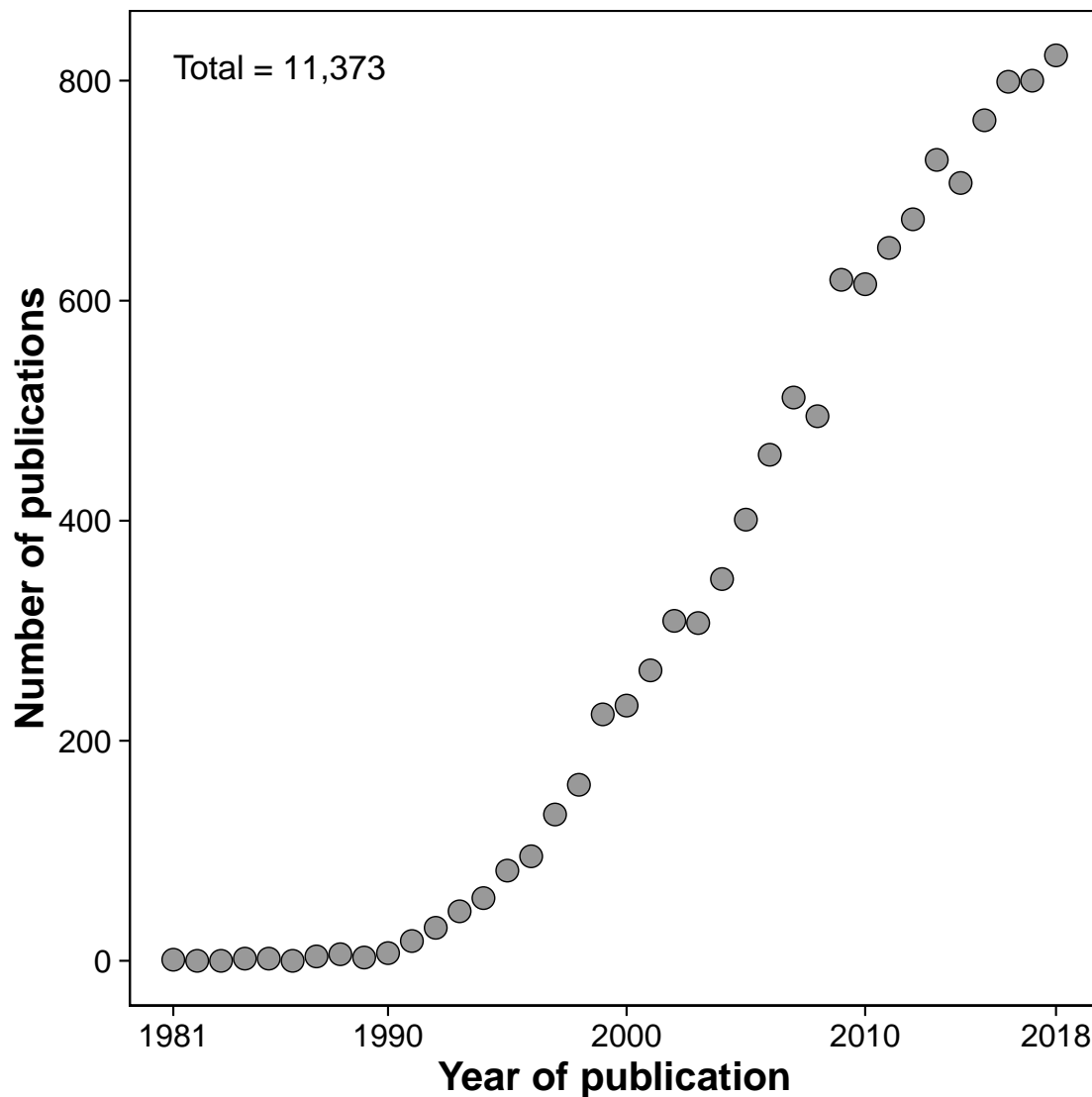
Figure 6.1: Number of articles and reviews on "habitat fragmentation" or "forest fragmentation" published from 1981 to 2018 based on a topic search in Web of Science.

In Chapter 2, the effect of habitat fragmentation induced by a hydroelectric dam (i.e. forest insularization) on community-level avifauna was assessed both at the patch and landscape scales. For the patch-scale analysis, I applied species-area relationships (SARs) and, as expected, found that the larger the forest island, the larger the number of bird species therein (Fig. 2.2). Thus, patch-scale habitat loss impaired the full complement of avifaunal species inhabiting forest islands. For the landscape-scale analysis, I first extrapolated the number of species to a hypothetical island with the combined area of all surveyed islands based on those SARs. Next, I compared the extrapolated value (alpha diversity) with the gamma diversity inventoried across the entire set of surveyed islands (Yaacobi et al. 2007). Contrary to *my* expectation, I did not find any apparent effect of habitat

fragmentation *per se* on the number of bird species at the landscape scale, since the alpha diversity extrapolated to the hypothetical largest island was statistically the same as the gamma diversity (Fig. 2.2). In other words, neither several small islands nor a single large island could be deemed as the best conservation strategy to safeguard the overall insular avifauna. Collectively, this means that patch-scale patterns cannot be extrapolated to whole landscapes (Fahrig 2019) or what may hold true within habitat patches does not necessarily hold true in the entire landscape (e.g. species losses).

The lack of convergence between the effects of patch-scale habitat loss and habitat fragmentation *per se* on species richness is due to the nested subset pattern of species composition. Consider two identical forest archipelagos (A and B), where the only difference is the species identities across islands. In archipelago A, the islands are perfectly nested such that all species on the largest island are also present on the smaller islands. In archipelago B, the islands are perfectly anti-nested such that no species is found on more than one island. Thus, alpha diversity of each island is the same in both A and B, but gamma diversity is much higher in B, resulting in a negative impact of forest fragmentation *per se* in archipelago A but a positive impact in archipelago B. In the case of the Tucuruí Hydroelectric Reservoir landscape, the avifauna was neither nested nor anti-nested, which explains the neutral effect of forest fragmentation on species richness, despite the clearly negative impact of patch-level habitat loss (Fig. 2.2).

Although difficult to quantify as non-significant results do not necessarily imply a lack of effect, neutral effects of habitat fragmentation *per se* on biodiversity are arguably more common than both positive and negative effects combined (Fahrig 2017). If this is the case, the amount of remaining habitat plays a pivotal role in explaining species distribution patterns in fragmented landscapes, relegating the spatial arrangement (i.e. size and isolation) of habitat remnants to a minor role. This idea led to the development of the habitat amount hypothesis (HAH; Fahrig 2013), which challenges the application of island biogeography theory (IBT; MacArthur & Wilson 1963, 1967) to terrestrial systems of habitat patches (see Haila 2002; Laurance 2008). According to the IBT, size and isolation (i.e. distance from the mainland source of immigrants) of oceanic islands are the two predictor variables of species richness therein, so islands behave as discrete spatial units. Conversely, the HAH posits that habitat patches are *not* discrete spatial units, and the habitat surrounding any given patch is the primary source of immigrants (Fahrig 2013). As a result, the HAH replaces two predictor variables, patch size and isolation, with a single predictor variable, habitat amount across the landscape.

In Chapter 3, bearing in mind the contrast between discrete (oceanic islands) and semi-contiguous spatial units (habitat patches) as the extremes of a *continuum*, I explored the degree to which either IBT or HAH were the most appropriate frameworks to explain patterns of species richness on forest islands within one of the world's largest hydroelectric reservoirs. In her publication on the HAH, Fahrig (2013) stated: "Although I focus here exclusively on habitat patches, the ideas I present may also apply to [...] islands within lakes". However, that speculation was yet to be confirmed.

I proposed a conceptual framework (Fig. 3.7) to test the hypothesis that the gradient of landscape functional connectivity – a combination of matrix permeability and species dispersal ability – would determine whether either IBT or HAH, or both, provide the most appropriate framework to assess the effects of habitat fragmentation on biodiversity (Fig. 3.7). Using primary (fieldwork) and secondary data (literature), I corroborated my hypothesis. Accordingly, increasing support for IBT should be expected for a species assemblage with low dispersal ability in patches surrounded by an impermeable matrix (Fig. 3.7c). Conversely, increasing support for HAH should be expected for a species assemblage in which high dispersal ability is prevalent, and habitat patches are surrounded by a permeable matrix (Fig. 3.7b). Under the intermediate scenarios, the most appropriate theoretical framework would depend on the relative contributions of matrix permeability and species dispersal ability. For instance, support for HAH should be expected if species successfully move among patches even if they are surrounded by a relatively impermeable matrix (Fig. 3.7a), whereas support for IBT should be expected if species fail to move among patches even if they are surrounded by a relatively permeable matrix (Fig. 3.7d). Thus, the most appropriate worldview in fragmentation ecology (IBT or HAH) is not only context-dependent but also dynamic as various processes leading to matrix habitat amelioration or degradation come about.

To accurately assess the impacts of habitat fragmentation – or any other type of anthropogenic stressor – on biodiversity, one should determine both the direction and magnitude of such impacts. Because 'impact' is essentially a relative measure, it is imperative to know the condition of ecosystems before significant human disruption occurs, which implies a comparison between control and impacted sites. This can be done employing three study designs: before-after, control-impact, and before-after-control-impact (Green 1979). In habitat fragmentation research, it is most common to employ a control-impact study design using the largest available habitat patch as a control site ("unfragmented landscape") to compare with smaller habitat patches ("fragmented landscape") (e.g. Daily et al. 2001; Uehara-Prado et al. 2007). However, the size of control sites can be rather

variable. For example, while Daily et al. (2001) used a 227-ha forest patch as a control site, Uehara-Prado et al. (2007) used a forest block larger than 10,000 ha.

In Chapter 4, I investigated the role of the size of control sites in assessing the impact of forest fragmentation on bird species richness. Given that larger control sites resemble pre-disturbance conditions to a larger extent than smaller control sites (Fig. 4.2), I hypothesised that for the same set of impacted sites (forest island smaller than 55 ha), the perceived (estimated) impact of forest fragmentation on bird species richness would be lowered by downsizing the control sites (Fig. 4.3). By corroborating this hypothesis, I provided quantitative evidence of widespread concerns raised by several tropical community ecologists: many of our "large" and "undisturbed" control sites are not anywhere nearly extensive enough to retain intact communities, and their use as control sites underestimates the impacts of anthropogenic stressors on biodiversity (Robinson 1999; Sigel et al. 2006; Escobar et al. 2008; Gardner et al. 2009; Meyer et al. 2015). Likewise, I highlighted the perils of using inappropriate control sites in comparative biodiversity studies, as often seen in the modern ecological literature.

Robust environmental impact assessments also depend on the suitability of the study design: (1) controlling for confounding factors (Ewers & Didham 2006) and (2) a sufficient number of sample units (3) spanning an ample range of the environmental gradient of interest (Kreyling et al. 2018). Not surprisingly, differences in study designs may lead to contrasting results in the literature (Lion et al. 2014; Almeida-Gomes et al. 2016) even for pervasive ecological patterns like the iconic species-area relationship (Rosenzweig 1995).

Despite a plethora of studies on species-area relationships (Matthews et al. 2016), the area *per se* effect on species richness is rarely, if ever, assessed in landscape-scale field studies. The fact that area correlates with the number of both individuals and habitat types masks the area *per se* effect, whenever sampling (Hill et al. 1994) and habitat diversity (Connor & McCoy 2001) effects are not controlled for. Moreover, matrix type (Kennedy et al. 2010), history of disturbance (e.g. clear-cut or burned forests; Stouffer & Bierregaard 1995), time since habitat patch isolation (Jones et al. 2016), and direct human disturbance (e.g. hunting pressure; Canale et al. 2012) all mediate the number of species in original habitat remnants embedded within human-modified landscapes, thereby modulating species-area relationships.

As a vast 'real-world' experimental landscape, the Balbina Hydroelectric Reservoir landscape is a unique setting to examine habitat area *per se* effects (Table 5.1) on species assemblages because (1) it provides over 3,500 replicated forest islands varying widely in size (Benchimol & Peres 2015); (2) all forest islands were created simultaneously *c.* 30 years ago (Fearnside 2016), and were therefore

subjected to an uniform and relatively long relaxation time; (3) the open-water matrix is equally hostile; (4) forest islands span similar elevations, are restricted to upland forest and lack perennial streams, ultimately reducing habitat diversity (Drucker et al. 2008; Schietti et al. 2014); (5) adjacent control sites in undisturbed continuous primary forest are widely available; and the (6) *de facto* protection from any human disturbance (from Brazil's largest Biological Reserve) has been ensured in islands spanning most of the archipelago since the time of reservoir formation.

In Chapter 5, I took advantage of autonomous recording units, which allows one to survey multiple sites at the same time, to obtain a large number of sample units spanning an ample range of the island size gradient. In doing so, I present the most extensive and best replicated fragmentation ecology study ever conducted to date on tropical anurans, in terms of the number of islands surveyed ($n = 74$), the range in island size (0.45-1,699 ha), and overall sampling effort (9,362 1-min recordings inspected aurally and visually). I also compiled 25 datasets from the literature to assess the generality of other field-based findings.

The results indicated that area *per se* has a strong effect on anuran species richness, but that sub-optimal sampling designs may weaken or mask our understanding of area effects (Figs. 5.3, 5.4, 5.6). To avoid such a constraint, at least 10 habitat patches spanning three orders of magnitude in size should be surveyed to yield reliable estimates of area-induced species losses, rendering our results of general interest for those interested in applying species-area relationships to real-world landscape-scale conservation problems, regardless of the taxonomic groups and type of patchy systems.

Altogether, the findings from the four data chapters indicate several key recommendations to habitat fragmentation research:

- Habitat fragmentation is a landscape-scale process, and its effects on biodiversity should be assessed at the scale of entire landscapes.

- The most appropriate guiding framework for biodiversity studies in fragmented landscapes depends on the functional connectivity among habitat remnants for the group of organisms of interest.

- Environmental impact assessments should focus on species assemblages derived from suitable control sites, if they are to be defined as reliable.

- Sub-optimal sampling designs jeopardise inferences on area-driven species losses.

6.1   References

1. Almeida-Gomes M, Vieira MV, Rocha CFD, Metzger JP, Coster G De. 2016. Patch size matters for amphibians in tropical fragmented landscapes. Biological Conservation 195:89–96.

2. Benchimol M, Peres CA. 2015. Edge-mediated compositional and functional decay of tree assemblages in Amazonian forest islands after 26 years of isolation. Journal of Ecology 103:408–420.

3. Canale GR, Peres CA, Guidorizzi CE, Gatto CAF, Kierulff MCM. 2012. Pervasive defaunation of forest remnants in a tropical biodiversity hotspot. PLOS ONE 7:e41671.

4. Connor EF, McCoy ED. 2001. Species-area relationships. Pages 397–411 in S. Levin, editor. Encyclopedia of Biodiversity, Volume 5. Academic Press, San Diego, California, USA.

5. Daily GC, Ehrlich PR, Sánchez-Azofeifa GA. 2001. Countryside biogeography: Use of human-dominated habitats by the avifauna of southern Costa Rica. Ecological Applications 11:1–13.

6. Drucker DP, Costa FRC, Magnusson WE. 2008. How wide is the riparian zone of small streams in tropical forests? A test with terrestrial herbs. Journal of Tropical Ecology 24:65–74.

7. Escobar F, Halffter G, Solís Á, Halffter V, Navarrete D. 2008. Temporal shifts in dung beetle community structure within a protected area of tropical wet forest: A 35-year study and its implications for long-term conservation. Journal of Applied Ecology 45:1584–1592.

8. Ewers RM, Didham RK. 2006. Confounding factors in the detection of species responses to habitat fragmentation. Biological Reviews 81:117–42.

9. Fahrig L. 2013. Rethinking patch size and isolation effects: the habitat amount hypothesis. Journal of Biogeography 40:1649–1663.

10. Fahrig L. 2017. Ecological responses to habitat fragmentation per se. Annual Review of Ecology, Evolution, and Systematics 48:1–23.

11. Fahrig L. 2019. Habitat fragmentation: A long and tangled tale. Global Ecology and Biogeography 28:33–41.

12. Fahrig L et al. 2019. Is habitat fragmentation bad for biodiversity? Biological Conservation 230:179–186.

13. Fearnside PM. 2016. Environmental and social impacts of hydroelectric dams in Brazilian Amazonia: Implications for the aluminum industry. World Development 77:48–65.

14. Fletcher RJ et al. 2018. Is habitat fragmentation good for biodiversity? Biological Conservation 226:9–15.

15. Gardner TA, Barlow J, Chazdon R, Ewers RM, Harvey CA, Peres CA, Sodhi NS. 2009. Prospects for tropical forest biodiversity in a human-modified world. Ecology Letters 12:561–582.

16. Green RH. 1979. Sampling Design and Statistical Methods for Environmental Biologists. John Wiley & Sons, Chichester, UK.

17. Haila Y. 2002. A conceptual genealogy of fragmentation research: From island biogeography to landscape ecology. Ecological Applications 12:321–334.

18. Hill JL, Curran PJ, Foody GM. 1994. The effect of sampling on the species-area curve. Global Ecology and Biogeography Letters 4:97–106.

19. Jones IL, Bunnefeld N, Jump AS, Peres CA, Dent DH. 2016. Extinction debt on reservoir land-bridge islands. Biological Conservation 199:75–83.

20. Kennedy CM, Marra PP, Fagan WF, Neel MC. 2010. Landscape matrix and species traits mediate responses of Neotropical resident birds to forest fragmentation in Jamaica. Ecological Monographs 80:651–669.

21. Kreyling J, Schweiger AH, Bahn M, Ineson P, Migliavacca M, Morel-Journel T, Christiansen JR, Schtickzelle N, Larsen KS. 2018. To replicate, or not to replicate - that is the question: how to tackle nonlinear responses in ecological experiments. Ecology Letters 21:1629–1638.

22. Laurance WF. 2008. Theory meets reality: How habitat fragmentation research has transcended island biogeographic theory. Biological Conservation 141:1731–1744.

23. Lion MB, Garda AA, Fonseca CR. 2014. Split distance: a key landscape metric shaping amphibian populations and communities in forest fragments. Diversity and Distributions 20:1245–1257.

24. MacArthur RH, Wilson EO. 1963. An equilibrium theory of insular zoogeography. Evolution 17:373–387.

25. MacArthur RH, Wilson EO. 1967. The Theory of Island Biogeography. Princeton University Press, Princeton, New Jersey, USA.

26. Matthews TJ, Guilhaumon F, Triantis KA, Borregaard MK, Whittaker RJ. 2016. On the form of species-area relationships in habitat islands and true islands. Global Ecology and Biogeography 25:847–858.

27. Meyer NFV, Esser HJ, Moreno R, van Langevelde F, Liefting Y, Ros Oller D, Vogels CBF, Carver AD, Nielsen CK, Jansen PA. 2015. An assessment of the terrestrial mammal communities in forests of Central Panama, using camera-trap surveys. Journal for Nature Conservation 26:28–35.

28. Miller-Rushing AJ, Primack RB, Devictor V, Corlett RT, Cumming GS, Loyola R, Maas B, Pejchar L. 2019. How does habitat fragmentation affect biodiversity? A controversial question at the core of conservation biology. Biological Conservation 232:271–273.

29. Robinson WD. 1999. Long-term changes in the avifauna of Barro Colorado Island, Panama, a tropical forest isolate. Conservation Biology 13:85–97.

30. Rosenzweig ML. 1995. Species Diversity in Space and Time. Cambridge University Press, Cambridge, UK.

31. Schietti J et al. 2014. Vertical distance from drainage drives floristic composition changes in an Amazonian rainforest. Plant Ecology & Diversity 7:241–253.

32. Sigel BJ, Sherry TW, Young BE. 2006. Avian Community Response to Lowland Tropical Rainforest Isolation: 40 Years of Change at La Selva Biological Station, Costa Rica. Conservation Biology 20:111–121.

33. Stouffer PC, Bierregaard RO. 1995. Use of Amazonian forest fragments by understory insectivorous birds. Ecology 76:2429–2445.

34. Uehara-Prado M, Brown KS, Freitas AVL. 2007. Species richness, composition and abundance of fruit-feeding butterflies in the Brazilian Atlantic Forest: comparison between a fragmented and a continuous landscape. Global Ecology and Biogeography 16:43–54.

35. Yaacobi G, Ziv Y, Rosenzweig ML. 2007. Habitat fragmentation may not matter to species diversity. Proceedings of the Royal Society B: Biological Sciences 274:2409–2412.

Chapter 7

Conclusions

Research is an exercise in the right asking questions and coming up with robust answers. Following up from the body of this doctoral dissertation, I present below the key questions and conclusions from each data chapter (Chapters 2 to 5).

Chapter 2: Ecological traits modulate bird species responses to forest fragmentation in an Amazonian anthropogenic archipelago

1. Do habitat generalists show a less steep decline in species richness as a function of diminishing island size compared to forest specialists?

   Yes, they do. Thus, local extinction rates of forest specialists on land–bridge islands are underestimated if habitat generalists are included in the species pool, because this functional group is less impacted by island area reduction than forest specialists.

2. Does forest fragmentation *per se* exacerbate or reduce the impact of forest loss on species richness for the overall species pool, forest specialists and habitat generalists?

   Forest fragmentation *per se* does not necessarily exacerbate the impact of forest loss on species richness.

3. Which suite of morpho-ecological traits best explains species rates of overall island occupancy within the forest archipelago?

   Rare species, especially those with low natural local abundance, are the most extinction-prone in fragmented landscapes.

4. How divergent are observed and detectability-corrected estimates of island occupancy for rarely detected species?

   Detectability-corrected estimates of island occupancy can be much higher than observed estimates for species with low detection probability, ultimately limiting the use of occupancy models for rare or otherwise elusive species.

- **Take-home message 1:** Forest islands within hydroelectric reservoirs are expected to typically harbour depauperate avian assemblages, mostly consisting of naturally abundant and habitat generalist species.

- **Take-home message 2:** Species occupancy modelling is not a 'magic bullet' that can be used indiscriminately, particularly in community-wide studies. Both research aims and species detectability should be carefully considered before using occupancy modelling.

Chapter 3: Patch-scale biodiversity retention in fragmented landscapes: Reconciling the habitat amount hypothesis with the island biogeography theory

1. Which is the most appropriate theoretical framework – Island Biogeography Theory (IBT) or Habitat Amount Hypothesis (HAH) – to explain patterns of species richness of understorey insectivorous birds on human-induced forest islands?

   The species richness of understorey insectivorous birds on forest islands is best explained by an island effect, which is consistent with the IBT. In other words, island size was a stronger predictor of species richness than habitat amount surrounding sample sites.

- **Take-home message:** The most appropriate theoretical framework – IBT or HAH – to explain patterns of species richness in fragmented landscapes depends on the degree of matrix permeability and species dispersal ability. Accordingly, low functional connectivity approximates to IBT, whereas high functional connectivity approximates to HAH.

Chapter 4: The role of baseline suitability in assessing the impacts of land-use change on biodiversity

1. Does the baseline intactness change the estimated impact of forest fragmentation on species richness?

   The higher the baseline intactness, the higher the estimated impact. This effectively means that the use of degraded baselines in environmental impact assessments may lead to over-optimistic estimates of the original impact, thereby masking the true rate of species loss induced by forest fragmentation.

2. Is the estimated impact of forest fragmentation on species richness lower for the overall species assemblage compared to that for the baseline species assemblage?

   Yes, it is. The use of the overall species assemblage further masks estimates of the most severe impacts of forest fragmentation on species, particularly those of higher conservation concern.

3. How different are the solutions for the minimum set problem targeting either the overall or the baseline species assemblage?

   Allocation of conservation investments (i.e. the number of sites to be protected) could be considerably reduced if one targets only those species that presumably occupied any given site prior to land-use change.

- **Take-home message:** Environmental impact assessments and conservation strategies can be severely biased by both the intactness of baseline sites and which set of focal species are considered, both of which can lead to severe underestimates of the effect size of the impact being assessed.

Chapter 5: Sampling design may obscure species-area relationships in landscape-scale field studies

1. What is the role of habitat area *per se* in explaining anuran species richness on Amazonian forest islands induced by river damming?

   Habitat area *per se* plays a decisive role in explaining anuran species richness on Amazonian forest islands within a hydroelectric reservoir.

2. How does sampling design affect the inferential power of island species-area relationships (ISAR) in the 'real world'?

   The fit and the slope of ISARs were affected to a large extent by the range in island size and to a much lesser extent by the number of islands surveyed. Remarkably, the lower the range in island size, the lower both the fit and the slope of ISARs.

3. What would be a suitable sampling design to depict ISARs?

   It is recommended that at least ten habitat patches spanning three orders of magnitude in size should be surveyed before reliable estimates of area-driven species losses in patchy landscapes can be derived.

- **Take-home message:** Habitat area plays a pivotal role well beyond that of sampling design and habitat diversity effects in explaining frog species richness on tropical forest islands. However, the inferential power of island species-area relationships is only reliable if derived from a suitable sampling design in the first place.

That's all folks!